Illumina-based identification of arbuscular mycorrhizal fungi from Karachay-Cherkessia soils for development of bio-fertilizers

Andrey Yurkov[1,*], *Alexey* Kryukov[1], *Yulia* Mikhaylova[2], and *Peter* Zhurbenko[2]

[1]All-Russian Research Institute for Agricultural Microbiology, Laboratory of Ecology of Symbiotic and Associative Rhizobacteria, 196608 Saint Petersburg, Russia
[2] Komarov Botanical Institute, Laboratory of Biosystematics and Cytology, 197376 Saint Petersburg, Russia

Abstract. The aim of the study was to investigate the species diversity of AM fungi in different parts of the North Caucasus, biodiversity hotspot, the center of the world's biological diversity. Samples were taken from 5 locations (stationary trial plots, STPs) in different ecosystems and at various altitudes. Identification was performed using sequencing for ITS1 and ITS2 regions, amplified with universal primers, Illumina MiSeq was employed. 19 genera of AM fungi were found on all STPs. The work did not reveal a correlation between the altitude and the species composition of AM fungi. At the same time, it should be assumed that a correlation could be found between the biodiversity of AM fungi and the type of ecosystem, which should be done in the future. The study shows it is necessary to use an analysis for both ITS regions, since the data obtained for each ITS region differ and complement each other. Analysis for the ITS2 region revealed 1.3 times more virtual taxa than for the ITS1, while the number of OTUs identified per species was similar for both regions. The highest biodiversity of AM fungi was found in STP #3 (with meadow flora). Only 4 species (*Rhizophagus irregularis, R. intraradices, Paraglomus laccatum, and Claroideoglomus claroideum*) were found on all five analyzed STPs. We found unexpectedly that with such a high biodiversity among the identified fungi, no different species were found in the Paraglomus genus, all the sequences of Paraglomus belonged to *Paraglomus laccatum*, whereas at least 9 species are distinguished in the genus by morphology. Further research will allow us to identify new strains of AM fungi, the efficiency of which may be higher than already studied ones. In the future this will make it possible to create more effective microbial biofertilizers for agriculture.

* Corresponding author: yurkovandrey@yandex.ru

1 Introduction

Arbuscular mycorrhiza (AM) is one of the oldest and most common types of symbiosis between plants and fungi. AM leads to an increase in the supply of nutrients to the plant, a decrease in the negative impact of drought and salinity, and an improvement in the water balance. The study of AM, its symbiotic efficacy and activity is an urgent line of research in modern biology. The comprehensive study of AM fungi can help us to understand their role in the evolution and maintenance of plant species diversity, the plant growth and the adaptation to environment. Isolated from soils strains of AM fungi could be used in agriculture to increase the crops productivity and obtain ecologically clean products. Therefore, the search for the new AM fungi strains and species receives priority in agricultural microbiology.

The identification of AM fungi species is a great challenge, and there are misidentifications in published works [1]. Closely related species are often undistinguishable by morphological features [2]. A significant number of cryptic taxa are described among AM fungi [3]. There is a lack of consensus in taxonomy classification of Glomeromycotina fungi [4]. Identification of AM fungi using molecular approaches faces also problems, such as a high level of intraspecific polymorphism and lack of universal DNA barcode, accepted for AM fungi [5-6]. Due to high level of intraspecific polymorphism Sanger sequencing demands specific primers and extensive cloning [7].

With introduction of Next Generation Sequencing (NGS) techniques and employment of universal primers the efficiency of identification AM fungi identification increased. NGS approaches allow to study successfully soil microbiomes. A weakness of NGS is a length of a DNA-barcode, which is shorter than in Sanger approach. Mostly used region for NGS-based AM fungi identification is ITS2, but number of authors believe, that separate analyses of both ITS1 and ITS2 regions are more informative. Different sets of fungal taxa can be discovered by using ITS1 and ITS2 as barcodes [8-9]. Many detected in experimental studies sequences of AM fungi are absent in genetic databases [2, 6, 10]. The greatest database GenBank NCBI contains less than 50% of found species of AM fungi [5-6]. So, it is important to investigate diversity of AM fungi and to submit in databases their new sequences, used as a barcode.

