

UDC 630*165.6+631.523.4

HIGHLY INHERITABLE VARIABLE COMPONENTS IN THE CLONAL PLANTATION OF SCOTS PINE

V. M. Efimov¹, V. V. Tarakanov², N. B. Naumova³, V. Y. Kovaleva⁴, **K. P. Kutsenogiy**⁵

¹ Institute of Cytology and Genetics, Russian Academy of Sciences, Siberian Branch
Prospekt Akademika Lavrent'eva, 10, Novosibirsk, 630090 Russian Federation

² West Siberian Branch of V. N. Sukachev Institute of Forest SB RAS –
Branch of the Federal Research Center «Krasnoyarsk Science Center»
Zhukovskiy str., 100/1, Novosibirsk, 630082 Russian Federation

³ Institute of Soil Science and Agrochemistry, Russian Academy of Sciences, Siberian Branch
Prospekt Akademika Lavrent'eva, 8/2, Novosibirsk, 630090 Russian Federation

⁴ Institute of Systematics and Ecology of Animals, Russian Academy of Sciences, Siberian Branch
Frunze str., 11, Novosibirsk, 630091 Russian Federation

⁵ Institute of Chemical Kinetics and Combustion, Russian Academy of Sciences, Siberian Branch
Institutskaya str., 3, Novosibirsk, 630091 Russian Federation

E-mail: vmefimov@ngs.ru, tarh012@mail.ru, nnaumova@mail.ru, vkova@ngs.ru, koutsen@kinetics.nsc.ru

Received 07.09.2019

The variability of the data on elemental composition of needles from the clonal population of Scots pine *Pinus sylvestris* L., established on the long-term field experiment, was studied by principal components extraction from the normalized data matrix, and broad-sense heritability H^2 , i.e. the contribution of clones to the total data variance was calculated both for the original variables and principal components. To find the linear combinations of variables with the highest heritability the discriminant analysis was performed. The results suggest the importance of multivariate statistics for forest genetics and selection in targeting search for genetic marker traits in the populations of woody plants, in assessing genetic differentiation among populations, identification of the best genotypes via their phenotypes, etc.

Keywords: *Pinus sylvestris* L., chemical elements, heritability, multivariate analysis.

How to cite: Efimov V. M., Tarakanov V. V., Naumova N. B., Kovaleva V. Y., Kutsenogiy K. P. Highly inheritable variable components in the clonal plantation of Scots pine // *Sibirskij Lesnoj Zhurnal* (Sib. J. For. Sci.). 2019. N. 6. P. 82–88 (in English with Russian abstract).

DOI: 10.15372/SJFS20190609

INTRODUCTION

Recently due to increased computational capacities the new statistical approaches have been employed to study various forms of variability using multivariate analysis. The multivariate analysis can be used to study the role of genetic factors in the variability of natural populations. The theory of multivariate genetic analysis of quantitative traits has been rapidly developing since the first half of the 20th century (Fisher, 1918, 1936; Wright, 1921; Smith, 1936; Hazel, 1943).

Smith and Hazel solved the problem of estimating the additive heritability of any linear combination of variables in the multidimensional space by means of phenotypic and genetic correlations. Besides that, they also posed and solved the problem of searching for linear combination that is maximally responsive to selection. R. Lande (1979) introduced the genetic matrix G , i. e. a multivariate analogue of heritability between parents and offspring. So in a matrix form the multivariate analogue of heritability $H = GP^{-1}$ was found where P is a phenotypic matrix of correlations between traits, as well as a

© Efimov V. M., Tarakanov V. V., Naumova N. B., Kovaleva V. Y., Kutsenogiy K. P., 2019

so called «breeder's equation» $\Delta\mu = GP^{-1}s$, where s stands for a selection differential and $\Delta\mu$ represents the response to selection. It is essential thing that the heritability of linear combination is usually significantly higher than that with the original traits it is created from (Klingenberg & Leamy, 2001; Efimov et al., 2005; Kovaleva et al., 2006).

