

Intraspecies variability of the 16S rRNA gene of the soil bacteria *Acinetobacter lwoffii* and *Paenibacillus taichungensis*

Anastasiya Nechayeva, Konstantin Boyarshin*, Olga Bespalova, Viktoriia Iatsenko, Evgeniy Seliverstov, Violetta Klyueva, and Olesya Makanina

Belgorod State National Research University, 308015, Belgorod, Russia

Abstract. The functioning of well-studied key groups of soil microorganisms depends on the microbial ecosystem in which they function in interaction with species belonging to other ecological groups. An analysis of the composition of the soil microbiota, and the possibility of analyzing the number of representatives of specific species, remains an urgent task of soil microbiology. Such analysis could be performed using modern and expensive methods such as metagenomics and mass-spectrometry, but some questions could also be answered using real-time PCR. This well-known approach is rather cheap for massive analysis and is ready to present reproducible results for practical agricultural applications. Understanding the variability of the primary structure of 16S rRNA is key to the reliable identification of bacterial species and provides an opportunity to choose the optimal pathways for their detection by PCR. In this work, analysis of the sequences of 16S rRNA of two species of soil bacteria, *Acinetobacter lwoffii* and *Paenibacillus taichungensis* is carried out. The most variable and most conservative areas on the level of species are detected. It was proved that conventional variable and conservative areas of the gene have on average nearly the same level of variability on the intraspecies level.

Keywords: *Acinetobacter lwoffii*, *Paenibacillus taichungensis*, variability of 16S rRNA

1 Introduction

Soil microbiome consists of thousands of species that have their special impact in soil fertility due to their participation in the sum of biochemical pathways leading to increase or decrease of bioavailability of the key nutritious chemical elements, maintenance of soil acidity, and suppression of other groups of bacteria or eukaryotes [1, 2]. The role of the soil microbiome as a whole remains underestimated against the background of several key groups, such as nitrogen fixers, nitrifiers, and the main groups of biodegraders.

Correcting the ratio of soil bacteria species is a promising method for improving soil fertility [3]. Such approaches as introducing nitrogen-fixing and nitrifying species into the soil are well known. This is not surprising, since the microorganisms involved in the cycle of nitrogen transformations are an important factor in ensuring the nitrogen content of plants [4]. The components of the natural soil microbial community that play roles related

* Corresponding author: boyarshin@bsu.edu.ru

to the transformations of nitrogenous compounds can have different effects on the level of bioavailability of nitrogenous compounds contained in the soil. In addition to the ecological groups that oxidize ammonium nitrogen to nitrites and nitrates, as well as assimilate molecular nitrogen from the atmosphere, there are widespread species that convert nitrogen into a volatile molecular state, as well as reduce it to the ammonium ion, which has toxic properties in significant concentrations [5 - 7].

Industrial cultivation of crops is associated with the removal of significant amounts of essential nutrients along with the harvest, and with the need to replenish the content of these elements in the soil by applying mineral and organic fertilizers. Ensuring the transfer of the key nutrient compounds contained in them into a form that is well absorbed by cultivated green plants is an urgent task of agricultural technology. This task is solved by the microorganisms present in the soil, either spontaneously developed in it, or additionally introduced in the form of biofertilizers [8, 9]. The effectiveness of both approaches depends on the usual structure of the microbial community of a given soil type. Environmental effects associated with the population density of key economically valuable groups of microorganisms are associated with the presence and number of other groups and their constituent specific species, which, in turn, affect the nutritional properties of the soil [10 - 12].

The detailed analysis of soil communities has become a separate extensive topic of scientific works in the field of ecology and agricultural soil science. Their usual subject was the parameters of the soil microbiota biodiversity, which generally reflect the general well-being of the soil microbial community, its ability to autoregulate and self-heal after exposure to natural and anthropogenic damaging factors [13, 14]. At the same time, a model that links the abundance of a wide range of different taxa and ecological groups with the provision of key biochemical processes necessary to maintain the efficiency of intensive agriculture remains to be expected [15].

The identification of an expanded list of key groups and taxa that have a significant impact on soil fertility, therefore, remains an urgent task, waiting for successive steps to solve it. In the arsenal of methods that are convenient for analyzing the presence and abundance of certain taxa in the soil microbiota, in addition to expensive high-performance methods such as metagenomic sequencing and proteomic approaches, there is such a common and effective analytical method as real-time PCR. Also, cultural methods, in particular, those related to the isolation of pure cultures of microorganisms and their determination by the 16S rRNA sequence (**Fig. 1.**) using PCR amplification and the Sanger method, do not lose their relevance. For the application of these approaches related to the determination and interpretation of the primary sequence of the 16S rRNA gene, it is of paramount importance to accumulate data on the variability of this sequence at various taxonomic levels [16].

