original scientific paper
Ratar. Povrt. 2025, 62(1): 1-11
doi: 10.5937/ratpov61-51638
Manuscript submitted: 16 June 2024
Manuscript accepted: 29 November 2024
Published online: 10 January 2025



# Identification of 16SrXII-A and 16SrXII-H Phytoplasma Subgroup in Russia, Using 16S rDNA NGS Analysis

**Natalya V. Girsova · Denis V. Erokhin · Dmitriy A. Vorobyov · Damir Z. Bogoutdinov · Tatyana B. Kastalyeva***

All-Russian Research Institute of Phytopathology, Bolshie Vyaziomy, Moscow Region, Russia 143050

*Corresponding author: kastalyeva@yandex.ru

**Summary:** The RFLP analysis of the 16S rDNA revealed two distinct phytoplasma groups, 16SrXII-A and 16SrXII-H, suggesting the presence of diverse phytoplasma strains within the analyzed isolates from the Russian State Collection of Pathogenic Microorganisms. Partial 16S rDNA sequences of four phytoplasma, '*Candidatus* Phytoplasma', isolates belonging to the 16SrXII group from the three plant species: common hop (*Humulus lupulus* L.), Indian datura (*Datura metel* L.) and grapevine (*Vitis vinifera* L.) are presented. DNA of gene coding 16S rRNA was amplified, using nested PCR with primers P1/16Sr-SR and R16F2n/R16R2.Fragment of 16S rDNA of the 1.2 Kb was eluted from the corresponding band after agarose gel electrophoresis, purified and used to generate libraries. Sequencing of DNA libraries was performed on a MiSeq sequencer, while assembly of sequences was carried out using the QIAGEN CIC Genomics Workbench software. To compare 16S rDNA sequences with reference strain, the MEGA 5.2 program was used. Two phytoplasma isolates from grapevine (NCBI acc. nos. OQ120463 and OQ120464) and one from datura (OQ130742) have shown 100% and 99.9% similarity to the reference strain and to previously deposited phytoplasma isolates from Russia infecting potato (*Solanum tuberosum* L.) and alfalfa (*Medicago sativa* L.), as well as 99.9-99.7% similarity to '*Ca.* P. solani' isolates from grapevine from Spain, Italy, Georgia, and from tobacco in Serbia. Phytoplasma isolate sequenced from common hop belonged to the 16SrXII-H subgroup, closely related to '*Ca.* P. convolvuli'. This type of phytoplasma was discovered and sequenced for the first time on the territory of the Russian Federation. Its similarity to the species '*Ca.* P. solani' did not exceed 97.2%, and it reached 99.8% relatively to the reference '*Ca.* P. convolvuli' isolate (JN833705) from field bindweed (*Convolvulus arvensis* L.).

**Key words**: 16S rDNA, stolbur phytoplasma, NGS sequencing, '*Candidatus* Phytoplasma solani', '*Candidatus* Phytoplasma convolvuli', *Humulus lupulus*, *Datura metel, Vitis vinifera*

## Introduction

Phytoplasmas, '*Candidatus* phytoplasma', are obligate intracellular microorganisms, plant pathogens that use a wide range of plant species as hosts, which they infect via phloem-feeding insects from the order Hemiptera. Phytoplasmas multiply in the cells of the vascular system of

higher plants or in the sieve cells of gymnosperms, and the development of vegetative and generative organs can be affected. Diseases caused by phytoplasmas have been known and studied in the former USSR since the early 1930s (as reviewed in Bogoutdinov et al., 2019) while infected agent was identified as a new pathogen more later (Doi et al., 1967). For many years, work with this phytopathogenic bacteria was hampered by the lack of a simple and reliable method for determining their presence in plants. However, attempts were made to set up a classification schemes of phytoplasmas, firstly on the basis of the biological properties of their hosts (the specificity of plants and insects and the symptoms of affected plants) (Chiykowski, & Sinha, 1989; Lee et al., 1992), then on the basis of molecular biological methods (DNA-DNA homology and serological data using specific molecular probes and monoclonal antibodies) (Bertaccini et al., 1990; Lee et al., 1991; Chen et al., 1992; Davies & Clark, 1992). The PCR method using universal and specific primers has become a turning point in the field of studying these pathogens and the diseases they cause (Deng & Hiruki, 1991; Davis et al., 1992; Lee et al., 1993). Phylogenetic analysis based on the 16S rRNA gene and ribosomal protein gene sequences showed that phytoplasmas form a large isolated monophyletic group within the class Mollicutes which opened up new opportunities for developing their classification (Lee et al., 1993; Schneider et al., 1993; Schneider et al., 1995, Lee, 1998). In the late 1990s, a comprehensive scheme for classifying phytoplasmas was developed based on the similarity of their patterns after electrophoresis in PAAG, when 16S rDNA fragments were amplified by PCR and separately subjected to restriction with 17 endonucleases using – a method of restriction fragment length polymorphism analysis (RFLP) reflecting the phylogenetic relationships of phytoplasma, and made possible identification of individual groups of phytoplasmas that corresponded to clusters previously identified on the basis of DNA-DNA homology (Lee et al., 1993). Subgroups within each group were determined by the presence of different 16S rDNA restriction sites detected by RFLP analysis. Criteria for assigning phytoplasmas to subgroups have evolved from considering a difference in a single restriction site as sufficient for classification (Lee, 1998) to now relying on similarity coefficients of RFLP patterns (Wei et al., 2007; Wei et al., 2008; Zhao et al., 2009).

Every year the number of identified groups, subgroups and species of phytoplasmas increased, so by 2022, there were 37 groups, more than 150 subgroups and 48 species of '*Ca.* Phytoplasma' (Wei & Zhao, 2022). As the availability of decoding DNA sequences increased, it became possible to perform RFLP analysis *in silico*, i.e. virtual cutting of DNA sequences with appropriate restriction enzymes. *In silico* RFLP analysis programs were developed and became publicly available, which began to replace RFLP analysis in real gels (Zhao et al., 2009).

