APPLICATION OF VARIOUS STATISTICAL METHODS TO ANALYZE GENETIC DIVERSITY OF AUSTRIAN (*Pinus nigra* Arn.) AND SCOTS PINE (*Pinus sylvestris* L.) Based On Protein Markers

Aleksandar LUČIĆ[1], Vasilije ISAJEV[2], Ljubinko RAKONJAC[1],
Milan MATARUGA[3], Vojka BABIĆ[4], Danijela RISTIĆ[4]
and Snežana MLADENOVIĆ DRINIĆ[4]

[1]Institute of Forestry, Belgrade, Serbia,[2]Faculty of Forestry, Belgrade, Serbia,
[3]Faculty of Forestry, Banja Luka, Bosnia and Hercegovina, [4]Maize Research
Institute, Zemun Polje, Belgrade, Serbia

Lučić A., V.Isajev, Lj.Rakonjac, M.Mataruga, V.Babić, D.Ristić
and S. Mladenović Drinić (2011): *Application of various statistical methods
to analyze genetic diversity of Austrian (Pinus nigra Arn.) and Scots pine
(Pinus sylvestris L.) based on protein markers.* - Genetika, Vol 43, No. 3,
477-486.
This paper presents the results of studies on protein polymorphism
in seeds of Scots pine (*Pinus sylvestris L.*) and Austrian pine (*Pinus nigra
Arn.*) as the most important economic species of the genus *Pinus* in Serbia.
Polymorphism of protein markers was determined in selected genotypes
originating from seven populations (Scots pine) and six populations
(Austrian pine). Analysis of protein markers was performed using two
statistical methods, NTSYS and correspondence analysis. Both methods
give the same arrangement of the analyzed populations, whereby, because
of a different view of genetic distances, they can and should be combined,

_____

Corresponding author: Aleksandar Lučić, Institute of forestry, Belgrade, Serbia,
phone: ++ 381 63 8413101, aleksandar.lucic@gmail.com

enabling easier and more precise understanding of mutual relationships of the observation units.

Key words: Austrian pine, correspondence analysis, NTSYS, proteins, Scots pine

INTRODUCTION

For more than a century, species of the genus *Pinus* (Scots and Austrian pine) have been in the focus of forestry scientific and professional public. The interest in this species is primarily the result of their characteristics: high genetic potential and genetic variation, as well as the taxonomic complexity and plasticity of the species and implementation in the afforestation of a wide range of site conditions. Its protective and reclamation function on the most severe terrains is the consequence of its extraordinarily low site demands. The fact that pines thrive over large areas affected by erosion processes, or on completely degraded and denuded areas, on which they produce excellent results, makes them economically the most important species used in forestry. All the above qualities, in addition to a vast natural range of distribution and a disjunctive range, led to their intensive introduction to the sites outside their natural range, which resulted in a great number of subspecies, varieties and transition forms. The result of this natural variation is a high genetic potential which provides the base and the potential for the successful breeding of the species.

Taking into account the advancement of the concepts of nursery production, i.e. the tendency of abandoning the planting material production at the level of the species (ISAJEV *et al., 1998*), and in the aim of the most efficient and fast realisation of the tasks, it is necessary to enhance the study of the genetic potential of Scots and Austrian pine population, because its directed implementation can improve the production of seed and planting material, and in this way also the success of afforestation and establishment of specific-purpose plantations of this species.

Consistent with the above problem which is assigned to modern forestry, the objective of the present study is to research and identify the population of Scots and Austrian pine in Serbia, as one of the most important commercial species, irreplaceable in afforestation.

In general, conifers are characterised by a very high level of genetic variability (HAMRICK, 1979), and the potential factors of which are: a) long life, free pollination with high fertility, and b) divergent selection for macro-micro geographical adaptation. According to SCALTSOYIANNES (1994), pine is characterised by high total variation due to intra-population genetic variability, which points to a significant genetic differentiation in local populations.