If the data on genetic correlations are absent, the analysis may be reduced to that of variables with high broad-sense heritability. R. A. Fisher (1936) and H. F. Smith (1936) laid the basics for such an approach. For the case of several groups Fisher suggested the discriminant analysis, i.e. searching for the linear combination of variables that maximizes the ratio of between-group and total variance. Earlier, we successfully tested multivariate approach for the discrimination of geographical populations (Tarakanov et al., 2011). If groups represent different clones, the problem merely becomes the one of maximizing broad-sense heritability.

Identification of highly heritable variable components may be employed for genetic selection of woody plants. A useful experimental model to test and develop the approach is provided with the data on elemental composition of needles and soils from the long-term Scots pine field plantation of different clones.

MATERIALS AND METHODS

The samples of 2-year old living needles were collected from the clonal plantation of Scots pine *Pinus sylvestris* L. in the forest-steppe zone of the Altai Krai, Russia (Kutsenogiy et al., 2007).

The elemental composition of phytomass samples was determined by X-ray fluorescence analysis

with synchrotron radiation (Baryshev et al., 1986, 1991). The elemental spectra measurements were carried out at the element analysis station of the Siberian Synchrotron Radiation Centre of the Institute of Nuclear Physics SB RAS (Novosibirsk, Russia) using VEPP-3. During measurements the electron energy was 2 GeV and the electric current in the electron pencil was 130 mA. Impulse registration, energy calibration and spectra recordings were carried out by Oxford WIN-MCA Ver. 03.80.26. The primary processing of the emission spectra (identification of elements and their peak area determination) were carried out using AXIL software. We used, as reference samples, the attested Russian standard specimens of grass phytomass, ash of biomaterials and sandy soil. Altogether 42 needle samples from 9 clones were analyzed for 20 chemical elements.

Due to strong asymmetrical distribution of all measured variables we normalized each original variable by summing it up with a randomly generated very small value of the variable ξ , $0 < \xi < 10^{-7}$, in order to rank in a random way equal values, for instance, zero. The obtained new values of the initial variable were substituted by their ranks, which, in their turn, were substituted by their normal scores, i. e. the mathematical expectations of order statistics in a set of N objects from the standard normal distribution. Fisher and Yates first suggested using normal scores as a universal normalization method in the introduction to their tables (Fisher & Yates, 1963; cited by Kendall & Stuart, 1979). If $N \rightarrow \infty$, then the correlation of normal scores with their initial ranks tends to 1.

The broad-sense heritability H^2 was estimated using ANOVA by the contribution of between-clone

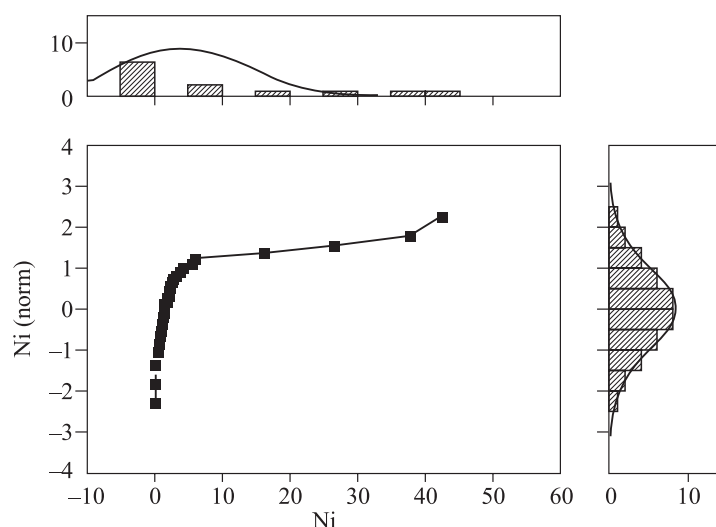


Fig. 1. Normalization of the variable describing Ni content in winter needles of Scots pine.

variability into the total variance of a variable according to the following formulas:

$$\lambda_F = F(k - 1)/(N - k), \quad (1)$$

$$H^2 = \lambda_F/(1 + \lambda_F), \quad (2)$$

where F is the Fisher's criterion, λ_F represents the ratio of between-clone variance to the intra-clone one, N is the number of objects, and k is the number of clones.

