Identification of Endophytic Microbiota of Phytoplasma-Infected Russian Olive Trees “Elaeagnus angustifolia L.” in the Northwest of Iran

Nadia Azizpour 1, Sevil Nematollahi 1,*, Reza Khakvar 2,*, Manizheh Jamshidi 1 and Mohammad Hossein Norouzi-Beirami 3

1 Department of Plant Protection, Tabriz Branch, Islamic Azad University, Tabriz 51579, Iran
2 Department of Plant Protection, Faculty of Agriculture, University of Tabriz, Tabriz 51664, Iran
3 Department of Computer Engineering, Osku Branch, Islamic Azad University, Osku 53518, Iran
* Correspondence: nematollahi2001@yahoo.com (S.N.); khakvar@tabrizu.ac.ir (R.K.)

Abstract: In this study, Russian olive trees exhibiting witches’-broom symptoms were collected from urban green areas in Tabriz, in the northwest of Iran. PCR analysis confirmed that phytoplasma caused the disease and, according to the resulting Sanger sequencing electropherogram, a mixed infection with two or more phytoplasma species within the Russian olive trees was revealed. Next-generation sequencing analyses, using the Illumina method, were performed on total DNA from the infected Russian olive plants to recognize the microbial genomic content and assemble the whole genome of the causative pathogen(s). The use of MetaPhAn and Kraken2 to analyze species composition revealed the very diverse and unique compositions of different Prokaryotic and Eukaryotic species within the infected plants. Several bacteria and fungi were discovered inside the samples, among which Mycoplasmatota was significantly dominating. Interestingly, the results also revealed a high level of endosymbiotic bacteria and Archaea (Methanobacteria) genome contents within the samples. Bowtie2, metaSPAdes, and CD-HIT pipelines were used to perform the initial genome assembly, and the whole genome of the notable phytoplasma species was assembled and submitted to Genbank.

Keywords: ‘Ca. Phytoplasma asteris’; Buchnera aphidicola; Candidatus Zinderia insecticola; MetaPhAn; species composition

1. Introduction

Phytoplasmas are non-culturable and cell wall lacking bacteria belonging to the Mollicutes class. Phytoplasmas are classified as ‘Candidatus Phytoplasma’ species, which have been associated with diseases in several hundred plant species. In Iran, where several phytoplasma diseases have been identified and studied in recent years, ‘Ca. Phytoplasma asteris’ is the most prominent species with a high rate of transmission to different plants, and the third most important phytoplasma after ‘Ca. Phytoplasma phoenicium’ and ‘Ca. Phytoplasma aurantifolia’ [1–9]. ‘Ca. Phytoplasma asteris’ is one of the aster yellows phytoplasma group members, which is the largest and the most diverse phytoplasma group. It is the causative agent of approximately 600 diseases of various herbaceous and perennial plants worldwide [10].

The Russian olive tree (Elaeagnus angustifolia) is indigenous to Iran, but grows as both cultivated and wild type tree in most countries in the Middle East. Apart from the thousands of hectares of wild-type Russian olive trees, in northwest Iran, there are over 6000 hectares of cultivated Russian olive trees, with a production rate of 25 tons per hectare, making the region one of the major Russian olive-producing regions in Iran and the Caucasus. In addition to edible Russian olive fruits, this tree is used in the wood, medicine,
and perfume industries, among others [11]. Similar to other perennial hosts in central and northwest Iran, Russian olive trees have been reported to be infected with ‘Ca. Phytoplasma asteris’ [1,12]. Moreover, infection of Russian olive trees with ‘Ca. Phytoplasma aurantifolia’ has been reported from central Iran [13]. Phytoplasma infection of the Russian olive has previously been reported only in Iran. The only symptoms that infected trees exhibit, whether infected with asteris or with aurantifolia species, are intense witches’-broom symptoms. Currently, no management is practiced for this disease in Iran and, as such, the infected Russian olive trees can serve as a source of infection to disseminate different phytoplasma to other trees in northwest Iran [6]; thus, further investigation of the phytoplasma types and their distributions in infected Russian olive orchards was necessary for disease management. Initial PCR results in the current study indicated that phytoplasma density in leaf midrib tissues is unusually high and appears to be a mixed contamination of several phytoplasmas or other pathogens. The current study was conducted to analyze the whole microbial genome content of Russian olive trees using Next-generation sequencing (NGS) to characterize the endophytic microbiota of infected Russian olive trees.

