DEVELOPMENT OF MOLECULAR GENETIC APPROACHES TO PREDICT THE OCCURANCE OF FOCI OF BACTERIAL DISEASES IN FOREST ECOSYSTEMS

Lyubov Ivashchenko[1,2], Stanislav Panteleev[2], Oleg Baranov[3]

[1] Belarusian State Technological University
[2] Forest Institute of the NAS, Belarus
[3] National Academy of Sciences, Belarus
lyba281997@mail.ru
stasikdesu@mail.ru
betula-belarus@mail.ru

Abstract

In the course of the study, specific primers were developed to identify 20 species of phytopathogenic bacteria from 12 genera associated with the occurrence of infectious diseases in seven forest-forming species of Belarus. The results showed that the developed oligonucleotides are specific for the 16S (mtDNA and cpDNA) and 23S rRNA genes of the studied phytopathogenic bacteria. At a further stage of research, in vitro tests will be carried out to establish the effectiveness of diagnosing bacterial marker loci, both using pure cultures of pathogen isolates and samples of affected tissues of woody plants.

Keywords: primers, bacterial diseases, marker locus

I. Introduction

In the international practice of plant protection on forest woody plants, several dozens bacterioses with varying degrees of damage are described, caused by representatives of the genera *Pseudomonas, Xanthomonas, Enterobacter, Erwinia, Agrobacterium, Brenneria, Xylella, Lelliottia, Rhizobium, Corynebacterium, Bacillus, Clostridium*, etc. At the same time, bacterioses equally affects forest woody plants in natural plantations, forest cultures, on field strips, in urban, park and forest park plantings, etc. Currently, certain bacterial diseases are known in almost all species of forest woody plants [1].

One of the modern highly effective methods for diagnosing bacterial diseases of woody plants is the use of molecular genetic analysis technologies [2]. Despite the obvious advantages of DNA markers over other groups of phytopathological analysis methods (such as early diagnosis of diseases, accuracy of detection and speed of analysis, etc.), work with tree species is characterized by certain specifics, the consideration of which, in most cases, determines the success of identifying pathogens [3]. At the same time, the choice of the optimal analysis approach is necessary for each element of diagnostics: sampling, selection of DNA marker regions and a method for their study.

The general principles of diagnosing pathogens of infectious diseases are reduced to the identification of the genetic material of the pathogen in the tissues of the host plant [2]. At the

same time, most phytodiseases of woody plants are of a complex nature and are associated not with a single pathogen, but with a multispecies complex, including pathogenic, conditionally pathogenic, and saprophytic microflora. At the same time, the obtained data on the species and quantitative composition of microorganism associations during direct assessment under in planta conditions are more accurate than indirect molecular genetic diagnostics through the stage of obtaining pure cultures of pathogen isolates [2]. The analysis of potential sources of infection is characterized by a similar basic principle: soil, water, air, scrapings, dust, plant debris, etc. It should also be noted that the most optimal approach to analysis is direct typing of associations, which subsequently allows them to be divided into separate types, to conduct identification and establishment of the species abundance of each taxon in the microflora.

The most reliable way to study pathogen associations is to use the metagenetic approach, which analyzes materials obtained as a result of the simultaneous diagnosis of marker loci of all microorganisms present in an infected sample [4]. The main difference when using the metagenetic approach is the accounting and quantitative assessment of microorganisms that are not cultivated under artificial conditions, which have a pronounced specialization in living tissues of microorganisms along with cultivated ones.

As shown by numerous molecular phytopathological studies, in most cases, plant tissues, regardless of the degree of infectious load, despite the multispecies nature of the identified associations, are characterized by the largest proportion of one or more dominant alternative types of pathogenic and opportunistic microflora [5]. At the same time, stable associations are noted for dominant infectious agents that correlate with certain types of diseases and the nature of plant damage: single, focal and epiphytotic forms, which indicates the leading role of species communities, rather than individual types of pathogens, in the development of plant diseases.

Over the past decade, for a large number of economically important phytopathogenic bacteria, some fragments of their genome, including barcoding loci based on rRNA genes and enzymes of primary metabolism, have been deciphered [6]. The advantage of barcoding loci is their conservatism within the same species, which makes it possible to determine the taxonomic affiliation of the infection. Currently, the nucleotide sequences of barcoding loci for various species are presented in the databases of the Gene Banks, and the amount of this information is updated daily [7]. However, despite the availability of an electronic reference resource for species identification, the use of only the presented genetic information as the main criterion for taxonomic typing has a number of limitations: for a number of species of forest phytopathogens, information in the database is missing or incomplete; some of the deposited specimens have a disputed specific designation; the presented variability of isolates is individual and does not reflect the entire spectrum of intraspecific polymorphism.