Then principal components were extracted from the whole set of data, and for each principal component obtained that was regarded as a new variable, the heritability was calculated according to the formulas (1) and (2).

To identify variables with maximal heritability we used discriminant analysis (MANOVA) for the whole data set. For each discriminant function the broad-sense heritability was calculated according to the following formula:

$$H^2 = \lambda/(1 + \lambda), \quad (3)$$

where $\lambda = \lambda_F$ represents the eigenvalue which is equivalent to calculation by the previous method. By definition, discriminant functions display maximal heritability no less than that of any source

variables included in the model. The contribution of variables into the principal components and discriminant functions were estimated by correlation coefficients.

RESULTS AND DISCUSSION

Chemical elements content in pine needles was shown to vary significantly between clones (Table 1).

Zn, K, Ca displayed the maximal between-clone variability, whereas such elements as Hg, Co, W, V displayed the minimal one, which agrees well with our previous findings (Tarakanov et al., 2007).

Five principal components (the first four and the sixth) differed significantly from all the others according to the degree of their between-clone variability (Table 1). This may indirectly indicate the number of independent genetic systems determining chemical elements' content in needles.

The nature of the components varied a lot. The first principal component, responsible for 24 % of the total variance, is derived from the opposition of K and Ca, on the one hand, and U, Se, Pb, Y, Bi, Hg, Th, Ni, on the other hand (Table 2).

However, the element with the highest heritability (Zn) did not correlate with this pattern. As

Table 1. Broad-sense heritability of chemical elements content in pine needles and of the principal components extracted from the respective data set

Elements	H^2	p	Principal components	H^2	p
Zn	0.862	0.000	PC1 (24) *	0.825	0.000
Ca	0.800	0.000	PC2 (13)	0.754	0.000
K	0.698	0.000	PC3 (11)	0.656	0.000
Br	0.671	0.000	PC4 (9)	0.428	0.010
Rb	0.611	0.000	PC5 (7)	0.100	0.876
Cu	0.553	0.000	PC6 (6)	0.424	0.011
Bi	0.537	0.001	PC7 (5)	0.145	0.691
U	0.527	0.001	PC8 (4)	0.250	0.242
Mn	0.478	0.003	PC9 (4)	0.239	0.280
Sr	0.467	0.004	PC10 (4)	0.146	0.683
Pb	0.464	0.004	PC11 (3)	0.082	0.930
Se	0.461	0.005	PC12 (3)	0.209	0.395
Ni	0.447	0.007	PC13 (2)	0.044	0.991
Th	0.442	0.007	PC14 (2)	0.059	0.975
Fe	0.399	0.020	PC15 (1)	0.195	0.456
Y	0.359	0.043	PC16 (1)	0.145	0.688
Hg	0.300	0.119	PC17 (1)	0.267	0.193
Co	0.295	0.130	PC18 (1)	0.213	0.379
W	0.243	0.265	PC19 (1)	0.119	0.804
V	0.185	0.503	PC20 (0)	0.065	0.966

Note. * Values in brackets indicate the portion of the total variance accounted for by the respective principal component.