In our previous work we performed a preliminary study of AM fungi from soil samples [6]. We optimized the sample preparation and DNA purification methodology, selected primers and showed, that both region ITS1 and ITS2 in a separated mode should be used monitoring of AM fungi diversity. The goal of this work is to analyze the AM fungi communities the in undisturbed soils, collected at different altitudes in the North Caucasus. We determined genetic polymorphism and species diversity of AM fungi on tested areas via Illumina MiSeq sequencing. The North Caucasus is a “hot spot” of global biodiversity, for this reason there are good chances to find symbiotically efficient strains of AM fungi. New obtained strains were included in the collection of ARRIAM and could be used for practical introduction in agriculture.

2 Materials and methods

2.1 Materials

The study areas were located in the Karachay-Cherkessia (the North Caucasus, Russia). The soil samples were collected in five sample plots in stationary trial places (STPs) at different altitudes. The studied STPs were as follow: 1) STP #1, a subalpine meadow at the Malaya Hatipara mountain (43°25'51.0"N 41°42'55.0"E; 2186 m); surface slope 20°,

dominant species: *Agasyllis latifolia, Angelica tatianae, Anthriscus sylvestris, Bupleurum polyphyllum, Heracleum* sp., *Pimpinella rhodantha*, and *P.saxifraga*; total 47 vascular plant species; alpine meadow soddy peaty soil; 2) STP #2, a mixed forest in the valley of the Bolshaya Hatipara river (43°24'56.0"N 41°42'49.0"E; 1507 m); surface slope 10°; tree species *Betula pendula, Fagus orientalis, Picea orientalis, Acer* sp., *Sorbus aucuparia, Corylus avellana*, and *Abies nordmanniana*; dominant species: *Petasites albus*; total 21 vascular plant species; brown forest humic-illuvial soil; 3) STP #3, a hay meadow in the valley of the Teberda river (43°25'12.0"N 41°43'45.0"E; 1342 m); dominant species *Medicago falcata* and *Trifolium repens*; total 12 vascular plant species; flood plain compacted soil; 4) STP #4, a pasture in the valley of the Teberda river, near with Novaya Teberda settlement (43°39'37.0"N 41°53'12.0"E; 1026 m); dominant species *Urtica dioica* and *Fragaria vesca*; total 40 vascular plant species; flood plain subacid/neutral soil; and 5) STP #5, a steppe fescue meadow in the valley of the Teberda river, Sadovoye settlement near Cherkessk (44°18'51.0"N 42°02'00.0"E; 479 m); dominant species *Festuca valesiaca*; total 18 vascular plant species.

2.2 DNA isolation, Illumina sequencing and data analysis

DNA isolation, PCR and sequencing are described in details in [6]. Briefly, soil samples were grinded with liquid nitrogen, after that DNA isolation were performed according a CTAB method [11]. To increase yield of PCR products the extracted DNA was purified in a Silica agarose gel. ITS1 and ITS2 regions were amplified separately. For amplification of ITS1 region we used the primers ITS-5 [12] and ITS-6RK [13]. For amplification of ITS2 region we used the primers ITS-3 and ITS-4 [12]. The primers with adapters for Illumina MiSeq were made in Evrogen (Evrogen, Russia). PCR products were purified in the Silica agarose gel. Amplicon libraries were sequenced on an Illumina MiSeq platform (Illumina, Inc., USA) using a MiSeq® Reagent Kit v3 (600-cycle) in pair-end (2x300 N) runs. Raw reads were treated to obtained sets of OTU using USEARCH software [14]. Phylogenetic analysis of obtained sequences was performed using the MEGA7 software [15] for ITS1 and ITS2 separately. We compared fungal taxa in samples, calculated p-distances and constructed phylogenetic trees via Unweighted Pair Group Method with Arithmetic mean (UPGMA), Neighbour Joining (NJ) and Maximum Parsimony (MP) methods. The biodiversity were analysed by the Shannon and Margalef indices.