Based on the foregoing, the purpose of this work was to develop a molecular genetic approach for the early diagnosis of bacterial infections of forest tree species, which ensures the prediction of the occurrence of disease foci in plantations.

II. Materials and methods

382 sequences were selected as starting material for bioinformatics studies from the NCBI GenBank international database. related to the 16S and 23S rRNA genes of 20 species of phytopathogenic bacteria from the genera *Brenneria, Erwinia, Gibsiella, Lelliota, Lonsdalea, Pantoea, Pectobacterium, Pseudomonas, Rahnella, Ralstonia, Xanthomonas, Xylella* associated with the formation of infectious diseases in seven forest-forming species of Belarus: pine, spruce, oak, birch, alder, ash, aspen. Deposits of orthologous mtDNA and cpDNA loci of these rocks were additionally analyzed.

Alignment and analysis of nucleotide sequences was performed using the Sequence Viewer

6.3 software package (Qiagen, USA); primer structure was developed using the NCBI Primer Blast online service (https://www.ncbi.nlm.nih.gov/tools/ blast).

The functionality of the designed primers was evaluated using the Primer Blast module in the NCBI GenBank nucleotide sequence database [https://blast.ncbi.nlm.nih.gov/Blast.cgi].

III. Results

The sizes of the 16S rRNA locus varied significantly among the studied bacterial species, from 1473 bp (*Agrobacterium tumefaciens*) before 1553 bp (*Bacilis pumilus*). The value of the homologous 16S rRNA gene in the mitochondrial genome of woody plants also changed significantly and ranged from 1788 (*Pinus sylvestris*) to 1950 (*Quercus robur*) bp. The increase in the size of the mt16S rRNA genes is due to the presence of five polynucleotide insertions, the size of some of them could reach 349 bp. The study of the 16S rRNA loci of the chloroplast genomes of the analyzed plants showed that their length did not depend on the taxonomic affiliation and was equal to 1491 bp. The decrease in the length of the cp16S rRNA gene compared to the bacterial homologue is due to the presence of four oligo(poly)nucleotide deletions of 9, 10, 11, and 22 bp in size in the inner part of the sequence.

In the structure of the 16S rRNA gene of all studied bacterial species, nine (V1–V9) variable regions with conserved regions located between them were identified. It should be noted that the marked polymorphic regions were clearly expressed when comparing bacterial species. At the same time, interspecies differences in cp16S rRNA did not exceed 0.5 per 100 nucleotides within divisions and 2.5 per 100 nucleotides between divisions. For mt16S rRNA, the level of differentiation between divisions could reach 24%, and within divisions, 4%.

The study of the results of the alignment of the nucleotide sequences of the 16S rRNA gene of bacteria and cpDNA of plants showed that in most cases, significant structural differences between the groups were related to the variable regions of the Bacteria locus, which did not allow the development of specific primers for the diagnosis of phytopathogenic microorganisms (Fig. 1).

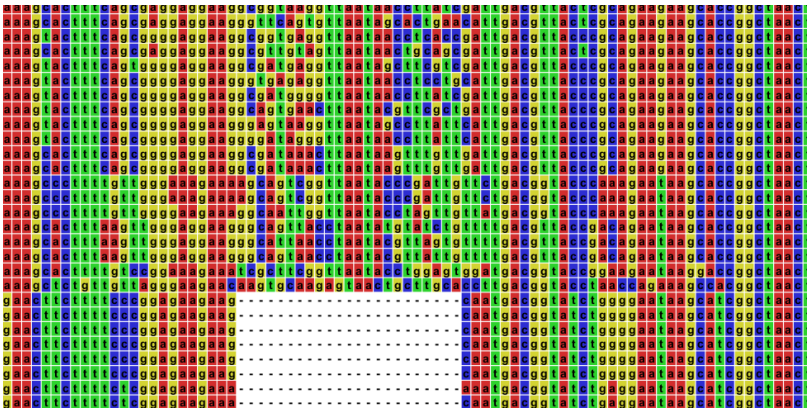

Fig. 1: *The deleted region of the 16S rRNA cpDNA gene common to plants (in bacteria, it is represented by a polymorphic region)*

Structural differences between bacterial 16S rRNA genes and plant mitochondrial genomes have been associated with large insertions in eukaryotic organisms. At the same time, the general nature of the revealed differences between plant organelles indicates their origin from systematically unrelated groups of microorganisms in the course of phylogenesis. The revealed high level of differences in the structures of the mtDNA 16S rRNA gene between gymnosperms and angiosperms may indicate a polyphyletic origin of mitochondria. The above structural features of the 16S rRNA gene in plant organelles did not allow us to identify discrete nucleotide regions that are structurally characteristic of eukaryotic organisms and distinguish them from

bacteria.