Seed proteins as genetic markers have been widely used in studying the genome of Austrian pine (MATARUGA, 2007; LUČIĆ *et al*. 2008). In these studies were explored the systematic affiliation of some provenances and genetic variation.

By analyzing the protein complex of Scots pine seeds LUČIĆ (2011) in the gene-environmental research, concludes that there is significant differentiation of Scots pine populations in Serbia.

WANG (1991) studied the isozymes of Scots pine of three northerly Swedish populations and four Chinese populations. SZMIDT (1992) in the analysis of isozymes of the above populations included the four population of Scots pine from Turkey.

PLUS-GLOWACKI (1994) analyzed 13 populations of Scots pine encompassing almost all major sites in Europe and East Asia. These authors suggest the existence of significant genetic differentiation of analyzed populations.

So far, in the analysis of protein complex often were used clusters obtained using the method of NTSYS (ROHLF, 2000). Cluster analysis is very useful for cultivators of plants and it is much used in genetic researches to define groups by their similarity-relatedness. However, this analysis has a drawback, meaning that there is no insight into relationships of units of observation-populations within and between clusters. More specifically, primarily for this analysis is arrangement of populations by similarity, where the distances of populations themselves are vaguely shown. One of the important elements for solving this is the experience of analysts in the context of research objective.

Correspondence analysis (GREENACRE, 1988) is a supplementary analysis to genetic distances and dendrograms, which gives a global view of relations between populations. It has the advantage because it does not assume that test units belong to certain groups, but more accurately represents the continuous variability. This type of analysis tends to give similar results like dendrograms of cluster analysis as it is expected from theory. At the same time correspondence analysis is more informative and more precise than dendrograms especially where there is considerable genetic exchange between geographically close populations.

MATERIALS AND METHODS

After the introduction of natural range and culture of Scots and Austrian pine in Serbia, it was decided that seed stands occurred in natural populations, seed stands occurred in older cultures and significant natural populations be the objects of research.

On the basis of their morphological (phenotypic) and physiological characteristics, seven populations of Scots pine (Table 1) and six populations of Austrian pine (Table 2) were selected for the research. Spatial distribution of the populations is shown in Figure 1 and 2.

In each population at the selected localities, the seeds were collected by felling 8 to 15 test trees. Averagely 100 cones were collected from each marked tree, and the seeds were processed and grouped by trees. The bulk sample for biochemical analysis was made by taking 15 seeds from each analysed pine tree. Protein isolation was performed according to *Wang et al.* (1994) and SDS-polyacrylamide gel electrophoresis following the method of *Leammli* (1970). Quantitative differences (number and distribution) in protein fractions were determined based on the electrophoregrams.

Similarity coefficients were calculated using *Sokal* and *Michener* (SM) (1958) method.

Sokal and Michener   **$GS_{ij} = a+d/a+b+c+d$**

Where:

a - band presence in both genotype *i* and *j* (1.1)
b - band presence in genotype *i* and absence in genotype *j* (1.0)
c - band presence in genotype *j* and absence in genotype *and* (0.1)
d - band absence in both genotype *i* and genotype *j* (0.0)

Analyses of genetic similarity coefficients were computed using NTSYS-PC software and correspondence analysis.

Figure 1: The studied populations of Scots pine   Figure 2: The studied populations of Austrian pine

Table 1. The studied populations of Scots pine (Pinus sylvestris L.) and Austrian pine (Pinus nigra Arn.)

	Population of Scots pine	Population of Austrian pine
I	MU „Dubočica Bare" 60 a.; FE „Golija" Ivanjica	MU "Divan-Breze", 27a.; FE "Šumarstvo" Raška
II	MU "Šargan", 25 b.; FE "Užice" Užice	MU "Divan-Lokva", 21a.; FE "Šumarstvo" Raška
III	MU "Radočelo-Crepuljnik" 4b.; FE "Stolovi" Kraljevo	MU "Crni vrh Ljeskovac", 69 c.; FE "Prijepolje" Prijepolje
IV	MU "Jablanička reka" 33 d.; FE "Rasina" Kruševac	MU "Šargan", 22 b.; FE "Užice" Užice
V	MU "Bukovik-Aleksinački", 23b, 24g.; FE "Niš", Niš	MU "Goč-Gvozdac", 92b.; Faculty of Forestry-Beograd-Goč
VI	MU "Kaluđerske bare",1a.; NP Tara, Bajina Bašta	MU "Studenica-Polumir", 17c, 26a.; FE "Stolovi" Kraljevo
VII	MU "Zlatar I", 22a.; FE "Prijepolje", Prijepolje	-

