The plastid and mitochondrial genomes of *Vavilovia formosa* (Stev.) Fed. and the phylogeny of related legume genera

N.V. Shatskaya1, V.S. Bogdanova1, O.E. Kosterin1,2, G.V. Vasiliev1, A.K. Kimeklis3, E.E. Andronov3, N.A. Provorov3

1 Institute of Cytology and Genetics, SB RAS, Novosibirsk, Russia
2 Novosibirsk State University, Novosibirsk, Russia
3 All-Russia Research Institute for Agricultural Microbiology, Pushkin, St. Petersburg, Russia

The plastid and mitochondrial genomes of *Vavilovia formosa* (Stev.) Fed. were assembled on the base of the data of high-throughput sequencing of DNA isolated from a sample from North Ossetia, Russia, using Illumina and PacBio platforms. The long PacBio reads were sufficient for reliable assembling organellar genomes while the short Illumina reads obtained from total DNA were unacceptable for this purpose because of substantial contamination by nuclear sequences. The organellar genomes were circular DNA molecules; the genome of mitochondria was represented by two circular chromosomes. A phylogenetic analysis on the basis of plastid genomes available in public databases was performed for some representatives of the tribes Fabeae, Trifolieae and Cicereae. As was expected, the *V. formosa* branch proved to be sister to the *Pisum* branch, and the tribe Fabeae was monophyletic. The position of *Trifolium* L. appeared sensitive to the phylogenetic reconstruction method, either clustering with Fabeae or with the genera *Medicago* L., *Trigonella* L. and *Mellilotus* Mill., but the internodes between successive divergences were short in all cases, suggesting that the radiation of *Trifolium*, other Trifolieae and Fabeae was fast, occurring within a small time interval as compared to further evolution of these lineages. The data on the relatedness of the plastid genomes of *Trifolium* and Fabeae correlate with the similarity of *N* 2-fixing symbionts in these legumes represented by *Rhizobium leguminosarum* biovars *trifoli* and *viciae*, while the symbionts of *Medicago*, *Mellilotus* and *Trigonella* belong to the *Sinorhizobium melloti* and 5. *medicae* species, which are distant from *Rhizobium*.

Key words: *Vavilovia formosa* (Stev.) Fed.; *Vavilovia* A. Fedorov; *Lathyrus* L.; *Vicia* L.; *Pisum* L.; *Lens* L.; *Trifolium* L.; *Medicago* L.; *Trigonella* L.; *Mellilotus* Mill.; *Cicer* L.; Fabeae; Trifolieae; Cicereae; crop wild relatives; pea; plastid genome; phylogeny; paraphyly; monophyly.

For citation: Shatskaya N.V., Bogdanova V.S., Kosterin O.E., Vasiliev G.V., Andronov E.E., Provorov N.A. The plastid and mitochondrial genomes of *Vavilovia formosa* (Stev.) Fed. and the phylogeny of related legume genera. Vavilovskii Zhurnal Genetiki i Selektsi = Vavilov Journal of Genetics and Breeding. 2019;23(8):972-980. DOI 10.18699/VJ19.574

Пластидный и митохондриальный геномы *Vavilovia formosa* (Stev.) Fed. и филогенез родственных родов бобовых

Н.В. Шацкая1, В.С. Богданова1, О.Э. Костерин1,2, Г.В. Васильев1, А.К. Кимеклис3, Е.Е. Андронов3, Н.А. Проворов3

1 Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
2 Новосibirский национальный исследовательский государственный университет, Новосибирск, Россия
3 Всероссийский научно-исследовательский институт сельскохозяйственной микробиологии, Пушкин, Санкт-Петербург, Россия

© Shatskaya N.V., Bogdanova V.S., Kosterin O.E., Vasiliev G.V., Kimeklis A.K., Andronov E.E., Provorov N.A., 2019

Пласти́дный и митохондриа́льный геномы

*Vavilovia formosa* (Stev.) Fed. и филогения родственных родов бобовых

N.V. Shatskaya1, V.S. Bogdanova1, O.E. Kosterin1,2, G.V. Vasiliev1, A.K. Kimeklis3, E.E. Andronov3, N.A. Provorov3

1 Federal Scientific Research Center Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
2 Novosibirsk State University, Novosibirsk, Russia
3 All-Russian Research Institute for Agricultural Microbiology, Pushkin, St. Petersburg, Russia