3 Results and discussion

After Illumina MiSeq sequencing, the average number of reads per sample was 77492 sequences, and after qualityfiltering it decreased to 20628 (73% of sequences less than 64 bp in length were excluded). Of these, 0.14% belonged to AM fungi. Data processing made it possible to identify AM fungi belonging to 19 different genera (Table 1). Using ITS2 were found 1.3 timesmore OTUsthan using ITS1 in terms of genera. But no significant differences in the number of defined OTUs were observed in terms of species. The most represented genus of AM fungi was Rhizophagus (62 OTU by species and 50 OTU by genus), the next was genus Dominikia (26 OTU by species and 71 OTU by genus), then Entrophospora (5 OTU by species and 40 OTU by genus), Paraglomus (6 OTUs by species and 35 OTUs by genus), Glomus (17 OTUs by species and 22 OTUs by genus), Acaulospora (8 OTUs by species and 18 OTUs by genus), Claroideoglomus (15 OTUs by species and 9 OTUs by genus). The rarest were Cetraspora (1 OTU), Innospora (1 OTU), Pacispora (1 OTU), Funneliformis (2 OTU), Otospora (3 OTU), and Kamienskia (3 OTU). STP # 3 (with meadow vegetation) had the greatest diversity of genera (1.4-3.1 times more)

and species (2.5-3.6 times more) than other STPs. STP # 2 (with forest vegetation) had the least diversity. In this regard, it can be concluded that the variety of AM fungi species does not havelinear correlation with the altitude, but it can be determined by the type of ecosystem and species composition of host plants.

Table 1. The number of identified OTU* for AM fungi genera and species levels on the STPs**

STP #		1				2				3				4				5			
Location, ecosystem		Subalpine meadow at the Malaya Hatipara mountain				Mixed forest in the valley of the Bolshaya Hatipara river				Hay meadow in the valley of the Teberda river				Pasture in the valley of the Teberda river				Steppe fescue meadow in the valley of the Teberda river			
Altitude above sea level (m)		2186				1507				1342				1026				479			
rDNA marker		ITS1		ITS2		ITS1		ITS2		ITS1		ITS2		ITS1		ITS2		ITS1		ITS2	
Difined to* **		G	S	G	S	G	S	G	S	G	S	G	S	G	S	G	S	G	S	G	S
Genera of AM fungi	*Acaulospora*	6		7	3	1	2		1	1	1	2	1			1					
	Ambispora	2		1	1	1		1	1		1		1	1		1					
	Archaeospora									7	5	3	1	1							
	Cetraspora																1				
	Claroideoglomus		1		2		1	1			2	4	3		1	4	3		1		1
	Diversispora						2	2			2					2					
	Dominikia	1		5						20	15	13		3		12	2	8	7	9	2
	Entrophospora			11				6	1			11	1			8	2			4	1
	Funneliformis										1						1				
	Glomus	2	1		1			1	1	4	5	4	3	2	2	9	4				
	Innospora															1					
	Kamienskia													1		2					
	Otospora									1				1	1						
	Pacispora									1											
	Paraglomus	2	1	3		3	1	3		8	2	6		1	1	3		6			1
	Rhizoglomus	1			1	1					1		2					1			
	Rhizophagus	12	2	4	6	4	4	5	10	7	11	4	9	7	1	6	8	1	9		2
	Septoglomus	1	1			1	1				4			1			1		1	1	
	Scutellospora	1				1										3		2			

*OTU – operational taxonomic units; **STP – stationary trial place; ***defined to genus (G, genus) or to species (S, species) according to 96% identity to the reference from the NCBI GeneBank

Analysis of the species diversity of STPs showed that STP # 3 included 23 out of 30 identified AM fungi species (Table 2). At the STPs of the Teberda Nature Reserve and adjacent territories in the Karachay-Cherkessia the genera Rhizophagus, Glomus, and Dominikia showed maximal observed species richness: 4, 5, and 5 species, respectively.