RESULTS AND DISCUSSION

Based on the electrophoregram were calculated the genetic similarity coefficients between studied populations, Table 2.

Table 2. The similarity coefficients after Sokal and Michener in the Scots and Austrian pine

Scots pine / Austrian pine	The similarity coefficients after Sokal and Michener						
	I	II	III	IV	V	VI	VII
I		0.73	0.60	0.67	0.53	0.60	0.73
II	0.91		0.47	0.53	0.53	0.87	0.60
III	0.77	0.77		0.80	0.67	0.47	0.87
IV	0.64	0.64	0.68		0.60	0.67	0.80
V	0.82	0.82	0.95	0.64		0.53	0.53
VI	0.95	0.95	0.82	0.68	0.86		0.60

        The genetic similarity coefficients for Austrian pine occur in the range of 0.64 to 0.95, while the genetic similarity coefficients for Scots pine from 0.47 to 0.87. This indicates the existence of greater genetic variability of Scots pine than the Austrian pine in Serbia.

        In order to determine the genetic similarity/distance among the study genotypes were performed analysis in NTSYS-pc software (ROHLF, 2000) and correspondence analysis.

        The matrices of genetic similarity coefficients presented through the shape of the dendrogram (NTSYS) or the graphics (correspondence) provide better visibility of similarity or distance of populations themselves with respect to the coefficients presented in Table 2. However, the reading of the very values from the charts is much more difficult, which justifies the practice that parallel with them are used tables of similarity coefficients.

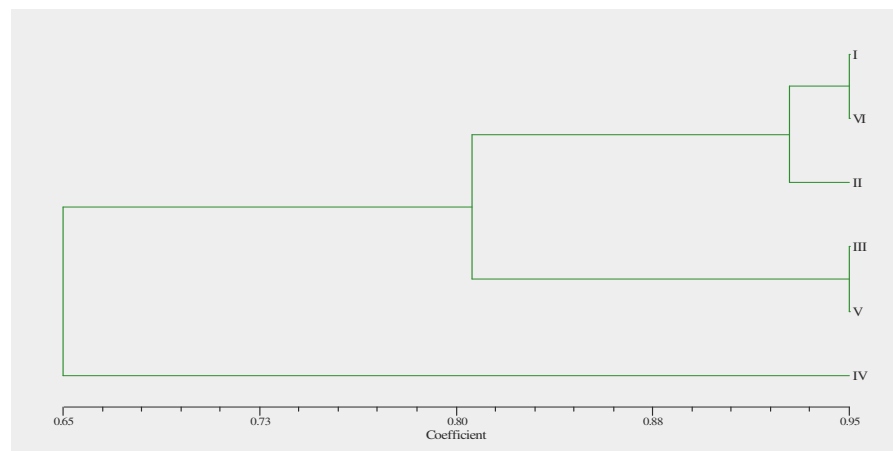**Austrian pine (*Pinus nigra* Arnold)**

Diagram 1. NTSYS-dendrogram of cluster analysis of Austrian pine populations after Sokal
          and Michener

        NTSYS dendrogram analysis-*Diagram 1*, shows clearly that the populations **I** and **VI**, as well **III** and **V** occur at the smallest genetic distances. Population **II** is more firmly tied to populations **I** and **VI** than to **III** and **V**. Austrian pine on Sargan-**IV** is at the largest genetic distance with respect to other populations.

