

DNA Barcoding: How Many Earthworm Species Are There in the South of West Siberia?

S. V. Shekhovtsov^{a, *}, N. E. Bazarova^{a, b}, D. I. Berman^c, N. A. Bulakhova^{c, d}, E. V. Golovanova^e,
S. V. Konyaev^f, T. M. Krugova^g, I. I. Lyubchanskii^a, and S. E. Peltek^a

^aInstitute of Cytology and Genetics, Siberian Branch, Russian Academy of Sciences, Novosibirsk, Russia

^bNovosibirsk State Agricultural University, Novosibirsk, Russia

^cInstitute of Biological Problems of the North, Far Eastern Branch, Russian Academy of Sciences, Magadan, Russia

^dTomsk State University, Tomsk, Russia

^eOmsk State Pedagogical University, Omsk, Russia

^fInstitute of Systematics and Ecology of Animals, Siberian Branch, Russian Academy of Sciences, Novosibirsk, Russia

^gTigirek State Natural Reserve, Barnaul, Russia

*e-mail: shekhovtsov@bionet.nsc.ru

Received September 25, 2015; in final form, October 20, 2015

Abstract—Earthworms are a widespread and ecologically important group of animals, which has the highest total biomass in some ecosystems and often defines the composition of soil fauna. Earthworms are known to have high cryptic genetic diversity. In this study we attempted to estimate earthworm species diversity in the south of West Siberia by DNA barcoding. This method employs short fragments of the genome to identify species, and allows one to work with specimens that cannot be identified by conventional techniques, as well as to search for new species and predict their phylogenetic affinities. As the target sequence we took a fragment of the mitochondrial cytochrome oxidase 1 (*cox1*) gene. The studied territory (Novosibirsk and Tomsk oblasts, Altai krai, and the Altai Republic) is known to contain 16 species and subspecies of earthworms. We analyzed 259 individuals from twelve locations and detected 27 genetic clusters. Ten of them correspond to known species (*A. caliginosa*, *E. fetida*, *O. tyraeum*, *D. rubidus tenuis*, *D. octaedra*, *E. balatonica*, *E. sibirica*, as well as three genetic lineages of *E. nordenskioldi nordenskioldi*). Seventeen of the 27 clusters do not have close sequence similarity to any known earthworm species. Representatives of some of these novel clusters are morphologically similar to the *Eisenia n. nordenskioldi/E. n. pallida* species complex and may belong to new genetic lineages of this complex. The rest of the novel clusters probably represent new earthworm species. Therefore, we can conclude that a large portion of earthworm biodiversity in the south of West Siberia is still unexplored.

Keywords: earthworms, Lumbricidae, West Siberia, DNA barcoding, *cox1*, cytochrome oxidase 1

DOI: 10.1134/S2079059717010130

INTRODUCTION

Earthworms are an ecologically important group of animals. The world's fauna contains relatively few species, about 3700 (Hendrix et al., 2008), only 47 of which were found in Russia (Vsevolodova-Perel', 1997). One reason for such a small diversity is the small number of morphological characters, due to their lifestyle, as subterranean life limits their morphological variation. The diagnostic traits include the position of the clitellum and genital openings, the shape of the cephalic lobe, the number and location (closeness) of the bristles, and sometimes the size and color of the body. These traits often evolve convergently in different species and groups; however, significant intraspecific variation is also possible.

The most complete and the most commonly used summary on Russian earthworms is the book "The

earthworms of the fauna of Russia: cadaster and key" (Vsevolodova-Perel, 1997). This reference for Asian Russia contains 28 species and subspecies of earthworms. The territory considered in this work (Novosibirsk and Tomsk regions, Altai krai, and the Altai Republic) is inhabited by 16 species of earthworms. All of them belong to the Lumbricidae family. Eight of them are invasive cosmopolitan species: *Allolobophora parva*, *Dendrodrilus rubidus tenuis*, *D. r. subrubicundus*, *Octolasion tyraeum* (= *O. lacteum*), *Aporrectodea caliginosa*, *Dendrobaena octaedra*, *Eiseniella tetraedra*, and *Eisenia fetida*. The species of Western Siberia include *Lumbricus terrestris*, *L. castaneus*, *L. rubellus*, and *Aporrectodea rosea* (Striganova and Poryadina, 2005); however, according to T.S. Vsevolodova-Perel', the last two species were misidentified local species. Four species of the genus *Eisenia* (*E. malevici*, *E. alta-*

