

Clonal variations in nutritional components of *Pinus koraiensis* seeds collected from seed orchards in Northeastern China

Zhen Zhang^{1,2} · Hanguo Zhang² · Chuanping Yang² · Lei Zhang² · Jia Du³ · Ying Jiang²

Received: 25 September 2014 / Accepted: 3 December 2014 / Published online: 24 October 2015
© Northeast Forestry University and Springer-Verlag Berlin Heidelberg 2015

Abstract From four Korean pine (*Pinus koraiensis*) orchards, 60 clones were selected and analyzed for the fatty acid and amino acid components of the seeds to reveal the variations and correlations of the seed characteristics among the clonal source orchards and clones. The nutritional components of the seeds of the *P. koraiensis* trees exhibited rich genetic variation; the variation coefficient of the fatty acids was 2.24–66.83 %, while the variation coefficient of the amino acids was 14.70–38.88 %. Relatively high genetic-improvement potential exists for the nutritional components of the seeds. The phenotypic differentiation of the fatty acid and amino acid components reveals that variation within the population (85.18 %) was the primary source for the variation of the fatty acid

components; variation among the orchards (63.08 %) was the primary source of the variation of the amino acid components. Data drawn from various clonal source orchards all showed that the seed characteristics were highly controlled by heritability ($h^2 > 80$ %), and the seed characteristics of the *P. koraiensis* trees exhibited a similar genetic gain trend. The principal components were analyzed to obtain the comprehensive principal component value for each clonal seed orchard. Twelve clones were selected based on a clonal selection rate of 20 %. Correlation and multiple stepwise-regression analyses were conducted, considering different location conditions, to reveal the stable correlations between the seed characteristics to facilitate improvements of the seed yield of *P. koraiensis* trees and the clonal selection. Species of real characteristics in *P. koraiensis* were controlled by higher heritability. Genetic gain was obtained by selecting of superior clones.

Keywords Fatty acid · Amino acid · Clone · Multi-characteristic selection

Project funding: This work was supported by the Special Research of Public Welfare Forestry Industry for Financial Support for “Korean pine Improved Type Selection and Directive Breeding Technology Research (201204320)” and the Fundamental Research Funds for the Central Universities (2572014AA23).

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

Corresponding editor: Hu Yanbo

✉ Hanguo Zhang
hanguozhang1@sina.com

¹ Research Institute of Subtropical Forestry, Chinese Academy of Forestry, 73, Daqiao Rd., Fuyang 31140, Zhejiang, People’s Republic of China

² School of Forestry, Northeast Forestry University, 26, Hexing Rd., Xiangfang Dist., Harbin 150040, People’s Republic of China

³ School of Municipal and Environmental Engineering, Harbin Institute of Technology, 73, Huanghe Rd., Nangang Dist., Harbin 150090, People’s Republic of China

Introduction

The tree species *Pinus koraiensis*, commonly called Korean pine, is in the genus *Pinus* in the family *Pinaceae*. Korean pine is a grade II national key preserved wild plant of China and is a native tree species in the Northeast China forest region (Lim 2012). *P. koraiensis* is a good source of timber, paper, and pine oil. In addition, *P. koraiensis* nuts are famous for their medical uses and nutritional values. *P. koraiensis* seeds are the most popular among current pine-seed products.

Pine nuts have very high nutritional value (Yoon et al. 1989; Wolff et al. 2000; Nergiz and Donmez 2004),

containing large amounts of crude protein, crude fat, polysaccharide, and crude fiber as well as vitamins, minerals, and trace elements (calcium, phosphorous, manganese, cobalt, copper, and zinc) (Nergiz and Donmez 2004; Zadernowski et al. 2009). The internal quality is a key criterion on which the fruit marketability and practicability of woody plants that are used for edible oil are determined. The content of the nutritional components in pine nuts, such as fat and protein, is equivalent to the fat content in walnuts, which is significantly higher than that in hazelnuts, wild apricots, yellow horns and oil palms. The unsaturated fatty acid components in pine nuts have an enhancing and protecting effect on cytokines. The range of amino acids in pine nuts is relatively complete and the essential amino acid content is abundant (Mensink and Katan 1990; Willett and Ascherio 1994; Savage 2001; Ferramosca et al. 2008).

Studies of the variety and nutritional quality of edible nuts, including walnuts, chestnuts and almonds, have already been conducted (Savage 2001; Nasri et al. 2005; Venkatachalam and Sathe 2006; Sathe et al. 2008; Fernandes et al. 2010; Sharma et al. 2010). *P. koraiensis* has been studied regarding its flowering and fruiting, seed orchard management and nutritional value; however, there have been very few reports dedicated to the assessment of *P. koraiensis*'s germplasm resources or comprehensive evaluations of the nutritional components of *P. koraiensis* nuts among different clones (Kang and Lindgren 1998; Mutke et al. 2005). Therefore, in the present study, we elucidate the nutritional values of *P. koraiensis* nuts through the analysis of the nutritional components of pine nuts from different seed orchards; discuss the variations of the nutritional components of *P. koraiensis* among clonal source orchards and among clones within the same population; identify closely related economic indices based on the correlation analysis; select high-quality clones using a multi-character combined selection strategy; provide theories and excellent germplasm resources for the selection of *P. koraiensis*; and lay the foundation for the construction of seed orchards and advanced-generation seed orchards.

Materials and methods

Materials

Experimental materials were collected from four *P. koraiensis* seed orchards, which were located in Weihe (WH) Township in Shangzhi City, Linkou (LK) County in Mudanjiang City, Tieli (TL) City and Hegang (HG) City in Heilongjiang Province, China. Fifteen clones were selected from each seed orchard (Table 1). For each clone, clones from three standard trees were randomly selected. The

clones collected from each tree were treated for seed production and were then used for sample preparation. The clones were collected in the fall of 2011 and were stored in the refrigerator (5 °C) in the experimental forest of the Northeast Forestry University for further use. The locations of the samples are shown in Fig. 1. The maternal plants in all four of the seed orchards were established after grafting in 1979. The parental origins for the seed orchards were the following: parent tree of Hebei (Weihe); parent tree of Wuying (Hegang); parent tree of the local plantation (Qingshan, Linkou); and parent tree of Wuying (Tieli).

Methods

The selected *P. koraiensis* seeds were first weighed following harvesting and then re-weighed after the shells were removed. The pine seed oil was extracted as follows. The pine nuts were ground into a fine powder and extracted using the Soxhlet extraction method at 70 °C for 8 h. This step was repeated three times. The solvent was dried in a rotary drier, and the oil was obtained, which was stored at −20 °C for further use. The extracted oil was treated using the simple methyl esterification method (Wolff et al. 1997) (KOH-CH₃) and was analyzed using an HP6890/HP5973 gas chromatograph-mass spectrometer (GC-MS) (Agilent, USA) under the following conditions. A DB-5 (60 m × 0.25 mm × 0.20 μm) chromatographic column was used, and the temperature of the injection port was 250 °C. The initial temperature was 140 °C, which was maintained for 5 min, then increased to 240 °C at 2.5 °C/min, and then maintained at 240 °C for 5 min: the total analysis time was 50 min. The injection volume was 1.0 μL. Helium (He) was used as the carrier gas, delivered at 1.0 mL/min. The diversion ratio was 20:1. The temperature of the transmission line was 250 °C. Electron ionization (EI) was used as the ionization mode. The temperature of the ion source was 230 °C. The multiplier voltage was 1.4 kV. The

Table 1 60 clone source samples *Pinus koraiensis* at Hegang, Linkou, Tieli, and Weihe in Heilongjiang Province, China

Location	Clone number							
Hegang	1	10	12	14	17	2	21	23
Linkou	11	13	15	16	18	19	20	24
Tieli	1024	1048	1061	1102	1104	1112	1131	1185
Weihe	008	019	028	042	056	057	063	065
Location	Clone number							
Hegang	39	41	43	44	49	6	8	
Linkou	26	27	3	32	6	8	79-36	
Tieli	1194	1209	1271	1357	1383	3083	3101	
Weihe	066	067	071	110	117	152	162	

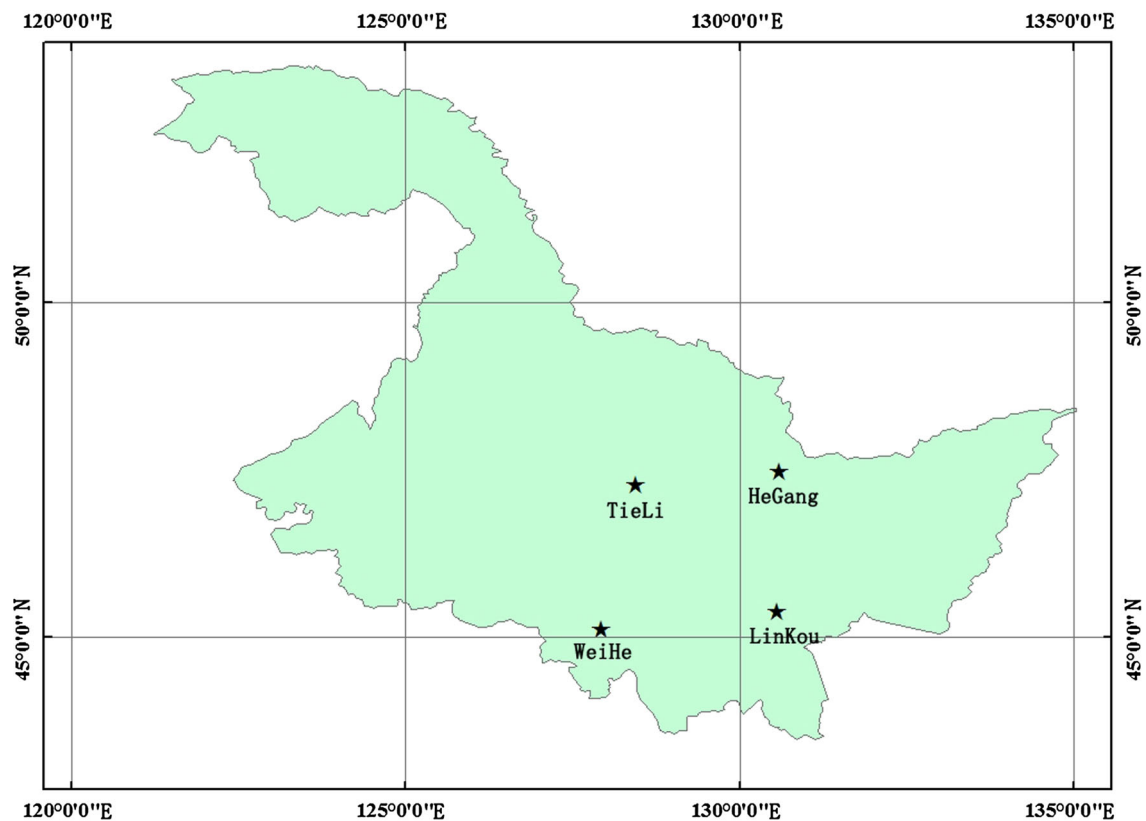


Fig. 1 Distribution map of sample collection of *Pinus koraiensis* at Hegang, Linkou, Tieli, and Weihe in Heilongjiang Province, China

electron energy was 70 eV. The molecular-weight scanning range was 50–550 amu. The scanning time was 0.5 s in the full-scan mode using the spectra library DATABASE/NIST98.L.

