J Mol Evol (1997) 45:332–336

JOURNAL OF MOLECULAR EVOLUTION

Restriction Endonuclease Analysis of Highly Repetitive DNA as a Phylogenetic Tool

V.V. Grechko,[1] L.V. Fedorova,[1] A.N. Fedorov,[2] S.Ya. Slobodyanyuk,[3] D.M. Ryabinin,[4] M.N. Melnikova,[5] A.A. Bannikova,[6] A.A. Lomov,[5] V.A. Sheremet'eva,[6] V.A. Gorshkov,[6] G.A. Sevostyanova,[7] S.K. Semenova,[4] A.P. Ryskov,[4] B.M. Mednikov,[5] I.S. Darevsky[8]

[1] Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Moscow 117984, Russia
[2] Institute of Molecular Genetics, Russian Academy of Sciences, Moscow 123182, Russia
[3] Institute of Limnology, Siberian Branch of the Russian Academy of Sciences, Irkutsk 664033, Russia
[4] Institute of Gene Biology, Russian Academy of Sciences, Moscow 117984, Russia
[5] Belozersky Institute of Physico-chemical Biology, Lomonosov Moscow State University, Moscow 118899, Russia
[6] Faculty of Biology, Lomonosov Moscow State University, Moscow 118899, Russia
[7] Lenin Pedagogical State University, Moscow 119243, Russia
[8] Zoological Institute, Russian Academy of Sciences, St. Petersburg, 199034, Russia

Received: 10 February 1997 / Accepted: 10 March 1997

Abstract. Multiple band patterns of DNA repeats in the 20–500-nucleotide range can be detected by digesting genomic DNA with short—cutting restriction endonucleases, followed by end labeling of the restriction fragments and fractionation in nondenaturing polyacrylamide gels. We call such band patterns obtained from genomic DNA ''taxonprints'' (Fedorov et al. 1992). Here we show that taxonprints for the taxonomic groups studied (mammals, reptiles, fish, insects—altogether more than 50 species) have the following properties: (1) All individuals from the same species have identical taxonprints. (2) Taxonprint bands can be subdivided into those specific for a single species and those specific for groups of closely related species, genera, and even families. (3) Each restriction endonuclease produces unique band patterns; thus, five to ten restriction enzymes (about 100 bands) may be sufficient for a statistical treatment of phylogenetic relationships based on polymorphisms of restriction endinuclease sites. We demonstrate that taxonprint analysis allows one to distinguish closely related species and to establish the degree of similarity among species and among genera. These characteristics make taxonprint analysis a valuable tool for taxonomic and phylogenetic studies.

Key words: Repetitive DNA sequences — Molecular phylogeny — Lacertidae — Mammalia — Taxonprint

Introduction

A study of repetitive DNA sequence variations promises to provide important information about genetic structure and dynamics of natural populations as well as determination of intra- and intertaxa relationships. A direct comparison of nucleotide sequences seems to be the most informative approach. However, sequencing is too expensive and time consuming for analysis of the vast numbers of individual DNAs often required for population and phylogenetic studies. Another traditional method of DNA analysis, restriction fragment length polymorphism (RFLP), is generally used for population studies, but its application as a taxonomic and phylogenetic tool was not developed because of difficulties in identifying taxa on the species level.

Earlier, several authors noticed the species specificity of highly repetitive sequences of genomic DNA and tried

Correspondence to: V.V. Grechko

to use this phenomenon for phylogenetic aims (Cooke 1975; Christie and Skinner 1979; Brown and Dover 1979, 1980; Elizur et al. 1982; Hembelen et al. 1982; Buckland 1983; Dandieu et al. 1984; Lima de Faria et al. 1984; Shubina and Mednikov 1986; Turner et al. 1991), but this approach has not been extensively applied. In this paper we have reexamined the problem on the basis of our own preliminary results, partly published in Russian journals, confirming the ideas about possible correlations between evolution of DNA repeats and speciation (Fedorov et al. 1992; Grechko et al. 1993; Potapov and Ryskov 1993).