Table 1. Locations sampled in this study

Sampling location	Number of individuals	Number of clusters
Podgornoe village, Chainsky district, Tomsk oblast	28	5
Novosibirsk, Novosibirsk oblast	56	7
Dubrovka village, Maslyaninsky district, Novosibirsk oblast	27	5
Temirtau village, Tashtagolsky district, Kemerovo oblast	12	2
Makarevka village, Soltonsky district, Altai krai	32	9
Mustag village, Tashtagolsky village, Kemerovo oblast	4	3
Mukhor-Cherga village, Shebalinsky district, Altai Republic	4	1
Artybash village, Turochaksky district, Altai Republic	38	3
Evrechala village, Turochaksky district, Altai Republic	8	3
Sarlyk village, Shebalinsky district, Altai Republic	5	2
Tigirek State Natural Reserve, Krasnoshchekovo district, Altai krai	31	7
Charyshskoe village, Charyshsky district, Altai krai	14	1

ica, *E. salairica*, and *E. tracta*) are endemic to Altai and the Salair Ridge. The remaining five species are a closely related group of *Eisenia n. nordenskioldi*/*E. n. pallida*/*E. atlavinyteae*, which is widespread throughout Siberia, as well as the rarer species *E. sibirica* and *E. balatonica*.

It is hard not to notice the sharp contrast between the very low species diversity (compared to most other groups of animals common in the same area) and the huge total biomass and ecological importance of this group.

Meanwhile, one of the first studies carried out on earthworms using DNA barcoding (King et al., 2008) showed that the majority of species is characterized by very high cryptic diversity. Within a few widespread species studied by King et al. (2008), two to five cryptic genetic lineages noticeably different from each other were found: the number of pairwise substitutions between these lineages amounted to 22%. A similar pattern was found in subsequent studies on both certain species of worms and their extensive samplings (Peréz-Losada et al., 2012; Porco et al., 2013).

DNA barcoding employs short genomic fragments for species identification (Hebert et al., 2003). The most commonly used sequence for DNA barcoding in animals is the mitochondrial cytochrome oxidase 1 (*cox1*, COI) gene. This method is very useful in the study of samples that cannot be reliably identified by conventional methods, such as juveniles and cocoons (Richard et al., 2010), fragments of organisms, feces, stomach contents, etc., or DNA isolated from soil samples (Bienert et al., 2012). It is also applied for groups containing a large number of species that are hard to identify, or for those with poorly developed taxonomy. DNA barcoding has more limited objectives than molecular phylogeny and does not include study of relationships among species or clarification of their systematic position (Waugh, 2007). Neverthe-

less, in some cases the method allows one to discover new species and tentatively establish their systematic affinity. Although DNA barcoding cannot be the basis for taxonomic studies by itself, it makes it possible to identify putative new species (Decaëns et al., 2013). For example, in the work on earthworms' fauna of New Zealand (Buckley et al., 2011), the authors detected 50 probable new species based on DNA barcoding results, which is about 1/4 of the total fauna of this group in that area.

The purpose of our work was to study genetic diversity of earthworms in the south of Western Siberia using DNA barcoding. For this, we sampled a collection of worms from several geographical locations and carried out sequencing of a fragment of the *cox1* mitochondrial gene. The obtained sequences were compared with our database of certain reliably identified species of earthworms, as well as the GenBank and BOLD databases.

MATERIALS AND METHODS

Samples of earthworms were collected in 2011–2015 in different locations of Western Siberia (Table 1, Fig. 1). The samples included both adult and juvenile individuals. Some adult animals were identified morphologically.