The data on plastid and mitochondrial genomes of *Vavilovia formosa* (Stev.) Fed. were assembled on the base of the data of high-throughput sequencing of DNA isolated from a sample from North Ossetia, Russia, using Illumina and PacBio platforms. The long PacBio reads were sufficient for reliable assembling organellar genomes while the short Illumina reads obtained from total DNA were unacceptable for this purpose because of substantial contamination by nuclear sequences. The organellar genomes were circular DNA molecules; the genome of mitochondria was represented by two circular chromosomes. A phylogenetic analysis on the basis of plastid genomes available in public databases was performed for some representatives of the tribes Fabeae, Trifolieae and Cicereae. As was expected, the *V. formosa* branch proved to be sister to the *Pisum* branch, and the tribe Fabeae was monophyletic. The position of *Trifolium* L. appeared sensitive to the phylogenetic reconstruction method, either clustering with Fabeae or with the genera *Medicago* L., *Trigonella* L. and *Mellilotus* Mill., but the internodes between successive divergences were short in all cases, suggesting that the radiation of *Trifolium*, other Trifolieae and Fabeae was fast, occurring within a small time interval as compared to further evolution of these lineages. The data on the relatedness of the plastid genomes of *Trifolium* and Fabeae correlate with the similarity of *N* 2-fixing symbionts in these legumes represented by *Rhizobium leguminosarum* biovars *trifoli* and *viciae*, while the symbionts of *Medicago*, *Mellilotus* and *Trigonella* belong to the *Sinorhizobium melloti* and 5. *medicae* species, which are distant from *Rhizobium*.

Key words: *Vavilovia formosa* (Stev.) Fed.; *Vavilovia* A. Fedorov; *Lathyrus* L.; *Vicia* L.; *Pisum* L.; *Lens* L.; *Trifolium* L.; *Medicago* L.; *Trigonella* L.; *Mellilotus* Mill.; *Cicer* L.; Fabeae; Trifolieae; Cicereae; crop wild relatives; pea; plastid genome; phylogeny; paraphyly; monophyly.

For citation: Shatskaya N.V., Bogdanova V.S., Kosterin O.E., Vasiliev G.V., Andronov E.E., Provorov N.A. The plastid and mitochondrial genomes of *Vavilovia formosa* (Stev.) Fed. and the phylogeny of related legume genera. Vavilovskii Zhurnal Genetiki i Selektsi = Vavilov Journal of Genetics and Breeding. 2019;23(8):972-980. DOI 10.18699/VJ19.574
**Introduction**

*Vavilovia formosa* (Stev.) Fed. is a small perennial herbaceous legume confined to highlands of the Caucasus and Anterior Asia (Davis, 1970; Vishnyakova et al., 2016). Although morphologically variable, it is traditionally considered the only member of the monotypic genus *Vavilovia* A. Fedorov. Morphological and molecular data suggest it to be the closest genus to *Pisum* L. (pea, annual plants), to which the important crop *Pisum sativum* L. belongs. Both genera belong to the tribe Fabae Rchb. Recently, H. Schaefer et al. (2012) reconstructed a phylogeny of this tribe and showed that *Pisum* and *Vavilovia* form a clade inside the speciose genus *Lathyrus* L., making it paraphyletic. They propose to subsume *Pisum* and *Vavilovia* to *Lathyrus* but the corresponding nomenclatural change for *Pisum* is not yet broadly accepted and that for *Vavilovia* has not been formally made.

At present, genomic research is extensively conducted in both fundamental (e.g., phylogenetics, phylogeography, evolutionary theory and taxonomy, comparative and functional genomics) and applied (e.g., QTL analysis, association mapping, and marker-assisted selection) aspects. Organellar genomes are among the most popular research objects, since their relatively small size allows their sequencing through the high-throughput approach rather easily. Since mitochondria and plastids have the apparently symbiotic origin (from α-proteobacteria and cyanobacteria, respectively), their research may shed light on the evolution of plants with microorganisms that genetically and functionally interact with these organelles. Specifically, the *N*₂-fixing symbionts of legumes, rhizobia, form organelle-like compartments, symbiosomes, inside the plant cells. Being metabolically integrated, they apparently coevolve with plastids and mitochondria (de la Peña et al., 2018).

Thus far, plastid genomes have been sequenced in many legume genera, including *Vicia, Lathyrus, Pisum, Lens, Cicer, Trifolium, Medicago*, etc., whereas complete mitochondrial genomes are available only for 14 species of Fabales (ncbi.nlm.nih.gov, accessed on March 5, 2018). This is due to the nature of plant mitochondrial DNA, which generally occurs as a set of interconverting subgenomic molecules as a result of homologous recombination between repeated regions (reviewed in Gualberto et al., 2014). For this reason, plant mitochondrial genomes are more difficult to assemble than those of plastids (Smith, Keeling, 2015).