Of particular importance is the intraspecific level of variation, which is difficult to correlate with interspecific and intergeneric diversity, and thus makes it difficult to identify species in cultural studies. Traditional approaches, often based on sequencing only a few variable regions of the 16S rRNA gene, are potentially vulnerable to aberrations associated with intraspecific variability. This study focuses on the analysis of the intraspecific variability of 16S rRNA of two soil bacteria species *Acinetobacter lwoffii* and *Paenibacillus taichungensis*, which are common in chernozem soils and are of interest as components of agricultural ecosystems.

2 Materials and methods

Ten pure cultures of bacteria were isolated on the devoid of carbon sources Vinogradsky's medium from the soil sample we took from a depth of 15 cm. Optimal growth was observed at a temperature of 30 °C.

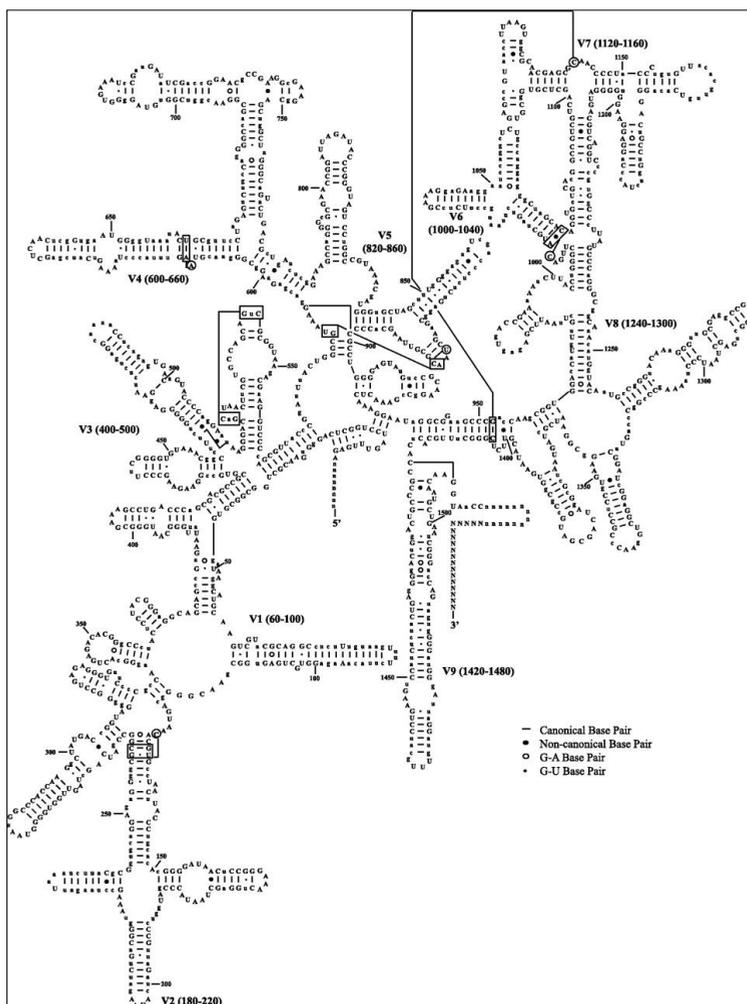


Fig. 1. Secondary structure of 16S rRNA. Variable regions are indicated as V with their numbers and positions. By Hall et al. 2020 [17].

Isolation of genomic DNA was carried out using a set of reagents "K-sorb" for DNA isolation on columns (JSC "Syntol", Moscow). The kit is based on the technology of nucleic acid isolation on spin columns with a sorbent membrane, containing silicon dioxide. Nucleic acids bind to the membrane due to the presence of a chaotropic salt or ethanol in the buffer. Then the bound nucleic acids are washed from impurities and eluted. All stages were carried out using a centrifuge. The concentrations of the obtained DNA were determined using the optical density of each sample determined by a Thermo Scientific Multiskan GO microplate spectrophotometer.