        The chart of correspondence analysis in three dimensions *Diagram 2*, shows clearly that the populations **I**, **II** and **VI**, as well as **III** and **V** occur at the smallest genetic distances. Population **IV** is at the largest genetic distances with respect to other populations.
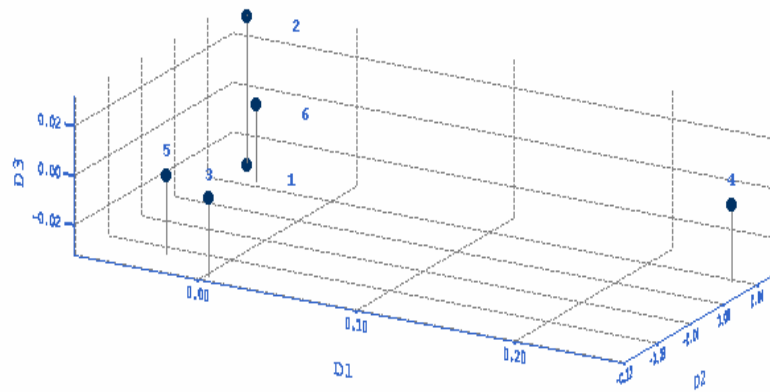
Diagram 2: Correspondence analyses of genetic similarity of Austrian pine analyzed populations in three dimensions after Sokal and Michener

With comparative analysis of both charts, we can see that we got almost the same results in terms of similarity and difference of the analyzed populations. Slightly higher genetic similarity in both charts among the populations **I**, **II** and **VI** was expected due to the small geographical distance of these populations. The specificity of the Austrian pine population on Sargan **IV** in comparison with other populations, confirms a study of WHELLER (1976) who recommends for breeders to look for seeds with characteristics that were collected close to the site of Sargan in Serbia.

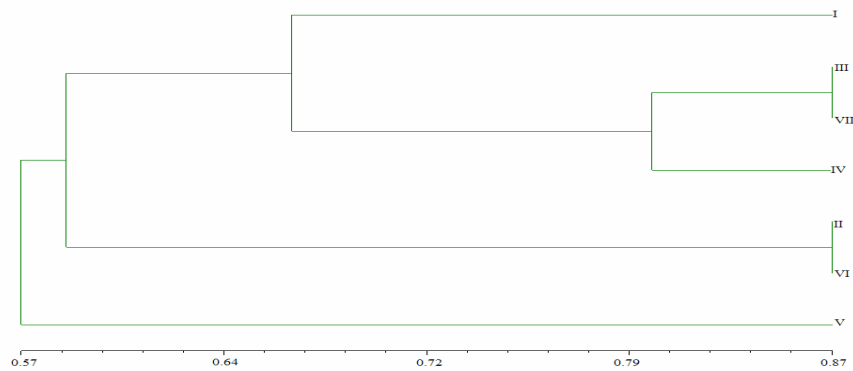**Scots pine (*Pinus sylvestris* L.)**

Diagram 3. NTSYS-dendrogram of cluster analysis of Scots pine populations after Sokal and Michener

NTSYS dendrogram analysis-*Diagram 3*, shows clearly that the populations **III** and **VII**, as well as **II** and **VI** occur at the smallest genetic distances. Populations **IV** and **I** are more firmly tied to populations **III** and **VII** than to **II** and **VI**, whereby the population **I** is characterized by greater distance. Population **V** is at the largest genetic distance with respect to other populations.
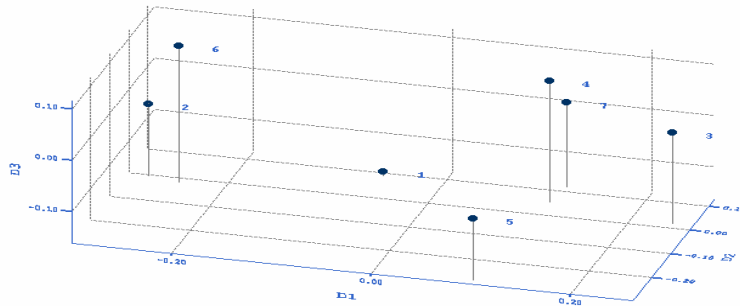

Diagram 4: Correspondence analyses of genetic similarity of Scots pine analyzed populations in three dimensions after Sokal and Michener

The chart of correspondence analysis in three dimensions *Diagram 4*, shows clearly that the populations **III, VII** and **IV**, as well as **II** and **VI** occur at the smallest genetic distances. Populations **I** and **V** are at the largest genetic distance with respect to other populations whereby it is most prevalent in the population **V**.

