Perspectives of using Illumina MiSeq for identification of arbuscular mycorrhizal fungi

A.A. Kryukov¹, A.O. Gorbunova¹, ², E.M. Machs³, Y.V. Mikhaylova³, A.V. Rodionov², ³, P.M. Zhurbenko³, A.P. Yurkov¹

¹ All-Russian Research Institute for Agricultural Microbiology, St. Petersburg, Russia
² Saint Petersburg State University, Biological Faculty, St. Petersburg, Russia
³ Komarov Botanical Institute, RAMS, St. Petersburg, Russia

Key words: Glomeromycotina; arbuscular mycorrhiza; Illumina MiSeq; species level.

Abstr.: Arbuscular mycorrhiza fungi (AMF) form one of the most common symbioses with the majority of land plants. AMF supply the plant with various mineral elements, primarily phosphorus, and improve the water supply. The search for the most effective AMF strains for symbiosis and the creation of microbial preparations on that basis is an important task for modern biology. Owing to the difficulties of cultivation without a host plant and their high genetic polymorphism, identifying AMF is very difficult. The high number of cryptic species often also makes morphological identification unreliable. Recent years have seen growth in the number of AMF biodiversity studies performed by modern NGS-based methods, Illumina MiSeq in particular. Currently, there are still many questions that remain for the identification of AMF. The most important are whether conservative or variable sequences should be used to select a marker for barcoding and whether universal primers or those specific to AMF should be used. In our work, we have successfully used universal primers ITS3 and ITS4 for the sequencing in Illumina MiSeq of the 5.8S rDNA – ITS2 region of the 35S rRNA genes, which contain both a conservative and variable regions. The molecular genetic approach for AMF identification was quite effective and allowed us to reliably identify eight of the nine isolates to the species level: five isolates of Rhizophagus irregularis, and one isolate of R. invermaius, Paraglomus laccatum, and Claroideoglomus etunicatum, respectively. For all five R. irregularis isolates high variability in the ITS region and the absence of ecotopic-related molecular characters in the ITS2 region were demonstrated. The NCBI data is still insufficient for accurate AMF identification of Acaulospora sp. isolates from the genus to the species level.

For citation: Kryukov A.A., Gorbunova A.O., Machs E.M., Mikhaylova Y.V., Rodionov A.V., Zhurbenko P.M., Yurkov A.P. Perspectives of using Illumina MiSeq for identification of arbuscular mycorrhizal fungi. Vavilovskii Zhurnal Genetiki i Selekttsii = Vavilov Journal of Genetics and Breeding. 2019. DOI 10.18699/VJ19.38-o.

© Kryukov A.A., Gorbunova A.O., Machs E.M., Mikhaylova Y.V., Rodionov A.V., Zhurbenko P.M., Yurkov A.P., 2019

Perspektivy использования Illumina MiSeq для идентификации грибов арбускулярной микоризы

А.А. Крюков¹, А.О. Горбунова¹, ², Э.М. Мачс³, Ю.В. Михайлова³, А.В. Родионов², ³, П.М. Журбенко³, А.П. Юрков¹

¹ Всесоюзный научно-исследовательский институт сельскохозяйственной микроbióологии, Санкт-Петербург, Россия
² Санкт-Петербургский государственный университет, биологический факультет, Санкт-Петербург, Россия
³ Комаровский ботанический институт РАН, Санкт-Петербург, Россия

e-mail: rainniar@rambler.ru

Грибы арбускулярной микоризы (AMГ) образуют один из наиболее распространенных симбиозов с большинством наземных растений. AMГ снабжают растения различными минеральными элементами, а в первую очередь, фосфором, а также улучшают водоснабжение. Поиск наиболее биотехнологически эффективных штаммов AMГ и создание на их основе микробных препаратов – важная задача современной биологии. Идентификация AMГ очень сложна. Это связано, прежде всего, с высоким генетическим полиморфизмом AMГ, а также с трудностями их выращивания без растения-хозяина. Морфологическая идентификация AMГ часто ненадежна из-за большого числа криптических видов. В последние годы увеличивается число работ по изучению биологического разнообразия AMГ, проводимых современными методами на основе NGS (Next Generation Sequencing), в частности Illumina MiSeq. В настоящее время остается много вопросов по идентификации AMГ. К наиболее важным из них относятся: выбор маркера для генетического штрих-кодирования AMГ – консервативных или вариабельных последовательностей, а также выбор приоритетов – специфичных для AMГ или универсальных. В нашей работе мы успешно использовали универсальные приоритеты ITS3 и ITS4 для секвенирования с Illumina MiSeq региона 5.8S рДНК – ITS2, содержащего как консервативные, так и вариабельные участки. Эффективность подхода для идентификации AMГ оказалась достаточным для идентификации 8 из 9 изолятов до уровня вида: 5 – Rhizophagus irregularis, 1 – R. invermaius, 1 – Paraglomus laccatum, 1 – Claroideoglomus etunicatum. Для всех изолятов R. irregularis показаны высо-
Introduction

Arbuscular mycorrhizal fungi (AMF) belonging to the sub-division Glomeromycotina are a relatively small yet diverse group. Various estimations report from 240 (Stockinger et al., 2014) to 348 species (Öpik et al., 2014). At the same time, AMF form mycorrhizal relationships with more than 200,000 species of land plants (Lee et al., 2013). Identification of the AMF strains with the highest symbiosis efficiency is of great value for agricultural applications. To do it, we isolated, identified (Kryukov et al., 2017; Kryukov, Yurkov, 2018; Yurkov et al., 2018b), and evaluated the symbiotic efficiency (Yurkov et al., 2017a, b; Yurkov et al., 2018a) for a number of strains from the collection.