Table 2. Correlation coefficients between chemical elements content in pine needles and the first six principal components obtained from the respective data set

Elements	PC1	PC2	PC3	PC4	PC5	PC6
K	<i>0.70 *</i>	0.34	0.18	-0.27	0.13	0.16
Ca	<i>0.56</i>	0.26	<i>0.52</i>	0.22	0.05	0.22
V	-0.27	-0.06	0.00	0.22	<i>0.75</i>	-0.06
Mn	-0.37	-0.29	0.22	<i>0.57</i>	-0.33	0.15
Fe	-0.30	<i>0.51</i>	-0.06	<i>0.41</i>	0.38	-0.35
Co	-0.20	<i>0.62</i>	-0.31	0.09	0.13	-0.10
Ni	-0.47	0.13	0.29	-0.04	0.39	0.39
Cu	-0.28	<i>0.46</i>	0.32	-0.32	0.09	<i>0.52</i>
Zn	-0.22	<i>0.56</i>	<i>0.59</i>	0.05	-0.23	-0.11
Se	-0.77	-0.03	-0.22	-0.26	0.17	0.13
Br	-0.21	<i>0.74</i>	-0.30	-0.03	-0.08	0.01
Rb	-0.31	<i>0.61</i>	-0.26	-0.18	-0.28	-0.19
Sr	-0.37	0.20	<i>0.60</i>	<i>0.49</i>	-0.16	-0.09
Y	-0.69	-0.24	0.18	-0.12	0.05	-0.16
W	0.08	-0.19	<i>0.59</i>	-0.15	0.22	-0.34
Hg	-0.59	-0.15	-0.03	-0.22	0.01	0.27
Pb	-0.69	0.01	-0.08	0.11	-0.36	0.19
Bi	-0.64	-0.24	0.33	-0.44	-0.02	-0.20
Th	-0.51	-0.18	-0.30	<i>0.56</i>	0.11	0.22
U	-0.71	-0.01	0.13	-0.26	-0.06	-0.25
$\lambda, \%$	24.26	13.11	10.85	8.97	6.95	5.68
$\Sigma\lambda, \%$	24.26	37.37	48.22	57.18	64.13	69.81

Note. * Coefficients in bold italics have $p < 0.05$.