The defatted powder of the pine nuts was extracted as follows: 50 mL of water was added to 0.5 g of powder for extraction. Ultrasonic-assisted extraction (60 °C, 500 W) was used for 2 h. After suction filtration, the constant volume of the extract was set at 50 mL. The extract was used to determine the protein content. The Bradford method was used for protein-content determination (Shen et al. 2008), with bovine serum albumin as the standard. The standard curve was $Y = 7.3943x + 0.0288$, where Y represents absorbance ($\lambda = 595$ nm), and the unit of x is mg/mL. The obtained protein amino acid sample from defatting the pine nuts was pre-treated using the hydrochloric acid-hydrolysis method (Bond et al. 2005). A Hitachi L-8800 amino acid analyzer was used for testing.

Data processing and analysis

The ArcGIS software was used to construct the geographic coordinates for the sample collection. Data Processing System 14.01 and Predictive Analytics Software 18.0 were used for the analysis of quantitative traits based on genetic

parameters. The experimental design used was hierarchical classification; the linear model was $X_{ijk} = \mu + S_i + T_{j(i)} + \varepsilon_{ijk}$, where each observed value, X_{ijk} , represents the k th observed value of the j th clone in the i th orchards; μ represents the orchards mean; the treatment effect among the seed orchards (fixed) is $S_i = (\mu_i - \mu)$; the clonal effect within the seed orchard (random) is $T_{j(i)} = (\mu_{ij} - \mu)$; and the random error is $\varepsilon_{ijk} = (X_{ijk} - \mu_{ij})$. The sum of squares of the total variation ($SS_T = \sum_{i=1}^a \sum_{j=1}^m \sum_{k=1}^n (x_{ijk} - \bar{x})^2$) consists of

sum of squares among the seed orchards ($SS_t = mn \sum_{i=1}^a (\bar{x}_i - \bar{x})^2$), the sum of squares among the clones within the orchards ($SS_d = n \sum_{i=1}^a \sum_{j=1}^m (\bar{x}_{ij} - \bar{x}_i)^2$), and the sum of squares of the testing error ($SS_e = \sum_{i=1}^a \sum_{j=1}^m \sum_{k=1}^n (x_{ijk} - \bar{x}_{ij})^2$).

The α represents that there are numbers of seed orchards, m represents there are numbers of m clones in each seed orchard, and n represents there are numbers of observations in each clone.

The phenotypic differentiation coefficient (V_{st}) was used to reflect the phenotypic differentiation among the seed orchards:

$$V_{st} = \delta_{i/s}^2 / (\delta_{i/s}^2 + \delta_s^2),$$

where $\delta_{i/s}^2$ represents the variance component among the seed orchards; δ_s^2 represents the variance component within the seed orchards.

A linear model, $X_{ij} = \mu + S_i + \varepsilon_{ij}$, was used to estimate the repeatability (R) of the clone in each orchard using the following formula: $R = V_C / (V_C + V_E / N)$, where V_C represents the variance of the clone, V_E represents the variance of the environment, and N represents the number of individuals measured for each clone. R was used as the upper limit of the estimate of the heritability (h^2). One standard deviation (σ_A)—by which the clonal mean surpassed the total mean of all the seed orchards—was used as the selection standard to estimate the genetic gain. The following characteristics are critical to model selection: genetic gain, $\Delta G(\%) = (h^2 \times S / \bar{x}) \times 100 \%$, $i = S / \sigma_A$, $\Delta G = (ih^2 \sigma_A / \bar{x}) \times 100 \%$, \bar{x} represents the mean of the characteristic; h^2 represents the clonal heritability; i represents the selection intensity; S represents the selection differential of the characteristic's value; and σ_A represents the standard deviation of the selection characteristic.

A Pearson correlation is used for the methods of correlation analysis. The following equation (Eq. 1) is used for measuring the relevance degree in traits, when data meets the normal distribution.

$$r_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} \quad (1)$$

The meaning of \bar{x} and \bar{y} represent x and y , respectively. Meanwhile, the meaning of x_i and y_i represent x and y . The method of multiple stepwise regression analysis is established by the optimal combination of the multiple independent variable x . The regression equation can predict the regression analysis of dependent variable y . The regression equation with $y = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + \dots + b_nx_n$. b_0 is the absolute term. b_1 , b_2 and b_n are partial regression coefficient of x_1 , x_2 and x_n . The multiple correlation coefficient is r , which represents the linear relationship between the independent variable (x_i) and the dependent variable (y).

To evaluate the information contained in our experimental data, principal component analysis (PCA) was applied. This is a chemometric method to visualize information contained in experimental data and to find the true dimensions of a dataset (Miloun et al. 1992). The variables measured for each sample describe each sample in a variables space. PCA generates a set of new orthogonal variables (linear combinations of the original ones) so that the maximal amount of variance contained in the dataset (information) is concentrated in the first principal

components. The loadings are the coefficients of the original variables defining each principal component.

Results

Variation of the nutritional components of Korean pine seeds

Korean pine nuts are the product of an edible woody oil tree species with oil content as high as 56.76 %. There are a variety of fatty acids in pine nuts. Gas Chromatograph-Mass Spectrometer-computer (GC-MS) was used to analyze the fat content of pine nuts from four different seed orchards (using a total of 60 clones). Table 2 shows that there were 11 types of fatty acids detected, including five saturated fatty acids—myristic acid (C14:0); palmitic acid (C16:0); margaric acid (C17:0); stearic acid (C18:0) and arachidic acid (C20:0)—as well as six unsaturated fatty acids—palmitoleic acid (C16:1 Δ 9c); oleic acid (C18:1 Δ 9c); gadoleic acid (C20:1 Δ 9c); linoleic acid (C18:2 Δ 9c,12c); α -linolenic acid (C18:3 Δ 9c,12c,15c) and eicosadienoic acid (C20:2 Δ 11c,14c). The mean level of saturated fatty acids (\sum SFA) was 12.05 %, and C16:0 had the highest content, with a mean of 7.12 %. The total content of unsaturated fatty acids that are beneficial to humans was 84.67 %. The unsaturated fatty acids present included monounsaturated fatty acids (\sum MUFA) (24.97 %) and polyunsaturated fatty acids (\sum PUFA) (59.70 %). Among the unsaturated fatty acids, linoleic acid (C18:2 Δ 9c, 12c) had the highest content (42.02 %), followed by oleic acid (C18:1 Δ 9c), α -linolenic acid (C18:3 Δ 9c,12c,15c), gadoleic acid (C20:1 Δ 9c), eicosadienoic acid (C20:2 Δ 11c,14c), and palmitoleic acid (C16:1 Δ 9c).

The fat content and fatty acid components of Korean pine nuts from different seed orchards and the multiple comparison between them (Table 2) show that—except for the C14:0, C16:0, C18:3, and C20:2 seeds—all the other indices significantly differed among different seed orchards. The genetic-variation coefficient reflects the genetic-variation capacity of the characteristic. Of the fatty acid components, C14:0 had the highest variation coefficient (66.83 %), and C18:2 had the lowest variation coefficient (2.24 %), indicating that there was significant variation between different fatty acid components. Therefore, it is imperative to collect as many fine tree species as possible and expand breeding orchards as much as possible when constructing advanced-generation seed orchards. The potential of using fat content from the following seeds—C14:0, C17:0, C18:0, C20:0, C16:1, and C20:2—as indices for improvement was relatively high. Stability was the best when using C16:0, C18:1, C20:1, C18:2, C18:3, \sum SFA, \sum MUFA, \sum PUFA, and \sum USFA as the selection standards.

Table 2 Analysis of the fatty acid composition and oil content from different seed orchards of *Pinus koraiensis* (%)

Location	C14:0	C16:0	C17:0	C18:0	C20:0	C16:1	C18:1	C20:1
Hegang	0.0349 (64.56)	6.8824 (7.93)	0.0636 (7.62)d	3.8409 (9.4)c	0.624 (11.27)c	0.13 (23.08)b	24.21 (6.57)a	2.03 (6.89)c
Linkou	0.0447 (51.54)	7.0602 (5.33)	0.0862 (10.55)a	4.153 (8.08)b	0.7416 (9.57)a	0.18 (16.67)a	22.47 (10.01)b	2.17 (8.29)b
Tieli	0.04 (76.13)	7.2698 (6.90)	0.0762 (9.28)b	5.2313 (10.67)a	0.7484 (11.91)a	0.17 (29.41)a	21.54 (8.91)d	2.19 (7.76)a
Weihe	0.0347 (74.46)	7.2491 (7.41)	0.0713 (7.07)c	4.2280 (9.68)b	0.7149 (9.05)b	0.17 (17.65)a	22.42 (6.07)c	2.20 (8.64)a
All populations	0.0386 (66.83)	7.1154 (7.24)	0.0743 (24.05)	4.1133 (10.22)	0.7072 (12.58)	0.1627 (25.15)	22.66 (9.01)	2.1497 (8.47)
Location	C18:2	C18:3	C20:2	\sum SFA	\sum MUFA	\sum PUFA	\sum USFA	Fat content
Hegang	41.80 (1.96)c	16.17 (4.02)	0.95 (30.53)	11.45 (6.36)d	26.37 (5.62)a	58.92 (2.29)c	85.29 (1.02)a	55.28 (10.25)c
Linkou	41.62 (3.03)d	17.13 (12.32)	1.08 (12.04)	12.09 (5.11)c	24.82 (8.56)b	59.83 (2.36)b	84.65 (0.95)b	63.44 (3.89)a
Tieli	42.55 (1.53)a	16.88 (4.74)	1.02 (16.67)	12.37 (6.85)a	23.9 (7.55)c	60.45 (1.92)a	84.36 (1.17)c	56.55 (6.76)b
Weihe	42.13 (1.54)b	16.37 (3.85)	1.11 (12.61)	12.3 (6.54)b	24.79 (4.8)b	59.61 (1.38)b	84.40 (1.26)c	51.77 (9.95)d
All populations	42.022 (2.24)	16.638 (7.61)	1.0422 (19.26)	12.05 (6.9)	24.97 (7.59)	59.7 (2.2)	84.674 (1.18)	56.76 (10.79)

Saturated fatty acids (\sum SFA): C14:0 + C16:0 + C17:0 + C18:0 + C20:0. Monounsaturated fatty acid (\sum MUFA): C16:1 + C18:1 + C20:1. Polyunsaturated fatty acids (\sum PUFA): C18:2 + C18:3 + C20:2. Unsaturated fatty acids (\sum USFA): \sum MUFA + \sum PUFA. The number in the parenthesis represents the variation coefficient. a, b, c and d indicate a statistically significant difference from each other at $p < 0.05$ level

Variation in the level of fatty acids in Korean pine seeds was shown in Table 2. The variation coefficient indicates the dispersing characteristics of the character value, and correspondingly, a larger variation coefficient indicating greater dispersal of the character value. Korean pine seeds have rich inherent genetic variations. In addition, environmental heterogeneity among different seed orchards enhances differences in the seeds. Performing breeding selection based on characteristics with a relatively high variation coefficient from all the seed orchards can provide excellent propagative materials for improving germplasm resources of Korean pine.