The DNA was isolated from several segments of the rear end of the body using the BioSilica kits (Novosibirsk). The *cox1* mitochondrial gene fragment was amplified using universal primers: direct LCO1490m (5'-TACTC-AACAA-ATCAC-AAAGA-TATTG-G-3') (Folmer et al. (2014) with modifications) and reverse HCO2198 (5'-TAAAC-TTCAG-GGTGA-CCAAA-AAATC-A-3') (Folmer et al., 1994) and COI-E- (5'-TATAC-TTCTG-GGTGT-CCGAA-GAATC-A-3') (Bely and Wray, 2004). The following amplification profile was used: 2 min at 94°C; 5 cycles: 20 s at 94°C;

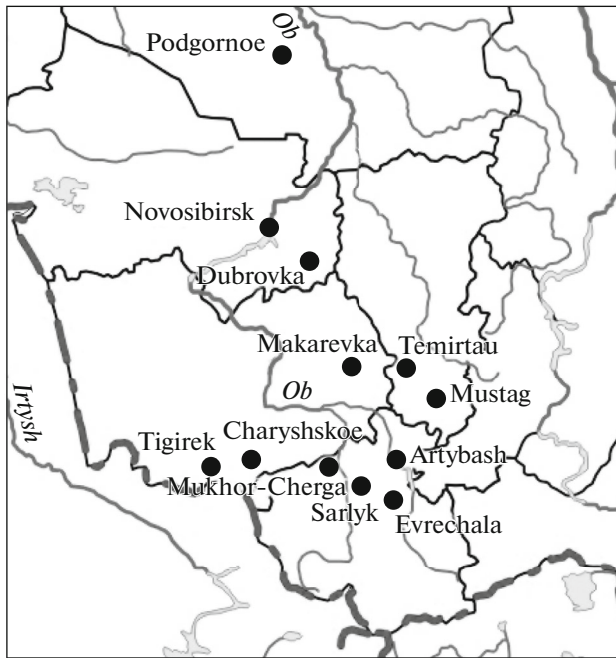


Fig. 1. Sampling locations.

20 s at 45°C; 50 s at 72°C; 33 cycles: 20 s at 94°C; 20 s at 55°C; and 50 s at 72°C.

The *coxI* gene fragments were sequenced using the BigDye 3.1 kit produced by Applied Biosystems (United States). Capillary electrophoresis was performed at the Collective Center for Sequencing of the Siberian Branch, Russian Academy of Sciences (SB RAS) in Novosibirsk. The obtained sequences were manually processed in the Chromas program. Phylogenetic trees were constructed using the Mega v. 5.0 program (Tamura et al., 2011). The GenBank (<http://www.ncbi.nlm.nih.gov>) and BOLD (<http://www.boldsystems.org>) databases were used to identify closely species. To construct the trees using the Minimum Evolution algorithm, we used the Kimura-2-parameter model of substitutions. Genetic distances for the K/θ analysis were calculated using Mega v. 5.0 program (Tamura et al., 2011).

This study combines the results of DNA barcoding of 259 earthworm individuals from 12 locations in the south of West Siberia. As a source of reference sequences, we used the BOLD and GenBank databases, as well as reliably identified specimens of certain species of earthworms from Russia (*E. n. nordenskioldi*, *E. n. pallida*, *E. sibirica*, *E. balatonica*, *E. fetida*, *D. r. tenuis*, *D. r. subrubicundus*, *O. tyrtaeum*, *D. octaedra*, *A. caliginosa*, *A. trapezoides*, *A. rosea*, and *L. rubellus*) (Shekhovtsov et al., 2013, 2014a, b, 2015).

RESULTS AND DISCUSSION

The tree constructed using the *coxI* gene is shown on Fig. 2. It is known that the trees constructed by