Table 2. Identified AM fungi species in STPs

Species of AM fungi*	STP #					Species of AM fungi*	STP #				
	1	2	3	4	5		1	2	3	4	5
Acaulospora paulinae		+	+			*Funneliformis mosseae*			+	+	
Ambispora fennica			+			*Glomus aggregatum*				+	
Ambispora gerdemannii	+					*Glomus hoi*			+	+	
Ambispora leptoticha	+	+				*Glomus indicum*	+	+	+	+	
Archaeospora europaea			+			*Glomus macrocarpum*			+	+	
Archaeospora spainiae			+			*Glomus tetrastratosum*	+		+		
Cetraspora pellucida				+		*Otospora bareae*				+	
Claroideoglomus claroideum	+	+	+	+	+	*Paraglomus laccatum*	+	+	+	+	+
Diversispora versiformis		+	+			*Rhizoglomus vesiculiferum*	+		+		
Dominikia bernensis			+	+	+	*Rhizophagus intraradices*	+	+	+	+	+
Dominikia disticha			+		+	*Rhizophagus invermaius*			+		
Dominikia indica			+		+	*Rhizophagus irregularis*	+	+	+	+	+
Dominikia iranica			+			*Septoglomus africanum*			+		
Dominikia litorea				+		*Septoglomus constrictum*				+	
Entrophospora infrequens		+	+	+	+	*Septoglomus nigrum*	+	+	+		+

*The analysis did not include virtual taxa defined at the genus level (such as: Innospora, Pacispora, Kamienskia, and Scutellospora), that may be due to the insufficient number of ITS sequences for these genera in the NCBI GeneBank.

Only 4 species (*Rhizophagus irregularis*, *R. intraradices*, *Paraglomus laccatum*, and *Claroideoglomus claroideum*) were found in all five analysed STPs. The four STPs contained the species *Entrophospora infrequens*, *Glomus indicum*, and *Septoglomus nigrum*. However, despite the significant distribution, the identification of genera such as Paraglomus and Entrophospora could be problematic. For these genera, there are still a small number of sequences presented in the NCBI GenBank. We constructed phylogenetic tree of ITS2 sequences assigned to the genus Paraglomus using the UPGMA method (Figure 1). For comparison and verification trees also were constructed using NJ and ME methods. All the obtained sequences belonging to the genus Paraglomus are grouped with good support with *P. laccatum* (GenBank accession number KY630227, sequencing author – Blaszkowski J.), which can be considered reliable. *P. laccatum* is widely represented on the analysed STPs (Table 2). On the figure 1 in OTUs names, the number before the dot indicates the STP number, and the number after the dot indicates the sample number. It can be noted that the sequences obtained from the same samples do not form a clade, which confirms the previously shown high intraspecific and intraorganismic polymorphism at the level of an individual AM fungus [6]. It is known that one AM fungus may have many nuclei and the ITS sequences in different nuclei may differ [2, 16]. It can be assumed that AM fungi of the same species collected at different altitudes (different STPs) could differ significantly. However, the results obtained for AM fungi of the genus Paraglomus indicate the opposite (Figure 1). At the same time, it should be noted that only 3 species of the genus Paraglomus are represented on the tree (there are only *P. laccatum*, *P. brasilianum*, *P. occultum* in the NCBI GenBank database). Meanwhile, three times more species in this genus are described by morphology according to the data of A. Schüßler [17], namely: *P. albidum*, *P. bolivianum*, *P. brasilianum*, *P. laccatum*, *P. lacteum*, *P. occidentale*, *P. occultum*, *P. pernambucanum*, *P. turpe*. Taking into account the fact that the surveyed

region – the North Caucasus – is one of the biodiversity hotspots, the probability of discovering new species of AM fungi was quite high. However, the phylogenetic analysis did not show the division of sequences attributed to the genus Paraglomus into subclades and did not show high genetic distances within the clade of *P. laccatum*, so all found AM fungi from Paraglomus genus was reliably identified as one species.