To amplify the 16S rRNA gene of each sample, we used a pair of primers specific for the entire Bacteria domain: 27F and 1525R. This pair of primers were chosen to trace intraspecific variation across all variable and conserved regions [4]. The program included

a preliminary stage of DNA melting and DNA polymerase activation at a temperature of 95 °C for 5 min. This was followed by 40 amplification cycles, including melting at 95 °C for 1 min, annealing of primers at 55 °C for 1 min, and elongation at 72 °C for 1.5 min. The reaction was carried out in the presence of a 2.5-fold ready-made reaction PCR mixture of ZAO “Syntol” (ZAO “Syntol”, Moscow) with SYBR Green fluorescent dye.

After amplification, the obtained and purified 16S rRNA amplicons were sequenced according to the Sanger method using the primers that were used in the amplification. Thus, we got two differently directed readings. At the same time, areas characterized by a low quality of reading were discarded based on the analysis of the chromatogram.

To test the reliability of the determination of these species from partial sequences of the 16S rRNA gene, we analyzed the intraspecific variability of these sequences. The search for 16S rRNA sequences was carried out in the Silva database, while sequences of at least 1500 bp were selected.

Using the created computer algorithm in the Python programming language, based on the sequences found, the consensus was obtained for both types, including designations characterizing the variability of certain positions [18].

Table 1. Designations of variable positions in DNA sequences, used to construct consensus sequences that carry information about the variability of each nucleotide position. On the right is an example of their application, with the alignment of three nucleotide sequences and obtaining a consensus. Variable positions are in bold.

R	A or G	B	C or G or T	1 GAGACTCAGGU 11 1 GAGACCGTAGU 11 1 GCTCTTTGGAU 11 GMKMYBDRRU
Y	C or T	D	A or G or T	
S	G or C	H	A or C or T	
W	A or T	V	A or C or G	
K	G or T	N	any nucleotide	
M	A or C			

These consensuses were then labeled according to the generally accepted division of 16S rRNA into variable and conserved regions [19, 20]. Conservation levels were calculated and compared for these sites.



Fig. 2. Traditionally allocated variable and conserved regions of 16S rRNA gene.

3 Results

After Sanger sequencing, sequences of two differently directed gene reads were obtained on the sequencing chromatogram. According to our analysis of the chromatogram, areas characterized by a low quality of reading were discarded. The results of sequencing of 16S amplicons showed that the soil microorganisms isolated by us are *Acinetobacter lwoffii* and *Paenibacillus taichungensis*.

All sequences, determined by sequencing by microorganisms, longer than 1500 bp were selected from the Silva database. After consensus was obtained for two species *Acinetobacter lwoffii* and *Paenibacillus taichungensis*, including variable positions, an analysis of the variability of the sequenced variable and conserved regions of 16S rRNA was carried out.

Comparison of the variability and conservatism of the regions of both species was carried out with the variable and conserved regions of the 16S rRNA gene of *Aquifex pyrophilus*, since the analysis of literature sources showed that this species is one of the oldest types of thermophilic bacteria inhabiting the soil [21, 22].

Variable and conserved regions of each sequence were labeled according to the conventional division of 16S rRNA sequences [18, 20, 22].

The levels of variability and conservatism were determined using our own computer algorithm in Python.

Table 2. Levels of the variability of variable and conservative regions of 16S rRNA gene.

Region	Positions (numeration by <i>Aquifex pyrophilus</i>)	Length	<i>Acinetobacter lwoffii</i> , % identity	<i>Paenibacillus taichungensis</i> , % identity
Beginning, C1	0 – 59	59	58,0	87,8
V1	60 – 100	40	57,4	86,0
C2	101 – 179	78	89,9	98,8
V2	180 – 220	40	88,2	84,6
C3	221 – 399	178	96,8	97,8
V3	400 – 500	100	80,8	94,9
C4	501 – 599	98	98,0	100,0
V4	600 – 660	60	83,3	100,0
C5	661 – 819	158	96,1	98,7
V5	820 -860	40	100,0	95,1
C6	861 – 999	138	100,0	98,6
V6	1000 – 1040	40	100,0	100,0
C7	1041 – 1119	78	100,0	98,6
V7	1120 – 1160	40	97,5	100,0
C8	1159 – 1239	80	97,5	96,2
V8	1240 – 1300	60	96,6	95,0
C9	1301 – 1419	118	100,0	99,2
V9	1420 – 1480	60	95,0	98,3
C10, End	1481 – 1564	83	86,0	81,0

Thus, it was determined that the most pronounced intraspecific variability of variable regions 1, 2, and 3 (in *A. lwoffii* also 4). Region 2 is sequenced with primer 27F completely, and 1 and 3 partially.

Variable regions 6, 7, 8, 9 sequenced using primer 1525R are characterized by less intraspecific variability, in particular, region 6 is conserved within both species.