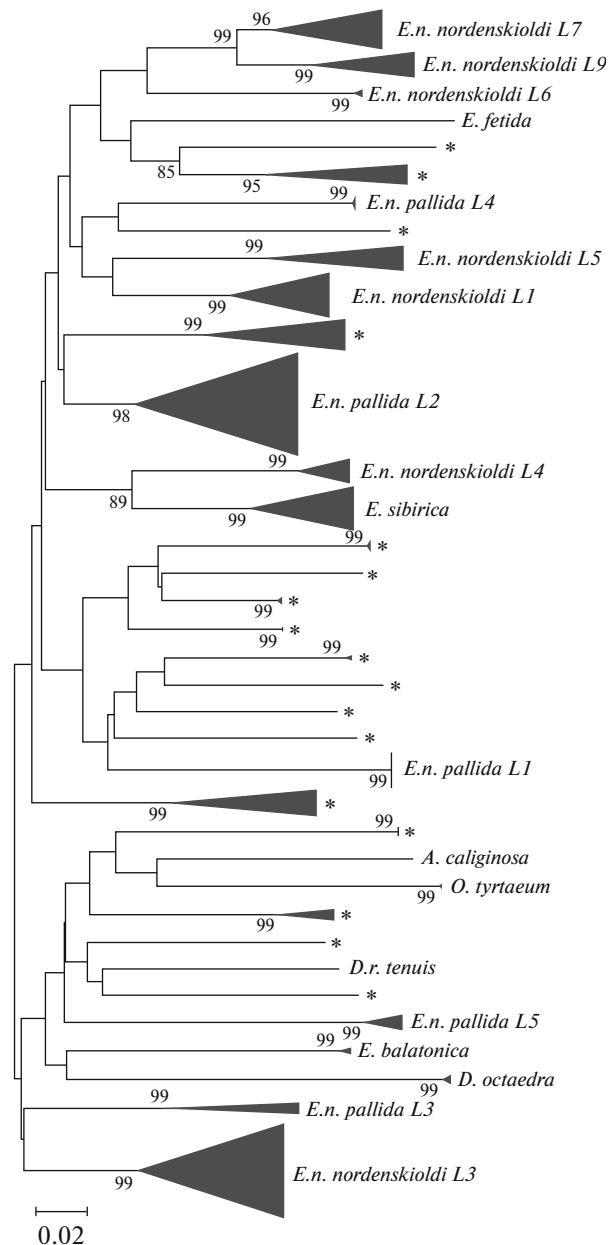


Fig. 2. Phylogenetic tree of *coxI* sequences constructed using the Minimum Evolution algorithm. Numbers above branches indicate bootstrap support. Clusters are collapsed into triangles; the base of the triangle is proportional to the number of sequences and its height, to nucleotide diversity within the cluster. Newly detected clusters are indicated by asterisks). For *E. n. nordenskioldi* and *E. n. pallida*, numbers of genetic lineages are given.

sequences of mitochondrial DNA for earthworms, as a rule, do not have bootstrap-supported branches corresponding to the super-specific taxa (Pop et al., 2007; Chang and James, 2011; James and Davidson, 2012), which is probably the result of the high rate of evolution of the mitochondrial DNA sequences in earthworms. Thus, genera and even families on such trees

Table 2. New *coxI* clusters detected in this study

<i>n</i>	Sampling location
1	Artybash
5	Artybash
11	Artybash, Makarevka, Evrechala
2	Makarevka
2	Makarevka
2	Makarevka
4	Tigirek, Makarevka
4	Tigirek
1	Tigirek
1	Tigirek
1	Tigirek
12	Tigirek, Sarlyk, Charyshsky
3	Tigirek, Mukhor-Cherga
1	Evrechala
1	Mustag
1	Mustag
1	Mustag

n, number of individuals with the *coxI* haplotype of the given cluster.

turn out to be polyphyletic. For this reason, it is assumed that mitochondrial DNA is useful for the identification of new species of earthworms but is unsuitable for studying phylogenetic relationships.

Clustering was performed as follows: a branch with bootstrap support of at least 95, which was not united with any other branch on the tree with bootstrap support >90 was considered a cluster corresponding to one operational taxonomic unit (OTU). An exception was made for genetic lineages 7 and 9 of *E. n. nordenskioldi*, as well as lineages 4 of *E. n. nordenskioldi* and *E. sibirica* (see Fig. 2), which have markedly different distributions, which suggests significant biological differences.