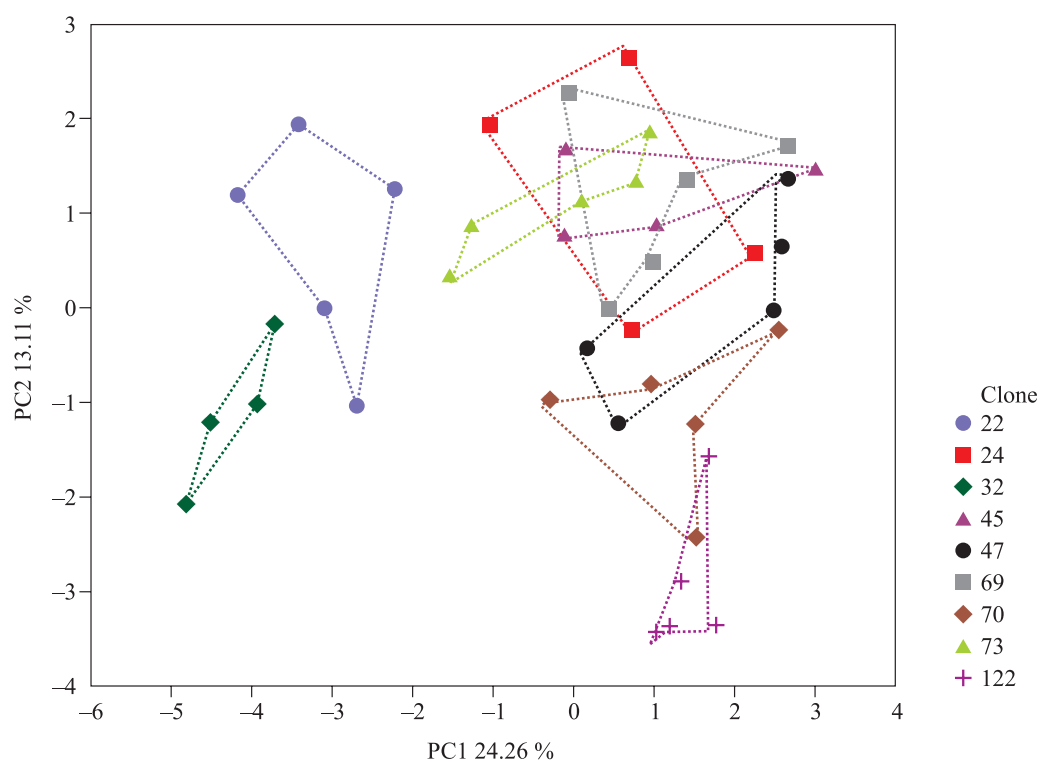


Fig. 2. Configuration of needle samples from different clones in the plane of the first two principal components.

Table 3. Heritability of the discriminant functions for the data set of chemical elements content in pine needles

Root		H^2	p -level
Root_1	(42)*	0.962	0.000
Root_2	(32)	0.952	0.000
Root_3	(10)	0.865	0.000
Root_4	(9)	0.842	0.015
Root_5	(3)	0.645	0.381
Root_6	(2)	0.595	0.710
Root_7	(1)	0.274	0.974
Root_8	(0)	0.229	0.908

Note. *Values in brackets indicate the portion of the total variance accounted for by the respective canonical roots.

(Fig. 2) shown, the decreased K and Ca contents and, correspondingly, the increased contents of the other eight elements are characteristic of clones 22 and 32, whereas other clones demonstrated the opposite patterns. The second principal component responsible for 13 % of the total variance, is mainly derived from Br, Co, Rb, Zn, Fe, Cu.

Fig. 2 shows that the positive pole of the second principal component is determined by clone 122. It is necessary to note that although the heritability of the first principal component is somewhat higher than that of K and Ca, it is still lower than the heri-

tability of Zn that determines the second principal component (Table 1 and 2).

However, the targeted search for the axes with maximal heritability by discriminant analysis resulted in the canonical root with heritability of 0.962, i. e. higher than the highest heritability of source variables (Table 3).

Table 3 shows that the first three canonical roots are formed by the pairs of the first three principal components with high heritability. In the plane of these roots various clones are well differentiated from each other (Fig. 3).

In this study we did not aim at biologically meaningful interpretation of the highly inheritable components obtained. We were mostly interested in that if multivariate statistics could reveal the components with higher heritability than the most inheritable original chemical characteristics of the studied objects, and the results obtained with elemental composition of pine needles obviously gave a positive answer. We have obtained similar results, i.e. higher heritability of canonical roots as compared to the original data, with other phytomass and soil characteristics from the same field experiments with different clones. Thus, the proposed approach may be used for targeting forest genetics research and selection work aimed at adaptively important traits, identification of genetically valuable trees in stands, and so on.