AMF are traditionally identified by more than 20 morphological characters (Schenk, Perez, 1990; Blaszkowski, 2019; INVAM, 2019; Schüßler, 2019). However, in some cases morphological identification fails to discriminate closely related species. A number of AMF species are morphologically indistinguishable. Furthermore, interspecific genetic polymorphism can sometimes be mistakenly treated as intraspecific polymorphism (Savy et al., 2017). Everything mentioned above may result in wrong estimations of AMF species. AMF usually cannot be cultivated on artificial media, thus making the identification very hard. As a result many AMF species are not studied at all, and so have not received the proper morphological description essential to their identification (Bruns et al., 2017). The type genus Glomus is a striking example of the taxonomic problems concerning species and genera identification. In the last twenty years, Glomus taxon has been revised a number of times, so that many Glomus species put under 14 other genera instead: Glomus, Entrophospora, Rhizophagus, Sclerocystis, Septoglomus (Schüßler, 2019). In this key, molecular-genetic identification of Glomeromycota is very significant as the exponential increase in the number of AMF DNA sequences deposited in databases over the last ten years attests (NCBI, 2018).

The ITS region is the main marker for AMF barcoding, it is well represented in different databases, among which are the UNITE (User-friendly Nordic ITS Ecotomycorrhiza) (UNITE, 2017) and MaarjAM (Öpik et al., 2010). It should be noted that the ITS is also often used for the barcoding of vascular plants and the construction of phylogeny (Rodionov et al., 2016). But the main advantage of the ITS region is the possibility of identifying AMF up to the species level. A majority of papers where less variable markers have been used report on identification up to the genus level or even to the order level (Schoch et al., 2012; Tedersoo et al., 2015).

In the case of Illumina MiSeq it is recommended that ITS2 or the full ITS region be used as the fungal barcode (Tedersoo et al., 2015). ITS2 provides a higher taxonomic resolution than SSU or LSU genes, which are suitable for identifying genera and higher-level taxa. ITS1 in fungi is usually shorter that ITS2, also ITS1 is considered as a hypervariable region and thus less suitable for barcoding fungi (Tedersoo et al., 2015). To increase efficiency of molecular-genetic identification, modifications of the universal primers ITS3 and ITS4 specific for various fungal divisions were proposed. In comparison with other primer combinations, the primer pair ITS3tagmix and ITS4tags gave a significantly larger number of sequencing reads as well as OTUs (Operational Taxonomic Units) (Tedersoo et al., 2015). At the same time, the contribution of AMF in the total OTUs pool was less than 3 %, whereas Agaricomycetes represented half of all OTUs. For AMF we proposed using the slightly modified primer ITS3 – CATC GATGAGAAGCGTAG (the modification is in bold) as a direct primer and the primer ITS4 without changes as reverse.

Various aspects of using specific primers for the AMF identification were reviewed earlier (Kryukov et al., 2017; Kryukov, Yurkov, 2018; Yurkov et al., 2018a). Identification using universal primers allows us to investigate the maximal broad range of AMF species and genera. But this then introduces the problem of foreign DNA admixture. Another problem with molecular genetic identification of AMF is the generation of chimeric sequences during the sequencing process (Senés-Guerrero et al., 2014). Phusion DNA-polymerase, which generates high accuracy PCR-products, serves to reduce the chances of introducing chimera (Senés-Guerrero et al., 2014). At the same time, special software has been developed for detecting chimeric sequences and excluding them from analysis, USEARCH for example (Edgar, 2010).

Before 2015, the main method of AMF molecular-genetic identification was cloning followed by Sanger sequencing. This method applied to AMF demands careful selection of efficient and highly specific primers, as well as the use of nested PCR. In addition, due to the high variability of marker sequences, only multiple sequence cloning can be used (Krüger et al., 2009), which is very time-consuming and labor-intensive.

NGS has turned into a powerful and attractive method of AM fungi identification since it can overcome the weak points in the Sanger-based identification. One of the earli-
est techniques of NGS was 454 pyrosequencing, which has been employed for AMF identification since 2009. Using the universal fungal primers NS31 and AM1 179,279 sequences were obtained, of which 77.5 % belonged to 47 taxa of AMF isolated from the roots of 10 plant species (Opik et al., 2009). However, these primers (NS31 and AM1) are not suitable for the analysis of the SSU region in Archaeosporaceae and Paraglomeraceae families (Helgason et al., 1998). Nonetheless, 454 pyrosequencing revealed in the roots of the *Hepatica nobilis* Mill. 1.5 times more fungal taxa than Sanger sequencing (Opik et al., 2009). This clearly showed the advantage of NGS methods over cloning-Sanger sequencing methods. 454 pyrosequencing following nested PCR was successfully used for AMF identification with a barcode in the LSU region (Senés-Guerrero, Schüßler, 2015). An interesting result of this work is that about 60 % of the studied plants each formed a symbiosis with at least 10 AMF taxa, and 2 % of plants had more than 25 AMF species in their root system. The authors used the LSU-D1fmodified primer (Senés-Guerrero, Schüßler, 2015) together with the LSUumBr primer (Krüger et al., 2009) in the second round of nested PCR. This modification allowed us to obtain 698,297 sequences, of which 0.17 % were the target AMF sequences, 41 taxa were detected, of which 15 are unknown, not registered in the databases (Senés-Guerrero, Schüßler, 2015). 454 pyrosequencing has its advantages over cloning-Sanger sequencing, but this technique is more expensive than the Illumina MiSeq technology that replaced pyrosequencing.

With the development of Illumina technology, Illumina MiSeq has been becoming more and more widely used due to its relatively low cost. Illumina MiSeq compared to HiSeq 2000 allows for processing sequences of reads 2.5 times longer, and each sequencing in this case is cheaper. The advantage of HiSeq is more deep sequencing, which allows us to obtain reads at a rate of three orders greater than MiSeq. But this is less significant for fungi identification than sequence read length (Razzauti et al., 2015).