A total of 27 clusters was detected. Seven of them belonged to well-known species: *A. caliginosa*, *E. fetida*, *E. balatonica*, *E. sibirica*, *O. tyrtaeum*, *D. r. tenuis*, and *D. octaedra*. We failed to find three cosmopolitan species known from Western Siberia (*E. tetraedra*, *D. r. subrubicundus*, and *A. parva*), most likely due to the fact that their characteristic habitats were not inspected. We also could not find any reliably identified Altai endemic. Three clusters matched the genetic lineages of *E. n. nordenskioldi* typical of this region (Shekhovtsov et al., 2013). The remaining 17 of the 27 clusters did not have close similarity to any of the previously studied DNA sequences (Table 2).

One of the methods to analyze the validity of specific identification based on DNA sequences is the K/θ method, or $4\times$ rule (Birky et al., 2010). It was

originally developed for organisms with asexual reproduction, but can also be used in cases where the populations under study are isolated and cannot exchange genes (which might be assumed for earthworms with their low mobility). This approach was used, for example, in the analysis of the species complex *A. caliginosa* (Fernández et al., 2012). The author of the K/θ method showed that genetic clusters can be considered different species with a 95% probability if they diverged $4\times N_e$ generations ago, where N_e is the effective population abundance. This parameter can be estimated by calculating the K/θ ratio, where K is the genetic distance between clusters, and θ is the genetic distance within the clusters. Thus, in order to consider clusters to be different species, the K/θ ratio should be greater than four.

According to our analysis, this rule was met for 20 of the 26 newly discovered clusters. The exceptions were lineages 2 and 3 within the subspecies of *E. n. pallida* and lineages 1, 5, and 9 of *E. n. nordenskioldi*, as well as one of the newly discovered clusters (Tigirek + Charyshskoe + Sarlyk). Obviously, the K/θ ratio depends on the extent of the detected genetic variation within a cluster, as this estimate is based on a very small sample from large populations (Birky et al., 2010). In our case, we studied several remote populations, and the result was the division of the clusters into several branches, for which the $K/\theta > 4$ rule was satisfied. For example, the Tigirek + Charyshskoe + Sarlyk cluster was divided into two: Tigirek + Charyshskoe and Sarlyk. Certain lineages of *E. n. pallida* and *E. n. nordenskioldi* were also divided into several subclusters. Thus, the use of K/θ leads to an increase in the number of allocated genetic clusters.

Worms belonging to 9 out of 17 clusters were morphologically similar to the group *E. n. nordenskioldi*/*E. n. pallida*. According to our data, Russia is inhabited by nine genetic lineages of *E. n. nordenskioldi* (four of them are in the region covered by this work) and five lineages of *E. n. pallida* (only one of them is found in Western Siberia). These lineages have differences both for the mitochondrial and nuclear DNA and are characterized by significantly different distributions. Taking into account high genetic variation of the *E. nordenskioldi* complex (Shekhovtsov et al., 2013), we can hypothesize that some of the detected clusters can be new lineages of this species complex.

Is it possible that the identified clusters correspond to earthworm species? According to the analysis of Hebert et al. (2003), annelids are characterized by higher level of mtDNA sequence divergence among species of the same genus compared to other animals. A divergence of the *coxI* sequences of different species of over 16% is characteristic for 70% of genera. According to the same authors, the degree of intraspecific variation rarely exceeds 2%. However, pairwise divergence values among *E. nordenskioldi* genetic lineages range from 16 to 29%. Genetic variation within

sufficiently sampled lineages of this species complex exceeds 5%, and in lineage 2 of *E. n. pallida* reaches 8.6%. Thus, judging by the degree of mtDNA variation, the clusters we found may be considered separate species.

However, representatives of significantly differing mitochondrial lineages of a species may in certain cases have no substantial differences in their nuclear genes, as was shown, for example, for *L. rubellus* (Giska et al., 2015), as well as for lineages 2 and 3 of *E. n. pallida*. Thus, existence of allegedly new earthworm species should also be confirmed using nuclear markers.

Although DNA barcoding cannot give a definitive answer to the question of how many species of earthworms inhabit the south of West Siberia, the obtained data show that real species diversity may be significantly higher than it is assumed according to conventional morphological analysis. Even a very limited sample of a small number of locations made it possible to confidently identify 17 new genetic clusters, which may well correspond to the new taxa (of specific or subspecific status), thus doubling the previously known number of those from this region (from 16 to 33). This work confirms the idea of the high cryptic variation of earthworms and indicates the prospects for further research in this direction.