We have demonstrated that the modified method designated ''DNA taxonprint'' (Fedorov et al. 1992) can be effectively used for revealing species—specific genomic markers in different taxonomic groups. These taxonprint markers are found to be helpful for evaluation of the degree of species similarity between species, genera, and families. More than 50 species from different taxa were studied.

Materials and Methods

Genomic DNA was purified from blood and other tissues using standard protocols with proteinase K digestion followed by phenol/chloroform extraction and then hydrolyzed with some restriction endonuclease. Instead of agarose gel fractionation and ethidium bromide staining, which have poor sensitivity, we have used an approach similar to that known as the mtDNA restriction mapping method (Brown 1980), which can also be applied to reveal some DNA repeats (Southern 1975; Donehower and Gillespie 1979). The ''sticky'' 3′—ends of the restriction fragments were labeled in a Klenow reaction with [α—^{32}P]dNTP and aliquots of the reaction mixture (about 0.1 μg per slot) were electrophoresed in 10% nondenaturing polyacrylamide gel in Tris—borate buffer, pH 8.3 (Sambrook et al. 1989). The gel was dried on the glass plate (Garoff and Ansorge 1981) and radioautographed.

Molecular data generated by taxonprints were analyzed by distance analysis (for lizards) and by parsimony analysis (for hedgehogs). The absence and presence of radioautograph bands with the same electrophoretic mobility were designated by 0 and 1, respectively. Two summary binary matrices based on lizard data (five restriction endonucleases) and on hedgehog data (nine restriction endonucleases) were calculated. The UPGMA cluster analysis (Sneath and Sokal 1973) was performed on the basis of pairwise D distances (Nei and Li 1979) with the NTSYS package, version 1.8 (Rohlf 1993). The construction of the parsimony tree and its analysis were performed by the program PHYLIP, version 3.5 (Felsenstein 1993), according to Bannikova et al. (1995).

Results

Figure 1 shows a typical taxonprint picture obtained after *Hinf*I restriction endonuclease digestion of several individual human DNA samples from representatives of populations of different world races (Table 1). Eight from the 29 total samples are presented. All these DNA samples give identical and enzyme-specific taxonprints. No differences between DNA samples were observed

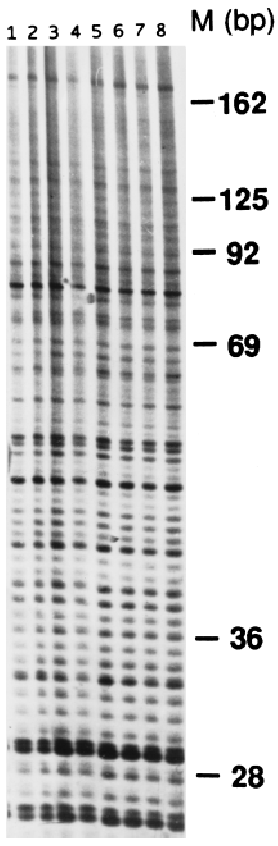

Fig. 1. *Hinf*I taxonprints of DNA of several human individuals from populations of different race. Mongoloids: Vietnamese *(1, 2),* Koryacs *(6),* Evenks *(7, 8),;* Negroids: Blacks from Central Africa *(3),* Blacks from Madagascar *(5);* Europeoids: Georgians *(4).* Each population is represented by five to six individuals (total number 29, all of them giving taxonprints identical to those shown here). *M*—markers in base pairs (pBR322 digested with *Msp*I).