A comparative study of the efficiency of 454 pyrosequencing and Illumina MiSeq showed a difference of five times in the diversity of sequences (in favor of the second method), but both approaches revealed the same species composition (Vasar et al., 2017). A comparative study of the Illumina MiSeq and Ion Torrent Personal Genome Machine (PGM) showed that the second method generated a 2–5 fold greater rate of error than Illumina MiSeq (Salipante et al., 2014).

The objectives of this study included the identification of AM fungi from the collection of the All Russian Research Institute for Agricultural Microbiology using the Illumina MiSeq approach and universal primers for the ITS2 region, supplemented by the morphological characteristics of the analyzed AMF spores.

Materials and methods

**AMF isolation.** AMF were isolated in 2015 from samples, collected in various habitats in two different regions (author of the analyzed AMF isolates: A.P. Yurkov). Only isolates with spores with an unambiguously identifiable morphology were used for molecular-genetic identification. Four strains were isolated from samples taken in the Rostov region (1.4 km NW from Zernograd): 46°52′2″ N, 40°16′8″ E, a tree belt area with oaks, maples and alders: 1) isolate (strain) 01-053 was isolated from *Ambrosia artemisiifolia* roots, isolate (strain) 01-056a and, 3) isolate 01-056b were both isolated from one ordinary chernozem soil sample, but differentiated via reinoculations of spores with a different morphology, 4) at 46°52′7″ N, 40°16′8″ E, a maize field, the isolate (strain) 02-060 was isolated from a *Zea mays* 282MB Zelenogradskij hybrid roots. Five isolates were obtained from samples taken in the Moscow region (Lobnya town, academic village in Lugovaya): 1) 56°02′33.80″ N, 37°29′13.70″ E, natural meadow, the isolate (strain) 03-097 was isolated from *Vicia sepium* roots, 2) 56°02′24.30″ N, 37°29′20.00″ E, a *Festuca rubra* field, isolate (strain) 04-067 was isolated from *Festuca rubra* roots, 3) isolate (strain) 04-068 was isolated from *Agrostis vulgaris* roots, 4) 56°02′31.40″ N, 37°29′17.60″ E, *Medicago × varia* field, isolate (strain) 05-077 was isolated from *Trifolium pratense* roots, 5) isolate 05-104 was isolated from a sod-podzol gleyic soil sample.

**AMF cultivation.** AMF collection at the All-Russian Research Institute for Agricultural Microbiology is cultivated in the *Plectranthus australis* R. Br. (taxonomical synonyms *P. verticillatus* (L.f.) Druce, *P. nuxmullarius* Briq.) line of Swedish ivy. For AMF-inoculated plants a soil-sand mix was used, described earlier (Yurkov et al., 2015). The substrate had a low level of plant-available phosphorus (3.0 mg P₂O₅/100 g). *Plectranthus* cuttings 12–15 cm length with two leaves were sterilized in 0.1 % sodium hypochlorite, then germinated in water. On the 7th day *Plectranthus* plants were inoculated by root fragments, containing AMF vesicles and arbuscules. Root fragments of mycorrhizal *Plectranthus* were selected by visual analysis in stereomicroscope MBS-10 (LZOS, Russia). For further reincultivations (each 6–8 months) sporocarps, or an arrangement of spores in an out-root area, or 5 mm length root fragments with large quantities of observable vesicles were used. Each inoculated plant was cultivated in an individual container with 350 g of sterile substrate at +24–26 °C, 18 h light day. Two luminescent lamp LB-40 (Russia) and OSRAM L36/77 Fluora (Germany) on a 1:1 ratio with output lumen ~4000 were used. Plants were watered every other day by 60 % of soil full water capacity. For the culture purification AMF spores were reincultivated at least three times.

**Morphological identification of AMF.** More that 20 features were used for morphological identification (Schenck, Pérez, 1990; Blaszkowski, 2003; INVAM, 2019): color, transparency, size and shape of extraradical (out-root) spores; shape of the place of attachment of spores to subtending hypha; number, thickness, density, elasticity or fragility, color in Melzer’s reagent of layers of spore walls and subtending hypha; presence/absence/disappearance/appearance of spore wall layers and subtending hypha during ontogenesis (from juvenile to mature spore); presence/absence of a septum in the place of attachment of the spores to subtending hypha; structural characteristics of AMF and intraradical spores.

In order to assess mycorrhization parameters and the type of mycorrhiza, the roots were macerated and stained by trypan blue according to the method developed by J.M. Phillips and D.S. Hayman (1970). Mycorrhization parameters were determined by light microscopy (Trouvelot et al., 1986) using a special computer program, developed earlier in our research group (Vorob’ev et al., 2016).
DNA extraction, PCR and sequencing. DNA extractions were carried out using the method of J.J. Doyle and J.L. Doyle (1987), with modifications. Micorrhizoroots of P australis were washed twice in distilled water, placed in 2 ml tubes, dried at +45 °C, and mechanically homogenized with glass beads - 2–4 mm in the FastPrep24 (MP Biomedicals, USA), followed by CTAB-protocol. The target region ITS2 was amplified with universal primers ITS3 (5'-GCTCATCGATGAAAGACG CAGC-3') and ITS4 (5'-TCTCCTCGTTATGATGTC-3') (White et al., 1990). Ready mix ScreenMix (Evrogen, Russia) was used for PCR. Amplicons were sliced from agarose gel and purified by the silica approach. Illumina library preparation was made according to MiSeq Reagent Kit Preparation Guide in the Core Center of “Genomic Technologies, Proteomics and Cell Biology” at the All-Russia Research Institute for Agricultural Microbiology (St. Petersburg, Russia). Libraries were sequenced on the Illumina MiSeq platform with MiSeq® Reagent Kit v3 (600 cycle) according to the manufacturer’s instruction (Illumina Inc., USA).