ACKNOWLEDGMENTS

The authors express their sincere gratitude to Tamara Semenovna Vsevolodova-Perel' for identifying part of our collections

The work was supported by government funding (project no. VI.58.1.3), state task 6.1957.2014/K, and grant MK-6685.2015.4 of the President of the Russian Federation.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

REFERENCES

Bely, A.E. and Wray, G.A., Molecular phylogeny of naidid worms (Annelida: Clitellata) based on cytochrome oxidase i, *Mol. Phylogenet. Evol.*, 2004, vol. 30, pp. 50–63. doi 10.1016/S1055-7903(03)00180-5

Bienert, F., De Danieli, S., Miquel, C., Coissac, E., Poillot, C., and Brun, J.-J., Tracking earthworm communities from soil DNA, *Mol. Ecol.*, 2012, vol. 21, no. 8, pp. 2017–2030. doi 10.1111/j.1365-294X.2011.05407.x

Birky, C.W., Jr., Adams, J., Gemmel, M., and Perry, J., Using population genetic theory and DNA sequences for species detection and identification in asexual organisms, *PLoS One*, 2010, vol. 5, no. 5, p. 10609. doi 10.1371/journal.pone.0010609

Buckley, T.R., James, S., Allwood, J., Bartlam, S., Howitt, R., and Prada, D., Phylogenetic analysis of New Zealand earthworms (Oligochaeta: Megascolecidae) reveals ancient clades and cryptic taxonomic diversity, *Mol. Phylogenet. Evol.*, 2011, vol. 58. doi 10.1016/j.ympev.2010.09.024

Chang, C.-H. and James, S., A critique of earthworm molecular phylogenetics, *Pedobiologia*, 2011, vol. 54, pp. S3–S9. doi 10.1016/j.pedobi.2011.07.015

Decaëns, T., Porco, D., Rougerie, R., Brown, G.G., and James, S.W., Potential of DNA barcoding for earthworm research in taxonomy and ecology, *Appl. Soil Ecol.*, 2013, vol. 65, pp. 35–42. doi 10.1016/j.apsoil.2013.01.001

Fernández, R., Almodóvar, A., Novo, M., Simancas, B., and Díaz Cosín, D.J., Adding complexity to the complex: New insights into the phylogeny, diversification and origin of parthenogenesis in the *Aporrectodea caliginosa* species complex (Oligochaeta, Lumbricidae), *Mol. Phylogenet. Evol.*, 2012, vol. 64, no. 2, pp. 368–379. doi 10.1016/j.ympev.2012.04.011

Folmer, O., Black, M., Hoeh, W., Lutz, R., and Vrijenhoek, R., DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates, *Mol. Mar. Biol. Biotech.*, 1994, vol. 3, pp. 294–299.

Giska, I., Sechi, P., and Babik, W., Deeply divergent sympatric mitochondrial lineages of the earthworm *Lumbricus rubellus* are not reproductively isolated, *BMC Evol. Biol.*, 2015, p. 217. doi 10.1186/s12862-015-0488-9

Hebert, P.D.N., Ratnasingham, S., and deWaard, J.R., Barcoding animal life: Cytochrome c oxidase subunit I divergences among closely related species, *Proc. R. Soc. B*, 2003, vol. 270, pp. 96–99. doi 10.1098/rsbl.2003.0025

Hendrix, P.F., Callahan, M.A., Drake, J.M., Huang, C.-Y., James, S.W., Snyder, B.A., and Zhang, W., Pandora's box contained bait: The global problem of introduced earthworms, *Ann. Rev. Ecol. Syst.*, 2008, vol. 39, pp. 593–613. doi 10.1146/annurev.ecolsys.39.110707.173426

James, S.W. and Davidson, S.K., Molecular phylogeny of earthworms (Annelida: Crassidellata) based on 28S, 18S and 16S gene sequences, *Invertebr. Syst.*, 2012, vol. 26, pp. 213–229. doi 10.1071/IS11012