Variation of the protein and amino acid components in pine nuts

Amino acids are the component units for protein. We should take the content and ratio of different amino acid components (especially essential amino acids) into consideration in the evaluation of the nutritional values of pine nuts. In the present study, the hydrochloric acid-hydrolysis method was used (tryptophan was damaged during the test; therefore, it was not detected). A total of 17 amino acids

were detected. The mass fraction mean of the total amount of acids (TAA) was 40.433 %. There were seven essential amino acids; the mass fraction mean of essential amino acids (EAA) was 9.79 %. The percentage of essential amino acids relative to total amino acids was 24.21 %. Glutamate (Glu) had the highest content among all the amino acid components of pine nuts, with a mass fraction percentage of 8.46 %. Leucine (Leu) had the highest content among all the essential amino acids, with a mass fraction percentage of 2.7828 %. Except for methionine, there were significant differences among all the other amino acid components from different seed orchards (Table 3). The variation coefficient of proline (Pro) (38.88 %) was the highest among the amino acid components, while the variation coefficient of methionine (Met) (14.70 %) was the lowest. The amino acid components had relatively high potential for genetic improvement.

Comparing the amino-acid components of pine nuts from different seed orchards, the protein content and the variation coefficients of aspartic (Asp) and glycine (Gly) were the highest in the seed orchard in Hegang. The variation coefficients of cysteine (Cys), Pro, Met, and phenylalanine (Phe) were the highest in the seed orchard in

Table 3 Analysis of protein and amino acid composition from different seed orchards of *Pinus koraiensis*

Location	Asp (%)	Ser (%)	Glu (%)	Gly (%)	Ala (%)	Cys (%)	Tyr (%)	His (%)	Arg (%)	Pro (%)
Hegang	3.1588 (16.93)c	2.27 (11.70)b	7.74 (10.89)c	2.2854 (18.08)b	1.932 (16.16)b	0.7349 (11.15)c	1.3229 (11.16)c	0.8173 (8.70)c	5.978 (12.00)c	1.8025 (17.5)c
Linkou	4.6118 (5.92)a	3.014 (8.14)a	10.603 (2.24)a	2.5958 (4.59)a	2.3623 (6.39)a	1.0079 (6.97)a	1.9083 (8.64)a	1.0884 (26.37)a	9.095 (19.77)a	2.988 (23.29)a
Tieli	3.4875 (10.15)b	2.2015 (12.52)c	7.943 (9.08)b	1.8238 (11.99)c	1.7321 (13.37)c	0.8241 (21.32)b	1.4272 (16.58)b	0.9681 (14.58)b	6.6545 (9.98)b	2.034 (42.6)b
Weihe	2.9733 (16.37)d	2.1601 (16.41)c	7.553 (16.90)d	1.8022 (17.93)d	1.6958 (17.67)d	0.7381 (13.89)c	1.2836 (19.44)d	0.7504 (15.29)d	5.726 (17.94)d	1.4913 (16.24)d
AP	3.5578 (21.89)	2.4114 (18.93)	8.460 (18.92)	2.1268 (21.31)	1.9280 (19.73)	0.8263 (20.13)	1.4855 (22.0)	0.9061 (18.70)	6.864 (23.15)	2.0788 (38.88)
Location	Thr (%)	Val (%)	Met (%)	Ile (%)	Leu (%)	Phe (%)	Lys (%)	TAA (%)	EAA (%)	Protein content (mg/g)
Hegang	1.0339 (8.63)d	0.9755 (9.36)d	0.6957 (10.83)	0.7482 (10.16)d	2.3846 (10.56)c	1.2020 (9.77)c	1.3934 (8.06)c	36.476 (10.34)c	8.433 (9.02)c	73.4 (34.12)c
Linkou	1.4655 (10.07)a	1.6637 (12.44)a	0.6830 (21.73)	1.3982 (14.21)a	3.6753 (5.22)a	1.7707 (5.58)a	2.0247 (19.51)a	51.956 (8.48)a	12.681 (9.04)a	96.0 (17.27)a
Tieli	1.1903 (14.52)b	1.1677 (16.37)b	0.6207 (19.06)	1.0552 (17.3)b	2.7472 (12.94)b	1.3728 (18.26)b	1.5727 (9.76)b	38.823 (7.91)b	9.727 (10.2)b	87.7 (22.92)b
Weihe	1.0632 (16.58)c	1.0151 (16.42)c	0.6663 (15.84)	0.8174 (19.50)c	2.3241 (16.53)d	1.1427 (14.53)d	1.2845 (15.13)d	34.477 (16.40)d	8.313 (15.67)d	62.8 (29.22)d
All population	1.1882 (19.13)	1.2055 (26.10)	0.6664 (14.70)	1.0047 (28.69)	2.7828 (22.90)	1.3721 (22.48)	1.5688 (20.67)	40.433 (20.13)	9.789 (20.83)	79.9 (29.82)

The number in the parenthesis represents the variation coefficient %. a, b, c and d indicate a statistically significant difference from each other at $p < 0.05$ level. Essential amino acids (EAA): Thr + Val + Met + Ile + Leu + Phe + Lys

Tieli. The variation coefficients of serine (Ser), Glu, alanine (Ala), tyrosine (Tyr), histidine (His), arginine (Arg), threonine (Thr), valine (Val), isoleucine (Ile), leucine (Leu), lysine (Lys), TAA, and EAA were the highest in the seed orchard in Weihe.

The protein content and all the amino acid components (except for methionine) from the seed orchard in Linkou were higher than the mean levels of the seeds from the other seed orchards, indicating that the breeding population of the seed orchard in Linkou was more suitable for the breeding of protein character-related propagative materials. The result may be associated with the annual precipitation and altitude of local seed orchard. The annual rainfall of linkou seed orchard is 520 mm, which is the lowest in the four locations. In addition, the elevation of linkou seed orchard is 400 m, which is the highest in the four locations. The different environment may effect the content of protein in tree. So future research should study the cause of this phenomenon.

Phenotypic differentiation among the fatty acid components and amino acid components

Nested analysis of variance was used to decompose the phenotypic variation. The percentages of variance

components relative to the total variation among orchards and the variation within the population were calculated. The mean phenotypic variance of the fatty acid components among the orchards was 14.28 % of the total variation (Table 4); the mean phenotypic variance of the fatty acid components within the orchards was 82.86 % of the total variation. The phenotypic differentiation coefficient of the fatty acid components among the orchards was 0.15–37.71 %.

The variation among the orchards (14.82 %) was far smaller than the variation within a given population (85.18 %). The variation within a population was the primary source of the variation of fatty acid components. Table 5 shows that the mean phenotypic variance of the amino acid components among orchards was 62.80 % of the total variation; the mean phenotypic variance of the amino acid components within orchards was 36.73 % of the total variation. The phenotypic variance coefficients of all the amino acid components were higher than 50 %, except for Met, with a relatively low phenotypic variance coefficient (5.10 %). The variation among the orchards (63.08 %) was greater than the variation within the orchards (36.92 %). The variation among the orchards was the primary source of the variation of amino acid components.

Table 4 Variance components composition of fatty acid and phenotypic differentiation coefficient between *Pinus koraiensis* seeds groups

Fatty acid	Mean square (degree of freedom)			variance component (δ^2)			Percentage of variance component (%)			V_{st} (%)
	$MS_{t/s}$ (df = 3)	MS_s (df = 56)	MS_e (df = 120)	$\delta_{t/s}^2$	δ_s^2	δ_e^2	$P_{t/s}$	P_s	P_e	
C14:0	0.00206	0.00201	0.0000006	0.000001	0.00067	0.0000006	0.15	99.76	0.09	0.15
C16:0	1.4854	0.6644	0.0491	0.0182	0.2051	0.0491	6.68	75.30	18.02	8.15
C17:0	0.0041	0.0008	0.000003	0.0001	0.0003	0.000003	20.41	78.72	0.87	20.59
C18:0	1.5432	0.4783	0.00202	0.0237	0.1588	0.0020	12.85	86.06	1.09	12.99
C20:0	0.1479	0.0153	0.0010	0.0029	0.0048	0.0010	33.53	55.38	11.10	37.71
\sum SFA	7.9137	1.6634	0.0581	0.1389	1.6634	0.0581	7.47	89.41	3.12	7.71
C16:1	0.0255	0.0038	0.0001	0.0005	0.0012	0.0001	27.47	67.03	5.49	29.07
C18:1	56.282	10.294	0.0029	1.0220	3.4303	0.0029	22.94	77.00	0.07	22.95
C20:1	0.2797	0.0901	0.0005	0.0042	0.0299	0.0005	12.16	86.48	1.36	12.33
\sum MUFA	47.218	8.938	0.003	0.8507	2.9784	0.003	22.20	77.72	0.08	22.22
C18:2	7.4683	2.4378	0.0014	0.1118	0.8122	0.0014	12.08	87.77	0.15	12.10
C18:3	8.8762	4.6476	0.0024	0.0940	1.5484	0.0024	5.71	94.14	0.15	5.72
C20:2	0.2269	0.1154	0.0006	0.0025	0.1154	0.0006	2.11	97.41	0.48	2.12
\sum PUFA	17.859	4.5681	0.0041	0.2954	1.5213	0.0041	16.22	83.55	0.23	16.26
\sum USFA	8.447	2.7274	0.0065	0.1271	0.907	0.0065	12.21	87.17	0.62	12.29
Mean							14.28	82.86	2.86	14.82

MS mean square, *MS t/s* mean square among the seed orchards, *MS s* mean square within the seed orchard, *MS e* mean square of environmental, *df* degree of freedom, δ^2 variance component, $\delta_{t/s}^2$ variance component among the seed orchards, δ_s^2 variance component within the seed orchard, δ_e^2 variance component of environmental, *P* Percentage of variance component, *P t/s* percentage of variance component among the seed orchards, *P s* percentage of variance component within the seed orchard, *P e* percentage of variance component in environmental

Evaluation of the clones within the population and parameter estimation

A single-factor, randomized linear model was used to analyze the genetic parameters of the source population of each clone. The upper limit of heritability that was estimated from clonal heritability was above 80 % (Table 6), indicating that the differences among the clones were primarily caused by genetic effects. The fatty acid and amino acid components were strongly controlled by genetics, and character selection could yield relative high genetic gains.