2. Material and Methods

In the period from September to the end of October 2020, Russian olive trees exhibiting witches’-broom were sampled in urban green spaces of Tabriz, northwest Iran (Figure 1). To obtain an overall and accurate view of the microbiome status of infected Russian olive trees, only one sample per hectare was taken. Total DNA was extracted from approximately 50 g of a mixture of the midrib tissues from ten infected trees (5 g from each tree) using the Murray and Thompson method [14]. The extracted total DNAs were diluted in sterile distilled water before storing at −20 °C until further use.

Figure 1. Witches’-broom symptoms of Phytoplasma-infected Russian Olive trees in Northwest of Iran.

2.1. PCR Assays and Sequencing of 16S rRNA Gene

PCR amplifications were performed using the phytoplasma universal primer pairs P1/P7 [15], which produce a 1784 bp fragment corresponding to 16S rDNA. This was followed by a nested PCR with the phytoplasma universal primer pair R16F2n/R16R2 [16], that produce a 1239 bp fragment from the 16S rDNA. The PCR reaction mixture in 15 μL contained 5 μL of template DNA, 1 μL each of primer pairs (20 pmol) and 9 μL Amplicon
Forests 2022, 13, 1684

Taq DNA Polymerase Master Mix RED (Ampliqon, Odense, Denmark). During the second run, 0.5 μL of the first-round product was used as the template. A 35 thermal-cycle program was carried out in each round of Nested-PCR, including 1 min (2 min for the first cycle) of denaturation at 94 °C, 2 min of annealing at 50 °C, and an extension for 3 min (10 min in final cycle) at 72 °C. Following amplification, each PCR product was electrophoresed through 1% agarose gel in 1 × TAE buffer, stained with ethidium bromide, and visualized in a UV gel documentation apparatus. Purification and sequencing of one of the 1239 bp nested PCR fragments of 16S rRNA gene from Russian olive phytoplasma isolates were performed at Bioneer Corporation (Daejeon, South Korea). The resulting nucleotide sequence was deposited in the NCBI GenBank and were compared with the entries GenBank database using the BLASTn program (online at http://www.ncbi.nlm.nih.gov/BLAST, access on 15 August 2022). Virtual RFLP analysis of 16S rRNA region was performed using web-based iPhyClassifier software on the obtained sequence of the Russian olive tree phytoplasma isolate to compare with the reference strain. Phylogenetic analysis was conducted with the software MEGAX using the Neighbor-joining method and bootstrapping 1000 times to estimate branching stability.

2.2. NGS Analyses, Species Composition Analysis and Genome Alignment

Three total-DNA samples were prepared for NGS analysis from three different infected trees. Each sample contained approximately 80 μL DNA solution, with concentration of 15 μg/μL of total-DNA. Extracted DNA from leaf midribs of infected Russian olive trees were purified and used for library preparation. A paired-end library was generated using the Illumina 1.9 Novaseq 6000 platform in Novogene Co. Ltd., Beijing, China. Finally, approximately 12 Gb raw data for each sample were ordered. Initially, the FastQC tool was used to ensure the readings were of high quality [17]. Based on the FastQC results, readings of lower quality were removed using the Trimmomatic tool. Using CDHit software, highly similar sequences were removed. Then, NGS data were entered into MetaPhlAn2 and Kraken2 software to analyze species composition. MetaPhlAn and Kraken are computational software for profiling the composition of microbial communities (Bacteria, Archaea, and Fungi) from metagenomic sequencing data (i.e., not 16S) with species-level. All reads were aligned to the various reference sequences using Bowtie2 v2.4.2 software [18] to extract the whole genome of the causative pathogen(s). Then, using metaSPAdes v3.9.0.1 software, all mapped reads were used in the assembly [19]. Average nucleic acid identity score (ANI) calculations were performed using the CJ Bioscience’s online Average Nucleotide Identity (ANI) calculator (www.ezbiocloud.net/tools/ani, accessed on 7 September 2022) [20].