Bioinformatics and fungal OTUs analysis. Two pipelines were used for sequencing data analysis. 1. Illumina reads processing were done by USEARCH software (Edgar, 2010). The key steps for USEARCH data treatment are described on https://www.drive5.com/usearch/. In further paragraphs we are briefly describing these steps. In the case of paired-end sequencing, Illumina sequencer makes sequences from both ends of a fragment and generates two files with forward and reverse reads. This data is written in FASTQ format, where each nucleotide corresponds with its quality score. For further treatment forward and reverse reads are merging using fastq_mergepairs command to give consensus sequences. This step includes resolving any mismatches found in the overlap alignments and calculation the posterior quality scores for the consensus sequences (Edgar, Flyvbjerg, 2014). To discard low-quality reads expected error filtering with fastq_filter command is used.

The next step is a dereplication, which means sorting unique sequences in order of decreasing abundance in the dataset. After this, singletons (sequences that are present exactly once) are discarded, since they are likely to have errors. However the remaining reads can still have errors. So the goal of the final step (denoising) is to identify a set of correct biological sequences. The denoising can be made by UPARSE algorithm which clusters sequences with 97% or more (Edgar, 2018) similarity and then chooses the most abundance sequence in each cluster. Also chimeric sequences, which occurs by combination parts of two or more biological sequences are detected and deleted at this step.

Raw forward and reverse reads were merged with minimal length parameter (“-fastq_minmergelen”) 130 bp and maximal difference parameter (“-fastq_maxdiffs”) 30 bp. Then low-quality read ends including primer sequences were trimmed, reads were filtered based on expected error value (E_max = 1). Singletons were removed from the dataset. Then data were divided on operational taxonomic units (OTUs) with a 97% similarity cut-off by UPARSE algorithm (Edgar, 2013). Chimeric sequences were removed. For further analysis the most represented sequence from each OTU was chosen. Data were checked for cross-talk errors. Sequences of AMF species were selected by BLAST+ (Altschul et al., 1990).

2. The primers we used are universal for a broad range of species; and after amplification in the ITS2 region the extracted plant DNA prevail over fungi DNA. A second pipeline was made for the selection of rare and unique reads with a high homology of AMF sequences. After quality control (FastQC) forward and reverse reads were trimmed and merged with minimal length parameter 230 bp by trimmomatic (Bolger et al., 2014) and fastq-join software (Aronesty, 2013). Then sequences were demultiplicated and sorted in the descending order of their frequency. AMF sequences were selected via character for AMF motifs, then aligned and checked via BLAST.

Obtained sequences were submitted to the GenBank database (https://www.ncbi.nlm.nih.gov/). Evolutionary analyses were conducted by using the Maximum Likelihood method in MEGA7 software (Kumar et al., 2016) with implementation of the Tamura-Nei model (Tamura, Nei, 1993) and 1,000 bootstrap analyses.

Results
Sequencing of 9 isolates yielded approximately 381,249 pair of reads, from 19,236 to 81,054 joined sequences for each isolate. The following OTUs of AMF isolates were identified via BLAST at the genera or species level and submitted to NCBI: MK948362-MK948371 (isolate number 01-053), MK948403-MK948404 (01-056a), MK968150 (01-056b), MK948427-MK948429 (02-060), MK948492-MK948496 (03-097), MK948434-MK948436 (04-067), MK948447 (04-068), MK948486-MK948491 (05-077), MK948503-MK948504 (05-104). The length of obtained sequences varied from 340 to 366 bp. Variability in GC content was distinct in different genera: a narrow range was shown in genera Claroideoglo- mus (38–39 %), Rhizophagus (36–39 %) and Paraglomus (42–46 %), whereas GC content in Acaulospora varied from 31 to 46 %. Owing to the significant variability of the ITS region in AM, alignment of sequences belonging to different genera and orders is ineffective, and in some cases impossible. As a result, four separate phylogenetic trees were constructed for the four genera mentioned above (Fig. 1–4).

Unusual deletion was determined by ITS2 sequence alignment. This deletion of 5–6 bp in alignment positions 97–102 was identified in various species of genus Rhizophagus, and in all OTUs of genus Paraglomus, and in Racocetra weresubiae (FR750135) (NCBI, 2019). This fact can serve to indicate the presence of a deletion-specific site related to secondary RNA structure. However, it can also indicate a relationship of sequences in this region, which is of greater interest because it is well known that AM mycelia contain a significant number of nuclei (Hosny et al., 1998), some of which carry this deletion in ITS.

Some samples (01-053, 01-056a, 02-060, 04-068, 05-077, and 03-097) demonstrated a high similarity with Rhizophagus (see Fig. 1). OTU isolates 01-053, 01-056a, 02-060, 04-068, 05-077 fell with high accuracy into the clade formed by Rhizophagus irregularis, while OTUs of 03-097 were placed in one clade with R. invermaius (bootstrap index = 97). This genus includes a small number of species, about 20 according to A. Schüßler (2019), but in the NCBI only 8 species are represented by ITS sequences. Thus one may consider that Rhizophagus is a genus that requires further sequencing.
Перспективы использования Illumina MiSeq для идентификации грибов арбускулярной микоризы
А.А. Крюков, А.О. Горбунова, Э.М. Мачс, Ю.В. Михайлова 
А.В. Родионов, П.М. Журбенко, А.П. Юрков
2019

Вавиловский журнал генетики и селекции / Vavilov Journal of Genetics and Breeding • 2019