As that was the case with the Austrian pine, both charts were almost identical in terms of similarity of the analyzed populations. With comparative analysis of both used methods, it can be concluded that the grouping of populations **III**, **VII** and **IV** in the zone of higher genetic proximity is marked with their geographical proximity. The same case is with populations **II** and **VI**. On the other hand, a little bit higher genetic distance of population **I** is specific in comparison with the other populations. This is explained not by a large geographic distance, but extreme isolation of population **I** on the Pester plateau and its gravitation to forests of Montenegro. Population **V** shows the greatest genetic distance compared to other populations. Exploring what is causing it, it was found out that this population was grown about 90 years ago on sample plots of the first forestry school in Serbia in Aleksinac, which was sponsored by Belgium and known as the "Belgian forest school". Large genetic distance takes us to the assumption that this population originated from seeds outside of Serbia, perhaps from Belgium.

CONCLUSION

The results obtained using different methods (NTSYS and correspondence analysis) in the analysis of protein markers, indicating the use of both methods in similar studies, as compatible. Diagrams of both methods are characterized by the same arrangement of the analyzed populations, whereby, because of a different view, they can and should be combined, enabling easier and more precise understanding of mutual relationships of the observation units.

Also, the results of this study suggest:

• The existence of significant genetic variability of Scots and Austrian pine populations.

• The strong influence of geographic distance and isolation of populations on genetic differentiation.

Also, the results provide a basis for better understanding of very variable and specific intraspecific genetic structure of Scots and Austrian pine in Serbia, which is the initial tool in planning of conservation processes and breeding of species.

REFERENCES

GREENACRE, M. (1988): Correspondence analysis of multivariate categorical data by weighted least-squares. Biometrika Vol. *75*, No. 3, 457-467.

HAMRICH, J.L., Y.B. LINHART, J.B. MITTON (1979): Relationships between life history characteristics and electrophoretically detectable genetic variation in plants. Annual review of ecology and systematics *10*: pp.173-200

ISAJEV, V., A. TUCOVIĆ, M. MATARUGA (1998): Advancement of technology of degrades site afforestation. Conference 'Some problems of forests and water and possible solutions'. JP Srbijašume, Belgrade (in Serbian)

LAEMLLI, V. K. (1970): Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature 227*: pp. 680 - 685.

LUČIĆ A., A. NIKOLIĆ, S. MLADENOVIĆ DRINIĆ, V. ISAJEV, V. LAVADINOVIĆ (2008): Genetic Characterisation Of Genotypes Of Austrian Pine *(Pinus Nigra* Arlnold) Populations Using Protein Markers. Genetika, Vol.*40*, No.2. Pp: 157-168

LUČIĆ A., V. ISAJEV, R. CVETIĆANIN, LJ. RAKONJAC, M. NOVAKOVIĆ, A. NIKOLIĆ, S. MLADENOVIĆ DRINIĆ (2011): Interpopulation genetic-ecological variation of Scots pine *(Pinus sylvestris L.)* in Serbia. Genetika, Vol.*43*, No.1. Pp: 1-18

MATARUGA, M., V. ISAJEV, K. KONSTANTINOV, S. MLADENOVIĆ-DRINIĆ, V. DANIČIĆ (2007): Proteins in seed and seedlings of selected Austrian pine (*Pinus nigra* Arnold) trees as genetic markers tolerant to drough. Genetika, Vol. *39*, No. 2, 259 -272.