Earlier, we showed that phytoplasmas belonging to the 16SrXII-A group, related to the species '*Ca.* Phytoplasma solani' were the cause of diseases in 51 plant species belonging to 21 families in 18 administrative units of eight Economic regions of the Russian Federation (Girsova et al., 2016; 2017, Kastalyeva et al., 2016, 2018, 2022). The largest number of species infected with this phytoplasma was found in economically important plants from the solanaceous (9), rose (7) and legume families (6). Among the plants infected with phytoplasma of the stolbur group 24 species of cultivated and ornamental plants, 13 species of wild herbaceous plants, including weeds, and 14 wild tree and shrub species were found (Bogoutdinov et al., 2019, 2020).

Considering that phytoplasmas are dependent by their life cycle on their hosts and that they cannot be cultured in cell-free media, a collection of these pathogenic microorganisms cannot exist in the form of a microbiological culture, rather DNA isolated from plants and insects infected with phytoplasmas is being stored at temperatures from -20° and below in specific collections. Over the years of work with phytoplasmas, the State Collection of Phytopathogenic Microorganisms (SCPM) created at the All-Russian Research Institute of Phytopathology has accumulated a significant number of DNA samples isolated from plants infected with phytoplasmas. Many phytoplasma strains isolated from solanaceous and leguminous plants were sequenced in 2007–2013, as a part of the ISTC (International Science and Technology Center) project (Girsova et al., 2016; Girsova et al., 2017). Of these, 12 were isolated and amplified from infected potatoes The

hosts of thirty-six were plants of the Fabacea family, and two individuals were insect species *Aphrophora alni* Fallen, 1805. In subsequent years, DNA was isolated not only from herbaceous plants, represented by food and feed crops (potatoes, clover, tomato), but also from cultivated and wild woody plants (oriental hornbeam, cultivated grapes, domestic apple), as well as ornamental plants (prickly pear, Indian datura) and weeds infected with phytoplasmas. Some of them were stored for more than 10 years. The 16S rRNA coding sequence was characterized by RFLP analysis but was not sequenced. In 2022, it became possible to carry out sequences of 12 samples, using a 2$^{nd}$ generation DNA MiSeq, Illumina. Ten of these were deponent in GenBank database.

The purpose of the work was to describe sequences of four phytoplasmas belonging to the 16SrXII group from three plant species and compare them with the 16S rDNA sequences of Russian phytoplasma isolates and phytoplasma isolates and clones from other countries previously deposited in the International NCBI GenBank database.

**Materials and Methods**

*Materials*. The work was carried out at the All-Russian Research Institute of Phytopathology in 2022. Four samples of three different plant species infected with phytoplasma were taken as plant material. DNA were isolated from them in different years and stored at −20°C for different time. So, isolate DNA Rus217 of *Humulus lupulus* (common hop) was kept from 2008, isolate DNA Rus-1999F of *Datura metel* L. (Indian datura) from 2017 and 2 isolates DNA Chardonnay-2175 and Chardonnay-2180 of *Vitis vinifera* L. (cultivated grape, variety Chardonnay) from 2018. Datura and grape leaves were collected in the North-Caucasian Economic Region of Russia, while hops were collected in the Volga Economic Region (Samara province).

Preparation of samples for sequencing included the following steps:

1. Isolation of total DNA from infected plants.
2. Nested PCR, consisting of two stages, allowing to produce the target DNA fragment in the required quantity (expected length of the target fragment was 1.2 Kb).
3. Purification of the obtained fragment using gel electrophoresis.
4. Evaluation of the quantity and quality of the obtained product.
5. Creation of a genomic library for each sample.
6. Evaluation of the obtained libraries.
7. Sequencing of libraries – obtaining nucleotide sequences.
8. Analysis of the obtained sequences.

*DNA isolation*. Total DNA was isolated from veins cut from fresh leaves, finely chopped and frozen at −18°C (hops) or from leaves dried flat under a slight press (grapes and datura) at room temperature (about 25–27°C). To isolate DNA, DNasy Plant mini kits from QIAGEN was used in accordance with the manufacturer's instructions or isolated according to the method described in Duduk et al. (2013). DNA of all isolates was subjected to RFLP analysis with several endonucleases to determine group belonging of the phytoplasma just after isolation.

*Amplification and purification of DNA samples before preparing DNA libraries*. A partial sequence (1.2 Kb) of the 16S rDNA gene was amplified using the nested PCR method, first with primers P1/16Sr-SR and then with primers R16F2n/R16R2 (Lee et al., 1998; Duduk et al., 2013). The quality and quantity of PCR products after amplification were checked by electrophoresis in a 1% agarose gel. The resulting PCR products were purified from non-target amplification products and other impurities, using the MinElute Gel Extraction Kit from QIAGEN in accordance with the manufacturer's instructions. The quality of the purified samples was checked by electrophoresis in 1% agarose gel, and the quantity was checked on a Qubit 4 fluorimeter (Thermo Fisher Scientific, Waltham, MA, USA).

*Creation of DNA libraries*. DNA libraries were obtained using the QIAseq FX Library Kit (QIAGEN, Germany), followed by analysis of the quality of the resulting libraries on a Bioanalyzer 2100 (Agilent, USA) using the High Sensitivity DNA Kit.

*Sequencing of DNA libraries* was performed on a MiSeq (Illumina, USA) on a MiSeq v2 Reagent Kit 300 Cycles PE flow cell according to the manufacturer's protocol.

*Assembling 16S rDNA sequences.* The QIAGEN CIC Genomics Workbench 22 V.22.0.1 program was used to assemble 16S rDNA sequences, and the MEGA 5.2 program was used to compare sequences.

**Results**

According to RFLP analysis, all four phytoplasma DNA samples belonged to the 16SrXII group. RFLP analysis with three endonucleases (*Alu*I, *Mse*I and *Taq*I) showed that the DNA of grape and datura phytoplasma belonged to the 16SrXII-A subgroup. The RFLP pattern obtained as a result of restriction with endonuclease *Taq*I of hop DNA infected by phytoplasma indicated that the latter belonged to the same group as the phytoplasmas from grapes and datura, but the pattern of restriction with endonuclease *Mse*I was clearly different (compare track No. 1 on Fig. 1C with that No. 3). After restriction of the DNA amplicon by endonuclease *Alu*I, the pattern type also had a difference, although subtle (Fig. 1B). Both differences became obvious after sequencing, which shown in bold in Table 1.