3. Results and Discussion
3.1. Russian Olive Phytoplasma Detection and Characterization

About 90% of the 250 Russian olive trees in sampling sites exhibited some degree of symptoms consistent with phytoplasma infection. Phytoplasma detection using nested PCR assays revealed that phytoplasma was present in all trees exhibiting witches’-broom symptoms. The blast results of 16S rRNA sequencing, which was deposited into NCBI under accession number MZ604438.1, showed around 98% similarity between the Russian olive phytoplasma and ‘Ca. Phytoplasma asteris’ reference strains. Constructed neighbor-joining tree also revealed that the current Russian olive tree’s phytoplasma was related to other identical strains of the ‘Ca. Phytoplasma asteris’ group (Figure 2). However, the overall similarity rate was less than the required threshold to prove this species belongs to the ‘Ca. Phytoplasma asteris’ species.
Figure 2. Neighbor-joining phylogenetic tree in Mega-X software based on analyses of partial 16SrDNA sequence of Russian olive trees phytoplasma and some other phytoplasma strains obtained from Genbank; *Acholeplasma oculi* was used as external species; Bootstrapping was performed to support the branches in 1000 replications.

On the other hand, using the iPhyClassifier web tool ([https://plantpathology.ba.ars.usda.gov/cgi-bin/resource/iphyclassifier.cgi; 7 September 2022](https://plantpathology.ba.ars.usda.gov/cgi-bin/resource/iphyclassifier.cgi)), virtual RFLP analyses of 16S rRNA gene revealed that this Phytoplasma had no identical pattern to the all-reference groups (Figure 3), which raised the idea that either we were dealing with a new phytoplasma species or there was some degree of mix-infection which led to some sort of sequencing error.
Figure 3. Virtual RFLP analyses 1239 bp nested PCR amplified fragments of representative Russian olive tree’s phytoplasma using *AluI* restriction enzyme.

A closer examination of the electropherogram files of the sequenced samples also strengthened the theory that the Russian olive trees might be co-infected with two or more species of phytoplasma species (Figure 4), which could only be separated using NGS sequencing.

Figure 4. Sanger sequencing electropherogram of Nested-PCR product of Infected sample of Russian Olive tree in Northwest of Iran.

3.2. Next-Generation Sequencing Results

For each sample, a total of $4.5 \times 10^7$ read pairs were obtained. FastQC analysis revealed that the length of raw reads was 150 bp, inserts size was 350 bp, and the GC content of raw sequences was on average 31% (±1%) with an appropriate Phred-score of >20.

MetaPhlAn2 and Kraken2 software analysis of the composition of microbial communities within samples revealed an extraordinary and unique species composition inside
infected trees (Table 1) (Figure 5). As shown in Table 1, after omitting the host plant’s DNA, the remaining DNA was composed of 58% bacteria, 40% viruses, 1.3% Eukarya, and 0.1% Archaea. Only the species found in all three samples with genome content above 0.01% are listed in Table 1.

Table 1. Species composition inside analyzed samples of Phytoplasma-infected Russian olive trees.