053-6 MK948367 Rhizophagus irregularis
053-8 MK948369 Rhizophagus irregularis
060-3 MK948429 Rhizophagus irregularis
053-5 MK948366 Rhizophagus irregularis
077-3 MK948488 Rhizophagus irregularis
077-6 MK948491 Rhizophagus irregularis
077-5 MK948490 Rhizophagus irregularis
FM8655610 R org Rhizophagus irregularis DAOM197198 W3182/Att1192-52
053-3 MK948364 Rhizophagus irregularis
077-2 MK948487 Rhizophagus irregularis
056a-1 MK948403 Rhizophagus irregularis
FM8655608 R org Rhizophagus irregularis DAOM197198 W3182/Att1192-52
053-7 MK948368 Rhizophagus irregularis
060-2 MK948428 Rhizophagus irregularis
053-1 MK948362 Rhizophagus irregularis
077-1 MK948486 Rhizophagus irregularis
053-9 MK948370 Rhizophagus irregularis
053-4 MK948365 Rhizophagus irregularis
056a-2 MK948404 Rhizophagus irregularis
FM865591 rh org Rhizophagus irregularis BEG195 W5272/Att1485-12
053-10 MK948371 Rhizophagus irregularis
060-1 MK948427 Rhizophagus irregularis
068-1 MK948447 Rhizophagus irregularis
FR750191 org Rhizophagus irregularis from Blaskowski
FR750187 org Rhizophagus irregularis W4682/Att857-12
077-4 MK948489 Rhizophagus irregularis
FR750078 R org Rhizophagus irregularis MUCL43195
GQ205076 Rhizophagus custos
JN847437 Rhizophagus custos
FM992402 R org Rhizophagus proliferus MUCL41827
FN547500 R org Rhizophagus proliferus MUCL41827 W4728/Att1296-0
FM865545 R org Rhizophagus intraradices MUCL49410 W5070/Att1102-9
FM865606 R org Rhizophagus intraradices FL208 W5166/Att4-38
FR750073 org Rhizophagus fasciculatus MUCL46100
FR750071 org Rhizophagus fasciculatus MUCL46100
FM865546 R org Rhizophagus intraradices MUCL49410 W5070/Att1102-9
097-3 MK948494 Rhizophagus invermaius
097-5 MK948496 Rhizophagus invermaius
097-1 MK948492 Rhizophagus invermaius
097-4 MK948495 Rhizophagus invermaius
HG969390 Rhizophagus invermaius
LN624112 Rhizophagus invermaius
097-2 MK948493 Rhizophagus invermaius
GQ205083 Rhizophagus clarus
FM865541 org Rhizophagus clarus W3776/Att894-7
HG964399 Rhizophagus melanus
AF413089 Funneliformis mosseae

**Fig. 1.** Phylogenetic tree of ITS-sequences from the genus Rhizophagus.

At the same time, it is beyond doubt that this study provides a molecular genetic identification at the species level of isolates 01-053, 01-056a, 02-060, 04-068, 05-077, 03-097, 03-097.

Due to the ambiguous position of OTUs of isolate 05-104 on a pre-built phylogenetic tree of the Acaulospora genus, as well as due to the possibility that OTUs of this isolate could incorporate other close genera from the Acaulosporaceae family, species from the Archaeospora, Ambispora, Diversispora, Gigaspora, Racocetra, Redeckera and Scutellospora genera were added to the tree. We found that both OTUs of 05-104 isolates were included in the Acaulospora clade with high bootstrap support (see Fig. 2).

However, owing to the significant differences from other species of Acaulospora, we define the MK948503 sequence...
Fig. 2. Common phylogenetic tree of ITS-sequences from Acaulospora, Archaeospora, Ambispora, Diversispora, Gigaspora, Racocetra, Redeckera and Scutellospora genera.
(05-104 isolate) as a virtual taxon. The MK948504 sequence (05-104 isolate), though it has a high similarity with the _A. sieverdingii_ and _A. paulinae_ species, did not show similarities with the indicated species according to morphological data, therefore it was also identified as a virtual taxon. At the same time, it can be reliably stated that these two OTUs belong to different taxa, since they have significant differences in GC content (31 and 46%). But differences did not appear in the case of morphological analysis, highlighting the obstacles of distinguishing among _Acaulospora_ species.

The OTUs of 01-056b isolate formed a well-supported subclade in the _Paraglomus_ genus (bootstrap index = 99) (see Fig. 3). Furthermore, the Paraglomerales order hosts two more genera (_Innospora_ and _Pervetustus_) in addition to the _Paraglomus_ genus, but the ITS sequences for them are unknown. Thus the outgroup for the phylogenetic tree was taken from another order. The 01-056b isolate was identified as _Paraglomus laccatum_. The sequences of the _P. laccatum_ 01-056b isolate turned out to be the shortest in comparison with isolates of other genera and ITS. Only three of eight species of this genera have ITS sequences in NCBI (Schüßler, 2019).

The _Claroideoglomus_ genus is the most studied of the above mentioned genera. The ITS data were reported for six out of eight species (Schüßler, 2019). The OTU of isolate 04-067 with high support belongs to the clade formed by the species _Claroideoglomus etunicatum_ (see Fig. 4).

To verify whether the molecular genetic identification of AMF was effective, we conducted a morphological identification of nine isolates (Supplementary). The principal morphological characteristics for comparison were: size, shape, and color of the spores in air by the CMYK standard; the shape and thickness of subtending hypha; the presence of a septum in subtending hypha; the wall thickness of subtending hypha; the presence of a septum in subtending hypha; the shape of the spores in Melzer’s reagent by the CMYK standard; the shape, color, and thickness of spore layers in Melzer’s reagent by the CMYK standard; the wall thickness of subtending hypha; the wall thickness of subtending hypha; the presence of a septum in subtending hypha; type of mycorrhiza in _P. australis_ standards; the environment of_ P. laccatum_ and _P. australis_ in the ARRIAM collection – are of the following taxa: 1) 01-053 isolate – _Rhizophagus irregularis_; 2) 01-056a – _R. irregularis_; 3) 01-056b – _Paraglomus laccatum_; 4) 02-060 – _R. irregularis_; 5) 03-097 – _R. invermaius_; 6) 04-067 – _Claroideoglomus etunicatum_; 7) 04-068 – _R. irregularis_.