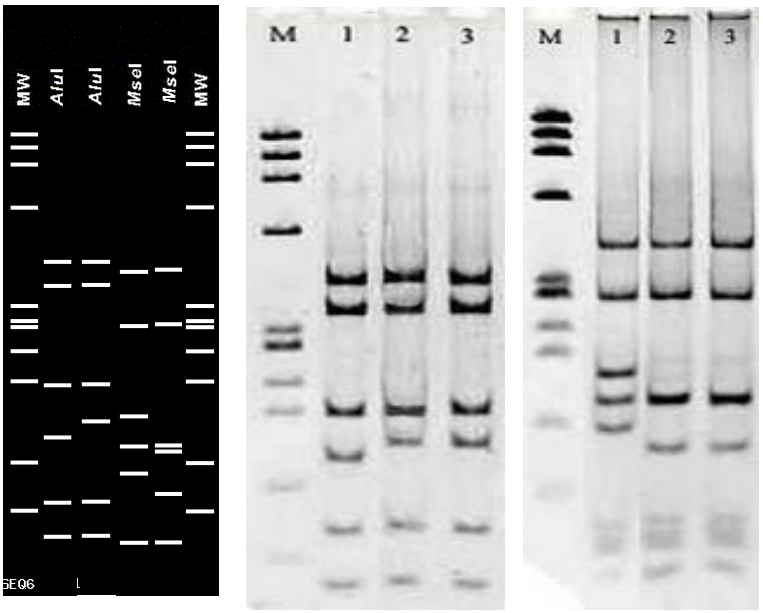


Figure. 1. **A** – virtual RFLP patterns, **B** and **C** after electrophoresis in 5% PAGE. Designations: **MW** and **M** – molecular weight markers ØX174 DNA/BsuRI (HaeIII) (Fermentas, Lithuania), fragment size (bp) (from top to bottom): 1) 1353, 2) 1078, 3) 872, 4) 603, 5) 310, 6) 281, 7) 271, 8) 234, 9) 194, 10) 118, 11) 72 (fragments 281 and 271 form one band on **B** and **C** patterns).
**A** (from left to right): M – marker DNA, "*Alu*I" first and "*Alu*I" second – restriction with enzyme *Alu*I of amplicon of isolate Chardonnay-2175 (acc. no. OQ120463) subgroup 16SrXII-A, and isolate Rus217 (acc. no. OQ103122) subgroup 16SrXII-H, respectively. "*Mse*I" – the same, using enzyme *Mse*I.
**B**: after *Alu*I restriction: 1 – DNA amplicon of isolate Rus-907c13 potato phytoplasma (16SrXII-A), 2 – DNA amplicon of isolate Rus216 field bindweed (*Convolvulus arvensis)* phytoplasma (16SrXII-H), 3 – DNA amplicon of isolate Russ217 (*Humulus lupulus)* phytoplasma.
C: the same as **B** after restriction with *Mse*I enzyme.

Table 1. Length of fragments (number of nucleotides) of phytoplasma subgroup 16SrXII-A and 16SrXII-H after restriction with endonucleases *Alu*I and *Mse*I

| *Alu*I | | Position of the band relative to the marker | *Mse*I | | Position of the band relative to the marker |
|---|---|---|---|---|---|
| 16SrXII-A | 16SrXII-H | | 16SrXII-A | 16SrXII-H | |
| 410 | 410 | between 4 and 5 | 386 | 386 | between 4 and 5 |
| 351 | 351 | between 4 and 5 | 273 | 273 | nearby 7 |
| 188 | 188 | nearby 9 | **156** | | between 8 and 9 |
| **137** | **150** | between 9 and 10 | 131, 130 | 131, 130, **126** | between 9 and 10 nearby 10 |
| 84 | 84 | between 10 and 1 | **109** | **91** | between 10 and 11 |
| 55 | 55 | below 11 | 45, 41 | 45, 41 | below 11 |

When assembling sequences, the Russian isolate Rus-PPT92 (Acc. No. EU344884) from the NCBI GenBank database was used as a reference for all four isolates of the 16SrXII group (Table 2, accession number 6). The sequences of the grape isolates (Acc. No. OQ120463 and OQ120464) began at 140 and 141 nucleotides, respectively, and for the Datura isolate (Acc. No. OQ130742) – at 143; the sequence read of these three isolates ended at nucleotide 1563. The nucleotide positions were determined in accordance with these of STOLL11 isolate (Acc. No. AF248959), which usually proposed to use as a reference (standard) isolate for the 16SrXII–A subgroup. However, sequence of STOLL11 isolate is not entirely justified to be a reference because there are 4 substitutions in its sequence characteristic only for this isolate (T628A, A719G, C1012T, T1208C), as well as there are 2 nucleotides (38A and 39T), which often are replaced by nucleotides 38G and 39C, respectively in other isolate. Therefore, we used the "consensus sequence" ('cons. ref. 16SrXII–A') as a reference, which was obtained by replacing the indicated six nucleotides in the sequence of strain STOLL11 with the following: 38R, 39Y, 628T, 719A, 1012C and 1208T. The percentage of identity of each of the sequenced isolates to the "consensus 16SrXII-A" was calculated using the formula:

Identity=100−N/L%,

where 'N' is the number of nucleotides different from the reference, 'L' is the number of 16S rDNA nucleotides read. L=1527−(n−1), $1527^{th}$ is the last nucleotide in the sequence encoding the 16S rRNA of the reference strain (AF248959), 'n' is the serial number of the first nucleotide read. If L <1527, then it is equal to the sequence number of the last nucleotide read.

The calculation was carried out within the sequence encoding only 16S rRNA; the nearby misc_RNA region was not affected.