King, R.A., Tibble, A.L., and Symondson, W.O.C., Opening a can of worms: Unprecedented sympatric cryptic diversity within British lumbricid earthworms, *Mol. Ecol.*, 2008, vol. 17, no. 21, pp. 4684–4698. doi 10.1111/j.1365-294X.2008.03931.x

Pérez-Losada, M., Bloch, R., Breinholt, J.W., Pfenninger, M., and Domínguez, J., Taxonomic assessment of Lumbricidae (Oligochaeta) earthworm genera using DNA barcodes, *Eur. J. Soil Biol.*, 2012, vol. 48, pp. 41–47. doi 10.1016/j.ejsobi.2011.10.003

Pop, A.A., Cech, G., Wink, M., Csuzdi, C., and Pop, V.V., Application of 16S, 18S rDNA and COI sequences in the molecular systematics of the earthworm family Lumbricidae (Annelida, Oligochaeta), *Eur. J. Soil Biol.*, 2007, vol. 43, pp. S43–S52. doi 10.1016/j.ejsobi.2007.08.007

Porco, D., Decaëns, T., Deharveng, L., James, S.W., Skarzynski, D., Erséus, C., Butt, K.R., Richard, B., and Hebert, P.D.N., Biological invasions in soil: DNA barcoding as a monitoring tool in a multiple taxa survey targeting European earthworms and springtails in North America, *Biol. Invasions*, 2013, vol. 15, pp. 899–910. doi 10.1007/s10530-012-0338-2

- Richard, B., Decaëns, T., Rougerie, R., James, S.W., Porco, D., and Hebert, P.D.N., Re-integrating earthworm juveniles into soil biodiversity studies: Species identification through DNA barcoding, *Mol. Ecol. Res.*, 2010, vol. 10, no. 4, pp. 606–614. doi 10.1111/j.1755-0998.2009.02822.x
- Shekhovtsov, S.V., Golovanova, E.V., and Peltek, S.E., Genetic diversity of the earthworm *Octolasion tyrtaeum* (Lumbricidae, Annelida), *Pedobiologia*, 2014a, vol. 57, pp. 245–250. doi 10.1016/j.pedobi.2014.09.002
- Shekhovtsov, S.V., Golovanova, E.V., and Peltek, S.E., Cryptic diversity within the Nordenskiöld's earthworm, *Eisenia nordenskiöldi* subsp. *nordenskiöldi* (Lumbricidae, Annelida), *Eur. J. Soil Biol.*, 2013, vol. 58, pp. 13–18. doi 10.1016/j.ejsobi.2013.05.004
- Shekhovtsov, S.V., Golovanova, E.V., and Peltek, S.E., Invasive lumbricid earthworms of Kamchatka (Oligochaeta), *Zool. Stud.*, 2014b, vol. 53, p. 52. doi 10.1186/s40555-014-0052-0
- Shekhovtsov, S.V., Berman, D.I., and Peltek, S.E., Phylogeography of the Earthworm *Eisenia nordenskiöldi* sub sp. *nordenskiöldi* (Lumbricidae, Oligochaeta) in Northeastern Eurasia, *Dokl. Biol. Sci.*, 2015, vol. 461, pp. 1–4. doi 10.1016/j.pedobi.2014.09.002
- Striganova, B.R. and Poryadina, N.M., *Zhivotnoe naselenie pochv boreal'nykh lesov Zapadno-Sibirskoi ravniny*: (The Animal Population of Boreal Forest Soils of the West Siberian Plain), Moscow: KMK, 2005.
- Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S., MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods, *Mol. Biol. Evol.*, 2011, vol. 28, pp. 2731–2739. doi 10.1093/molbev/msr121
- Vsevolodova-Perel', T.S., *Dozhdevye chervi Rossii: Kadastr i opredelitel'*: (Earthworms of Russia: The Inventory and Identifier), Moscow: Nauka, 1997.
- Waugh, J., DNA barcoding in animal species: Progress, potential and pitfalls, *BioEssays*, 2007, vol. 29, pp. 188–197. doi 10.1002/bies.20529

Translated by K. Lazarev