1 Supplementary are available in the online version of the paper: http://www.bionet.nsc.ru/vogis/download/pict-2020-24/appx0.pdf
Perspectives of using Illumina MiSeq for identification of arbuscular mycorrhizal fungi

Discussion

The Illumina MiSeq approach allowed us to determine eight out of nine AMF isolates from the All Russian Research Institute for Agricultural Microbiology collection. It is a very powerful method, which enables the identification of a large number of AMF taxa in fungal communities, especially if the proportion of targeted marker sequences in a sample is small. There are other effective methods of NGS, but several comparative studies indicate that MiSeq in some cases is more efficient as it can provide longer reads and fewer errors in comparison to HiSeq and IonTorrent techniques, respectively (Salipante et al., 2014; Razzaauti et al., 2015).

These comparisons show the clear advantage of the sequence-targeted NGS approach as contrasted with the alternative cloning-sequencing method for AMF species identification. However, the main disadvantage of the Illumina MiSeq are relatively short reads (250 bp × 2), which do not allow the use of long markers, such as the entire cloned SSU-ITS1-5.8S-ITS2-LSU region that was most often used for AMF barcoding from 2009 to 2012 (Krüger et al., 2009; De Castro et al., 2018). A prerequisite for the correct AMF identification by the MiSeq method is the employment of a short marker region for sequencing (400–500 bp). The most commonly used length-appropriate marker is ITS2. The advantage of this region is that it provides sufficient variability for identification at the species level. Other less variable regions such as D1–D2 of the LSU allow identification only at the genus level (Krüger et al., 2009).

The use of universal primers for ITS2 region is the optimal choice for the identification of poorly studied taxa, which in recent years has helped to identify a significant number of virtual taxa of AMF (Öpik et al., 2014). These taxa can subsequently receive species names in the presence of individual morphological features and stable maintenance of isolates/strains in culture. For example, we studied the 05-104 sample in which there are two OTUs that belong to two different virtual Acaulospora sp. taxa. Another advantage of using the ITS2 region for identification is that it is a substantially representative sequence database has been stored in NCBI in comparison to other marker regions, for example, for the SSU region (NCBI, 2018).

It is important to note that OTUs related to one R. irregularis species and collected in different ecotopes (isolates 01-053, 01-056a, 02-060, 04-068 and 05-077) did not cluster separately on the phylogenetic tree (see Fig. 1). This suggests that all studied R. irregularis isolates shared one ribotype. Thus, it was not possible to identify the ecotype-related features of R. irregularis isolates, perhaps they are objectively missing.

Also, it should be noted that NGS methods produce huge data arrays. We paid special attention to choose the correct tools for data treatment. Several pipelines have been developed to process rDNA sequences that are generated using Illumina’s MiSeq platform. Among them mothur (Schloss et al., 2016), QIIME (Caporaso et al., 2010), USEARCH and VSEARCH (Rognes et al., 2016) are the most popular. Generally, all of them have similar steps in sequences treatment. We used USEARCH since it has a large set of tools and detailed documentation. However, two approaches were used to search for AMF OTUs among a great pool of sequences. Both approaches were equally effective since they allowed us to identify the
same set of taxa. At the same time the use of the USEARCH software has significantly reduced the time spent on data processing, which makes it possible to recommend it as the main tool to identify AMF from NGS data.

Combining the results of molecular genetics and morphological identification, we assert that the effectiveness of the Illumina MiSeq method as applied to AMF identification is not inferior to morphological methods that are significantly more labor intensive. However, the NCBI database is still insufficient for identification of some AMF species. The reason for this is that more than half of the known AMF taxa are still absent in the database. For example, according to A. Schüßler (2019), there are up to 56 species in the genus *Acaulospora*. The NCBI contains 37 species just for *Acaulospora* sp., and ITS sequences are presented for only 28 species. Thus, the obtained OTUs of the 05-104 isolate may belong to one of *Acaulospora* species that has yet to be studied.

**Conclusion**

We have determined that the most effective method for AMF identification is Illumina MiSeq supplemented by application of universal primers for the ITS2 region. Considerable efforts of morphologists in collaboration with molecular geneticists are required to establish a reliable taxonomy of the Glomeromycotina subdivision and to improve the efficiency of the molecular genetic AMF identification as a key method in the future.

**References**

Altschul S.F., Gish W., Miller W., Myers E.W., Lipman D.J. Basic local alignment search tool. J. Mol. Biol. 1990;215:403-410. DOI 10.1016/S0022-8028(93)70772-7.

Aronesty E. Comparison of sequencing utility programs. TOBioiJ. 2013;7:1-8. DOI 10.2174/187506201307010001.

Blaszkowzki J. Definitions of morphological characters of spores of arbuscular fungi. 2003. Available at http://www.zor.zut.edu.pl/Glomeromycota/Definitions%20of%20spore%20characters.htm

Blaszkowzki J. Species descriptions and illustrations. 2019. Available at http://www.zor.zut.edu.pl/

Bolger A.M., Lohse M., Usadel B. Trimmomatic: A flexible trimmer for Illumina Sequence Data. Bioinformatics. 2014;30(15):2114-2120. DOI 10.1093/bioinformatics/btu170.

Bruns T.D., Corradi N., Redecker D., Taylor J.W., Öpik M. Glomeromycotina subdivision and to improve the efficiency of the molecular genetic AMF identification as a key method in the future.

**References**

Edgar R.C. Search and clustering orders of magnitude faster than BLAST. Bioinformatics (Oxford, England). 2010;26(19):2460-2461. DOI 10.1093/bioinformatics/btq461.

Edgar R.C. UPARSE: Highly accurate OTU sequences from microbial amplicon reads. Nat. Methods. 2013;10:996-998. DOI 10.1038/nmeth.2604.

Edgar R.C. Updating the 97 % identity threshold for 16S ribosomal RNA OTUs. Bioinformatics. 2018;34(14):2371-2375. DOI 10.1093/bioinformatics/bty113.