As starting material for sequencing we used amplified DNA fragments obtained after nested PCR (F2nR2 fragment). The presence of a band containing these fragments after agarose gel electrophoresis was visualized with an intercalating dye, excised, and the DNA eluted. The size of the DNA fragments was expected to be approximately 1244 bp. It was assumed that the synthesized DNA fragments after assembly would also be 1244 bp long. The reads of three of the sequences (one from Datura phytoplasma and the other two from grapes) began with the primer for the forward reaction R16F2n, but did not end at the 1387th nucleotide, as expected given the use of the R16R2 primer, but included an additional 176 nucleotides, i.e. the reaction was completed at the place where synthesis began in the first reverse PCR (primer 16Sr-SR), namely, at the 1563rd nucleotide. The read of the phytoplasma DNA amplicon from hops (group 16SrXII-H) began at 81 nucleotides and also ended at 1563. This result can be explained by the fact that the DNA fragments in the band excised from the gel and corresponding to approximately 1.2 Kb could include incomplete fragments formed in the first PCR during the synthesis of DNA with a length of 1.8 Kb, which was then diluted 20 times and used in as a template in nested PCR.

Table 2. Comparison of the sequence of 16S rDNA fragments of Russian isolates belonging to the 16SrXII group, related to *'Ca.* Phytoplasma solani' and *'Ca.* P. convolvuli', with isolates of the same group from other countries

| № | GenBank Acc. no.[1] | Isolate (strain), country | Host-plant | Start and finish of reading [2] | Number of nucleotides read in 16S rRNA[3] | Single mutations or number of differebces[4] | Phytoplasma subgroup | Date (year) | Similarity with 'consensus reference'[5] (%) |
|---|---|---|---|---|---|---|---|---|---|
| Reference isolate *Ca.* Phytoplasma solani[5] | | | | | | | | | |
| 1 | AF248959 | STOL11 | *Catharanthus roseus* | 1…1783 | 1527 | A628T, G719A, T1012C, C1208T | 16SrXII-A | 2001 | 99,6 |
| Russian isolates of *Ca.* Phytoplasma from the State Collection of Phytopathogen Microorganisms belonging to the 16SrXII group | | | | | | | | | |
| 2 | OQ120463 | Chardonnay -2175 | *Vitis vinifera* | 140…1563 | 1388 | C1361G, T1362A | 16SrXII-A | 2022 | 99,9 |
| 3 | OQ120464 | Chardonnay -2180 | *Vitis vinifera* | 141…1563 | 1387 | C1361G, T1362A | 16SrXII-A | 2022 | 99,9 |
| 4 | OQ130742 | Rus-1999F | *Datura metel* | 143…1563 | 1385 | C1361G, T1362A | 16SrXII-A | 2023 | 99,9 |
| 5 | OQ103122 vs consensus reference | Rus217 | *Humulus lupulus* | 81… 1561 | 1447 | 41 | 16SrXII-H (*Ca.* P. convolvuli) | 2022 | 97,2 |
| | OQ103122 vs JN833705 | | | 50…1530 | 1478 | C1330G, T1331A, G1526C | | | 99,8 |
| 6 | EU344884 | Rus-PPT92 | *Solanum tuberosum* | 26...1754 | 1502 | A1002del | 16SrXII-A | 2010 | 99,9 |
| 7 | KP864674 | Rus-907c13 | *Solanum tuberosum* | 26...1503 | 1478 | no | 16SrXII-A | 2016 | 100,00 |
| 8 | KY587524 | Rus-AWB803F | *Medicago sativa* | 26... 1538 | 1502 | A1090del | 16SrXII-A | 2017 | 99,9 |
| *Ca.* phytoplasma solani isolates of other countries from NCBI GenBank | | | | | | | | | |
| 9 | AJ964960 | 2642BN, Spain | *Vitis vinifera* | 6…1684 | 1522 | C237N, G1018A | 16SrXII-A | 2010 | 99,9 |
| 10 | JQ730740 | 284/09, Serbia | tobacco | 39…1706 | 1489 | G221A | 16SrXII-A | 2012 | 99,9 |
| 11 | KF996535 | GE07, Georgia | *Vitis vinifera*, Chardonnay variety | 144...1387 | 1244 | T1362A | 16SrXII-A | 2014 | 99,9 |
| 12 | EU836651 | BN-Op121, Italy | *Vitis vinifera,* cv. Chardonnay | 144...1387 | 1244 | T186C, T1057C, T1362G[6] | 16SrXII-F | 2009 | 99,8 |
| 13 | EU836652 | BN-Op30, Italy | *Vitis vinifera*, cv. Croatina | 144...1387 | 1244 | T416C, T1057C[7], T1362A | 16SrXII-F | 2009 | 99,8 |
| 14 | EU836647 | BN-Fc3, Georgia | *Vitis vinifera,* cv. Chardonnay | 144...1387 | 1244 | T186C, T244C[8], T468C, T1189C | 16SrXII-G | 2009 | 99,7 |
| Other species of 16SrXII group Ca. Phytoplasma No. 6 from NCBI GenBank | | | | | | | | | |
| 15 | L76865 vs consensus reference | *Ca* P. australiense | *Vitis vinifera* | 151…1532 | 1377 | 28 | 16SrXII-B | 1997 | 98,0 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 16 | AJ243045 vs consensus reference | SLY, *Ca.* P. australiense | *Fragaria* × *ananassa* | 38… 1705 | 1490 | 27 | 16SrXII-C | 2003 | 98,2 |
| | AJ243045 vs L76865 | | | | | 5 | | | 99,7 |
| 17 | AB010425 vs consensus reference | *Ca.* P. japonicum | *Hydrangea* | 5…1531 | 1523 | 57 | 16SrXII-D | 2007 | 96,3 |
| 18 | DQ086423 vs consensus reference | Straw Y *Ca.* P. fragariae, Lithuania | *Fragaria* × *ananassa* | 119…1333 | 1215 | 30 | 16SrXII-E | 2006 | 97,5 |

Notes: 1 – Accession number in the international GenBank database. 2 - Nucleotide sequence number corresponding to that in the sequence of the STOLL11 isolate (Acc. no. AF248959). 3 – number of read nucleotides of 16S DNA encoding 16S rRNA. 4 – The entry A628T means that the adenine at position 628 is replaced by thymine. 5 – 'Consensus reference sequence' of 16S rDNA corresponds to the sequence of isolate STOLL11 (Acc. no. AF248959) with amendments: A38R, T39Y, T628A, A719G, C1012T, T1208C. 6 - Substitution T1362G adds 2 *Tha*I restriction enzyme sites; 7 - Substitution T1057C adds 1 *Tha*I restriction enzyme site; 8 - Substitution T244C adds 1 *Alu*I restriction enzyme site.