Table 1. Species studied

Erinaceidae: *Erinaceus europaeus, E. concolor, E. amurensis, E. dauuricus; Hemiechinus auritus, H. hypomelas.* Talpidae: *Talpa europea; Mogera robusta.* Cottidae: *Cottus kessleri; Cottocomephorus inermis; Paracottus kneri.* Asprocottidae: *Asprocottus herzensteinii.* Abissocottidae: *Limnocottus megalops.* Comephoridae: *Comephorus dibowskii.* Lacertidae: *Lacerta mixta, L. valentini, L. portschinskii, L. raddei, L. nairensis, L. armeniaca, L. dahli, L. rostombekovi, L. uzzelli, L. rudis, L. saxicola, L. caucasica, L. praticola, L. derjugini, L. agilis, L. strigata, L. viridis; L. vivipara; Podarcis taurica, P. muralis; Eremias velox; Ophisops elegans.* Teiidae: *Cnemidophorus uniparens.* Muridae: *Mus musculus* (populations from Moscow, Cuba, Kazakhstan, Uzbekistan, Peru, and BalbC); *Apodemus sylvaticus* (populations from Moscow Region, East and West Caucasus, and Bulgaria); *A. agrarius* (populations from Moscow and Novgorod Regions, East Mongolia, and Tuva). Cricetidae: *Phodopus sungorus, Ph. campbelli, Ph. roborovski* (populations from East Mongolia and Tuva). Bovidae: *Bos taurus, B. indicus; Bison bison, Bis. bonansus; Ovis orientalis, O. ammon;* domestic sheep (six pedigrees). Cervidae: *Capreolus capreolus; Cervus nippon, C. elephus.* Bombicidae: *Bombyx mori* (clones ''Bukhara,'' ''Margelanskaya,'' Ukrainian'' 14, 15, and New, ''Mereffa'' 6 and 7, parthenoclones PC9, 4P29, and PCFbxFb7). Humans: Mongoloids (Vietnamese, Koryacs, Evenks); Europeans (Georgians); Negroids (from Central Africa and Madagascar)