The plastid and mitochondrial genomes of *V. formosa* are not available yet, in spite of explosive interest to this species in the recent decade (Akopian, Gabrielyan, 2008; Mikić et al., 2009, 2013, 2014; Sinjushin et al., 2009; Akopian et al., 2010, 2014; Atlagić et al., 2010; Oskouieyan et al., 2010; Sinjushin, Belyakova, 2010; Zemerski-Škorić et al., 2010; Zorić et al., 2010; Vishnyakova et al., 2013, 2016; Safronova et al., 2014, 2015). This interest was motivated by *V. formosa* being although the most distant but still a pea crop wild relative, which may harbor some genes useful for pea pre-breeding and somehow transferrable to pea.

The phylogenetic tree of most species of the tribe Fabae has been extensively and reliably reconstructed by H. Schaefer et al. (2012), but the positions of the genera evolutionary closest to this tribe are problematic. According to the traditional taxonomy, the tribes most related to Fabae are Ciceraceae Alefeld, with the only genus *Cicer* L., and then Trifolieae (Bonn) Benth., with the genera *Trifolium* s.l., *Medicago* L. s.l., *Trigonella* L., *Mellilotus* Mill., *Cicer* L.; Fabaeae; Trifolieae; Ciceraceae; dicke sorodniche kulturnykh rastenij; горю; пластидный геном; филогения; парафилия; монофилия.

**Ключевые слова:** *Vavilovia formosa* (Stev.) Fed.; *Vavilovia A. Fedorov*; *Lathyrus L.; Vicia L.; Pisum L.; Lens L.; *Trifolium* L.; *Medicago L.; Trigonella L.; Mellilotus Mill.; *Cicer* L.; Fabaeae; Trifolieae; Ciceraceae; дике сородние культурных растений; горю; пластидный геном; филогения; парафилия; монофилия.
In view of this controversy of the phylogenetic position of *Trifolium*, it was interesting to consider a phylogenetic tree reconstructed from complete or nearly complete plastid genomes.

In this work we (i) for the first time report the complete DNA sequence of both plastid and mitochondrial genomes of *Vavilovia formosa* and (ii) use the plastid genome to reconstruct the phylogeny of several legume genera. The obtained data allow us to address the correlation between the plastid-based phylogeny of legumes and their symbiotic affinities presumably reflecting the tight functional and coevolutionary interactions of plastids with temporal nitrogen-fixing organelles, symbiosomes (de la Peña et al., 2018).

**Materials and methods**

**Plant material.** *Vavilovia* seeds were provided by the Gorsky State Agrarian University in Vladikavkaz. They represent a *V. formosa* population in the North Ossetian State Natural Reserve, North Ossetia, the Caucasus, Russia.

**DNA isolation and high throughput sequencing.** DNA from *Vavilovia* plant tissues was isolated with AxyPrep™ Multisource Genomic DNA Miniprep kit according to manufacturer’s recommendations. Whole genome sequencing was done on the Illumina and PacBio platforms in the Macrogen genome sequencing company (Korea).

**Organelar genome assembly.** To assemble the plastid genome of *Vavilovia*, the reads were filtered with the Mirabait utility of the MIRA4.0 package (Chevreux et al., 1999) using the sequence of the *Pisum sativum* chloroplast genome (NC_014057) as a probe. A subset of sequences longer than 10 kb was searched to find a read containing the starting point of the assembly, the *trnH* gene. Then a read overlapping the initial read was selected, the reads were merged, and the dataset was searched for a next read to elongate the assembly. The assembly was elongated in such a manner, until it closed into a circle. It was used as a reference sequence for mapping the *Vavilovia* plastid genome with MIRA4.0 (Chevreux et al., 1999).

Two assemblies were made, one starting from long PacBio reads and the other from short Illumina reads. It is commonly accepted that the best results are gained by combination of these two types of reads (see e.g. Gnerre et al., 2011). The comparison of the two assemblies of *Vavilovia* plastid genome revealed that there appeared various regions with a lot of discrepancies. While the assembly of long reads corresponded well to the reference sequence, the assembly of short reads had a number of mismatches, such as nucleotide substitutions and short indels.

To understand the origin of the discrepancy, some of such regions were checked more carefully. For example, the region corresponding to nucleotide positions 39,400–40,000 of the reference sequence contained 17 mismatches within about 120 bp of alignment, possibly due to the nuclear origin of the reads involved into the assembly. Since short Illumina reads (up to 150 bp) do not permit to investigate their genomic environment, all long reads of our dataset that shared homology with that region were checked whether they belonged to the chloroplast genome indeed. A sample of 939 PacBio reads longer than 10 kb was filtered with the Mirabait utility (Chevreux et al., 1999) using the above-mentioned stretch of 600 bp of the reference sequence. In total, 89 reads were filtered out. Of them, 80 lay entirely in the plastid genome, while the other 9 matched the assembly partially, sharing with it DNA stretches of variable lengths, 300 to 16,000 bp. These 9 reads were used as a query for a BLAST search of the nonredundant database at ncbi.nlm.nih.gov (Altschul et al., 1990). Three of them appeared to correspond entirely to the mitochondrial genome. This is quite natural, since the mitochondrial genome shares about 2.5 kb of homologous DNA stretches with the chloroplast genome, as evidenced from the comparison of the *Vicia faba* L. mitochondrial genome (KC189947) with the *Pisum sativum* plastid genome (NC_014057).