Based on the analysis of the structural features of the nucleotide sequences of the 16S rRNA mtDNA and cpDNA genes of plants, a design of oligonucleotide primers was developed that allows selective diagnosis of the genetic material of phytopathogenic bacteria in infected tissues of the studied species of forest woody plants: cpF AGATACCCTGGTAGTCCAC and cpR ATTACTAGCGATTCCRRCTT (the size of the amplified fragment is ≈ 560 bp.), as applied to nucleic acid preparations containing plant cpDNA; mtF TGARATGTTGGGTTAAGTCCCG and mtR TACAAGGCCCGGGAACG (amplified fragment size ≈ 320 bp), as applied to nucleic acid preparations containing plant mtDNA.

The study of the 23S rRNA gene in various bacterial species showed that its size also varied to a large extent, from 2877 bp (*Ralstonia solanacearum*) up to 3037 n.b. (*Brenneria goodwinii*). The value of the homologous 23S rRNA gene in the mitochondrial genome of woody plants also changed significantly and ranged from 3149 (*Pinus sylvestris*) to 3416 (*Quercus robur*) bp. For chloroplast genomes, within the studied angiosperm species, the value of the 23S rRNA gene was constant and equaled 2810 bp. For coniferous plants, the gene size was slightly smaller and amounted to 2802 bp. for *P. sylvestris* and 2806 bp for *Picea abies*.

The main differences in the structure of bacterial and plant (chloroplast) 23S rRNAs were associated with the edge regions of the genes. At the same time, in the case of most bacterial genes (when compared with similar ones in cpDNA), the presence of structural regions at the 5' end of the sequence and the absence at the 3' end of the sequence were observed (Fig. 2).

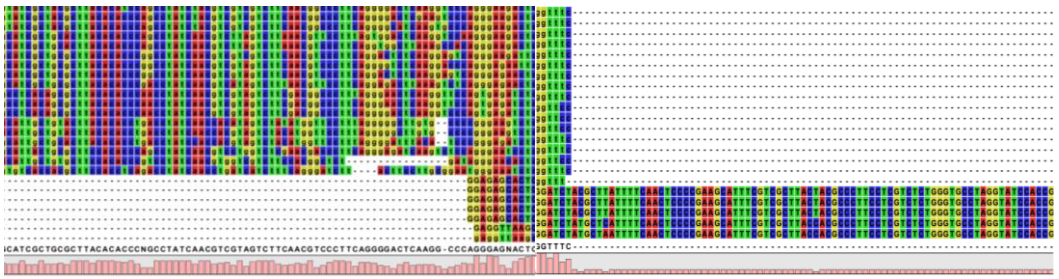

Fig. 2: *Alignment results for the 23S rRNA gene of bacteria and plant cpDNA (5'- region on the left, 3'- region on the right)*

In addition, oligo (up to 5 bp) and polynucleotide (up to 30 bp) deletions in relation to bacterial species were found in the inner part of plant genes.

A comparative analysis of the 23S rRNA gene of bacterial species and plant mtDNA showed that the main differences in the size of these loci are associated with the presence of a significant number of oligo- and polynucleotide deletions/insertions located in the central part of the sequences.

Based on the analysis of the structural features of the nucleotide sequences of the 23S rRNA gene, 20 different variants of oligonucleotide primers were developed for diagnosing the genetic material of bacteria in infected tissues of forest woody plants.

The functionality of the developed primers with respect to the 16S and 23S rRNA genes was tested in silico using the Primer Blast module in the NCBI GenBank database of nucleotide sequences. On fig. 3 shows the results for the mtF and mtR oligonucleotide pair.