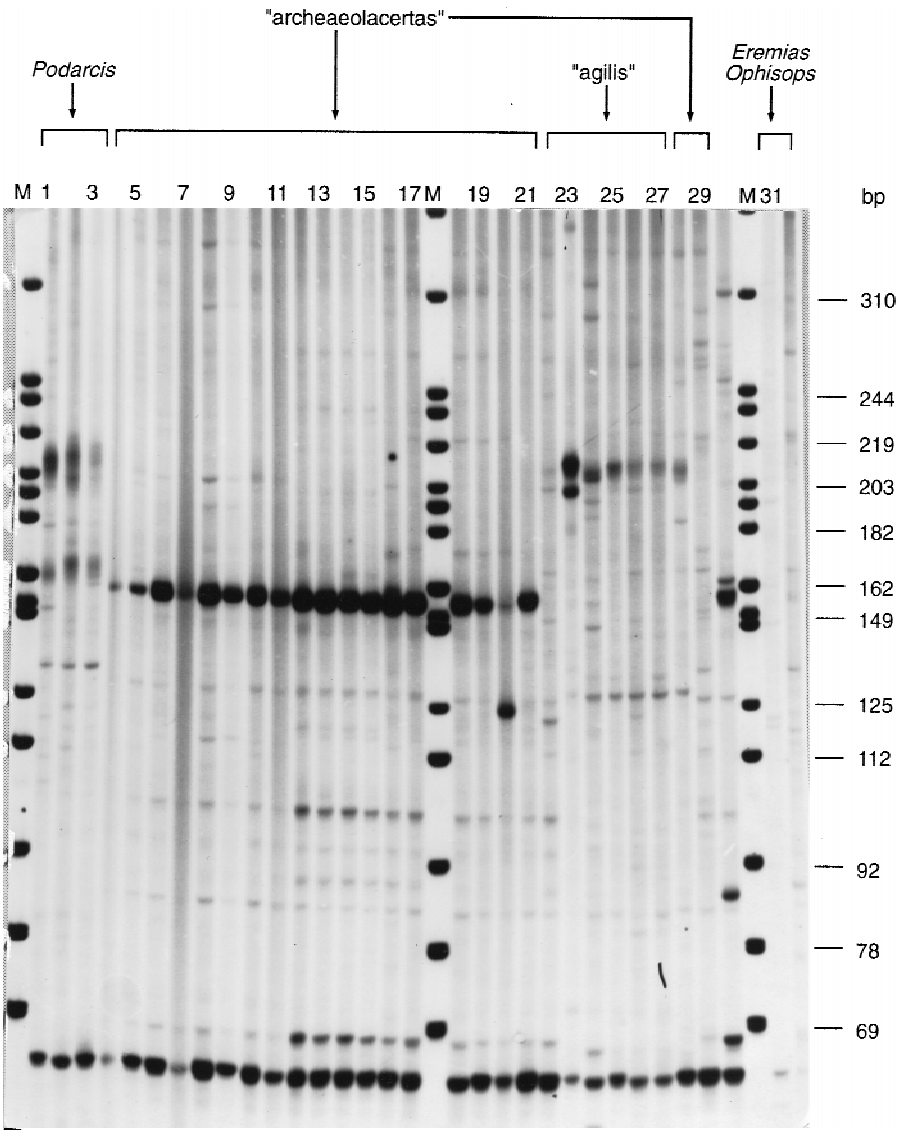

Fig 2. *Taq*I taxonprints of family Lacertidae lizards DNA. *1—Podarcis taurica; 2, 3—P. muralis; 4, 5—Lacerta mixta; 6, 7—L. valentini; 8, 9—L. portschinskii; 10—L. armeniaca* (parthenoclone); *11—L. dahli* (parthenoclone); *12—L. raddei* (Egegnadzor); *13—L. raddei* (Gosh); *14—L. raddei* (Chosrov); *15—L. nairensis; 16—L. unisexualis* (parthenoclone); *17—L. rostombekovi* (parthenoclone); *18—L. rudis chechenica; 19—L. rudis obscura; 20—L. saxicola lindholmi; 21—L. saxicola darevskii; 22—L. caucasica daghestanica; 23—L. strigata; 24—L. viridis; 25—L. agilis agilis; 26—L. agilis boemica; 27—L. agilis chersonensis; 28—L. vivipara; 29—L. praticola; 30—L. derjugini; 31—Eremias velox; 32—Ophisops elegans.*

when four other endonucleases (*Taq*I, *Csp*6I, *Msp*I, and *Bsu*RI) were tested.

Our study of individual DNA samples from different populations of lizards, hedgehogs, moles, mice, hamsters, the silk worm *Bombyx mori*, and other taxa (see Table 1), revealing species specificity also, confirmed the taxonprint's intraspecies identity. This allows one to use small numbers of individuals for species comparisons and also does not require any intrapopulational statistical analysis.

Figure 2 shows *Taq*I taxonprints of Caucasian rock lizards species (fam. Lacertidae, genus *Lacerta*, see Table 1). Among them are five parthenogenetic clones and five bisexual species which are supposed to be the parental species in the hybridogeneous origin of these parthenoclones (lanes 4–17) (Darevsky 1993). DNA of the representatives of other *Lacerta* (lanes 18–30) and other genera are shown also. A substantial similarity is seen within the parthenogenetic and some bisexual species groups. Another group of similar taxonprints is formed by ''green'' lizards (lanes 23–28) and by two species of *Podarcis genus* (lanes 1–3). Other genera (lanes 31, 32) have only rare family—specific bands that are common for all *Lacerta.*

On the taxonprints obtained with *Msp*I, *Hinf*I, *Tru*9I, *Sau*3A, *Hin*6I, *Eco*130I, *Asu*I, *Csp*6I, *Eco*130I + *Hind* III, and *Eco*88I, some other differences and similarities were observed (not shown here). About 100 bands from some of these taxoprints were used for statistical treatment by the UPGMA method. The genetic distance tree obtained (Fig. 3A) corresponds in principle to general ideas of phylogenetic relationships within the family Lacertidae (Arnold 1989; Moritz et al. 1992; Murphy et al. 1996). A detailed description of this study is being prepared for publication.