The remaining six reads contained a 300–500 bp region of homology with plastid/mitochondrial DNA, but the rest part had either no homology in the nonredundant database or 1000–1500 bp stretches of homology with genomic clones of some leguminous plants. Most probably, these reads represented nuclear copies of plastid genes. The DNA stretch corresponding to the region 54,400–55,200 of the reference sequence had 15 mismatches per 600 bp of the assembly. A total of 88 reads (longer than 10 kb) that had homology to this region were filtered out. Four of them matched the plastid genome partially, with 8–16 kb corresponding to the plastid genome and 600–2700 bp with no significant similarity in the nonredundant database. Other two randomly taken regions had no obvious discrepancies in the assembly made of short reads with the reference sequence. Seventy-eight reads (longer than 10 kb) were filtered out that passed across the region 60,000–60,500. One of them contained a stretch of 2700 bp that did not match the plastid DNA. All of the 119 long reads passing across the region 80,000–80,500 entirely matched the plastid DNA.

Based on the above observations, a conclusion was inferred that discrepancies in the assemblies made from long vs. short reads arose due to the presence of nuclear copies of plastid DNA of various lengths, from about 300 to 16,000 bp, with the mean of about 7,000 bp. Therefore, the assembly of long PacBio reads was considered more appropriate for plastid genome reconstruction.

The resulting assembly was reasonably consistent. The total amount of mismatches was 0.25 %, and the average coverage depth was 78. These values suggested that the PacBio reads were sufficient for reliable assembling an organellar genome, while the short Illumina reads obtained from total DNA were unacceptable for this purpose because of substantial contamination by nuclear sequences.

The mitochondrial genome was assembled in a similar manner, with the original filtering of reads using the *V. faba* mitochondrial genome (KC189947). The assembly consisted of two ring chromosomes with average coverage depth 84 and 59, and the total number of mismatches was 0.37 %.

The plastid genome of *V. formosa* was assigned the accession number MK604478, and the two chromosomes of its mitochondrial genome got the accession numbers MK48602 and MK48603 in public databases.

**Alignment of plastid genomes for phylogenetic analysis.** We undertook phylogenetic analysis of the plastid genomes available in public databases of some representatives of the tribes Fabaeae, Trifolieae and Cicereae. The plastid genome sequence of *Vavilovia formosa* in general was not collinear
Accession numbers, coverage information, and percentages of similarity to the *V. formosa* plastid genome in the plastid genome reconstructions studied

| Accession number | Species                  | Tribe     | Representation of the reconstruction in the original plastid genome, % | Coverage of the *V. formosa* plastid genome, % | Identity to the *V. formosa* plastid genome, % |
|------------------|--------------------------|-----------|------------------------------------------------------------------------|-----------------------------------------------|-----------------------------------------------|
| HG966674         | *Pisum sativum* voucher WL1238 | Fabeae    | 92.9                                                                   | 93                                            | 95.5                                          |
| MG458702         | *Pisum fulvum* voucher WL2140 | Fabeae    | 93.0                                                                   | 92                                            | 95.5                                          |
| KJ850235         | *Lathyrus clymenum*       | Fabeae    | 90.0                                                                   | 89                                            | 95.7                                          |
| KJ850239         | *Lens culinaris*          | Fabeae    | 84.2                                                                   | 85                                            | 92.7                                          |
| KF042344         | *Vicia faba*              | Fabeae    | 86.8                                                                   | 87                                            | 95.4                                          |
| KJ850242         | *Vicia sativa*            | Fabeae    | 84.7                                                                   | 84                                            | 93.8                                          |
| KJ788292         | *Trifolium strictum*      | Trifolieae| 81.0                                                                   | 83                                            | 93.9                                          |
| EU849487         | *Trifolium subterraneum*  | Trifolieae| 66.2                                                                   | 77                                            | 93.4                                          |
| KC894706         | *Trifolium repens*        | Trifolieae| 76.8                                                                   | 83                                            | 94.0                                          |
| KU321683         | *Medicago sativa*         | Trifolieae| 77.7                                                                   | 81                                            | 94.0                                          |
| JX512024         | *Medicago truncatula*      | Trifolieae| 78.3                                                                   | 78                                            | 93.9                                          |
| NC_041419        | *Melilotus albus*         | Trifolieae| 79.2                                                                   | 82                                            | 94.1                                          |
| MK460508         | *Trigonella foenum-graecum* voucher I.S. Choi MD025 | Trifolieae| 81.9                                                                   | 84                                            | 94.0                                          |
| EU835853         | *Cicer arietinum* voucher ICCV 10 | Ciceraceae | 70.2                                                                   | 79                                            | 92.1                                          |
| DQ317523         | *Glycine max* cultivar PI 437654 | Phaseoleae| 57.2                                                                   | 70                                            | 87.2                                          |