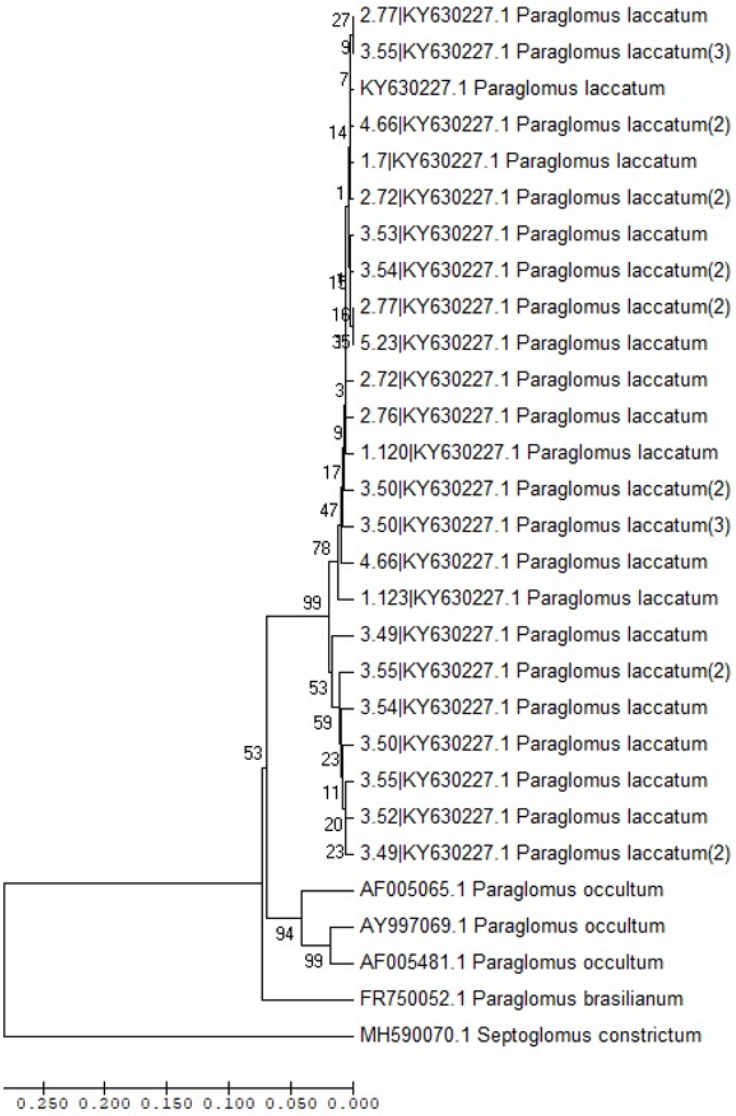

Fig. 1. Phylogentic tree of AM fungi of the genus Paraglomus, constructed for ITS2 by UPGMA method.

The assessment of the biodiversity indices (calculated by means of OTU numbers for each AMF species detected for both ITS1 and ITS2) showed the highest indices were for STP #3 (the Shannon index was 4.19, and the Margalef index was 6.12); the lowest indices were for STP #2 (the Shannon index was 2.85, and the Margalef index was 3.04). The data obtained are consistent with the above-mentioned results of estimation of AM fungi species

distribution in analized STPs. The possibility for detection of high AM fungal biodiversity (for example, for STP #3) was determined by the fact the identification was carried out not for one ITS regions, but simultaneously for both ITS1 and ITS2. A comparison of the different STPs suggests that we have discovered a new region with high AM fungal biodiversity, STP #3 (the number of AM fungal species is 23). The previously studied fungal diversity for three other STPs in Karachay-Cherkessia showed significantly lower AM fungal diversity with the highest number of species equal up to 8 per STP [13]. It should be noted that such a high diversity of AM fungi species is also characteristic of other mountain region, such as the Andes [18]. Thus, the distribution of AM fungal species and their ability to form AM symbiosis with different plant species are still insufficiently studied. Systematic researches of AM fungal biodiversity are required in the future, and the use of Illumina MiSeq can be considered one of the most effective methods for identification of AM fungi.

4 Conclusion

In the study in Karachay-Cherkessia a significant number of AM fungi was revealed: 30 different species from 19 genera per 5 STPs. The most widespread genus was Rhizophagus: species R. irregularis and R. intraradices were found on all five analyzed STPs. Identification by Illumina MiSeq with universal primers showed high efficiency this method. At the same time, it should be noted the analysis of ITS1 and ITS2 complements each other, since for different AM fungi the effectiveness of PCR for these regions was different.

Samples were taken in various STPs ranging from 479 to 2186 meters above sea level, however, no direct correlations were found between the altitudes of locations and the number of detected OTU or the number of AM fungal species. On the other hand, the meadow ecosystems had the greatest biodiversity of AM fungi, and forest ecosystem had the least one that may indicate a possible dependence of AM fungal species composition on the plant community and type of ecotope.