Taxonprint analysis has been applied to the examination of six species of hedgehogs (20 specimens) and two species of moles (see Table 1). The taxonprints were

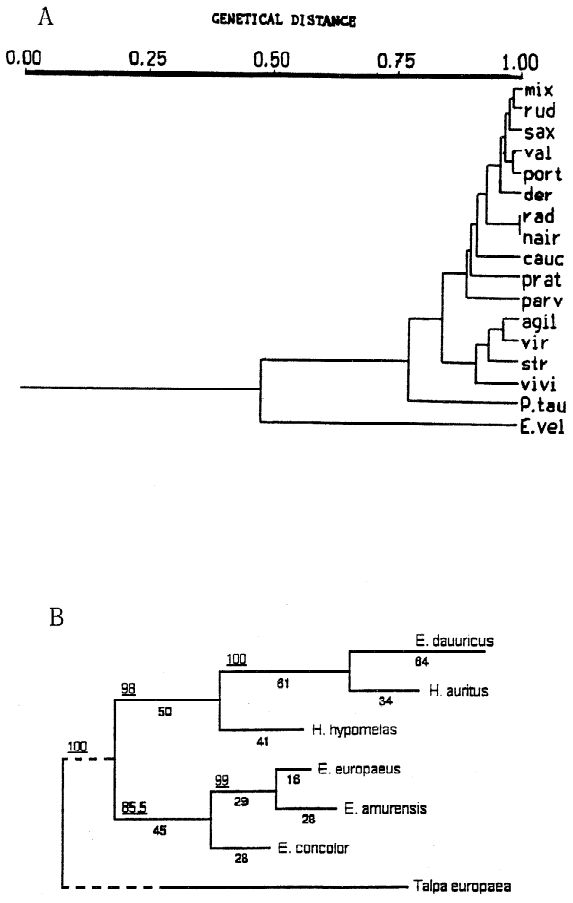

Fig. 3. Evolutionary trees produced by distance analysis of lizards **(A)** and by parsimony analysis of hedgehogs **(B).** *Underlined figures* are the values of bootstrap indices. *Bold figures* are the values of genetic distance according to Fitch (see Materials and Methods).

produced with nine restriction endonucleases (*Bsp*143I, *Csp*6I, *Taq*I, *Hinf*I, *Msp*I, *Eco*130I, *Bcn*I, *Bsu*RI, and *Sau*96I).

The relationships illustrated by the hedgehog dendrogram (Fig. 3B) correspond in the main to the ideas of zoologists about the division of hedgehogs into two genera, *Erinaceus* and *Hemiechinus,* but at the same time they allow us to offer some finer definitions. According to the taxonprint analysis, the Daurian hedgehog *(E. dauuricus)* appears to be closer to the eared hedgehog group, namely, to *H. auritus,* than to hedgehogs of the genus *Erinaceus.* Some common genus-specific bands were found in DNA hydrolyzates of Daurian and eared hedgehogs *(H. auritus)* obtained with all nucleases used. Genus-specific bands were observed only in some cases in comparisons with the long—prickled hedgehog *(H. hypomelas),* and only in rare cases when Daurian hedgehogs were compared with the other *Erinaceus* species.

Species specificity of three taxonprints of six fish species from the Cottoid complex from Lake Baikal reveals specific bands when restriction endonucleases *Taq*I, *Ava*II, and *Sty*I are used (not demonstrated). All species have at least one locus in common in each of the taxonprints, and five of them have four other loci in common. The species form two subgroups based on the similarity of all other loci. As we have studied only six from 29 species of this complex, we can only mention that UPGMA cluster analysis shows *Cottus kessleri* and *Cottocomephorus inermis* to be in closer relation than the latter species with other species of *Cottocomephorus* genus. This result correlates with mtDNA analysis of Grachev et al. (1992). The existing subdivision of the Cottoid complex into taxonomic groups has to be justified (Kirilchik and Slobodyanyuk 1997).

Discussion

Similar experiments carried out with representatives of some other species (see Table 1) confirmed that taxonprint markers are conservative species-specific characters and could be useful for the study of interspecies variability and similarity (Skurikhina et al. 1993; Potapov et al. 1994; Ryskov et al. 1994; Melnikova et al. 1995; Mednikov et al. 1995; Bannikova et al. 1995; Chelomina et al. 1995).