DNA amplicons of phytoplasma isolates isolated as part of total DNA from grapes and datura, sequenced in 2022, had 99.9% similarity to the "consensus isolate" '*Ca.* P. solani', belonging to the 16SrXII-A subgroup (Table 2). The identity of Russian isolates from potato (2 samples) and alfalfa (1 sample), the sequences of which were previously deposited in the GenBank database (Acc. No. EU344884, KP864674, KY587524), was equally high (99.9%, 100% and 99.9%, respectively). The same 99.9% similarity to consensus DNA was demonstrated by isolates of the 16SrXII-A subgroup isolated and sequenced from grapes in Spain and Georgia (AJ964960 and KF996535, respectively) and from tobacco leaves in Serbia (Acc. No. JQ730740). Isolates belonging to the 16SrXII-F subgroup (Acc. No. EU836651, EU836652) from Italy showed 99.8% identity with the consensus 16S rRNA sequence, and isolate 16SrXII-G from Georgia (Acc. No. EU836647) – 99.7%.

DNA phytoplasma isolated from hop had 41 differences from the consensus 16S rDNA (16SrXII-A). The identity was 97.17%. Accessing the interactive online program iPhyClassifier to compare the sequence under study with all other 16S rDNAs available in the GenBank database allows you to almost instantly obtain an answer about group and subgroup membership (Zhao et al., 2009) (https://plantpathology.ba.ars.usda. gov/cgi-bin/resource/iphyclassifier.cgi). To our great surprise, the hop phytoplasma DNA was identified as belonging to the 16SrXII-H subgroup, a species related to '*Candidatus* Phytoplasma convolvuli' (Fig. 1A, Table 1). The similarity with the reference strain of the subgroup (Acc. No. JN833705) was 99.8% (Table 2).

Other species of 'Candidatus Phytoplasma' from the International NCBI GenBank database, which belong to 16SrXII group, namely: 'Ca. P. australiense', 'Ca. P. japonicum' and 'Ca. P. fragariae' presented four subgroups: 16SrXII-B, 16SrXII-C, 16SrXII-D and 16SrXII-E (Table 2). Phytoplasmas of the 16SrXII-B and 16SrXII-C subgroups belonged to the same species ('Ca. P. australiense') and had 99.7% similarity with each other and 98.0–98.2% with "cons. ref. 16SrXII-A", which meets the modern criteria for classification as a new species. In Ca. P. japonicum (subgroup 16SrXII-D) percent identity with "cons. ref. 16SrXII-A" – 96.3% – was even lower than the criterion established for a new species in 2004 and, apparently, it could claim to belong to a new group (IRPCM, 2004).

## Discussion

The phytoplasma classification system based on RFLP analysis takes into account mutations in the restriction sites of 17 specific endonucleases of a 16S rRNA fragment slightly

larger than 1200 bp. In routine and preparatory (rough) work, it is sufficient to use 4 to 8 of the most frequently used restriction endonucleases (*Alu*I, *Mse*I, *Taq*I, *Hha*I, *Hha*II, *Hae*III, *Rsa*I, *Sau*3AI) to determine the phytoplasma belonging to a particular group and subgroup. Most researchers prefer to perform amplicon rDNA sequencing and to establish the group/subgroup belonging using the automatic iPhyClassifier system (Zhao et al., 2009). A typical example is the sequences presented in Table 2 under sequence numbers 11–14 (KF996535, EU836647, EU836651 and EU836652, respectively). All of them start from the 144th rRNA nucleotide and end at the 1387th (relative to the standard). However, researchers note a number of disadvantages, which often leads to uncertainty when assigning some phytoplasma strains to a subgroup and group (Kirdat et al., 2023). It was concluded about the need to revise the criteria for determining the species '*Candidatus* phytoplasma' (Bertaccini et al., 2022). For this purpose, it is proposed to increase the 16S rRNA gene identity threshold from 97.5% to 98.65%, while increasing the length of the analyzed DNA fragment to 1500 nucleotide pairs. Status of the new species '*Ca.* Phytoplasma' can be assigned if the sequence identity of the 16S rRNA gene is ≥98.65%, and threshold of 95% genome-wide average nucleotide identity is suggested.

Thus, sequencing and analysis of longer rDNA sequences, housekeeping genes, ribosomal, membrane and other genes is becoming relevant. The next-generation sequencing (NGS) based on the simultaneous sequencing of thousands of DNA fragments allows reading and is used to read the complete genomes of phytoplasmas and other bacteria. This technology can be implemented in several ways (Fernandez et al., 2020, Zhang et al., 2023). Thus, the Illumina MiSeq system uses sequencing by synthesis (SBS) technology combined with an innovative workflow.

In our work we did not set such ambitious tasks, as genome sequencing. This was the first test of the possibility of obtaining a result on a 2nd generation sequencer. More than 2 thousand contigs were obtained for each sample, after the assembly of which, the coincidence of the sequence with the reference strains approached 100% (Table 1, Nos. 2-5).

Sequencing of the 16S rRNA gene from various regions of the Russian Federation made it possible to determine their relationship and differences in the sequence of 16S rRNA fragments with those for phytoplasma from other countries. Thanks to sequencing, it became possible to identify 16SrXII-H subgroup of phytoplasma, which was difficult to establish using only RFLP analysis because we had deal with atypical infected plant-host.