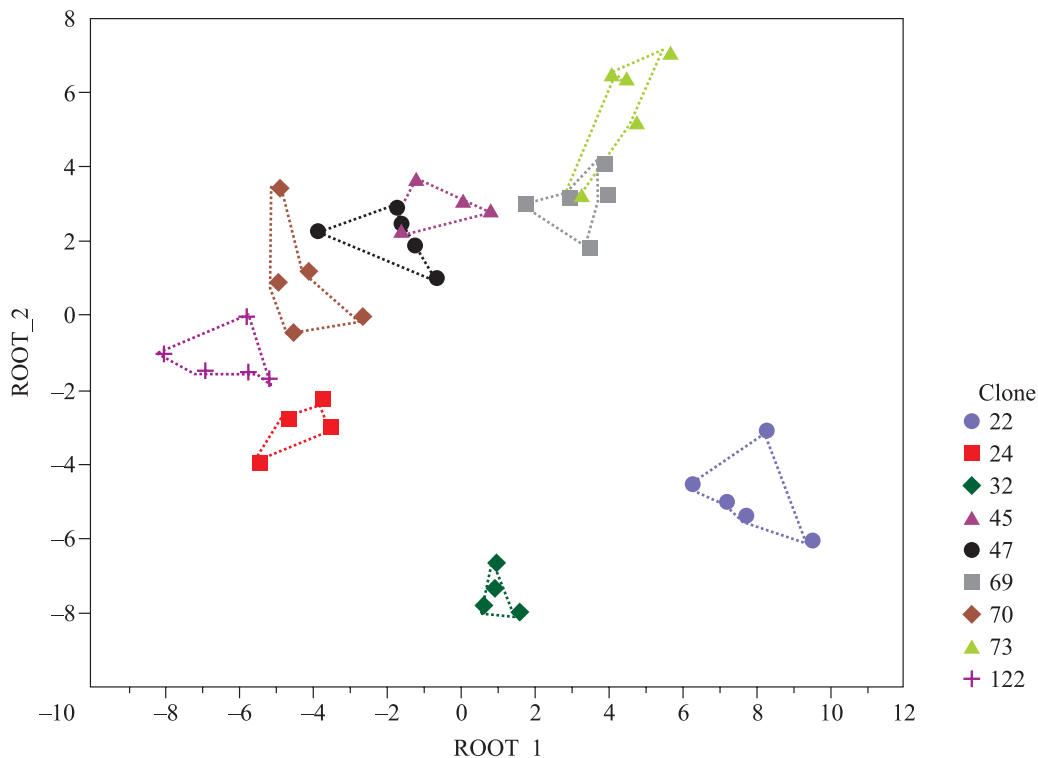


Fig. 3. Configuration of needle samples from different clones in the plane of the first two canonical roots.

We should note that the studied problem is at the initial stage of investigation challenging us with a number of questions, such as:

a) reproducibility of highly inheritable variable components related to environmental conditions and genetic composition of populations;

b) studying the genetic «architecture» of the variables-components; ideally, localization of concrete genes responsible for the genetic variability of the given complex traits;

c) development of the algorithm of the targeted search for the set of variables that will provide highly inheritable components.

In particular, one of the possible ways to solve the last problem is to include into the analyzed set a number of variables, with some of them regarded as background ones in relation to the rest (Dragavtsev, 2002; Naumova et al., 2005). Another way is to use a number of variables characteristic of the environment of the analyzed set of trees (microrelief, soil fertility etc.).

Also, of great interest is a question opposite to the one we tried to answer in the present study: the search of variable components with zero heritability and maximal dependence on the environmental factors, such as moisture, soil fertility, etc.). We are planning to try and give answers to these questions in our further studies.

CONCLUSION

As a result of the study, we developed a procedure to reveal highly inheritable – variable components using of discriminant analyses. The procedure applied for the analysis of the data on elemental composition of needles from the clonal population of Scots pine established on the long-term field experiment showed broad-sense heritability H^2 , i. e. the contribution of clones into the total data variance, to exceed 90 % for some of the components ed. The results suggest the importance of multivariate statistics for revealing genetic marker traits in the populations of woody plants for forest genetic and selection, in particular for assessing genetic differentiation among populations, identification of the best genotypes via their phenotypes, etc.

Authors thank G. A. Kovalskaya for assistance in analyzing the elemental composition of pine needles. The study was supported by the SBRAS Integration Grant No. 5.23 and RFBR Grant No. 07-04-01714-a.