In essence, the DNA taxonprinting is presumed to represent the mapping of all repetitive regions of a genome considered together as a separate genome entity. What is being compared are the mutational events that can be detected by endonucleases used over a large part of genomic DNA. It is quite interesting that the distribution and number of mutations thus registered are constant among the members of a population. Therefore, this feature can be considered as a real species ''character.''

Thus the main result of this work is the demonstration that a modified RFLP method named ''DNA taxonprint'' can identify taxa at the species level. There is a good correlation between taxonprint band pattern similarity and the proximity of taxonomic groups based on morphological criteria. These characteristics, together with the lack of destinations among individuals and populations and the availability of statistically significant numbers of markers, permit one to consider the DNA taxonprint as an important tool for molecular phylogeny and taxonomy.

It should be noted that the method is rather simple, reproducible, and cheap; needs small amounts of DNA; and can be applied in a short time with many samples simultaneously. We anticipate that the DNA taxonprinting together with more traditional approaches can give new insight into the ''species problem.''

Acknowledgments. We are grateful to A.S. Antonov, who has supplied much constructive criticism.

References

Arnold EN (1989) Towards a phylogeny and biogeography of Lacertidae: relationships within an Old-World family of lizards derived from morphology. Bull Br Mus Nat Hist (Zool) 55:209–257

Bannikova AA, Fedorova LV, Fedorov AN, Troitzky AV, Grechko VV, Dolgov VA, Lomov AA, Mednikov BM (1995) A comparison of repetitive DNA sequences of the family Erinaceidae (Mammalia) by restriction endonuclease analysis. Genetica (Russ) 31:1498–1506

Brown WM (1980) Polymorphism of mitochondrial DNA of humans as revealed by restriction endonuclease analysis. Proc Natl Acad Sci USA 77:3605–3609

Brown SDM, Dover GA (1979) Conservation of sequences in related genomes of *Apodemus:* constraints on the maintenance of satellite DNA sequences. Nucleic Acids Res 6:2423–2434

Brown SDM, Dover GA (1980) Conservation of segmental variants of satellite DNA of *Mus musculus* in related species: *Mus spretus.* Nature 285:47–79

Buckland RA (1983) Comparative structure and evolution of goat and sheep DNAs. Nucleic Acids Res 11:1349–1360

Chelomina GN, Lyapunova EA, Vorontsov NN, Suchia K (1995) Molecular genetic typing of three forms of wood and field mice belonging to transpalearctic genus (*Apodemus;* Muridae, Rodentia). Genetica (Russ) 31:820–824

Christie NT, Skinner DM (1979) Evidence for nonrandom alteration in the fraction of the highly repetitive DNA of eucaryote. Nucleic Acids Res 6:781–796

Cooke HJ (1975) Evolution of the long range structure of satellite DNA in the genus *Apodemus.* J Mol Biol 94:87–99

Dandieu S, Rahuel C, Ruffie J, Lucotte G (1984) Comparison of the highly repetitive DNA sequences of humans and different species of primates. Biochem Syst Ecol 12:237–244

Darevsky IS (1993) Evolution and ecology of parthenogenesis in reptiles. Current research on biology of amphibians and reptiles. Proceedings of the First World Congress of Herpetology, Society for Study of Amphibians and Reptiles, Oxford, OH pp 21–39

Donehower L, Gillespie D (1979) Restriction sites periodicities in highly repetitive DNA in Primates. J Mol Biol 134:805–834

Elizur A, Dennis ES, Peacock WJ (1982) Satellite DNA sequences in the red kangaroo *(Macropus rufus).* Austral J Biol Sci 35:313–325

Fedorov AN, Grechko VV, Slobodyanyuk SYa, Fedorova LV, Timokhina GI (1992) Taxonomic analysis of DNA repeated sequences. Mol Biol (Russ) 26:464–469

Felsenstein J (1993) PHYLIP version 3.5c. Distributed by author. Department of genetics, University of Washington, Seattle