<table>
<thead>
<tr>
<th>Super-Kingdom</th>
<th>Species</th>
<th>Average Percentage</th>
<th>Variance % (±) between 3 Samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bacteria</td>
<td>Candidatus Phytoplasma asteris</td>
<td>34.1566</td>
<td>±3.0%</td>
</tr>
<tr>
<td></td>
<td>Candidatus Phytoplasma unclassified</td>
<td>14.2817</td>
<td>±5.5%</td>
</tr>
<tr>
<td></td>
<td>Buchnera aphidicola</td>
<td>4.24049</td>
<td>±1.0%</td>
</tr>
<tr>
<td></td>
<td>Candidatus Zinderia insecticola</td>
<td>3.27798</td>
<td>±10%</td>
</tr>
<tr>
<td></td>
<td>Salinivibrio costicola</td>
<td>0.89778</td>
<td>±8.0%</td>
</tr>
<tr>
<td></td>
<td>Candidatus Carsonella ruddii</td>
<td>0.80322</td>
<td>±4.4%</td>
</tr>
<tr>
<td></td>
<td>Onion yellows phytoplasma</td>
<td>0.55685</td>
<td>±1.0%</td>
</tr>
<tr>
<td></td>
<td>Candidatus Phytoplasma australiense</td>
<td>0.26511</td>
<td>±2.5%</td>
</tr>
<tr>
<td></td>
<td>Candidatus Regiella insecticola</td>
<td>0.14813</td>
<td>±0.5%</td>
</tr>
<tr>
<td></td>
<td>Brevundimona sp.</td>
<td>0.13313</td>
<td>±2.0%</td>
</tr>
<tr>
<td></td>
<td>Candidatus Sulcia muellieri</td>
<td>0.09308</td>
<td>±1.0%</td>
</tr>
<tr>
<td></td>
<td>Idiomarina loihiensis</td>
<td>0.09286</td>
<td>±1.5%</td>
</tr>
<tr>
<td>Viruses</td>
<td>Vicia cryptic virus</td>
<td>26.07742</td>
<td>±8.0%</td>
</tr>
<tr>
<td></td>
<td>Malvastrum leaf curl Philippines betasatellite</td>
<td>10.20115</td>
<td>±1.2%</td>
</tr>
<tr>
<td></td>
<td>Dasheen mosaic virus</td>
<td>03.06538</td>
<td>±2.0%</td>
</tr>
<tr>
<td></td>
<td>Bombyx mori nucleopolyhedrovirus</td>
<td>0.712181</td>
<td>±1.0%</td>
</tr>
<tr>
<td>Eukaroyta</td>
<td>Eimeria tenella</td>
<td>0.3625</td>
<td>±7.0%</td>
</tr>
<tr>
<td></td>
<td>Eremothecium sp.</td>
<td>0.2817</td>
<td>±1.0%</td>
</tr>
<tr>
<td></td>
<td>Schizosaccharomyces sp.</td>
<td>0.2231</td>
<td>±5.5%</td>
</tr>
<tr>
<td></td>
<td>Plasmodium yoelii</td>
<td>0.0589</td>
<td>±2.4%</td>
</tr>
<tr>
<td></td>
<td>Dictyostelium sp.</td>
<td>0.0511</td>
<td>±2.0%</td>
</tr>
<tr>
<td>Archaea</td>
<td>Methanobrevibacter sp.</td>
<td>0.1012</td>
<td>±10%</td>
</tr>
</tbody>
</table>

The results indicated that there was a mixed infection of various phytoplasma species inside the infected Russian olive trees. Mycoplasmatota DNA, regardless of the host genome, accounted for 50% of the total DNA of the microbiota of the infected Russian olive trees, with Candidatus Phytoplasma asteris identified as the main Phytoplasma species. The whole-genome of the main identified phytoplasma was assembled and deposited into NCBI (under accession number JAHFWK000000000.1). This submitted genome showed the average nucleotide identity (ANI) score of ≥97.2% for previously submitted ‘Candidatus Phytoplasma asteris’ species. Interestingly, the results also revealed a high genome content of insect endosymbiotic bacteria, like Buchnera aphidicola, Candidatus Zinderia insecticola, and Candidatus Regiella insecticola, within the samples. This is the first report of the presence of these entomopathogenic bacteria within Russian olive tissues. Perennial plant tissues are a rich source of insect symbiotic or pathogenic bacteria [21,22]. In this study, the eukaryotic genome was detected in trace amounts (1.3% of the total) in all three samples, primarily of Ascomycota phyla (such as Eremothecium sp. and Schizosaccharomyces sp.), and Apicomplexa phyla (such as Eimeria tenella and Plasmodium yoelii). Additionally, significant amounts of Methanobrevibacter sp. genome from the Euryarchaeota phyla (Archaea super-kingdom) were detected in all three tested samples.
Figure 5. Overview of the average percentage of different microbial groups found inside the Phytoplasma-infected Russian olive trees in the northwest of Iran.