The selection standard considered was a difference of one standard deviation by which the clonal mean surpassed the total mean of the population character of all the seed orchards (to analyze unsaturated fatty acid components with the desired economic characteristics and essential amino acid components in pine nuts). The component index selection had the expected genetic gains. Among the fatty acid components, C16:1, C18:1, C20:1, C18:3, C20:2, \sum MUFA, and \sum SFA from all four of the seed orchards achieved over 5 % genetic gains. Among the amino acid components, Thr, Val, Leu, Phe, TAA, and EAA had the most prominent selection effects, and all achieved over 10 % genetic gains.

Among the fatty acid components, C20:2 had the best genetic gains in all four of the seed orchards, with a genetic gain of variable amplitude (ΔG) of 21.56 % to 55.77 %.

Among the amino acid components, Met had the highest genetic gains in the seed orchard in Hegang (18.69 %); the genetic gains of Phe were the greatest in the seed orchards in Linkou (29.5 %) and Tieli (48.36 %); and the genetic gain of Ile was the highest in the seed orchard in Weihe (43.38 %).

Correlation between seed characteristics

Because different orchards have different genetic backgrounds and different location conditions, the clonal orchards were analyzed separately for characteristic indices. The linear relationships between the protein content and oil content of the seeds from the four different locations were investigated. The correlations for the locations were as follows: Pearson_(HG) = 0.604, Pearson_(LK) = 0.226, Pearson_(TL) = 0.227, and Pearson_(WH) = 0.435. It is clear that the correlations were not consistent. Therefore, linear correlation analysis was performed separately on the fatty acid components and amino acid components.

Tables 7 and 8 list the analysis of the correlations among fatty acid components. In all four of the seed orchards, \sum SFA exhibited a positive correlation with C16:0, C18:0, and C20:0, and a negative correlation with C18:1 and \sum MUFA. \sum MUFA exhibited a negative correlation with C20:0, C20:1, C18:3, and \sum PUFA and a positive correlation with C18:1 and \sum USFA. \sum PUFA

Table 5 Amino acid composition variance components and phenotypic differentiation coefficient between *Pinus koraiensis* seeds groups

Amino acid	Mean square (degree of freedom)			variance component (δ^2)			Percentage of variance component (%)			V_{st} (%)
	$MS_{t/s}$ (df = 3)	MS_s (df = 56)	MS_e (df = 120)	$\delta_{t/s}^2$	δ_s^2	δ_e^2	$P_{t/s}$	P_s	P_e	
Asp	24.2498	0.6359	0.0018	0.5248	0.2113	0.0019	71.11	28.64	0.25	71.29
Thr	1.7448	0.0707	0.0004	0.0372	0.0234	0.0004	60.93	38.37	0.70	61.36
Ser	7.3552	0.2685	0.0018	0.1575	0.0889	0.0018	63.45	35.82	0.73	63.92
Glu	93.0101	3.1988	0.003	1.9958	1.0653	0.0030	65.13	34.77	0.10	65.20
Gly	6.6329	0.2980	0.00139	0.1408	0.0989	0.0014	58.41	41.02	0.58	58.75
Ala	4.2848	0.2282	0.0022	0.0901	0.0753	0.0022	53.75	44.92	1.33	54.46
Cys	0.7368	0.0468	0.0010	0.0153	0.0153	0.0010	48.52	48.29	3.20	50.07
Val	4.5074	0.0744	0.0003	0.0985	0.0247	0.0003	79.77	20.00	0.23	79.95
Met	0.0484	0.0280	0.000044	0.0005	0.0093	0.000044	4.62	94.93	0.45	5.10
Ile	3.8734	0.0578	0.0001	0.0848	0.0192	0.0001	81.42	18.47	0.11	81.51
Leu	17.5038	0.359	0.0006	0.3810	0.1195	0.0006	76.03	23.85	0.13	76.12
Tyr	3.74	0.1398	0.0006	0.0800	0.0464	0.0006	63.00	36.52	0.48	63.30
Phe	3.6071	0.1102	0.0003	0.0777	0.0366	0.0003	67.79	31.95	0.26	67.97
Lys	4.7924	0.079	0.0002	0.1047	0.0263	0.0002	79.83	20.01	0.15	79.95
His	1.0380	0.0360	0.0001	0.0223	0.0120	0.0001	64.87	34.90	0.23	65.05
Arg	106.5909	2.3573	0.0001	2.3163	0.7858	0.0001	74.67	25.33	0.00	74.67
Pro	18.7393	1.0846	0.0001	0.3923	0.3615	0.0001	52.04	47.95	0.01	52.04
TAA	2797.72	61.77	0.0129	60.7989	20.5882	0.0129	74.69	25.29	0.02	74.70
EAA	185.759	3.338	0.001	4.0538	1.1123	0.0011	78.45	21.53	0.02	78.47
Mean							62.80	36.73	0.48	63.08

MS mean square, *MS t/s* mean square among the seed orchards, *MS s* mean square within the seed orchard, *MS e* mean square of environmental, *df* degree of freedom, δ^2 variance component, $\delta_{t/s}^2$ variance component among the seed orchards, δ_s^2 variance component within the seed orchard, δ_e^2 variance component of environmental, *P* percentage of variance component, *P t/s* percentage of variance component among the seed orchards, *P s* percentage of variance component within the seed orchard, *P e* Percentage of variance component in environmental

exhibited a positive correlation with C20:1 and C18:3; C18:1 exhibited a negative correlation with C20:0, C20:1, C18:3 and \sum PUFA; and \sum USFA exhibited a negative correlation with C16:0. Therefore, a combined selection of characteristics can be carried out between \sum MUFA and C18:1/ \sum USFA as well as between \sum PUFA and C20:1/C18:3.

Tables 9 and 10 list the correlations among the amino acid components. In the four different seed orchards, TAA exhibited a positive correlation with Thr, Ser, Glu, Gly, Val, Met, Leu, Tyr, Phe, Lys, His, and Arg. EAA exhibited a positive correlation with Thr, Ser, Glu, Gly, Val, Met, Ile, Leu, Tyr, Phe, Lys, Arg, and TAA. Arg exhibited a positive correlation with Ser, Glu, Leu, Tyr, Phe, Lys and His. Thr exhibited a positive correlation with Ser, Val and Tyr. Phe exhibited a positive correlation with Glu, Gly, Leu, and Tyr. Val exhibited a positive correlation with Ser and Val, and Gly exhibited a positive correlation with Ile and Leu. Therefore, combined selection of characteristics can be carried out among TAA, EAA, Thr, Ser, Phe, Arg, and Val.

Under different location conditions, correlations among the fatty acid components and the amino acid components of Korean pine nuts were very stable, indicating that the inherent genetic bases of these characteristics were closely related, and the characteristics exhibited broad adaptability to the heterogeneity of the environment. All the other characteristics performed differently under different location conditions, indicating that different location conditions altered the allele frequencies of specific sites through natural selection, and these sites controlled the adaptability-related characteristics.

Multiple stepwise-regression analysis

Multiple stepwise-regression analysis of fatty acid components on the oil content of seeds We investigated the fatty acid components that affected oil content. The fatty acid components of Korean pine trees from all four of the seed orchards were used as independent variables: (X), X_1 (C14:0), X_2 (C16:0), X_3 (C17:0), X_4 (C18:0), X_5 (C20:0), X_6 (\sum SFA),

Table 6 Genetic parameters and genetic gain in clone population of *Pinus koraiensis*

Composition	Clonal repeatability (%)				Selection intensity				Genetic gain ΔG (%)			
	h_{HG}^2	h_{LK}^2	h_{TL}^2	h_{WH}^2	i_{HG}	i_{LK}	i_{TL}	i_{WH}	ΔG_{HG}	ΔG_{LK}	ΔG_{TL}	ΔG_{WH}
\sum SFA	80.92	98.57	97.91	98.31	1.38	1.89	1.53	3.10	5.64	9.74	10.80	20.78
C16:1	98.04	90.74	96.15	80.17	1.65	1.13	1.46	1.37	36.90	15.27	37.48	21.29
C18:1	97.68	97.98	98.97	97.93	1.14	1.19	1.51	1.39	7.62	12.19	13.78	8.61
C20:1	96.09	98.47	98.62	96.65	1.31	1.86	1.27	1.43	8.42	15.16	10.04	12.52
\sum MUFA	99.71	99.96	99.97	97.92	1.10	1.28	1.55	1.29	6.30	11.24	11.99	6.33
C18:2	99.24	93.64	98.64	94.55	1.40	1.05	1.67	1.78	2.79	3.04	2.59	2.82
C18:3	99.65	96.94	97.82	99.71	1.33	3.38	1.58	1.59	5.42	13.26	7.64	6.26
C20:2	98.35	96.98	97.91	94.93	1.44	1.10	3.32	1.66	44.37	42.67	55.77	21.56
\sum PUFA	97.42	96.90	95.89	95.73	1.78	2.46	1.66	1.46	4.14	5.95	3.25	2.05
\sum USFA	97.89	97.51	99.73	96.71	1.23	1.05	1.31	1.12	1.23	1.01	1.57	1.41
Thr	96.52	93.68	94.48	97.26	1.39	1.39	1.23	1.51	11.91	13.14	17.36	24.81
Val	98.69	95.55	95.36	95.19	1.55	1.38	1.28	2.22	14.88	12.31	20.74	35.57
Met	97.12	97.64	96.57	97.29	1.73	0.68	1.24	1.98	18.69	6.98	23.62	31.08
Ile	93.55	96.81	97.68	93.79	1.22	0.88	1.65	2.35	11.51	6.05	28.28	43.38
Leu	95.88	97.34	93.98	96.77	1.39	2.10	1.67	2.18	14.46	19.98	20.36	35.75
Phe	96.22	95.89	97.66	95.38	1.43	2.72	2.64	2.06	13.40	29.50	48.36	29.41
Lys	95.49	97.54	97.15	95.79	1.06	1.01	1.32	1.97	8.45	8.17	12.80	28.76
TAA	98.88	97.68	95.44	96.58	1.57	1.30	1.58	2.06	16.60	12.03	12.23	33.35
EAA	97.68	98.25	96.28	97.31	1.41	1.27	1.77	2.17	12.77	10.12	17.86	33.74

h_{HG}^2 , clonal repeatability in Hegang seed orchard; h_{LK}^2 , clonal repeatability in Linkou seed orchard; h_{TL}^2 , clonal repeatability in TL seed orchard; h_{WH}^2 , clonal repeatability in Weihe seed orchard; i_{HG} , selection intensity in Hegang seed orchard; i_{LK} , selection intensity in Linkou seed orchard; i_{TL} , selection intensity in Tieli seed orchard; i_{WH} , selection intensity in Weihe seed orchard; ΔG_{HG} , genetic gain in Hegang seed orchard; ΔG_{LK} , genetic gain in Linkou seed orchard; ΔG_{TL} , genetic gain in Tieli seed orchard; ΔG_{WH} , genetic gain in Weihe seed orchard