to those of other Fabaceae, differing from them by a large number of structural rearrangements. To make alignment, homologous DNA stretches were found by Blastn software at ncbi.nlm.nih.gov and manually put in the order and orientation corresponding to the *Vavilovia* plastid genome. Each next stretch of homology was sought in the portion of the plastid genome of a species to be aligned that was not yet included in the reconstruction. Then the plastid genome of *V. formosa* and reconstructions of the plastid genomes of the other species were aligned with ClustalW (Larkin et al., 2007) incorporated into the MEGA6 package (Tamura et al., 2013) using the Kimura 2-p parameter, GTR+I+G model was chosen using jModelTest 2.1.10 (Guindon, Gascuel, 2003; Darriba et al., 2012). The GTR+I+G model was chosen using the program FigTree 1.4.3 (http://tree.bio.ed.ac.uk/software/figtree/) by A. Rambaut. The Maximum Likelihood reconstruction of the phylogeny was made with the aid of the MEGA6 package (Tamura et al., 2013) using the Kimura 2-p parameter, GTR+I+G model of mutation rates, and bootstrap test with 100 replications.

**Results and discussion**

**The structure of the *Vavilovia formosa* mitochondrial genome.** The mitochondrial genome of *Vavilovia* was assembled into two non-overlapping circles, 264,766 bp and 88,581 bp, totaling 353,347 bp. This is close to the mitochondrial genome size of *Lotus japonicus* L. (JN872551), 380,861 bp; or *Arabidopsis thaliana* (L.) Heynh. (NC_037304), 367,808 bp. It is larger than in *Medicago truncatula* (KT971339), 271,618 bp, and smaller than in *Vicia faba* (KC189947), 588,000 bp.

Interestingly, it appeared impossible to construct a single master molecule of the *Vavilovia* mitochondrial genome. Instead, two ‘chromosomes’ were obtained (Fig. 1). However, this is quite consistent with the dynamic nature of plant mitochondrial genomes (Gualberto et al., 2014). Another curious fact concerns the *Nad5* gene, which appeared to belong to both ‘chromosomes’, since its exons 1–3 reside in the first, larger ‘chromosome’, whereas exons 4–5 are in the second ‘chromosome’, thus requiring trans-splicing to produce the entire coding sequence. A similar situation has been described in *Silene L.*, where some species possess up to 128 mitochondrial ‘chromosomes’, with exons of many genes present in more than one ‘chromosome’ (Sloan et al., 2012).

Phylogenetic analysis of the mitochondrial genomes of *Vavilovia* and related genera is impossible yet, as of the studied genera (see Table) complete mitochondrial genomes are presently available only for *M. truncatula*, *V. faba* and *G. max* (ncbi.nlm.nih.gov accessed on 22 August 2019).

**The structure of the plastid genome of *Vavilovia formosa***. The content of the plastid genome of *V. formosa* is shown schematically in Fig. 2.

The total length is 122,196 bp, which is similar to the plastid genome size of *Pisum*, 122,180 bp in *P. sativum* (HG966674) and 120,837 bp in *P. fulvum* (MG458702). Expectedly, the gene content appeared very similar to that of *Pisum*. A notable difference is that the *Vavilovia* plastid genome has a tandem triplication of the tRNA gene for methionine. The three copies differ by nucleotide substitutions and a 5 bp long insertion/
The plastid and mitochondrial genomes of *Vavilovia formosa* (Stev.) Fed. and the phylogeny of related legume genera

Fig. 1. Schematic presentation of *V. formosa* mitochondrial genome (assembled as two circles) drawn with OGDRAW (Lohse et al., 2013).
deletion. It is not known whether all the copies are functional. In addition, the gene order in Vavilovia differs from that of Pisum by 10 rearrangements.