Russian olive trees grow wild in most of the middle east including Iran, and are primarily planted in these areas to combat drought. Recently, Russian olive trees exhibiting witches’-broom symptoms have become widespread in Iran’s northwest; however, infected trees receive little attention due to their low economic value. Such infected trees can serve as a source of disease for other plants. ‘Ca. Phytoplasma asteris’ has recently been reported to be widespread among almond trees in this region [6]. Several herbaceous plants in Iran have also been found infected with ‘Ca. Phytoplasma asteris’ exhibiting a wide variety of symptoms [4,7–9]. However, this phytoplasma infects only a few trees. The only tree diseases associated with ‘Ca. Phytoplasma asteris’ in the northwest of Iran are sweet cherry witches’-broom (5), apple yellowing, apple rosetting [2] and pear yellowing (3) in the central part of Iran. Almond little leaf [6] and Russian olive witches’ broom disease in the northwest of Iran are the only reported tree diseases associated with ‘Ca. Phytoplasma asteris’ [1,12]. Based on these reports, Russian olive trees in Iran are an undoubtedly significant, but underrated, host for ‘Ca. Phytoplasma asteris’ in the northwest. However, the NGS approach revealed novel aspects of this tree in regard to hosting plant pathogens. Using an NGS approach, this study determined the genetic content of two additional Phytoplasma species found inside the Russian olive tree: —Candidatus Phytoplasma australiense and Onion yellows phytoplasma-. While conventional PCR or other methods failed to detect this mixed infection, NGS revealed that this tree hosts various phytoplasmas. Recently, Trivellone and co-workers [23] reported similar results in the capability of NGS in the identification of mix-infections in crop plants.

NGS results indicated that the viral-related genome content consisted of 40% of total DNA and that the primary viral infection inside Russian olive trees was identified as “Vicia cryptic virus” (Table 1). Most of the viral genomes were found to belong to plant viruses, and infection of host plants by these viruses is natural and possible. However, the high amounts of identified viral genomes may be due to the high degree of similarity between the phytoplasma genome (main chromosome plus plasmids) and viral genomes, and a significant proportion of these viral infections belong to the phytoplasma genome. According to previous reports, the proteins encoded by phytoplasma plasmids have domains similar to those found in eukaryotic proteins and viruses, especially Circoviruses and Geminiviruses (plant ssDNA twin viruses). Regarding this finding, the ancestral phytoplasmas plasmid, which is surrounded by the host cell cytoplasm, may have interacted with an ancestral eukaryotic ssDNA virus. Extracellular DNA recombination interactions are critical for evolution because they result in the genetic diversity of phytoplasmas, which enables microorganisms to rapidly adapt to new environmental conditions [24,25].
Apart from the mixed-infection with certain phytoplasma, discovering more than four different types of insect symbiotic or entomopathogenic bacteria within Russian olive tissues was novel. Evidently, the examined plant samples were internally infested with nymphs or larvae of various insects at the time of collection.

4. Conclusions

Characterization of microorganisms’ genomic properties is a time-consuming and expensive process. The whole-genome sequences of an Iranian strain of ‘Ca. Phytoplasma asteris’ (main Phytoplasma species) and some other endophytes were obtained in this study using an inexpensive NGS analysis. Following the initial assembly and gene clustering, a total of 557 genes and 477 proteins were recognized. The main molecular functions of these genes/ proteins could be attributed to 43 transport proteins, 222 proteins with catalytic activity, 90 hypothetical proteins, 32 tRNA, 24 rRNA, 21 pseudogenes, 15 polymerases, 15 ATP-binding proteins, seven DNA and RNA-binding proteins, six elongation factors, six ribosomal proteins, three DNA recombination/repair protein and seven other proteins (the details are accessible on the NCBI Genbank under the accession number JAHFWK0000000000.1). Attempts are being made to compile gene catalogs and extract the general chromosome structure of the ‘Ca. Phytoplasma asteris’ and other identified bacterial species.

Author Contributions: Lab work & Methodology, N.A.; Conceptualization and Supervision, S.N. and M.J.; Software, M.H.N.-B.; Writing, Editing, Validation and Project Administration, R.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: We thank the Plant Bacteriology Lab of the University of Tabriz for providing laboratory equipment for this research.

Conflicts of Interest: The authors of this article confirm that there are no known conflicts of interest associated with this publication and there has been no significant financial support for this work that could have influenced its outcome.

References