Genome sequencing data from phytopathogens, including phytoplasmas, can be decisive for the development of integrated approaches to protecting crops from infection by these phytopathogens. Molecular genetic analyzes confirmed the vector role of bindweed planthopper (*Hyalesthes obsoletus* Sign.) and root planthopper (*Pentastiridius leporinus* L.), respectively, both belonging to the family Cixiidae (Hemiptera: Fulgoromorpha). It should be noted, that the first was known as a vector of this pathogen from 1945 (Sukhov & Vovk, 1946), i.e. long before its nature became known (Bogoutdinov et al., 2019 & 2020). More complete data on the sequence of 16S rDNA obtained by sequencing make it possible to associate diseases of cultivated plants with plants reserving infection, and insect vectors. To control effectively spread of the disease in the agroecosystems, it is necessary to limit the number of precisely those reserving-plants, which the vector is bound to. In the recent past, controversial issues on infectivity of certified vegetatively propagated planting material with phytoplasma arose when purchasing of agricultural crops abroad (for example grapes) or in remote regions of the Russian Federation (potatoes). Knowledge of gene sequences from the areas of origin of planting material and from the region of planting allows us effectively establish the root cause of the disease. Sequencing can also clarify the microhabitats of populations with specific gene sequences in the general distribution area of a certain group of phytoplasma, which can also be used in organizing protection methods and applying advanced foreign experience from countries with similar populations of the pathogen, reserving-plants and vectors.

## Conclusions

Sequencing of 16S rDNA of phytoplasma isolated as a part of total DNA from the grape variety Chardonnay (*Vitis vinifera*) (Acc. No. OQ120463, OQ120464) and *Datura metel*, (Acc. No. OQ130742), collected in the North-Caucasus Economical Region, showed a high degree of identity of their sequences (99.9–100%) with the 16S rDNA sequences of phytoplasma from potato (*S. tuberosum*) previously deposited in the GenBank database (Acc. no. EU344884, KP864674) and alfalfa (*M. sativa*) (Acc. No. KY587524), as well as with 16S rDNA from grapes from Italy (EU836651, EU836652), Georgia (EU836647, KF996535), Spain (AJ964960) and tobacco from Serbia (JQ730740). All these strains were related to phytoplasma species '*Candidatus* Phytoplasma solani', belonging to 16SrXII-A subgroup. The 16S rDNA sequence of the phytoplasma found in common hop (*H. lupulus*) from the Samara region (Acc. OQ103122) revealed 99.8% similarity with the similar sequence of the reference isolate of bindweed phytoplasma (*C. arvensis*) (Acc. JN833705) from the GenBank database. Thus, for the first time in the Russian Federation, the 16S rDNA of phytoplasma infecting hop plants was sequenced and shown as belonging to the 16SrXII-H subgroup. A close relationship with the phytoplasma of the species '*Candidatus* Phytoplasma convolvuli' was established.

**Author contributions:** Conceptualization: T.B. Kastalyeva, N.V. Girsova; Data curation: D.A. Vorobyov, D.V. Erokhin; Formal analysis: D.A. Vorobyov, T.B. Kastalyeva; Resources: D.Z. Bogoutdinov, N.V. Girsova; Software: D.A. Vorobyov; Validation; D.V. Erokhin, D.A. Vorobyov, T.B. Kastalyeva; Investigation: N.V. Girsova, D.V. Erokhin, D.A. Vorobyov; Visualization: T.B. Kastalyeva, N.V. Girsova; Writing - original draft: T.B. Kastalyeva, D.Z. Bogoutdinov; Writing - review & editing: N.V. Girsova, D.V. Erokhin, D.A. Vorobyov, D.Z. Bogoutdinov, TB. Kastalyeva

**Competing interests' statement:** No competing interests were disclosed.

**Funding statement:** The study was carried out within the framework of the State assignment # 0598-2019-0002.

**Data availability statement**: Data referring in this study were deposited in NCBI GenBank database under numbers: 0Q103122.1, OQ120463.1, OQ120464.1 and OQ130742.1.

## References

Bertaccini, A., Davis, R. E., & Lee, I.-M., et al. (1990). Detection of chrysanthemum yellows mycoplasmalike organism by dot hybridization and Southern blot analysis. *Plant Dis.*, *74*, 40–43.

Bertaccini, A.; Arocha-Rosete, Y., Contaldo, N., Duduk, B., Fiore, N., Montano, H. G., Kube, M., Kuo, C.-H., Martini, M., Oshima, K., Quaglino, F., Schneider, B., Wei, W. & Zamorano, A. (2022) Revision of the "Candidatus Phytoplasma" Species Description Guidelines. *Int. J. Syst. Evol. Microbiol. 72*(4), 005353. https://doi.org/10.1099/ijsem.0.005353

Bogoutdinov, D. Z., Kastalyeva, T. B., Girsova, N. V., Samsonova, L. N. (2019). Phytoplasma diseases: A review of 50-year history and current advances. *Agricultural biology, 54*(1), 3–18. https://doi.org/10.15389/agrobiology.2019.1.3eng

Bogoutdinov, D. Z., Girsova, N. V., & Kastalyeva, T. B. (2020). Diversity of plant species affected by phytoplasma of the stolbur group in Russia. *Tauride Bulletin of Agrarian Science, 3*/23, 26–42. Russian. https://doi.org/10.33952/2542-0720-2020-3-23-26-42

Bogoutdinov, D., Girsova, N. & Kastalyeva, T. (2021). Danger of phytoplasma diseases for fodder crop cultivation. *IOP Conf. Ser.: Earth Environ. Sci.*, *663*, 012033 https://doi.org/10.1088/1755-1315/663/1/012033

Chen, T.-A., Lei, J. D., & Lin, C. P. (1992). Detection and identification of plant and insect mollicutes. In: *The Mycoplasmas*. R. F. Whitcomb and J. G. Tully, eds. New York. Academic Press, 5, 393–424.

Chiykowski, L. N., & Sinha, R. C. (1989). Differentiation of MLO disease by means of symptomatology and vector transmission. *Zenthl. Bakteriol. Hyg. Suppl., 20*, 280–287.

Davies, D. L., & Clark, M. F. (1992). Production and characterization of polyclonal and monoclonal antibodies against peach yellow leafroll MLO-associated antigens. *Acta Hortic.*, *383*, 275–283. https://doi.org/10.17660/ActaHortic.1992.309.40

Davis, R. E., Princ, J. P., Hammond, R. W., et al. (1992). Polymerase chain reaction detection of Italian periwinkle virescence mycoplasmalike organism (MLO)

and investigation of genetic relatedness with other MLOs. *Phytopathol. Mediterr., 31*, 5–12.