**Phylogenetic analysis involving plastid genomes of some related legume genera including Vavilovia.** Figure 3 shows the obtained Bayesian phylogeny reconstruction for representatives of the tribes Fabeae, Trifolieae and Cicerarieae based on the plastid genome reconstructions and using the soybean plastid genome reconstruction as an outgroup. As expected for so long sequences, all nodes of the obtained tree are well supported by high posterior probabilities. The tree topology is also expectable, and it corresponds to the expected for so long sequences, all nodes of the obtained tree are well supported by high posterior probabilities. The tree are well supported by high posterior probabilities. The tree topology is also expectable, and it corresponds to the expected for so long sequences, all nodes of the obtained tree are well supported by high posterior probabilities. The tree are well supported by high posterior probabilities. The tree topology is also expectable, and it corresponds to the expected for so long sequences, all nodes of the obtained tree are well supported by high posterior probabilities. The tree are well supported by high posterior probabilities. The tree topology is also expectable, and it corresponds to the expected for so long sequences, all nodes of the obtained tree are well supported by high posterior probabilities. The tree are well supported by high posterior probabilities. The tree topology is also expectable, and it corresponds to the expected for so long sequences, all nodes of the obtained tree are well supported by high posterior probabilities. The tree are well supported by high posterior probabilities. The tree topology is also expectable, and it corresponds to the expected for so long sequences, all nodes of the obtained tree are well supported by high posterior probabilities. The tree are well supported by high posterior probabilities.

As already mentioned, the positions of Medicago and Trifolium in the phylogenetic reconstructions by M.F. Wojciechowski et al. (2004) (not involving Vavilovia) and H. Schaefer et al. (2012) contradicted the traditional taxonomy as showing Medicago, Trigonella and Melilotus to be a sister branch to that uniting Trifolium and Fabeae, thus making the traditional tribe Trifolieae paraphyletic. The phylogeny reconstructed here by the Bayesian analysis of the complete (or nearly complete) plastid genomes is expected to be more reliable, and it is consistent with the aforementioned results by M.F. Wojciechowski et al. (2004) and H. Schaefer et al. (2012). However, one can notice that although the node uniting Trifolium with Fabeae has a robust support of the posterior probability of 0.86 (see Fig. 3), the branch leading to it after the divergence from Medicago is very short. The same is seen in the trees by H. Schaefer et al. (2012).

At the same time Trifolium and other representatives of the traditional Trifolieae – Medicago, Melilotus, Trigonella, formed a united branch in the Maximum Likelihood tree with the highest possible bootstrap support (100), which is sister to Fabeae (Fig. 4). A similar pattern, with Medicago and Trifolium forming a branch sister to Pisum, was constructed by K. Kreplak et al. (2019), who made a phylogenetic reconstruction based on 28 nuclear sequences using the same Maximum Likelihood method. However, the branch leading to the traditional Trifolieae, including Trifolium, is again very short, both in our tree (see Fig. 4) and in the tree by (Kreplak et al., 2019, Fig. 2, b).

The fact that the positions of Fabeae, Trifolium, and other Trifolieae in the tree depend on the method of phylogeny re-
The plastid and mitochondrial genomes of *Vavilovia formosa* (Stev.) Fed. and the phylogeny of related legume genera

**Conclusion**

Thus, phylogenetic analysis of a sample of the available plastid genomes of representatives of related legume genera, including *Vavilovia*, reported here, confirmed the expected
phylogenetic position of *Vavilovia* itself but challenged the presumed position of *Trifolium* and conjectured a certain coevolution between the plastids and bacterial symbionts of legumes, possibly because of their functional interaction.

**References**

Akopian J.A., Gabrielyan I.G. On high-mountain pea, *Vavilovia formosa* (Stev.) Fed. in Armenia. Crop Wild Relative. 2008:6: 26-27.

Akopian J., Sarukhanyan N., Gabrielyan I., Vanyan A., Mikić A., Smykal P., Kenicer G., Vishnyakova M., Ambrose M. Reports on establishing an ex situ site for “beautiful” vavilovia (*Vavilovia formosa*) in Armenia. Genet. Resour. Crop Evol. 2010:57:1127-1134. DOI 10.1007/s10722-010-9606-0.

Akopian J.A., Sinjushin A.A., Gabrielyan I.G., Shabayog A. On some biomorphological peculiarities of seedlings of *Vavilovia formosa* (Stev.) Fed. (Fabaceae). Legum Perspect. 2014:5:34-35.

Altshul S.F., Gish W., Miller W., Myers E.W., Lipman D.J. Basic local alignment search tool. J. Mol. Biol. 1990:215:403-410. DOI 10.1016/S0022-2836(05)80360-2.

Atlagić J., Mikić A., Sarukhanyan N., Vanyan A., Akopian J., Gabrielyan I., Smykal P., Kenicer G., Vishnyakova M. Ambrose M. Contributions to the characterization of *Vavilovia formosa* (syn. *Pisum formosum*). II. Morphology of androecium and gynoecium and mitosis. Pisin Genet. 2010:41:21-24.

Biondi E.G., Giuntini P.E., Roumiantseva M.L., Andronov E.E., Onchotchouk O.P., Kurchak O.N., Simarov B.V., Dzyubenko N.I., Mengoni A., Bazzicalupi M. Genetic relationship of *Sinorhizobium melloti* and *Sinorhizobium medicarum* strains isolated from Caucasian region. FEBS Microbiol. Lett. 2003;220(2):207-213. DOI 10.1016/S0378-1097(03)00083-3.