REFERENCES

- Baryshev V. B., Kolmogorov Yu. P., Kulipanov G. N., Skrynsky A. N. Rentgenofluorestsenty elementny analiz s ispol'zovaniem sinkhrotronnogo izlucheniya (X-ray fluorescence elemental analysis using synchrotron radiation) // Zhurnal Analiticheskoy Khimii (J. Anal. Chem.) 1986. V. 41. N. 3. P. 389–401 (in Russian with English abstract).
- Baryshev V. B., Kulipanov G. N., Skrynsky A. N. Handbook of synchrotron radiation. V. 3 / G. Brown & D. Moncton (Eds.). Amsterdam: Elsevier, 1991. 639 p.
- Dragavtsev V. A. Ekologo-geneticheskiy skринing genofonda i metody konstruirovaniya sortov sel'skokhozyaystvennykh rasteniy po urozhaynosti, ustoychivosti i kachestvu. Dopolnennoe izdanie (Ecological-genetic screening of the gene pool and methods for constructing varieties of agricultural plants by yield, resistance and quality. Updated edition). Saint-Petersburg: Vserossiyskiy Institut Rasteniyevodstva (All-Russian Institute for Plant Cultivation), 2002. 41 p. (in Russian).
- Efimov V. M., Kovaleva V. Y., Markel A. L. A new approach to the study of genetic variability of complex characters // Heredity. 2005. V. 94. P. 101–107.
- Fisher R. A. The correlation between relatives under the supposition of mendelian inheritance // Trans. R. Soc. Edinburgh. 1918. V. 52. P. 399–433.
- Fisher R. A. The use of multiple measurements in taxonomic problems // Ann. Eugen. 1936. V. 7. Iss. 2. P. 179–188.
- Fisher R. A., Yates F. Statistical tables for biological, agricultural and medical research. 6th Ed. Edinburgh and London: Oliver & Boyd, 1963. 146 p.
- Hazel L. N. The genetic bases for constructing selection indexes // Genetics. 1943. V. 28. P. 476–490.
- Kendall M. G., Stuart A. The advanced theory of statistics. V. 2. Inference and relationship. N. Y.: John Wiley & Sons, 1979. 521 p.
- Klingenberg C. P., Leamy L. J. Quantitative genetics of geometric shape in the mouse mandible // Evolution. 2001. V. 55. Iss. 11. P. 2342–2352.
- Kovaleva V. Yu., Efimov V. M., Faleev V. I., Galaktionov Yu. K., Abramov S. A., Nazarova G. G., Muzyka V. Yu., Epifantseva L. Yu. Role of genetic factors in landscape-geographic variation of the water vole (*Arvicola terrestris* L.) // Rus. J. Ecol. 2006. V. 37. N. 6. P. 431–436 (Original Russian Text © V. Yu. Kovaleva, V. M. Efimov, V. I. Faleev, Yu. K. Galaktionov, S. A. Abramov, G. G. Nazarova, V. Yu. Muzyka, L. Yu. Epifantseva. Publ. in Ekologiya. 2006. N. 6. P. 469–474).
- Kutsenogiy K. P., Makarikova R. P., Milyutin L. I., Naumova N. B., Tarakanov V. V., Chankina O. V. The use of X-ray fluorescence analysis with synchrotron radiation to measure elemental composition of phytomass and soils // Nuclear Instruments and Methods in Physics Research. Section A: Accelerators, Spectrometers, Detectors and Associated Equipment. Proc. XVI Int. Synchrotron Radiation Conf. SR 2006, 10–14 July, 2006, Novosibirsk, Russia / M. Z. Kuzin, V. F. Pindyurin (Eds.). Elsevier BV., 2007. V. 575. Iss. 1–2. P. 214–217.