Edgar R.C., Flyvbjerg H. Error filtering, pair assembly and error correction for next-generation sequencing reads. Bioinformatics. 2015;31(21):3476-3482. DOI 10.1093/bioinformatics/btv401.

Heldagson T., Daniell T.J., Husband R., Fitter A.H., Young J.P.W. Ploughing up the wood-wide web? Nature. 1998;394(6692):431-431. DOI 10.1038/28764.

Hosny M., Gianninazzi-Pearson V., Dulieu H. Nuclear DNA content of 11 fungal species in Glomales. Genome. 1998;41(3):422-426. DOI 10.1139/g98-038.

INVMAM. Species descriptions from reference cultures. 2019. Available at http://fungi.invmam.vwu.edu/thefungi/species-descriptions.html

Krüger M., Stockinger H., Krüger C., Schüßler A. DNA-based species level detection of Glomeromycota: one PCR primer set for arbuscular mycorrhizal fungi. New Phytol. 2009;183(1):212-223. DOI 10.1111/j.1469-8137.2009.02835.x.

Kryukov A.A., Yurkov A.P. Optimization procedures for molecular-genetic identification of arbuscular mycorrhizal fungi in symbiotic phase on the example of two closely kindred strains. Mikologiya i Fitotopatiya = Mycology and Phytopathology. 2018;52(1):38-48. (in Russian)

Kryukov A.A., Yurkov A.P., Shishova M.F. Features of the molecular phylogeny of arbuscular mycorrhizal fungi based on the analysis of the SSU – ITS1-5.8S rRNA – ITS2 – LSU marker region. Estestvennuye i Tekhnicheskie Nauki = Natural and Technical Sciences. 2017; 2:15-21. (in Russian)

Kumar S., Stecher G., Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. Mol. Biol. Evol. 2016;33:1870-1874. DOI 10.1093/molbev/msw054.

Lee E.H., Jo K.K., Ko H.K., Eom A.H. Diversity of arbuscular mycorrhizal fungi and their roles in ecosystems. Mycobiology. 2013; 31(2):347-348. DOI 10.1007/s11259-013-0310-5.

Lind K., Limpens E., Zhang Z., Ivanov S., Saunders D.G.O., Mu D., Pang E., Cao H., Cha H., Tao Lin T., Qian Zhou Q., Yi Shang Y., Ying Li Y., Trupti Sharma T., Van Velzen R., de Ruijter N., Aanen D.K., Win J., Sophien Kamanon S., Ton Bisseling T., Geurts R., Huang S. Single nucleus genome sequencing reveals high similarity among nuclei of an endomycorrhizal fungus. PLoS Genet. 2014;10(1):e1004078. DOI 10.1371/journal.pgen.1004078.

NCBI. National Center for Biotechnology Information. 2018. Available at https://www.ncbi.nlm.nih.gov

Öpik M., Davison J., Moora M., Zobel M. DNA-based detection and identification of Glomeromycota: the virtual taxonomy of environmental sequences. Botany. 2014;92(2):135-147. DOI 10.1139/bot-2013-0110.

Öpik M., Metsis M., Daniell T.J., Zobel M., Moora M. Large-scale parallel 454 sequencing reveals host ecological group specificity of arbuscular mycorrhizal fungi in a boreonemoral forest. New Phytol. 2009;184(2):424-437. DOI 10.1111/j.1469-8137.2009.02920.x.

Öpik M., Vanatoa A., Vanatoa E., Moora M., Davison J., Kalwij M.J., Reier Ü., Zobel M. The online database MaarjAM reveals global and ecosystemic distribution patterns in arbuscular mycorrhizal fungi (Glomeromycota). New Phytol. 2010;188(1):223-241. DOI 10.1111/j.1469-8137.2010.03334.x.

Phillips J.M., Hayman D.S. Improved procedures for clearing roots and staining parasitic and vesicular-arbuscular mycorrhizal fungi for rapid assessment of infection. Transact. Brit. Mycol. Soc. 1970;55:158-161. DOI 10.1016/S0007-5531(70)80110-3.
A.A. Kryukov, A.O. Gorbunova, E.M. Machs, Y.V. Mikhaylova
A.V. Rodionov, P.M. Zhurbenko, A.P. Yurkov

Perspectives of using Illumina MiSeq for identification of arbuscular mycorrhizal fungi

Razzaqui M., Galan M., Bernard M., Maman S., Klopp C., Charbonnel N., Vassies-Tauass M., Eloit M., Cosson J.-F. A comparison between transcriptome sequencing and 16S metagenomics for detection of bacterial pathogens in wildlife. PLoS Negl. Trop. Dis. 2015; 9(8):e0003929. DOI 10.1371/journal.pntd.0003929.

Rodionov A.V., Gnatikov A.A., Kotsinyan A.R. ITS1–5.8S rDNA–ITS2 sequence in the 35S rRNA genes as a marker in the reconstruction of cereal (Poaceae) phylogeny. Uspekhi Sovremennoy Biologii = Advances in Current Biology. 2016;136(5):416. (in Russian)

Rognes T., Flouri T., Nichols B., Quince C., Mahé F. VSEARCH: a versatile open source tool for metagenomics. PeerJ. 2016;4:e2584. DOI 10.7717/peerj.2584.

Salipante S.J., Kawashima T., Rosenthal C., Hoogestraat D., Cummings L.A., Sengupta D.J., Harkins T.T., Cookson B.T., Hoffman N.G. Performance comparison of Illumina and ion torrent next-generation sequencing platforms for 16S rRNA-based bacterial community profiling. Appl. Environ. Microbiol. 2014;80(24):7583-7591. DOI 10.1128/aem.02206-14.

Savary J.C., Guerres­Guerrero C., Torres­Cortés G., Pfeiffer S., Rojas M., Schüßler A. Glomeromycota species list. 2015. Available at http://www.amf­phylogeny.com/amphylo_species.html

Schloss PD. Application of a database-independent approach to assess the quality of operational taxonomic unit picking methods. mSys. 2016;1(2):e00027-16.