X_7 (C16:1), X_8 (C18:1), X_9 (C20:1), X_{10} (\sum MUFA), X_{11} (C18:2), X_{12} (C18:3), X_{13} (C20:2), X_{14} (\sum PUFA), X_{15} (\sum USFA), and X_{16} (\sum), while the oil content of the seeds was used as the dependent variable (Y). The dependent variable Y was normally distributed, and therefore, multiple stepwise-regression analysis was performed. After rejecting the independent variables with an insignificant regression coefficient, the best regression equations were as follows:

$$\begin{aligned}
 Y_{HG} &= -0.30 - 2.41X_1 + 4.04X_3 + 0.09X_4 - 2.33X_7 \\
 &\quad + 0.03X_{12} + 0.09X_{13} \quad (R^2 = 0.64) \\
 Y_{LK} &= 94.39 + 71.95X_1 + 39.29X_2 + 45.53X_4 - 36.64X_6 \\
 &\quad - 1.28X_{11} - 4.7X_{13} \quad (R^2 = 0.8549) \\
 Y_{TL} &= -620.48 - 61.42X_1 - 1.05X_4 + 2.5X_{14} + 5.51X_{16} \\
 &\quad (R^2 = 0.5718) \\
 Y_{WH} &= 285.38 - 8.38X_2 - 812.75X_3 - 133.4X_5 + 12.24X_6 \\
 &\quad + 59.4X_9 - 53.8X_{13} - 4.05X_{14} \quad (R^2 = 0.6764)
 \end{aligned}$$

Even though the oil content (Y) was affected by different fatty acid components in the four seed orchards, C14:0(X_1), C18:0 (X_4) and C20:2 (X_{13}) were the most stable and had a linear effect with the oil content (Y) in more than three of the seed orchards.

Multiple stepwise-regression analysis of amino acid components with regard to the protein content

We investigated the important factors of amino acid components that affected the protein content. The amino acid components of the Korean pines from all four of the seed orchards were used as independent variables: (Z), Z_1 (Asp), Z_2 (Thr), Z_3 (Ser), Z_4 (Glu), Z_5 (Gly), Z_6 (Ala), Z_7 (Cys), Z_8 (Val), Z_9 (Met), Z_{10} (Ile), Z_{11} (Leu), Z_{12} (Tyr), Z_{13} (Phe), Z_{14} (Lys), Z_{15} (His), Z_{16} (Arg), Z_{17} (Pro), Z_{18} (TAA), and Z_{19} (EAA). The protein content of the seeds was used as the dependent variable (Y), which was normally distributed. Therefore, multiple stepwise-regression analysis was performed and the best regression equations were as follows:

Table 7 The correlation of fatty acid composition in *Pinus koraiensis* seeds between Hegang and Linkou seed orchards

Fatty acid	C14:0 X ₁	C16:0 X ₂	C17:0 X ₃	C18:0 X ₄	C20:0 X ₅	∑SFA X ₆	C16:1 X ₇	C18:1 X ₈
C14:0 X ₁	1	0.094	0.109	0.249	0.248	0.262	0.254	-0.394
C16:0 X ₂	0.672**	1	0.075	0.12	0.543*	0.745**	0.011	-0.526*
C17:0 X ₃	0.258	0.433	1	0.769**	0.743**	0.593*	-0.323	-0.4
C18:0 X ₄	0.427	0.234	0.366	1	0.811**	0.748**	0.044	-0.636*
C20:0 X ₅	0.416	0.28	0.482	0.97**	1	0.924**	-0.051	-0.846**
∑SFA X ₆	0.685**	0.703**	0.509	0.855**	0.862**	1	0.02	-0.791**
C16:1 X ₇	-0.739**	-0.493	0.167	-0.024	0.057	-0.285	1	-0.04
C18:1 X ₈	-0.241	-0.318	-0.681**	-0.459	-0.557*	-0.513	-0.157	1
C20:1 X ₉	0.033	0.323	0.589*	0.291	0.367	0.384	0.254	-0.777**
∑MUFAX ₁₀	-0.269	-0.32	-0.673**	-0.466	-0.563*	-0.521*	-0.126	0.998**
C18:2 X ₁₁	-0.128	0.043	0.452	0.35	0.408	0.277	0.292	-0.795**
C18:3 X ₁₂	-0.029	-0.123	0.359	0.1	0.197	0.02	0.437	-0.659**
C20:2 X ₁₃	0.016	-0.14	0.451	0.128	0.137	0.027	0.194	-0.198
∑PUFAX ₁₄	-0.088	-0.064	0.545*	0.288	0.373	0.184	0.43	-0.844**
∑USFA X ₁₅	-0.608*	-0.659**	-0.31	-0.356	-0.39	-0.615*	0.46	0.403
∑ X ₁₆	-0.198	-0.247	0.033	0.264	0.227	0.056	0.344	0.082
Fatty acid	C20:1 X ₉	∑MUFA X ₁₀	C18:2 X ₁₁	C18:3 X ₁₂	C20:2 X ₁₃	∑PUFA X ₁₄	∑USFA X ₁₅	∑ X ₁₆
C14:0 X ₁	0.105	-0.404	0.079	0.231	0.247	0.438	-0.302	-0.226
C16:0 X ₂	0.405	-0.523*	-0.385	0.492	0.341	0.424	-0.639*	-0.159
C17:0 X ₃	0.089	-0.42	-0.072	0.272	0.096	0.351	-0.495	-0.095
C18:0 X ₄	0.184	-0.656**	-0.136	0.455	0.337	0.589*	-0.703**	-0.295
C20:0 X ₅	0.507	-0.853**	-0.416	0.733**	0.548*	0.776**	-0.897**	-0.429
∑SFA X ₆	0.41	-0.802**	-0.354	0.648**	0.464	0.695**	-0.904**	-0.315
C16:1 X ₇	0.03	-0.028	-0.121	0.102	0.148	0.057	0.028	0.096
C18:1 X ₈	-0.707**	0.998**	0.454	-0.878**	-0.748**	-0.977**	0.927**	0.719**
C20:1 X ₉	1	-0.664**	-0.666**	0.778**	0.915**	0.654**	-0.610*	-0.660**
∑MUFAX ₁₀	-0.738**	1	0.423	-0.863**	-0.713**	-0.978**	0.930**	0.707**
C18:2 X ₁₁	0.805**	-0.774**	1	-0.806**	-0.609*	-0.371	0.468	0.441
C18:3 X ₁₂	0.432	-0.659**	0.425	1	0.768**	0.848**	-0.795**	-0.67**
C20:2 X ₁₃	0.481	-0.165	0.271	0.155	1	0.696**	-0.666**	-0.693**
∑PUFAX ₁₄	0.802**	-0.823**	0.871**	0.772**	0.456	1	-0.832**	-0.672**
∑USFA X ₁₅	-0.016	0.439	0.031	0.074	0.434	0.149	1	0.691**
∑ X ₁₆	0.3	0.121	0.271	0.111	0.573*	0.343	0.752**	1

Upper right was the correlation of fatty acid composition in Linkou *Pinus koraiensis* seed orchard; Left lower was the correlation of fatty acid composition in Hegang *Pinus koraiensis* pine seed orchard. ∑: ∑SFA + ∑USFA. * indicate a statistically significant correlation at p < 0.05 level; ** indicate a statistically significant correlation at p < 0.01 level

$$\begin{aligned}
 Y_{HG} &= -9.87 - 207.95Z_2 + 100.3Z_3 - 48.21Z_6 - 23.08Z_7 \\
 &\quad + 139.87Z_8 + 40.66Z_9 - 19.14Z_{12} - 25.09Z_{13} \\
 &\quad + 38.03Z_{14} + 64.2Z_{15} - 9.21Z_{16} + 3.734Z_{17} - 6.12Z_{19} \\
 &\quad (R^2 = 0.9996)
 \end{aligned}
 \qquad
 \begin{aligned}
 Y_{LK} &= 12.38 - 2.49Z_2 + 14.03Z_3 + 3.64Z_4 - 10.12Z_5 \\
 &\quad + 3.91Z_6 - 3.59Z_8 + 6.76Z_9 + 7.63Z_{11} \\
 &\quad - 12.11Z_{12} - 11.99Z_{15} - 3Z_{16} - Z_{17} \quad (R^2 = 0.9985)
 \end{aligned}$$

Table 8 The correlation of fatty acid composition in *Pinus koraiensis* seeds between Tieli and Weihe seed orchards

Fatty acid	C14:0 X ₁	C16:0 X ₂	C17:0 X ₃	C18:0 X ₄	C20:0 X ₅	∑SFA X ₆	C16:1 X ₇	C18:1 X ₈
C14:0 X ₁	1	-0.429	-0.415	-0.003	0.039	-0.255	0.123	-0.134
C16:0 X ₂	0.26	1	0.576*	0.288	0.645**	0.850**	0.045	-0.435
C17:0 X ₃	0.201	0.236	1	0.368	0.524*	0.600*	-0.294	-0.197
C18:0 X ₄	-0.112	0.292	0.187	1	0.779**	0.746**	0.18	-0.666**
C20:0 X ₅	-0.077	0.631*	0.36	0.705**	1	0.896**	0.159	-0.693**
∑SFA X ₆	0.124	0.827**	0.295	0.776**	0.855**	1	0.133	-0.677**
C16:1 X ₇	0.284	0.508	0.504	0.038	0.29	0.367	1	-0.475
C18:1 X ₈	-0.276	-0.760**	-0.618*	-0.448	-0.688**	-0.778**	-0.523*	1
C20:1 X ₉	0.047	0.416	0.307	-0.069	0.497	0.268	0.188	-0.563*
∑MUFAX ₁₀	-0.282	-0.756**	-0.616*	-0.482	-0.677**	-0.792**	-0.513	0.997**
C18:2 X ₁₁	0.298	0.172	0.613*	-0.297	-0.156	-0.055	0.339	-0.408
C18:3 X ₁₂	0.269	0.447	0.454	0.457	0.568*	0.582*	0.325	-0.832**
C20:2 X ₁₃	0.285	0.485	0.221	-0.136	0.194	0.248	0.106	-0.465
∑PUFAX ₁₄	0.395	0.476	0.691**	0.128	0.333	0.406	0.431	-0.871**
∑USFA X ₁₅	-0.051	-0.822**	-0.315	-0.729**	-0.846**	-0.969**	-0.432	0.799**
∑ X ₁₆	0.207	-0.411	-0.226	-0.224	-0.408	-0.401	-0.43	0.479