Our study contributed to the solution of a well-known problem related to the deficit of AM sequences of fungi in the NCBI GenBank for correct taxa identification. So for example the database contains only 3 of 9 morphologically identified species of Paraglomus. Since the study was conducted at biodiversity ‘hot-spot’, North Caucasus, we expected to find new species, virtual taxa among the obtained sequences of the genus Paraglomus. However, on the resulting phylogenetic tree, all the sequences appeared in one well-supported clade and with minimal genetic distances between them, that indicated an accurate identification. However, the addition of new sequences to the NCBI GenBank database will facilitate the isolation of new strains of AM fungi and their identification. In the future, the evaluation of the symbiotic efficiency of new AM fungal isolates will allow us to create more effective biofertilizers for agriculture.

This work was supported by the Russian Foundation for Basic Research grant No. 19-29-05275 MK and partially grant No. 20-016-00245 A and partially the state task No. 0482-2021-0004. The research was carried out using the equipment of the Core Center “Genomic Technologies, Proteomics and Cell Biology” in ARRIAM.

References

1. A. Schußler, M. Kruger, C. Walker, PloS One **6** (2011) DOI: 10.1371/journal.pone.0023333

2. A. A. Kryukov, A. P. Yurkov, Mikologiya i Fitopatologiya **52**, 38 (2018)
3. R. Savary, F. G. Masclaux, T. Wyss, G. Droh, J. C. Corella, A. P. Machado, J. B. Morton, I. R. Sanders, ISME J. **12**, 1 (2017) doi: 10.1038/ismej.2017.153
4. J. W. Spatafora, Y. Chang, G. L. Benny, K. Lazarus, M. E. Smith, M. L. Berbee, G. Bonito, N. Corradi, I. Grigoriev, A. Gryganskyi, T. Y. James, K. O'Donnell, R. W. Roberson, T. N. Taylor, J. Uehling, R. Vilgalys, M. M. White, J. E. Stajich, Mycologia **108**, 1028 (2016)
5. C. Senes-Guerrero, A. Schußler, Fungal Divers **77**, 317 (2016) DOI: 10.1007/s13225-015-0328-7
6. A. Kryukov, A. O. Gorbunova, E. M. Machs, Y. V. Mikhaylova, A. V. Rodionov, P. M. Zhurbenko, A. P. Yurkov, Vavilov journal of genetics and breeding **24**, 158 (2020)
7. M. Kruger, H. Stockinger, C. Kruger, A. Schubler, New Phytol. **183**, 212 (2009) DOI: 10.1111/j.1469-8137.2009.02835.x
8. A. A. Kichko, T. S. Aksenova, V. M. Shapkin, A. O. Zverev, A. V. Khutti, E. E. Andronov, Sel'skokhozyaistvennaya Biologiya **54**, 990 (2019) DOI: 10.15389/agrobiology.2019.5.990rus
9. L. Sommermann, J. Geistlinger, D. Wibberg, A. Deubel, J. Zwanzig, D. Babin, A. Schlüter, I. Schellenberg, PLoS One **13**, (2018) DOI: 10.1371/journal.pone.0195345
10. A. P. Yurkov, A. A. Kryukov, A. O. Gorbunova, A. P. Kozhemyakov, G. V. Stepanova, E. M. Machs, A. V. Rodionov, M. F. Shishova, Ecological Genetics **16**, 11 (2018) DOI: 10.17816/ecogen16211-23
11. J. J. Doyle, J. L. Doyle, Phytochem. Bull. **19**, 11-15 (1987)
12. T. J. White, T. Bruns, S. Lee, J. Taylor, *PCR Protocols: a Guide to Methods and Applications* (Academic Press San Diego, 1990)
13. A. P. Yurkov, A. A. Kryukov, A. O. Gorbunova, A. V. Shcherbakov, P. M. Zhurbenko, BIO web of Conferences **23**, 02013 (2020)
14. R. C. Edgar, Bioinformatics **26**, 2460-2461 (2010)
15. S. Kumar, G. Stecher, K. Tamura, Mol. Biol. Evol. **33**, 1870 (2016)
16. E. Tisserant, M. Malbreil, A. Kuo, A. Kohler, A. Symeonidi, R. Balestrini, et al., PNAS, **110**, 50 (2013) DOI: 10.1073/pnas.1313452110
17. A. Schüßler. AMF species list, last updated 18 Mar 2021. URL: http://www.amf-phylogeny.com.
18. C. Senes-Guerrero, A. Schußler, Fungal Divers **77**, 317 (2016). DOI: 10.1007/s13225-015-0328-7