Chevreux B., Wetter T., Suahi S. Genome sequence assembly using trace signals and additional sequence information. In: Computer Science and Biology: Proc. German Conference on Bioinformatics (GCB). 1999:99-45-56.

Darriba D., Taboada G.L., Doallo R., Posada D. JModelTest 2: more models, new heuristics and parallel computing. Nat. Methods. 2012;9(8):772. DOI 10.1038/nmeth.2109.

Davis P.H. *Vavilovia* A. Fed. In: Davis P.H. (Ed.) Flora of Turkey and East Aegean Islands, Vol. 3. Royal Botanical Gardens Edinburgh, Edinburgh. 1970;44-45.

de la Peña T.C., Fedorova E., Puyeo J.J., Lucas M.M. The symbiosome: legume and rhizobia co-evolution toward a nitrogen-fixing organanelle? Front. Plant Sci. 2018;8:2229. DOI 10.3389/fpls.2017.02229.

Drummond A.J., Rambaut A. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. 2007;7:1. DOI 10.1186/1471-2148-7-214.

Dudeja S.S., Nidhi. Molecular diversity of rhizobia and nonrhizobia bacteria from nodules of cool season legumes. In: Salar R.K., Gahla P., Ambrose M. Reports on legumes, possibly because of their functional interaction. In: Davis P.H. (Ed.) Flora of Turkey and East Aegean Islands, Vol. 3. Royal Botanical Gardens Edinburgh, Edinburgh. 1970;44-45.

de la Peña T.C., Fedorova E., Puyeo J.J., Lucas M.M. The symbiosome: legume and rhizobia co-evolution toward a nitrogen-fixing organanelle? Front. Plant Sci. 2018;8:2229. DOI 10.3389/fpls.2017.02229.

Drummond A.J., Rambaut A. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. 2007;7:1. DOI 10.1186/1471-2148-7-214.

Dudeja S.S., Nidhi. Molecular diversity of rhizobia and nonrhizobia bacteria from nodules of cool season legumes. In: Salar R.K., Gahla P., Ambrose M. Reports on legumes, possibly because of their functional interaction. In: Davis P.H. (Ed.) Flora of Turkey and East Aegean Islands, Vol. 3. Royal Botanical Gardens Edinburgh, Edinburgh. 1970;44-45.

de la Peña T.C., Fedorova E., Puyeo J.J., Lucas M.M. The symbiosome: legume and rhizobia co-evolution toward a nitrogen-fixing organanelle? Front. Plant Sci. 2018;8:2229. DOI 10.3389/fpls.2017.02229.

Drummond A.J., Rambaut A. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. 2007;7:1. DOI 10.1186/1471-2148-7-214.

Dudeja S.S., Nidhi. Molecular diversity of rhizobia and nonrhizobia bacteria from nodules of cool season legumes. In: Salar R.K., Gahla P., Ambrose M. Reports on legumes, possibly because of their functional interaction. In: Davis P.H. (Ed.) Flora of Turkey and East Aegean Islands, Vol. 3. Royal Botanical Gardens Edinburgh, Edinburgh. 1970;44-45.

de la Peña T.C., Fedorova E., Puyeo J.J., Lucas M.M. The symbiosome: legume and rhizobia co-evolution toward a nitrogen-fixing organanelle? Front. Plant Sci. 2018;8:2229. DOI 10.3389/fpls.2017.02229.

Drummond A.J., Rambaut A. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. 2007;7:1. DOI 10.1186/1471-2148-7-214.

Dudeja S.S., Nidhi. Molecular diversity of rhizobia and nonrhizobia bacteria from nodules of cool season legumes. In: Salar R.K., Gahla P., Ambrose M. Reports on legumes, possibly because of their functional interaction. In: Davis P.H. (Ed.) Flora of Turkey and East Aegean Islands, Vol. 3. Royal Botanical Gardens Edinburgh, Edinburgh. 1970;44-45.

de la Peña T.C., Fedorova E., Puyeo J.J., Lucas M.M. The symbiosome: legume and rhizobia co-evolution toward a nitrogen-fixing organanelle? Front. Plant Sci. 2018;8:2229. DOI 10.3389/fpls.2017.02229.

Drummond A.J., Rambaut A. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. 2007;7:1. DOI 10.1186/1471-2148-7-214.

Dudeja S.S., Nidhi. Molecular diversity of rhizobia and nonrhizobia bacteria from nodules of cool season legumes. In: Salar R.K., Gahla P., Ambrose M. Reports on legumes, possibly because of their functional interaction. In: Davis P.H. (Ed.) Flora of Turkey and East Aegean Islands, Vol. 3. Royal Botanical Gardens Edinburgh, Edinburgh. 1970;44-45.

de la Peña T.C., Fedorova E., Puyeo J.J., Lucas M.M. The symbiosome: legume and rhizobia co-evolution toward a nitrogen-fixing organanelle? Front. Plant Sci. 2018;8:2229. DOI 10.3389/fpls.2017.02229.