Fatty acid	C20:1 X ₉	∑MUFA X ₁₀	C18:2 X ₁₁	C18:3 X ₁₂	C20:2 X ₁₃	∑PUFA X ₁₄	∑USFA X ₁₅	∑ X ₁₆
C14:0 X ₁	0.071	-0.139	0.165	0.293	0.011	0.36	0.125	-0.207
C16:0 X ₂	0.548*	-0.409	-0.416	-0.306	0.533*	-0.478	-0.840**	-0.507
C17:0 X ₃	0.5	-0.154	-0.082	-0.613*	0.508	-0.451	-0.530*	-0.18
C18:0 X ₄	0.665**	-0.650**	-0.007	0.001	0.667**	0.11	-0.654**	-0.21
C20:0 X ₅	0.843**	-0.654**	-0.308	-0.054	0.698**	-0.168	-0.877**	-0.509
∑SFA X ₆	0.765**	-0.648**	-0.302	-0.203	0.742**	-0.27	-0.951**	-0.49
C16:1 X ₇	0.359	-0.458	0.157	0.398	0.308	0.486	-0.141	-0.104
C18:1 X ₈	-0.838**	0.996**	-0.089	-0.329	-0.845**	-0.47	0.766**	0.674**
C20:1 X ₉	1	-0.788**	-0.232	0.144	0.913**	0.082	-0.833**	-0.673**
∑MUFAX ₁₀	-0.499	1	-0.134	-0.342	-0.812**	-0.511	0.738**	0.660**
C18:2 X ₁₁	0.216	-0.404	1	-0.213	-0.135	0.614*	0.329	0.268
C18:3 X ₁₂	0.472	-0.832**	0.099	1	0.127	0.623*	0.099	-0.164
C20:2 X ₁₃	0.562*	-0.438	0.445	0.172	1	0.162	-0.798**	-0.623*
∑PUFAX ₁₄	0.529*	-0.866**	0.697**	0.771**	0.515*	1	0.203	-0.02
∑USFA X ₁₅	-0.292	0.811**	0.077	-0.614*	-0.197	-0.409	1	0.736**
∑ X ₁₆	-0.228	0.477	0.112	-0.421	0.062	-0.219	0.615*	1

Upper right was the correlation of fatty acid composition in Weihe *Pinus koraiensis* seed orchard; Left lower was the correlation of fatty acid composition in Tieli *Pinus koraiensis* seed orchard. ∑: ∑SFA + ∑USFA. * indicate a statistically significant correlation at p < 0.05 level; ** indicate a statistically significant correlation at p < 0.01 level

$$\begin{aligned}
 Y_{TL} = & 3.53 - Z_1 - 6.35Z_2 - 6.52Z_3 + 2.34Z_4 - 1.31Z_7 \\
 & + 10.04Z_8 + 1.34Z_{10} + 0.44Z_{11} - 4.63Z_{12} \\
 & - 9.23Z_{13} + 14.32Z_{15} - 1.35Z_{16} + 1.35Z_{19} \\
 & (R^2 = 0.9989)
 \end{aligned}$$

$$\begin{aligned}
 Y_{WH} = & 21.64 + 13.43Z_3 + 43.61Z_4 + 104.61Z_6 \\
 & - 62.38Z_7 - 51.74Z_8 + 44.38Z_9 - 13.42Z_{10} \\
 & + 115.09Z_{11} + 6.38Z_{14} + 173.72Z_{15} + 16.53Z_{16} \\
 & + 21.31Z_{17} - 28.84Z_{18} \quad (R^2 = 0.9937)
 \end{aligned}$$

Table 9 The correlation of amino acid composition in *Pinus koraiensis* seeds between Hegang and Linkou seed orchards

Amino acid	Asp Z ₁	Thr Z ₂	Ser Z ₃	Glu Z ₄	Gly Z ₅	Ala Z ₆	Cys Z ₇	Val Z ₈	Met Z ₉	Ile Z ₁₀
Asp Z ₁	1	-0.161	0.291	0.162	0.098	0.185	0.024	0.312	-0.154	-0.241
Thr Z ₂	0.181	1	0.676**	0.726**	0.827**	0.801**	0.509	0.639*	0.483	0.681**
Ser Z ₃	0.331	0.956**	1	0.72**	0.887**	0.916**	0.549*	0.756**	0.572*	0.596*
Glu Z ₄	0.394	0.930**	0.962**	1	0.726**	0.805**	0.821**	0.647**	0.519*	0.594*
Gly Z ₅	0.176	0.650**	0.773**	0.603*	1	0.912**	0.495	0.777**	0.615*	0.796**
Ala Z ₆	0.28	0.743**	0.856**	0.749**	0.952**	1	0.544*	0.797**	0.713**	0.699**
Cys Z ₇	0.228	0.713**	0.835**	0.803**	0.763**	0.822**	1	0.549*	0.294	0.593*
Val Z ₈	0.019	0.905**	0.909**	0.820**	0.843**	0.891**	0.816**	1	0.532*	0.673**
Met Z ₉	0.461	0.684**	0.847**	0.773**	0.883**	0.929**	0.870**	0.776**	1	0.640*
Ile Z ₁₀	-0.024	0.936**	0.887**	0.813**	0.731**	0.768**	0.727**	0.965**	0.644**	1
Leu Z ₁₁	0.29	0.962**	0.989**	0.940**	0.785**	0.857**	0.817**	0.929**	0.825**	0.924**
Tyr Z ₁₂	0.233	0.955**	0.982**	0.938**	0.762**	0.825**	0.851**	0.926**	0.805**	0.926**
Phe Z ₁₃	0.131	0.967**	0.917**	0.880**	0.607*	0.658**	0.648**	0.843**	0.626*	0.899**
Lys Z ₁₄	0.231	0.935**	0.844**	0.900**	0.394	0.563*	0.559*	0.754**	0.502	0.810**
His Z ₁₅	0.192	0.972**	0.920**	0.891**	0.607*	0.684**	0.616*	0.830**	0.636*	0.879**
Arg Z ₁₆	0.245	0.928**	0.901**	0.933**	0.458	0.557*	0.695**	0.747**	0.597*	0.812**
Pro Z ₁₇	0.326	0.677**	0.798**	0.737**	0.834**	0.902**	0.829**	0.792**	0.906**	0.660**
TAAZ ₁₈	0.429	0.930**	0.992**	0.968**	0.767**	0.864**	0.845**	0.880**	0.877**	0.848**
EAAZ ₁₉	0.217	0.987**	0.982**	0.940**	0.746**	0.827**	0.789**	0.943**	0.777**	0.948**
Amino acid	Leu Z ₁₁	Tyr Z ₁₂	Phe Z ₁₃	Lys Z ₁₄	His Z ₁₅	Arg Z ₁₆	Pro Z ₁₇	TAA Z ₁₈	EAA Z ₁₉	
Asp Z ₁	0.195	0.132	0.499	0.397	-0.153	0.277	-0.058	0.4	0.22	
Thr Z ₂	0.823**	0.761**	0.6*	0.575*	0.887**	0.617*	0.674**	0.712**	0.838**	
Ser Z ₃	0.887**	0.898**	0.867**	0.752**	0.706**	0.906**	0.708**	0.928**	0.910**	
Glu Z ₄	0.773**	0.840**	0.624*	0.423	0.754**	0.792**	0.708**	0.881**	0.756**	
Gly Z ₅	0.939**	0.908**	0.815**	0.624*	0.828**	0.706**	0.641*	0.841**	0.945**	
Ala Z ₆	0.902**	0.940**	0.813**	0.669**	0.820**	0.838**	0.738**	0.922**	0.935**	
Cys Z ₇	0.523*	0.661**	0.423	0.423	0.624*	0.619*	0.549*	0.670**	0.568*	
Val Z ₈	0.853**	0.733**	0.682**	0.743**	0.682**	0.622*	0.479	0.789**	0.897**	
Met Z ₉	0.631*	0.715**	0.355	0.118	0.591*	0.534*	0.508	0.540*	0.590*	
Ile Z ₁₀	0.703**	0.735**	0.452	0.372	0.827**	0.439	0.437	0.562*	0.735**	
Leu Z ₁₁	1	0.874**	0.798**	0.665**	0.839**	0.788**	0.701**	0.901**	0.972**	
Tyr Z ₁₂	0.986**	1	0.765**	0.591*	0.818**	0.803**	0.728**	0.902**	0.891**	
Phe Z ₁₃	0.923**	0.935**	1	0.731**	0.555*	0.767**	0.557*	0.863**	0.850**	
Lys Z ₁₄	0.842**	0.825**	0.876**	1	0.572*	0.630*	0.519*	0.713**	0.777**	
His Z ₁₅	0.931**	0.912**	0.968**	0.904**	1	0.655**	0.786**	0.758**	0.846**	
Arg Z ₁₆	0.891**	0.913**	0.937**	0.890**	0.917**	1	0.759**	0.921**	0.790**	
Pro Z ₁₇	0.766**	0.779**	0.607*	0.535*	0.588*	0.540*	1	0.764**	0.685**	
TAAZ ₁₈	0.979**	0.968**	0.882**	0.830**	0.890**	0.880**	0.828**	1	0.913**	
EAAZ ₁₉	0.990**	0.982**	0.949**	0.887**	0.950**	0.904**	0.753**	0.964**	1	

Upper right was the correlation of amino acid composition in LK Korean pine seed orchard; Left lower was the correlation of amino acid composition in HG Korean pine seed orchard; * indicate a statistically significant correlation at $p < 0.05$ level; ** indicate a statistically significant correlation at $p < 0.01$ level

Thr(Z₂), Ser(Z₃), Glu(Z₄), Ala(Z₆), Cys(Z₇), Val(Z₈), Met(Z₉), Leu(Z₁₁), Tyr(Z₁₂), His(Z₁₅), Arg(Z₁₆) and Pro(Z₁₇) were the most stable and had a significant linear relation to the protein content (Y).