- Lande R. Quantitative genetic analysis of multivariate evolution, applied to brain: body size allometry // *Evolution*. 1979. V. 33. N. 1. Pt. 2. P. 402–416.
- Naumova N. B., Makarikova R. P., Tarakanov V. V., Kutsenogiy K. P., Milyutin L. I., Chankina O. V., Savchenko T. I. Vliyaniye razlichnykh klonov sosny obyknovennoy na svoystva dernovo-podzolistoy pochvy (Influence of various clones of Scots pine on the properties of sod-podzol soil) // *Sibirskiy Ekologicheskiy Zhurnal (Sib. J. Ecol.)*. 2005. V. 12. N. 5. P. 915–924 (in Russian with English abstract).
- Smith H. F. A discriminant function for plant selection // *Ann. Eugen.* 1936. V. 7. Iss. 3. P. 240–250.
- Tarakanov V. V., Chankina O. V., Kutsenogiy K. P., Naumova N. B., Makarikova R. P., Milyutin L. I., Rogovtsev R. V., Efimov V. M. Influence of geographic populations on the elemental composition of pine phytomass and soil // *J. Surface Investigation X-ray, Synchrotron Neutron Techn.* 2011. V. 5. Iss. 6. P. 1091–1097 (Original Russian Text © V. V. Tarakanov, O. V. Chankina, K. P. Kutsenogiy, N. B. Naumova, R. P. Makarikova, L. I. Milyutin, R. V. Rogovtsev, V. M. Efimov. Publ. in *Poverkhnost', Rentgenovskie, Sinkhrotronnye i Neitronnye Issledovaniya*. 2011. N. 11. P. 72–78).
- Tarakanov V. V., Milyutin L. I., Kutsenogiy K. P., Kovalskaya G. A., Ignat'ev L. A., Samsonova A. E. Elementny sostav khvoi v raznykh klonakh sosny obyknovennoy (The elemental composition of needles in different Scots pine clones) // *Lesovedenie (Rus. J. For. Sci.)*. 2007. N. 1. P. 28–35 (in Russian with English abstract).
- Wright S. Systems of mating // *Genetics*. 1921. N. 6. P. 111–178.

УДК 630*165.6+631.523.4

ВЫСОКОНАСЛЕДУЕМЫЕ ПРИЗНАКИ-КОМПОНЕНТЫ В КЛОНОВОЙ ПОПУЛЯЦИИ СОСНЫ

В. М. Ефимов¹, В. В. Тараканов², Н. Б. Наумова³, В. Ю. Ковалева⁴, **К. П. Куценогий**⁵

¹ *Институт цитологии и генетики СО РАН
630090, Новосибирск, просп. Академика Лаврентьева, 10*

² *Западно-Сибирское отделение Института леса им. В. Н. Сукачева СО РАН –
филиал Федерального государственного бюджетного научного учреждения
«Федеральный исследовательский центр «Красноярский научный центр
Сибирского отделения Российской академии наук»
630082, Новосибирск, ул. Жуковского, 100/1*

³ *Институт почвоведения и агрохимии СО РАН
630090, Новосибирск, просп. Академика Лаврентьева, 8/2*

⁴ *Институт систематики и экологии животных СО РАН
630091, Новосибирск, ул. Фрунзе, 11*

⁵ *Институт химической кинетики и горения СО РАН
630090, Новосибирск, ул. Институтская, 3*

E-mail: vmefimov@ngs.ru, tarh012@mail.ru, nnaumova@mail.ru, vkova@ngs.ru, koutsen@kinetics.nsc.ru

Поступила в редакцию 07.09.2019 г.

Разработана процедура выявления высоконаследуемых признаков-компонент, заключающаяся в поэтапном использовании метода главных компонент и дискриминантного анализа. При анализе изменчивости по элементному составу хвои на клоновой плантации сосны обыкновенной *Pinus sylvestris* L. коэффициенты наследуемости в широком смысле слова H^2 (доли влияния клонов в общей дисперсии признаков) по некоторым признакам-компонентам могут превышать 90 %. Подчеркивается актуальность применения методов многомерной статистики для идентификации в популяциях древесных видов генетически маркерных признаков, что может найти применение при решении таких задач лесной генетики и селекции, как оценка уровня генетической дифференциации популяций, идентификация лучших генотипов по фенотипам, и др.

Ключевые слова: *Pinus sylvestris* L., химические элементы, наследуемость, многомерный анализ.