Drummond A.J., Rambaut A. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. 2007;7:1. DOI 10.1186/1471-2148-7-214.

Dudeja S.S., Nidhi. Molecular diversity of rhizobia and nonrhizobia bacteria from nodules of cool season legumes. In: Salar R.K., Gahla P., Ambrose M. Reports on legumes, possibly because of their functional interaction. In: Davis P.H. (Ed.) Flora of Turkey and East Aegean Islands, Vol. 3. Royal Botanical Gardens Edinburgh, Edinburgh. 1970;44-45.

de la Peña T.C., Fedorova E., Puyeo J.J., Lucas M.M. The symbiosome: legume and rhizobia co-evolution toward a nitrogen-fixing organanelle? Front. Plant Sci. 2018;8:2229. DOI 10.3389/fpls.2017.02229.
from morphological and molecular data. Pism Genet. 2009;41:15-20.

Sloan D.B., Alverson A.J., Chuckalovcak J.P., Wu M., McCauley D.E., Palmer J.D., Taylor D.R. Rapid evolution of enormous, multichromosomal genomes in flowering plant mitochondria with exceptionally high mutation rates. PLoS Biol. 2012;10(1):e1001241. DOI 10.1371/journal.pbio.1001241.

Smith D.R., Keeling P.J. Mitochondrial and plastid genome architecture: Reoccurring themes, but significant differences at the extremes. Proc. Natl. Acad. Sci. USA. 2015;112:10177-10184. DOI 10.1073/pnas.1422049112.

Tamura K., Stecher G., Peterson D., Kumar S. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol. Biol. Evol. 2013;30:2725-2729. DOI 10.1093/molbev/mst197.

Vishnyakova M., Burlayeva M., Akopian J., Murtazaliev R., Mikić A. Reviewing and updating the detected locations of beautiful vavilovia (Vavilovia formosa) on the Caucasus sensu stricto. Genet. Res. Crop Evol. 2016;63:1085-1102. DOI 10.1007/s10722-016-0440-x.

Vishnyakova M.A., Burlayeva M.O., Seferova I.V., Bagmet L.V., Semenov V.A. Expedition collections of the tribe Vicieae representatives in the Russian Federation and the adjacent territories of the North Caucasus. Bull. Appl. Bot. Genet. Plant Breed. 2013;172:82-86.

Wojciechowski M.F., Lavin M., Sanderson M.J. A phylogeny of legumes (Leguminosae) based on analysis of the plastid matK gene resolves many well-supported subclades within the family. Am. J. Bot. 2004;91(11):1846-1862. DOI 10.3732/ajb.91.11.1846.

Yakovlev G.P. Bobovye Zemnogo Shara = Legumes of the Globe. Leningrad: Nauka Publ., 1991. (in Russian)

Zemerski-Škorić T., Mikić A., Sarukhanyan N., Vanyan A., Akopian J., Gabrielyan I., Smykal P., Kenicer G., Vishnyakova M., Ambrose M. Contributions to the characterization of Vavilovia formosa (syn. Pisum formosum). III. Contents of macro- and microelements. Pism Genet. 2010;41:28-30.

Zorić L., Luković J., Mikić A., Akopian J., Gabrielyan I., Sarukhanyan N., Vanyan A., Smykal P., Kenicer G., Vishnyakova M., Ambrose M. Contributions to the characterization of Vavilovia formosa (syn. Pisum formosum). I. Anatomy of stem, leaf and calyx. Pism Genet. 2010;41:21-24.

ORCID ID
O.E. Kosterin orcid.org/0000-0001-5955-4057

Acknowledgements. The work by N.V. Shatskaya, V.S. Bogdanova, O.E. Kosterin and G.V. Vasiliev on the assembly of the plastid and mitochondrial genomes and phylogenetic analysis was supported by State Budgeted Project 0324-2019-0041 for the Institute of Cytology and Genetics SB RAS, Novosibirsk, and the Russian Foundation for Basic Research, project 19-04-00162. The work by E.A. Andronov, A.K. Kimeklis, and N.A. Provorov on DNA isolation and high-throughput sequencing was supported by the Russian Science Foundation, project 19-16-00081. The assembly of organellar genomes and phylogenetic analysis were performed with the use of the Computational Facility of the Siberian Supercomputer Center SB RAS and the Computational Facility of the Novosibirsk State University.

Conflict of interest. The authors declare no conflict of interest.

Received April 15, 2019. Revised September 20, 2019. Accepted September 20, 2019.