Multi-characteristic selection

The clones were from four different orchards and were therefore analyzed separately for their principal

Table 10 The correlation of amino acid composition in *Pinus koraiensis* seeds between Tieli and Weihe seed orchards

Amino acid	Asp Z ₁	Thr Z ₂	Ser Z ₃	Glu Z ₄	Gly Z ₅	Ala Z ₆	Cys Z ₇	Val Z ₈	Met Z ₉	Ile Z ₁₀
Asp Z ₁	1	0.768**	0.921**	0.960**	0.951**	0.949**	0.914**	0.897**	0.894**	0.905**
Thr Z ₂	0.211	1	0.935**	0.868**	0.890**	0.885**	0.886**	0.825**	0.829**	0.826**
Ser Z ₃	0.228	0.652**	1	0.976**	0.980**	0.970**	0.977**	0.911**	0.942**	0.903**
Glu Z ₄	0.414	0.339	0.746**	1	0.981**	0.988**	0.979**	0.934**	0.931**	0.927**
Gly Z ₅	0.792**	0.189	0.439	0.394	1	0.987**	0.973**	0.951**	0.928**	0.951**
Ala Z ₆	-0.066	0.476	0.446	0.157	-0.046	1	0.979**	0.972**	0.921**	0.960**
Cys Z ₇	-0.054	-0.193	0.34	0.467	0.051	-0.317	1	0.941**	0.926**	0.928**
Val Z ₈	0.254	0.605*	0.558*	0.285	0.436	-0.136	0.217	1	0.879**	0.978**
Met Z ₉	0.388	0.173	0.447	0.516*	0.441	0.16	0.338	0.353	1	0.851**
Ile Z ₁₀	0.473	0.249	0.336	0.363	0.529*	-0.113	-0.101	0.306	0.26	1
Leu Z ₁₁	0.671**	0.292	0.432	0.352	0.762**	0.033	0.015	0.393	0.216	0.29
Tyr Z ₁₂	0.527*	0.756**	0.607*	0.44	0.432	0.208	0.074	0.535*	0.267	0.297
Phe Z ₁₃	0.737**	0.223	0.462	0.521*	0.824**	-0.147	0.171	0.465	0.619*	0.578*
Lys Z ₁₄	0.720**	0.225	0.423	0.582*	0.465	-0.111	0.273	0.272	0.302	0.306
His Z ₁₅	0.303	0.331	0.445	0.153	0.634*	0.109	-0.048	0.241	-0.048	0.492
Arg Z ₁₆	0.563*	0.51	0.737**	0.652**	0.745**	0.101	0.152	0.467	0.25	0.449
Pro Z ₁₇	-0.194	-0.041	0.078	-0.349	0.066	0.504	-0.092	-0.092	0.025	-0.143
TAA Z ₁₈	0.644**	0.548*	0.840**	0.691**	0.797**	0.309	0.225	0.529*	0.522*	0.488
EAA Z ₁₉	0.751**	0.551*	0.670**	0.587*	0.813**	0.013	0.121	0.695**	0.544*	0.610*
Amino acid	Leu Z ₁₁	Tyr Z ₁₂	Phe Z ₁₃	Lys Z ₁₄	His Z ₁₅	Arg Z ₁₆	Pro Z ₁₇	TAA Z ₁₈	EAA Z ₁₉	
Asp Z ₁	0.947**	0.933**	0.947**	0.892**	0.946**	0.935**	0.872**	0.963**	0.935**	
Thr Z ₂	0.900**	0.918**	0.917**	0.881**	0.896**	0.880**	0.804**	0.898**	0.923**	
Ser Z ₃	0.978**	0.984**	0.974**	0.909**	0.971**	0.980**	0.873**	0.989**	0.978**	
Glu Z ₄	0.972**	0.982**	0.965**	0.896**	0.959**	0.979**	0.858**	0.994**	0.969**	
Gly Z ₅	0.991**	0.965**	0.980**	0.916**	0.955**	0.952**	0.840**	0.988**	0.988**	
Ala Z ₆	0.986**	0.978**	0.966**	0.910**	0.937**	0.950**	0.837**	0.988**	0.986**	
Cys Z ₇	0.971**	0.974**	0.937**	0.884**	0.928**	0.956**	0.817**	0.978**	0.967**	
Val Z ₈	0.961**	0.920**	0.911**	0.856**	0.850**	0.865**	0.734**	0.933**	0.958**	
Met Z ₉	0.921**	0.932**	0.903**	0.766**	0.890**	0.929**	0.815**	0.936**	0.911**	
Ile Z ₁₀	0.966**	0.916**	0.930**	0.895**	0.862**	0.850**	0.749**	0.930**	0.965**	
Leu Z ₁₁	1	0.974**	0.980**	0.936**	0.951**	0.943**	0.849**	0.986**	0.997**	
Tyr Z ₁₂	0.508	1	0.967**	0.915**	0.961**	0.974**	0.872**	0.989**	0.977**	
Phe Z ₁₃	0.712**	0.534*	1	0.941**	0.979**	0.952**	0.866**	0.982**	0.985**	
Lys Z ₁₄	0.4	0.49	0.452	1	0.934**	0.877**	0.859**	0.924**	0.945**	
His Z ₁₅	0.393	0.502	0.417	0.127	1	0.976**	0.904**	0.976**	0.952**	
Arg Z ₁₆	0.661**	0.612*	0.574*	0.604*	0.663**	1	0.883**	0.981**	0.939**	
Pro Z ₁₇	0.026	-0.051	-0.253	-0.307	0.255	-0.062	1	0.884**	0.849**	
TAA Z ₁₈	0.722**	0.714**	0.705**	0.578*	0.603*	0.874**	0.161	1	0.985**	
EAA Z ₁₉	0.798**	0.710**	0.880**	0.592*	0.45	0.761**	-0.15	0.869**	1	

Upper right was the correlation of amino acid composition in Weihe *Pinus koraiensis* seed orchard; Left lower was the correlation of amino acid composition in Tieli *Pinus koraiensis* seed orchard; * indicate a statistically significant correlation at p < 0.05 level; ** indicate a statistically significant correlation at p < 0.01 level

Table 11 Comparing mean of nutrients between superior clones and seed orchards

Superior clone		Composition									
		C16:1 (%)	C18:1 (%)	C20:1 (%)	\sum MUFA (%)	C18:2 (%)	C18:3 (%)	C20:2 (%)	\sum PUFA (%)	\sum USFA (%)	Oil content %
HG6, HG8, HG39	The mean of superior clone	0.13	25.1	2.14	26.79	42.41	16.4	1.03	59.84	85.54	56.65
		1.59	3.68	5.07	1.51	1.46	1.38	8.32	1.55	0.26	2.48
LK79-36, LK15, LK32	The mean which is higher than it in Linkou seed orchards (%)	0.19	23.14	2.27	25.39	42.22	17.22	1.09	60.06	84.86	63.58
		9.62	2.96	4.44	2.33	1.46	0.53	1.18	0.4	0.27	0.22
TL3101, TL1194, TL1131	The mean of superior clone	0.2	21.54	2.23	23.97	43.1	16.99	1.06	60.45	84.37	56.79
	The mean which is higher than it in Tieli seed orchards (%)	13.9	0.02	1.67	0.28	1.31	0.65	3.85	0.01	0.03	0.43
WH28, WH19, WH117	The mean of superior clone	0.17	23.48	2.22	25.61	42.44	16.45	1.13	59.86	85.16	54.48
	The mean which is higher than it in Weihe seed orchards (%)	1.96	4.79	1.03	3.36	0.75	0.51	1.82	0.43	0.93	5.22

Superior clone		Composition									
		Thr (%)	Val (%)	Met (%)	Ile (%)	Leu (%)	Phe (%)	Lys (%)	TAA (%)	EAA (%)	Protein mg/g
HG6, HG8, HG39	The mean of superior clone	1.16	1.13	0.79	0.86	2.73	1.33	1.52	41.77	9.52	10.65
	The mean which is higher than it in Hegang seed orchards (%)	11.77	15.74	13.11	15.2	14.5	10.65	9.48	13.35	12.88	45.04
LK79-36, LK15, LK32	The mean of superior clone	1.65	1.84	0.73	1.47	4.16	2.08	2.22	58.25	14.15	9.81
	The mean which is higher than it in Linkou seed orchards (%)	12.65	10.7	6.44	5.55	13.12	17.41	9.62	11.72	11.59	2.16
TL3101, TL1194, TL1131	The mean of superior clone	1.21	1.29	0.62	1.17	3.18	1.6	1.65	41.7	10.7	8.8
	The mean which is higher than it in Tieli seed orchards (%)	1.36	10.16	0.4	11.2	15.47	16.78	5.18	6.42	10.03	0.32
WH28, WH19, WH117	The mean of superior clone	1.29	1.28	0.82	1.08	2.96	1.42	1.59	43.8	10.43	6.74
	The mean which is higher than it in Weihe seed orchards (%)	20.87	26.5	23.14	32.14	27.16	23.88	23.73	25.92	25.46	7.41

components. There are ten fatty acids indicators in oil content, including C16:1, C18:1, C20:1, C18:2, C20:2, C18:3, \sum MUFA, \sum PUFA, and \sum USFA. Ten indicators were selected as the variable factors for the principle components. We chose the protein content from protein indicators, including Thr, Val, Met, Ile, Leu, Phe, Lys, TAA, and EAA. The ten indicators were selected as the variable factors for the principle components (Table 11). The factors with an initial eigenvalue greater than 1 were extracted from each population, and the cumulative variance for each population was 89.08 % for HG, 83.39 % for LK, 84.99 % for TL and 91.38 % for WH.

The extracted principal components included most of the information of the quality characteristics of the seeds.

The ratio of the eigenvalue associated with each principle component to the sum of the eigenvalues of the total extracted principle components was used as the weight to obtain a comprehensive model for the principle components. The comprehensive principle component values for all the orchards were calculated and sorted.

A selection rate of 20 % was used for each population to select clones. HG6, HG8, and HG39 were selected from the seed orchard in Hegang; LK79-36, LK15 and LK32 were selected from the seed orchard in Linkou; TL3101, TL1194, and TL1131 were selected from the seed orchard in Tieli; and WH28, WH19, and WH117 were selected from the seed orchard in Weihe. The nutritional components of the selected clones had higher values than the

clonal source population for all the seed orchards (Table 11). The selected clones had good development values for seeds, and the selected clones could also provide excellent propagative materials for the construction of clonal seed orchards.

Discussion

Pine nuts are woody oil plants with high nutritional value. The content of oil and protein in pine nuts are equaled with that in walnut, *armeniaca sibirica* and chestnut (Chen et al. 2000; Hao et al. 2002; Mahesh and Shridhar 2006; Seeram et al. 2008; Li et al. 2009). In the present study, GC-MS was used to determine the fatty acid content of pine seed oil. In total, 11 fatty acids were detected, and the unsaturated fatty acid content (\sum USFA) was 84.67 %; 17 amino acids were detected in the pine nuts, 7 of which were essential amino acids for humans. The essential amino acid content was 24.03 % of the total amino acid content.

The phenotypic characteristics of the plants reflect the adaptability of genotypes to environmental change. Irreversible changes occur to phenotypic characteristics under long-term selective pressure. New phenotypes are generated after stable inheritance. Therefore, phenotypic variations are very important for adaptability and evolution (LePage 2001; Pigliucci et al. 2006). Genetic variation is a product of gradual adaptation to the environment and continuous evolution of organisms and is an important characteristic of genetic information.

Phenotypic variation of seeds not only determines the dispersity of the species and the distribution pattern of the population but is also an important economic characteristic that can be used for development. It also is very important for the selection of varieties of forest trees (Greipsson and Davy 1995). Currently, genetic analysis is the focus of the study of the variation of seed characteristics of plants, especially the study of important economic characteristics. Seeking genetic materials with high-quality characteristics lays the foundation for genetic improvement.

The nutritional components of the Korean pine nuts contained rich genetic variation; the variation coefficient of fatty acids was 2.24–66.83 %, while the variation coefficient of amino acid components was 14.70–38.88 %, indicating that there was relatively large potential for genetic improvement among the nutritional components of Korean pine nuts.