Schoch C.L., Seifert K.A., Huhndorf S., Robert V., Spouge J.L., Stajich J.E., Wacholder S., Geiser D.T., Hoffmann F., Tamura K., Warnock D.W., Wurtele E.S., de Andrade G.M.C., Bruns T. Towards a universal DNA barcode marker for fungi. Proc. Natl. Acad. Sci. USA. 2012;109:6241-6246. DOI 10.1073/pnas.1117018109.

Schüßler A. Glomeromycota species list. 2019. Available at http://www.amf­phylogeny.com/amphylo_species.html

Senés-Guerrero C., Schüßler A. A conserved arbuscular mycorrhizal fungal core­species community colonizes potato roots in the Andes. Fungal Divers. 2015;77(1):317-333. DOI 10.1007/s11235-015-0328-7.

Senés-Guerrero C., Torres-Cortés G., Pfeiffer S., Rojas M., Schüßler A. Potato-associated arbuscular mycorrhizal fungal communities in the Peruvian Andes. Mycorrhiza. 2014;24(6):405-417. DOI 10.1007/s00572-013-0549-0.

Stockinger H., Krüger M., Schüssler A. DNA barcoding of arbuscular mycorrhizal fungi. New Phytol. 2010;187(2):461-474. DOI 10.1111/j.1469-8137.2010.03262.x.

Stockinger H., Peytet-Guzzon M., Koegel S., Bouffaud M.-L., Redeker D. The largest subunit of RNA polymerase II as a new marker gene to study assemblages of arbuscular mycorrhizal fungi in the field. PLoS One. 2014;9(10):e107783. DOI 10.1371/journal. pone.0107783.

Tamura K., Nei M. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol. Biol. Evol. 1993;10:512-526. DOI 10.1093/oxfordjournals.molbev.a040023.

Tedeross L., Anslan S., Bahram M., Põllme S., Riit T., Liiv I., Köljalg U., KIisand V., Nilsson H., Hildebrand F., Bork P., Abarenkov K. Shotgun metagenomes and a new primer pair­barcode combination of amplicons reveal biases in metabarcoding analyses of fungi. MycoKeys. 2015;10:1-43. DOI 10.3897/mycokeys.10.4852.

Trouvelot A., Kough J.L., Gianinazzi-Pearson V. Mesure du taux de mycorrhization V A d’un système radiculaire. Recherche de méthodes ayant une signification fonctionnelle. In: Gianinazzi-Pearson V., Gianinazzi S. (Eds.). Physiological and Genetic Aspects of Mycorrhizae. Paris: INRA-Press. 1986. (In French)

UNITE. Unified system for the DNA based fungal species linked to the classification Ver. 7.2. 2017. Available at https://unite.ut.ee

Vasar M., Andreson R., Davison J., Jairus T., Moora M., Rennm M., Young J.P.W., Zobel M., Öpik M. Increased sequencing depth does not increase captured diversity of arbuscular mycorrhizal fungi. Mycorrhiza. 2017;27(8):761-773. DOI 10.1007/s00572-017-0791-y.

Vorob‘ev N.I., Yurkov A.P., Provorov N.A. Programma Vychisleniya Indeksov Mikorizatsii Kornei Rastenii = The Program for Calculation of Indexes of Plant Root Mycorrhization. Registration certificate of computer software Svid-vo No. 2016612112 ot 12.02.16 ogs. reg. programmy EVM. FGBNUI VNIIKSHM (RF). Razavil. 21.12.2015. Opubl. 20.03.16 No. 2016612112, 12.02.16. FSBSI ARRIAM (RU). Appl.: December 21, 2015. Publ. March 20, 2016. (In Russian)

White T.J., Bruns T., Lee S., Taylor J. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis M.A., Gelfand D.H., Sninsky J.J., White T.J. (Eds.). PCR Protocols: a Guide to Methods and Applications. San Diego: Academic Press, 1990.

Yurkov A.P., Jacobi L.M., Gapeeva N.E., Stepanova G.V., Shishova M.F. Development of arbuscular mycorrhiza in highly responsive and mycotrophic host plant – black medic (Medicago lupulina L.) Journal of the Russian Academy of Biological Science. 2015;46(5):263-275. DOI 10.1134/S1062360415050082.

Yurkov A.P., Kozhemjakov A.P., Stepanova G.V. The effectiveness of some microbial bioproducts based on bacteria and fungi arbuscular mycorrhiza. Multifunctional Adaptive Fodder Production: Collection of Proceedings (Mnogofunktsional’noe Adaptivnoe Kormoproizvodstvo: Sbornik Trudov). 2018a;19:20-28. (In Russian)

Yurkov A.P., Kryukov A.A., Gorbunova A.O., Kožhемjakов А.Н., Stepanova G.V., Machs E.M., Rodionov A.V., Shishova M.F. Molecular genetic identification of arbuscular mycorrhiza fungi. Eco­logicheskaya Genetika = Ecological Genetics. 2018b;16(11):23. DOI 10.17866/ecogen16211-23. (In Russian)

Yurkov A.P., Kryukov A.A., Jacobi L.M., Kožhемjakov A.P., Shishova M.F. The relationship between the activity indicators and the effectiveness of arbuscular mycorrhiza fungi strains of various origin. Tavricheskiy Vestnik Agrarnoy Nauki = Tavrichesky Bulletin of Agrarian Science. 2017a;4(12):31-41. (In Russian)

Yurkov A.P., Laktionov Yu.V., Kožhемjakov A.P., Stepanova G.V. Analysis of the symbiotic efficiency of bacterial and fungal preparations on feed crops according to seed yield. Kormoproizvodstvo = Fodder Production. 2017b;3:16-21. (In Russian)