4 Discussion

We have obtained pure cultures of two types of soil bacteria capable to grow on Vinogradsky's medium. Genomic DNA was isolated from these obtained cultures, and amplicons of the 16S rRNA gene were isolated and purified. The cultivated species are identified by the 16S rRNA sequencing as *Acinetobacter lwoffii* and *Paenibacillus taichungensis*. Using the created computer algorithm, consensus sequences of 16S rRNA genes of both types were formed. Using another created algorithm, a comparative analysis of the variability of the traditionally distinguished variable and conservative gene regions was carried out.

The applied method for obtaining 16S rRNA amplicons, based on the use of primers 27F and 1525R, makes it possible to obtain data on the primary structure of regions characterized by different levels of intraspecific variability. The traditionally isolated

variable and conserved regions of 16S rRNA weakly reflect the distribution of the levels of intraspecific variability over the regions of the molecules of the studied species.

In conclusion, the difference in genetic variability between traditionally distinguished conservative and variable regions in 16S rRNA gene may not correspond to their theoretic partition on the species level. It's not surprising if consider the overall low variability of the gene among the strains of one species. Mutations, that sporadically occur at some positions even in regions that contain very important elements of structure participating in tertiary structure formation and intermolecular interactions, occur with a nearly same low frequency out of conserved regions also. These circumstances don't neither devalue the partition of 16S rRNA regarding higher taxonomic levels of variability nor possess any methodological problems for species recognition by 16S rRNA gene sequencing.

References

1. K.M. Scow, O. Somasco, N. Gunapala, S. Lau, R. Venette, H. Ferris, R. Miller, C. Shennan, *California Agriculture*. **48**, 5 (1994)
2. D.H. Wall, R.D. Bardgett, V. Behan-Pelletier, J.E. Herrick, T.H. Jones, K. Ritz et al., Oxford Press University. (2012)
3. D.C. Coleman, W.B. Whitman, *Pedobiologia*. **49**, 6 (2005)
4. D.H. Buckley, T.M. Schmidt, *Environ. Microbiol.* **5**, 6 (2003)
5. R.G. Morgado, S. Loureiro, M.N. González-Alcaraz, *Soil Pollution*. (2018)
6. S. Thijs, M. op de Beeck, B. Beckers, S. Truyens, V. Stevens, J. D. van Hamme, N. Weyens, J. Vangronsveld, *Frontiers in Microbiology*. **8**, 28 (2017)
7. I. Zarraonaindia, S.M. Owens, P. Weisenhorn, K. West, J. Hampton-Marcell, S. Lax, N.A. Bokulich, D.A. Mills, M. Gilles, S. Taghavi, D. van der Lelie, J.A. Gilbert, *mBio*. **6**, 2 (2015)
8. P.E. Brown, *Circular*. (1913)
9. P. Lavelle, A.V. Spain, Springer Science & Business Media. (2001)
10. P.S. Giller, *Biodivers. Conserv.* **5** (1996)
11. T. Decaens, *Global Ecol. Biogeogr.* **19**, 3 (2010)
12. E. Barrios, *Ecol. Econ.* **64**, 2 (2007)
13. S. Thiele-Bruhn, J. Bloem, F.T. De Vries, K. Kalbitz, C. Wagg, *Curr. Opin. Environ. Sustainability*. **4**, 5 (2012)
14. P. Lavelle, C. Lattaud, D. Trigo, I. Barois, *Plant Soil*. **170** (1995)
15. R.D. Bardgett, *Zoology* **105** (2002)
16. M. Schloter, O. Dilly, J.C. Agric. Ecosyst. Environ. **98**, 1-3 (2003)
17. Justine R. Hall, Kendra R. Mitchell, Olan Jackson-Weaver, Ara S. Kooser, Brandi R. Cron, Laura J. Crossey, Cristina D. Takacs-Vesbach, *Applied and Environmental Microbiology*. **74**, 15 (2020)
18. Nomenclature Committee of the International Union of Biochemistry (NC-IUB). *Nucleic Acids Research*. **13**, 9 (1984)
19. C.E. Corless, M. Guiver, R. Borrow, *J Clin Microbiol.* **38** (2000)
20. Peter H. Janssen, *Applied and Environmental Microbiology*. **72**, 3 (2006)
21. G. Deckert, P. Warren, T. Gaasterland, *Nature*. **392** (1998)
22. Z. Shao, W. Mages, R. Schmitt, *Journal of bacteriology*, **176** (1994)