The phenotypic differentiation among the fatty acid and amino acid components of pine nuts reveals that the variation within the population (85.18 %) was the primary source of the variation of the fatty acid components; the variation among the orchards (63.08 %) was the primary source of the variation of the amino acid components. The

amino acid and fatty acid components had a negative or insignificant Pearson correlation, which might explain the inconsistent results of the phenotypic differentiation of the amino acid and fatty acid components. In addition, the heterogeneity of different seed orchards enhances the variations of seed characteristics. Korean pine trees are monoecious and cross-pollinated and are naturally pollinated in heterozygous seed orchards for a very long time. Thus, these trees have formed extensive hybrid orchards (Politov et al. 1999; Goroshkevich 2004), which provide advantageous conditions for the selection of germplasm resources of Korean pine. Genetic markers and seeding tests are useful for further revealing the genetic differences among the orchards and exploring the variation patterns (Zheng et al. 2009; Sharma et al. 2010; Gomes et al. 2014).

Under the same condition as other factors, the greater standard deviation yields greater gain in traits expected. Low phenotypic variation means that there is little opportunity to find trees much bigger than the population average. When selecting strength equals with heritability in characters, a greater gain is obtained in the greater variability characteristics of groups. Traits with bigger heritability can obtain greater genetic gain, with different characters with different genetic rates.

We studied the characteristics of Korean pine seeds from different seed orchards and revealed the breeding potential of the germplasm resource. We also estimated the genetic parameters and the expected genetic gains for the Korean pine seed characteristics. The seed characteristics were all highly controlled by heritability ($h^2 > 80$ %) in the different clonal source orchards. The Korean pine seed characteristics all exhibited a similar trend for genetic gains; the fatty acid components (C16:1, C18:1, C20:1, C18:3, C20:2, \sum MUFA, and \sum SFA) achieved more than 5 % genetic gains in the four seed orchards, and the amino acid components (Thr, Val, Leu, Phe, TAA, and EAA) all achieved more than 10 % genetic gains in the orchards, which provided a basis for the quality selection of seed characteristics. Through the analysis of the principal components, the comprehensive principal component values for all the seed orchards were calculated and sorted. A clonal selection rate of 20 % was used as the standard, and 12 clones were selected. The desired characteristics of the selected clones were generally higher than the means of the clonal orchards, indicating that the clones could provide good clonal propagative materials.

The level of variation was bigger in the seed orchards due to the different sampling-site conditions and fewer hits, which caused the different between phenotypic and genetic variation (Sun et al. 2005). The differences inevitably led to different correlation between real properties in different seed orchards. In comparison, the habitat

differences between individuals were smaller than the differences between seed orchards. The proportion of differences was extremely small in terms of environmental impact, which was more likely determined by genetic factors of plant. In order to fully reveal the real cause of phenotypic variation in seed orchards, we need to conduct further research into internal patterns of genetic variation, spatial distribution patterns, and the relationship in real properties.

Correlation analysis was conducted on clonal orchards separately. There was no significant correlation between the oil and protein contents of the pine nuts. Therefore, correlation analyses were conducted separately among the fatty acid and amino acid components. Under different location conditions, the correlation between the partial characteristics of the fatty acid and amino acid components was very stable; the inherent genetic bases of the characteristics were closely related and had a broad adaptability to the heterogeneity of the environment. Furthermore, based on the multiple stepwise-regression analysis, the factors that affected the oil and protein contents provided a basis for improving the seed yield and clonal selection of Korean pine trees.

In the present study, we conducted an initial evaluation of the germplasm resources of the Korean pine trees. However, we only used annual variation data from 1 year for analysis. Thus, the analysis of annual variation data from multiple years is necessary. Further studies will be focused on the management of field seedling cultivation and the morphology, phenophase, and patterns of physiological and environmental habitat factors.

Conclusion

Through the study of the seed characteristics of four clonal seed orchards of Korean pine, we reveal that the nutritional components of the Korean pine nuts contained rich genetic variation. Through selection, expected genetic gains could be achieved among the nutritional components of the pine nuts, and high-quality clones could be selected as superior clones, including HG6, HG8, HG39, LK79-36, LK15, LK32, TL3101, TL1194, TL1131, WH28, WH19, and WH117. The composition of fatty acid and amino acid in superior clones was significantly higher than the average level of clonal seed orchards. We also studied the phenotypic differentiation of clonal orchards under different locations and the correlations between population characteristics. The results reveal the stability among the seed characteristics, which provides a basis for improving the seed yield of Korean pine trees and the clonal selection.

References

- Bond B, Fernandez DR, VanderJagt DJ, Williams M, Huang YS, Chuang LT, Millson M, Andrews R, Glew RH (2005) Fatty acid, amino acid and trace mineral analysis of three complementary foods from Jos, Nigeria. *J Food Compos Anal* 18:675–690
- Chen ZX, Lei ZX, Liu HN, Wu GY, Wang XB (2000) Analysis of chestnut ingredients and its integrated quality evaluation. *J Fruit Sci* 17(4):286–289 (In Chinese)
- Fernandes DC, Freitas JB, Czeder LP, Naves MMV (2010) Nutritional composition and protein value of the baru (*Dipteryx alata* Vog.) almond from the Brazilian Savanna. *J Sci Food Agri* 90:1650–1655
- Ferramosca A, Savy V, Einerhand AWC, Zara V (2008) Pinus koraiensis seed oil (*PinnoThin (TM)*) supplementation reduces body weight gain and lipid concentration in liver and plasma of mice. *J Animal Feed Sci* 17:621–630
- Gomes RA Jr, Gurgel FdL, Peixoto Lda, Bhering LL, da Cunha RNV, Lopes R, Pina AJda, Veiga AS (2014) Evaluation of interspecific hybrids of palm oil reveals great genetic variability and potential selection gain. *Ind Crops Prod* 52:512–518
- Goroshkevich SN (2004) Natural Hybridization between Russian Stone Pine (*Pinus sibirica*) and Japanese Stone Pine (*Pinus pumila*). USDA Forest Service Proceedings pp. 169–171
- Greipsson S, Davy AJ (1995) Seed mass and germination behavior in orchards of the dune-building grass *Leymus-Arenarius*. *Ann Bot* 76:493–501
- Hao YB, Wang KJ, Wang SL (2002) analysis of fatty acid composition in several early fruit *Walnut* protein. *Food Sci* 23(10):123–125 (In Chinese)
- Kang KS, Lindgren D (1998) Fertility variation and its effect on the relatedness of seeds in *Pinus densiflora*, *Pinus thunbergii* and *Pinus koraiensis* clonal seed orchards. *Silvae Genetica* 47:196–201
- LePage BA (2001) New species of *Picea* A. Dietrich (*Pinaceae*) from the middle Eocene of Axel Heiberg Island, Arctic Canada. *Bot J Linn Soc* 135:137–167
- Li M, Liu Y, Sun C, Meng YN, Yang KQ, Hou LQ, Wang JY (2009) Research advance about nutrients and medicinal value of *Walnut*. *J Chinese Cereals Oils Assoc* 24(6):166–169 (In Chinese)
- Lim TK (2012) *Pinus koraiensis*. Edible Medicinal and Non-Medicinal Plants. pp 297–303
- Mahesh V, Shridhar KS (2006) Chemical composition of selected edible nut seeds. *Food Feed Chem* 54(13):4705–4714
- Mensink RP, Katan MB (1990) Effect of dietary trans-fatty-acids on high-density and low-density-lipoprotein cholesterol levels in healthy-subjects. *N Engl J Med* 323:439–445
- Miloun M, Miltky J, Forina M (1992) Chemometric for analytical chemistry: PC-aided statistical data analysis, 1. Ellis Harwood *New York*
- Mutke S, Gordo J, Gil L (2005) Variability of Mediterranean stone pine cone production: yield loss as response to climate change. *Agric For Meteorol* 132:263–272
- Nasri N, Khaldi A, Fady B, Triki S (2005) Fatty acids from seeds of *Pinus pinea* L.: composition and population profiling. *Phytochemistry* 66:1729–1735
- Nergiz C, Donmez I (2004) Chemical composition and nutritive value of *Pinus pinea* L. seeds. *Food Chem* 86:365–368
- Pigliucci M, Murren CJ, Schlichting CD (2006) Phenotypic plasticity and evolution by genetic assimilation. *J Exp Biol* 209:2362–2367
- Politov DV, Belokon MM, Maluchenko OP, Belokon YS, Molozhnikov VN, Mejnartowicz LE, Krutovskii KV (1999) Genetic evidence of natural hybridization between siberian stone pine, *Pinus Sibirica* du tour, and dwarf Siberian pine, *P. Pumila* (Pall.) Regel. *Forest Genetics* pp 41–48

- Sathe SK, Seeram NP, Kshirsagar HH, Heber D, Lapsley KA (2008) Fatty acid composition of California grown almonds. *J Food Sci* 73:607–614
- Savage GP (2001) Chemical composition of walnuts (*Juglans regia* L.) grown in New Zealand. *Plant Foods Hum Nutr* 56:75–82
- Seeram NP, Kshirsagar HH, Heber D, Lapsley KA (2008) Fatty acid composition of California grown Almonds. *J Food Sci* 73(9):604–607
- Sharma OC, Murkute AA, Kanwar MS (2010) Assessing genetic divergence in seedling trees of Persian walnut (*Juglans regia*). *Indian J Agric Sci* 80:360–363
- Shen L, Wang X, Wang Z, Wu Y, Chen J (2008) Studies on tea protein extraction using alkaline and enzyme methods. *Food Chem* 107:929–938
- Sun HQ, Li A, Ban W (2005) Morphological variation and its adaptive significance for *Changnienia amoena*, an endangered orchid. *Biodivers Sci* 13(5):376–386 (In Chinese)
- Venkatachalam M, Sathe SK (2006) Chemical composition of selected edible nut seeds. *J Agri Food Chem* 54:4705–4714
- Willet WC, Ascherio A (1994) Trans-fatty-acids—are the effects only marginal. *Am J Public Health* 84:722–724
- Wolff RL, Deluc LG, Marpeau AM, Comps B (1997) Chemotaxonomic differentiation of conifer families and genera based on the seed oil fatty acid compositions: multivariate analyses. *Trees Struct Funct* 12:57–65
- Wolff RL, Pedrono F, Pasquier E, Marpeau AM (2000) General characteristics of *Pinus spp.* seed fatty acid compositions, and importance of Delta 5-olefinic acids in the taxonomy and phylogeny of the genus. *Lipids* 35:1–22
- Yoon TH, Im KJ, Koh ET, Ju JS (1989) Fatty-acid compositions of *Pinus Koraiensis* seed. *Nutr Res* 9:357–361
- Zadernowski R, Naczek M, Czaplicki S (2009) Chemical composition of *Pinus sibirica* nut oils. *Eur J Lipid Sci Technol* 111:698–704
- Zheng YL, Sun WB, Zhou Y, Coombs D (2009) Variation in seed and seedling traits among natural orchards of *Trigonobalanus doichangensis* (*A. Camus*) Forman (*Fagaceae*), a rare and endangered plant in southwest China. *New Forest* 37:285–294