Deng, S., & Hiruki, C. (1991). Amplification of 16s rRNA genes from culturable and non-culturable mollicutes. *J. Microbiol. Methods, 14*(1), 53–61.

Doi, Y., Teranaka, M., Yora, K., Asuyama, H. (1967). Mycoplasma or PLT grouplike microrganisms found in the phloem elements of plants infected with mulberry dwarf, potato witches' broom, aster yellows or pawlownia witches' broom. *Japanese Journal of Phytopathology, 33*(4), 259–266. https://doi.org/10.3186/jjphytopath.33.259

Duduk, B., Paltrinieri, S., Lee, I.-M., & Bertaccini, A. (2013). Nested PCR and RFLP Analysis Based on the 16S rRNA Gene. In: *Phytoplasma: Methods and Protocols Methods in Molecular Biology*. M. Dickinson, & J. Hodgetts, eds. Humana Press, *938*, 159–170. https://doi.org/10.1007/978-1-62703-089-2_14

Fernandez F. D., Zubert, C., Huettel, B., Kube, M., Conci, L. R. (2020). Draft genome sequence of "*Candidatus* Phytoplasma pruni" (X-disease group, subgroup 16SrIII-B) strain ChTDIII from Argentina. *Microbiol. Resour. Announc.*, *9*. Article. e00792-20. https://doi.org/10.1128/MRA.00792-20

Girsova, N. V., Bottner-Parker, K. D., Bogoutdinov, D. Z., Meshkov, Y. I., Mozhaeva, K. A., Kastalyeva, T. B., Lee, I.-M. (2016). Diverse phytoplasmas associated with potato stolbur and other related potato diseases in Russia. *European Journal of Plant Pathology, 145,* 139–153. https://doi.org/10.1007/s10658-015-0824-3

Girsova, N. V., Bottner-Parker, K. D., Bogoutdinov, D. Z., Kastalyeva, T. B. Meshkov, Yu. I., Mozhaeva, K. A., & Lee, Ing-Ming. (2017). Diverse phytoplasmas associated with leguminous crops in Russia. *European Journal of Plant Pathology, 149*(3), 599–610. https://doi.org/10.1007/s10658-017-1209-6

Girsova, N. V, Bogoutdinov, D. Z., Kastalyeva, T. B. & Yengalycheva, I. A. (2021). Phytoplasma diseases of vegetable crops in Russia. *IOP Conf. Ser.: Earth Environ. Sci.,* Article *937*, 022133. https://doi.org/10.1088/1755-1315/937/2/022133

Girsova, N. V., Bogoutdinov, D. Z., Molchanov, A. G., Kastalyeva, T. B. (2024). Detection of Phytoplasma belonging to the 16SrXXI group in Scott pine and mountain pine. *Forestry* **(**Lesovedenie), Issue 2**,** 214–220. Russian. https://doi.org/10.31857/S0024114824020096

IRPCM (2004) 'Candidatus Phytoplasma', a taxon for the wall-less, non-helical prokaryotes that colonise plant phloem and insects. International Journal of Systematic and Evolutionary Microbiology, 54, 1243-1255. http://dx.doi.org/10.1099/ijs.0.02854-01

Kastalyeva, T. B., Bogoutdinov, D. Z., Bottner-Parker, K. D., Girsova N. V., & Lee, I.-M. (2016). Diverse phytoplasmas associated with disease in various crops in Russia— Pathogens and Vectors, *Agricultural Biology* (Sel'skokhozyaistvennaya Biologiya), *51*(3), 367–375. https://doi.org/10.15389/agrobiology.2016.3.367eng

Kastalyeva, T. B., Girsova, N. V., & Bogoutdinov, D. Z. (2018). Taxonomic affiliation of phytoplasma, identified in potato in the Russian Federation. *Achievements of Science and Technology in Agribusiness* (Dostizheniya nauki i tekhniki APK). *32*(3), 40–44. Russian. https://doi.org/10.24411/0235-2451-2018-10308

Kastalyeva, T. B., Bogoutdinov, D. Z., & Girsova N. V. (2022). Danger of phytoplasma diseases for crops determining the food independence of the Country. *Achievements of Science and Technology in Agribusiness* (Dostizheniya nauki i tekhniki APK). *36*(2), 14–19. Russian. https://doi.org/10.53859/02352451_2022_36_2_14

Kirdat, K., Tiwarekar, B., Sathe, S., & Yadav, A. (2023). From sequences to species: Charting the phytoplasma classification and taxonomy in the era of taxogenomics. *Front. Microbiol. 14*, 1123783. https://doi.org/10.3389/fmicb.2023.1123783

Lee, I.-M., Davis, R. E., & Hiruki, C. (1991) Genetic relatedness among clover proliferation mycoplasmalike organisms (MLOs) and other MLOs investigated by nucleic acid hybridization and restriction fragment length polymorphism analyses. *Appl. Environ. Microbiol., 57*(12), 3565–3569. https://doi.org/10.1128/aem.57.12.3565-3569.1991

Lee, I.-M., Davis, R. E., Chen, T.-A., Chiykowski L. N., Fletcher J., Hiruki C., & Schaff D. A. (1992). A genotype-based system for identification and classification of mycoplasmalike organisms (MLOs) in the aster yellows MLO strain cluster. *Phytopathology, 82*, 977–986. https://doi.org/10.1094/Phyto-82-977

Lee, I.-M., Mogen, B. D., & Davis R. E. (1992). 16S rRNA primer pairs designed for general detection and strain cluster identification among plant pathogenic mycoplasmalike organisms (MLOs) by PCR. (Abstr.) *Phytopathology, 82*, 1094.