PRUS-GROWACKI, W., E. BERNARD (1994): Allozyme variation in populations of *Pinus sylvestris* L. from a 1912 provenance trial in Pulawy (Poland). Silvae Genetica *43*:pp.132-138

ROHLF, F. J. (2000): NTSYS-pc. Numerical taxonomy and multivariate analysis system. Version 2.0 Exeter Software, Setaket, N.Y.

SCALTSOYIANNES, A., R. ROHR, K. PANETSOS (1994):Allozyme frequency distribution in five European populations of black pine (*Pinus nigra* Arnold) Silvae genetica *43*. 1 pp. 20-25.

SOKAL, R.R., C.D. MICHENER (1958): A statistical method for evaluating systematic relationships. *Univ Kans Sci Bull* *38*: pp. 1409-1438

SZMIDT, A. E., X. R. WANG, M. Y. LU (1992): Empirical assessment of allozyme and RAPD variation in *Pinus sylvestris* (L.) using haploid tissue analysis. Heredity *76*: pp. 412-420

WANG, C., K. BIAN, H. ZHANG, Z. ZHOU, J. WANG (1994): Polyacrylamide gel electrophoresis of salt soluble proteins for maize variety identification and genetic purity assessment. Seed Sci. Tech. *21* (51).

WANG, X-R., A. SZMIDT, D. LINDGREN (1991): Allozyme differentiation among populations of *Pinus sylvestris* from Sweden and China. Hereditas *114*: pp. 219-226

WHEELER, N. C., H. B. KRIEBEL, C. H. LEE, R. A. READ, J. W. WRIGHT (1976). 15-year performance of European black pine in provenance test in North Central United States. *Silv. Gen. 25,* 1-6.

PRIMENA RAZLIČITIH STATISTIČKIH METODA ZA ANALIZU GENETIČKOG DIVERZITETA CRNOG (*Pinus nigra* ARN.) I BELOG BORA (*Pinus sylvestris* L.) NA OSNOVU PROTEINSKIH MARKERA

Aleksandar LUČIĆ[1)] Vasilije ISAJEV[2], Ljubinko RAKONJAC[1],
Milan MATARUGA[3], Vojka BABIĆ[4], Danijela RISTIĆ[4]
i Snežana MLADENOVIĆ DRINIĆ[4]

[1]Institut za šumarstvo, Beograd, Srbija, [2]Šumarski fakultet, Beograd, Srbija,
[3]Šumarski fakultet, Banja Luka, Bosna and Hercegovina, [4]Institut za kukuruz,
Zemun Polje, Beograd, Srbija

I z v o d

U radu su prikazani rezultati proučavanja polimorfizma proteina u semenu belog (*Pinus sylvestris* L.) i crnog bora (*Pinus nigra* Arn.) kao najvažnijih ekonomskih vrsta iz roda *Pinus* u Srbiji. Polimorfizam proteinskih markera utvrđen je u odabranim genotipovima poreklom iz sedam populacija (beli bor) i šest populacija (crni bor). Analiza proteinskih markera vršena je upotrebom dva statistička metoda, NTSYS i korespodencijalna analiza.

Dobijeni rezultati upotrebom različitih statističkih metoda (NTSYS i korespodencionalna analiza) u analizi proteinskih markera, upućujući na upotrebu oba metoda u sličnim istraživanjima, kao kompatibilnih. Grafikoni obe metode se karakterišu istim rasporedom analiziranih populacija, pri čemu se usled različitog prikaza trebaju kombinovati, što omogućuje lakše i preciznije sagledavanje međusobnih odnosa jedinica posmatranja.

Takođe, dobijeni rezultati u ovom radu ukazuju na:

* Postojanje značajne genetičke varijabilnosti populacija belog i crnog bora.
* Snažan uticaj geografske distance i izolovanosti populacija na genetičku izdiferenciranost.

Ujedno, rezultati istraživanja pružaju osnovu za bliže upoznavanje veoma varijabilne i specifične unutarvrsne genetičke strukture belog i crnog bora u Srbiji, što predstavlja polazno sredstvo u planiranju procesa konzervacija i oplemenjivanja vrsta.