Garoff H, Ansorge W (1981) Improvements of DNA sequencing gels. Anal Biochem 115:450–457

Grachev MA, Slobodyanyuk SYa, Kholodilov NG, Fyodorov SP, Belicov SI, Sherbakov DYu, Sideleva VG, Zubin AA, Kharchenko VV (1992) Comparative study of two protein coding regions of mtDNA from three endemic Sculpins (Cottoidei) of Lake Baikal. J Mol Evol 34:85–90

Grechko VV, Ryabinin DM, Fedorova LV, Fedorov AN, Darevsky IS, Ryskov AP (1993) Taxoprint analysis of DNA from some species of the fam. Lacertidae. Mol Biol (Russ) 27:1404–1414

Hembelen V, Leweke B, Roth A, Stadler K (1982) Organization of highly repetitive satellite DNA of two Cucurbitaceae species *(Cucumis melon* and *C. sativus).* Nucleic Acids Res 10:631–644

Kirilchik SV, Slobodyanyuk SYa (1997) Evolution of the cytochrome *b* fragment of some Cottoidea species mitochondrial DNA of Lake Baikal and other localities. Mol Biol (Russ) 31:168–175

Lima de Faria A, Arnasson U, Widegren B, Essen—Miller J, Isacsson M, Olsson E, Jaworska M (1984) Conservation of repetitive DNA sequences in deer species studied by Southern blot transfer. J Mol Evol 20:17–24

Mednikov BM, Bannikova AA, Lomov AA, Melnikova MN, Shubina EA (1995) Restriction analysis of repetitive DNA sequences: species criteria and species origin. Mol Biol (Russ) 29:1308–1319

Melnikova MN, Grechko VV, Mednikov BM (1995) Investigation of genetic divergence and polymorphism of nuclear DNA in species and populations of domestic and wild sheep. Genetica (Russ) 31: 1120–1131

Moritz C, Uzzell T, Spolsky C, Hotz H, Darevsky I, Kupriyanova L, Danielyan F (1992) The maternal ancestry and approximate age of parthenogenetic species of Caucasian rock lizards (*Lacerta:* Lacertidae). Genetica 87:53–62

Murphy RW, Darevsky IS, MacCulloch RD, Fu J, Kupriyanova LA (1996) Evolution of the bisexual species of Caucasian rock lizards: a phylogenetic evaluation of allozyme data. Russ J Herp 3:18–31

Nei M, Li W-H (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc Natl Acad Sci USA 76:5269–5273

Potapov SG, Ryskov AP (1993) Analysis of variability of repetitive elements of DNA of some rodent genomes at the taxonomical level. Genetica (Russ) 29:869–872

Potapov SG, Vassilyev VA, Samarina OP, Ryskov AP (1994) Molecular genetic fingerprinting of *Phodopus* genus representatives. Genetica (Russ) 30:6150621

Rohlf FJ (1993) NTSYS—pc. Numerical taxonomy and multivariate analysis system. Department of Ecology and Evolution, State University of New York, Stony Brook

Ryskov AP, Kudryavtsev PI, Vassilyev SG, Sirko IP (1994) Diagnostic potentialities of molecular-genetic approaches in the taxonomy of the tribe Bovini. Genetica (Russ) 73:115–123

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY

Shubina EA, Mednikov BM (1986) The repetitive families in DNA of Far East salmon *(Onchorhinchus).* Mol Biol (Russ) 20:947–956

Skurikhina LA, Oleynik AG, Nevshchupov SV, Brykov VA (1993) Phylogeny of Far East salmon by restriction endonuclease analysis data. Genetica (Russ) 29:1508–1518

Sneath PHA, Sokal RR (1973) Numerical taxonomy: the principles and practice of numerical classification. Freeman, San Francisco, p 573

Southern EM (1975) Long range periodicities in mouse satellite. J Mol Biol 94:51–69

Turner BJ, Edler JF, Laughlin TF (1991) Repetitive DNA sequences and divergence of fish populations: some hopeful beginnings. J Fish Biol 39:131–142