Lee, I.-M., Hammond, R. W., Davis, R. E., Gundersen, D. E. (1993). Universal amplification and analysis of pathogen 16S rDNA for classification and identification of mycoplasma-like organisms. *Phytopathology, 83*, 834-842. https://doi.org/10.1094/Phyto-83-834

Lee, I.-M., Gundersen-Rindal, D. E., Davis, R. E., & Bartoszyk, I. M. (1998). Revised classification scheme of phytoplasmas based on RFLP analyses of 16S rRNA and ribosomal protein gene sequences. *International Journal of Systematic and Evolutionary Microbiology, 48*(4), 1153–1169. https://doi.org/10.1099/00207713-48-4-1153

Schneider, B., Ahrens, U., Kirkpatrick, B. C., & Seemuller, E. (1993). Classification of plant-pathogenic mycoplasmalike organisms using restriction-site analysis of PCR amplified 16S rDNA. *J. Gen. Microbiol., 139*(3), 519–527. https://doi.org/10.1099/00221287-139-3-519

Schneider, B., Seemuller, E., Smart, C. D., & Kirkpatrick. B. C. (1995). Phylogenetic classification of plant pathogenic mycoplasma-like organisms or phytoplasmas. In: *Moleculur and Diagnostic Procedures in Mycoplasmology*. R. Razin and J. G. Tully, eds. San Diego. Academic Press, *I*, 369–380.

Sukhov K. S., Vovk A. M. (1946). The cicada *Hyalesthes obsoletus* is a vector of Stolbur of Solenaceous. *Doklady AN SSSR*, 1946, *53*(2), 153–156. Russian.

Wei, W., Davis R. E., Lee I.-M., & Zhao Y. (2007). Computer-simulated RFLP analysis of 16S rRNA genes: identification of ten new phytoplasma groups. *International Journal of Systematic and Evolutionary Microbiology, 57*(8), 1855–1867. https://doi.org/10.1099/ijs.0.65000-0

Wei, W., Lee, I.-M., Davis, R. E., Suo, X., & Zhao, Y. (2008). Automated RFLP pattern comparison and similarity coefficient calculation for rapid delineation of new and distinct phytoplasma 16Sr subgroup lineages. *International Journal of Systematic and Evolutionary Microbiology*, *58*(10), 2368–2377. https://doi.org/10.1099/ijs.0.65868-0

Wei, W., Zhao, Y. (2022). Phytoplasma Taxonomy: Nomenclature, Classification, and Identification. *Biology, 11*(8), 1119. https://doi.org/10.3390/biology11081119

Zhang, R.-Y., Wang, X.-Y., Li, J, Shan, H.-L., Li, Y.-H., Huang, Y.-K. & He, X.-H. (2023) Complete genome sequence of "Candidatus Phytoplasma sacchari" obtained using a filter-based DNA enrichment method and Nanopore sequencing. Front. Microbiol. 14. Article. 1252709. https://doi.org/10.3389/fmicb.2023.1252709

Zhao, Ya., Wei, W., Lee, I.-M., Shao, J., Suo, X., & Davis, R. E. (2009). Construction of an interactive online phytoplasma classification tool, iPhyClassifier, and its application in analysis of the peach X-disease phytoplasma group (16SrIII). *International Journal of Systematic and Evolutionary Microbiology, 59*(10), 2582–2593. https://doi.org/10.1099/ijs.0.010249-0

## Identifikacija podgrupe fitoplazme 16SrKSII-A i 16SrKSII-H u Rusiji korišćenjem analize 16S rDNA NGS

**Natalya Girsova · Denis Erokhin · Dmitriy Vorobiov · Damir Bogoutdinov · Tatyana Kastalyeva**

**Sažetak:** RFLP analiza 16S rDNK otkrila je dve različite grupe fitoplazme, 16SrKSII-A i 16SrKSII-H, što ukazuje na prisustvo različitih sojeva fitoplazme unutar analiziranih izolata iz Ruske državne zbirke patogenih mikroorganizama. Delimične sekvence 16S rDNK četiri fitoplazme, *'Candidatus* Phitoplasma', izolata koji pripadaju grupi 16SrKSII iz tri biljne vrste: običnog hmelja (*Humulus lupulus* L.), indijske dature (*Datura metel* L.) i vinove loze (*Vitis vinifera* L.) su predstavljeni u radu. DNK gena koji kodira 16S rRNA je amplifikovana korišćenjem nested PCR-a sa prajmerima P1/16Sr-SR i R16F2n/R16R2. Fragment 16S rDNK od 1,2 Kb je eluiran iz odgovarajuće trake nakon elektroforeze u agaroznom gelu, prečišćen i korišćen za generisanje librarija. Sekvenciranje DNK biblioteka je obavljeno na MiSek sekvenceru, dok je sastavljanje sekvenci obavljeno korišćenjem KIAGEN CIC Genomics Vorkbench softvera. Da bi se uporedile sekvence 16S rDNK sa referentnim sojem, korišćen je program MEGA 5.2. Dva izolata fitoplazme iz vinove loze (NCBI po br. OK120463 i OK120464) i jedan iz dature (OK130742) pokazali su 100% i 99,9%, sličnost sa referentnim sojem i sa prethodno deponovanim izolatima fitoplazme iz Rusije (S infecting potaroso tuberos) i lucerke (*Medicago sativa* L.), kao i 99,9-99,7% sličnosti sa 'Ca. P. solani' izolati iz vinove loze iz Španije, Italije, Gruzije i iz duvana u Srbiji. Izolat fitoplazme sekvencioniran iz običnog hmelja pripadao je podgrupi 16SrKSII-H, bliskoj vezi sa 'Ca. P. convolvuli'. Ova vrsta fitoplazme je prvi put otkrivena i sekvencionirana na teritoriji Ruske Federacije. Sličnost sa vrstom 'Ca. P. solani' nije premašila 97,2%, a dostigla je 99,8% u odnosu na referencu 'Ca. Izolat P. convolvuli' (JN833705) iz poljskog vijuga (*Convolvulus arvensis* L.).

**Ključne reči:** 16S rDNK, stolbur fitoplazma, sekvenciranje sledeće generacije, 'Candidatus *Phytoplasma solani*', 'Candidatus *Phytoplasma convolvuli*', *Humulus lupulus*, *Datura metel*, *Vitis vinifera*