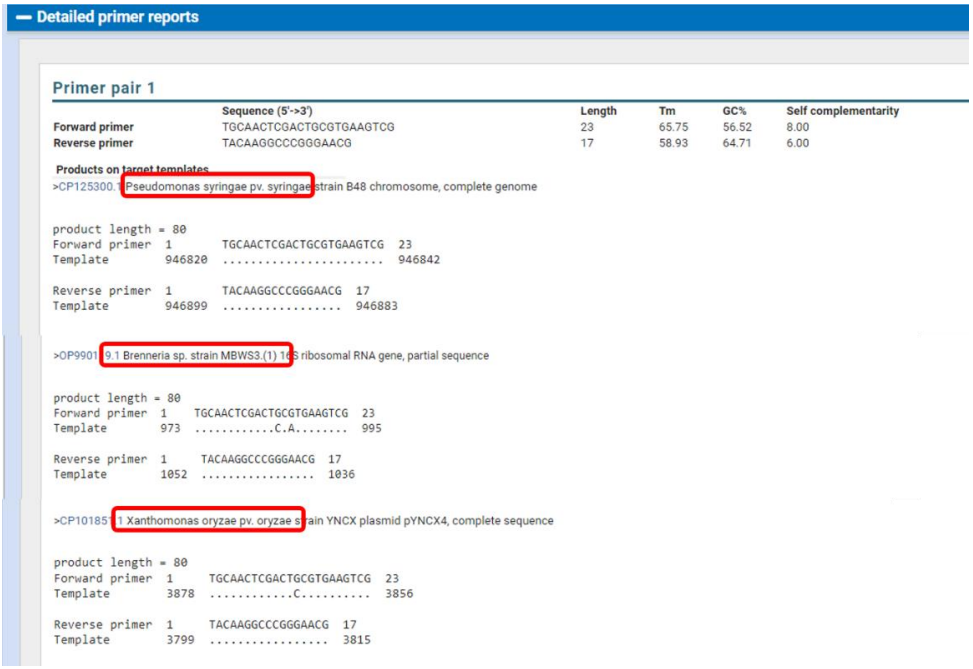

Fig. 3: *Results of testing a pair of primers mtF and mtR in NCBI Primer Blast*

As can be seen from fig. 3, the primers are specific for representatives of the genera mentioned above and do not work for plants, which makes it possible to use them for the identification of bacteria in samples of diseased tissues of woody plants. Similar results were obtained when testing a pair of primers cpF and cpR.

As a result of the testing of primers designed on the basis of the nucleotide structure of the 23S rRNA gene, two pairs of oligonucleotides were selected in silico, allowing amplification of 331 and 589 bp fragments: 23S-BacSpF1 (AAACACGAAAGTGGACGTAT) and 23S-BacSpR1 (CTCCCCACCTATCCTACACA); 23S-BacSpF2 (GAAGTGCGAATGCTGACA) and 23S-BacSpR2 (ATACGTCCACTTTCGTGTTT).

The results obtained showed that the primer pairs are specific for the 16S (mtDNA and cpDNA) and 23S rRNA genes of representatives of the genera *Bacillus, Brenneria, Erwinia, Gibsiella, Lelliota, Lonsdalea, Pantoea, Pectobacterium, Pseudomonas, Rahnella, Ralstonia, Xanthomonas, Xylella*. The specificity for representatives of the genera mentioned above and the inability to amplify homologous loci in plants makes it possible to use them for the identification of bacteria in samples of affected tissues of woody plants. At the same time, parallel sequencing or t-RFLP analysis are possible for carrying out metagenetic analysis.

IV. Conclusion

The article proposes a modern approach for predicting the occurrence of foci of bacterial diseases in forest plantations, which allows for rapid and efficient identification of bacterial species and their communities in the tissues of affected woody plants based on the use of specific primers for the 16S and 23S rRNA genes of phytopathogenic bacteria. At the next stage of the research, *in vitro* tests will be carried out to study the effectiveness of diagnosing marker loci of phytopathogenic bacteria, both using pure cultures of pathogen isolates and samples of affected plant tissues of the main forest-forming species of Belarus, characterized by symptoms of bacterial diseases.

References

[1] Goychuk, A. et al. (2023). Bacterial Diseases of Bioenergy Woody Plants in Ukraine. *Sustainability,* 15(5): 4189. doi: 10.3390/su15054189.

[2] Dyakov, Yu.T. (2001). General and molecular phytopathology. *Moscow: Society of Phytopathologists*, 301 p.

[3] Padutov V.E., Baranov O.Yu., Voropaev E.V. (2007). Methods of molecular genetic analysis. *Minsk: Unipol, 176 p.*

[4] Charles T. (2010). Metagenomics: Theory, Methods and Applications. *Caister Academic Press*, 212 p.

[5] Baranov O.Yu., Panteleev S.V. (2012). Molecular genetic diagnosis of fungal diseases in forest nurseries. *Forestry and hunting,* 6: 21–29.

[6] White, T.J. (1990). Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. *PCR protocols: a guide to methods and applications*, 315–322.