

# ASSEMBLY OF MITOCHONDRIAL GENOME OF SIBERIAN CEDAR (*Pinus sibirica*)

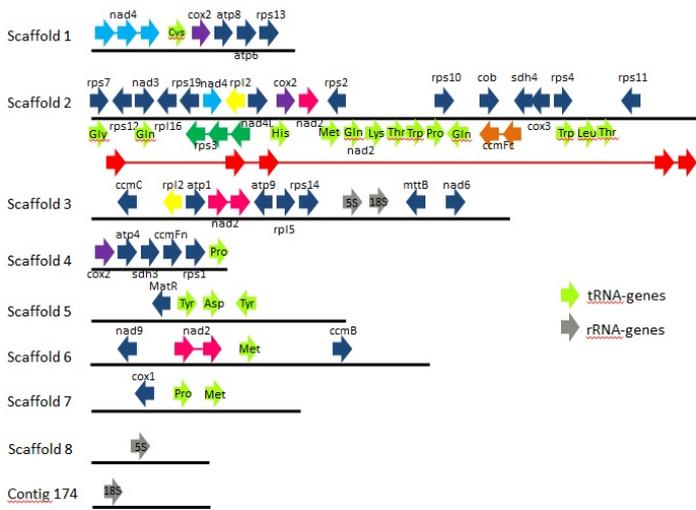


Figure 1: *P. sibirica* mitochondrial genome annotation

At the Siberian Federal University, the mitochondrial sequences of *Pinus sibirica* were obtained. Sequencing was performed on the IlluminaHiSeq 2000 platform. The contigs were assembled using the CLC BIO. 879595 scaffolds with a total length of 283.4 Mb were obtained.

Mitochondrial scaffolds were selected by aligning the genomic assembly with BLAST against a base of 3489 (143.5 Mb) complete and partial sequences of mitochondrial genomes of seed plants. As a result, 97 scaffolds with a total length of 3.1 Mb were selected.

**Results:** Protein-coding genes were searched using BLAST in the Genomic Workbench CLC software by aligning the assembly against well-annotated cycad and ginkgo genomes, tRNA genes were detected using tRNA-scan and Aragorn, rRNA genes were detected using RNAmmer. 40 protein-coding genes were found located in 7 scaffolds, also 18 tRNA genes and 2 rRNA genes were found. The total length of the repeats was 0.2 Mb, which is 6.2% of the total assembly length using RepeatModeler, TEclass.

To compare with, the total assembly length for *Picea glauca* is 5.9 Mb, *Pinus taeda* is 1.2 Mb, *Picea abies* is 4.3 Mb. We believe the difference in the size of the mitochondrial genome results both from the difference in sizes of the non-coding regions, and the number of genes.

**Conclusion:** Sequences of the mitochondrial genome of Siberian pine were identified and annotated. The cedar assembly was within the length of the coniferous mitochondrial genome assemblies, which vary in a wide range: from 1.2 Mb in *Pinus taeda* and 3.9 Mb in *Pinus lambertiana* to 5.9 Mb in *Picea glauca*. We believed that the difference in the size of the mitochondrial genome is mainly due to different sizes of non-coding regions.

<sup>1</sup>Laboratory of Forest Genomics, Genome Research and Education Center, Institute of Fundamental Biology and Biotechnology, Siberian Federal University, 660036 Krasnoyarsk, Russia. <sup>2</sup>Laboratory of Genomic Research and Biotechnology, Federal Research Center “Krasnoyarsk Science Center”, Siberian Branch, Russian Academy of Sciences, 660036 Krasnoyarsk, Russia. <sup>3</sup>Institute of Systematics and Ecology of Animals, Siberian Branch, Russian Academy of Sciences, 630091 Novosibirsk, Russia. <sup>4</sup>Department of High Performance Computing, Institute of Space and Information Technologies, Siberian Federal University, 660074 Krasnoyarsk, Russia. <sup>5</sup>Laboratory of Forest Genetics and Selection, V.N. Sukachev Institute of Forest, Siberian Branch of Russian Academy of Sciences, Krasnoyarsk, 660036, Russia. <sup>6</sup>Laboratory of Plant Genetic Engineering, Siberian Institute of Plant Physiology and Biochemistry, Siberian Branch, Russian Academy of Sciences, 664033 Irkutsk, Russia. <sup>7</sup>Institute of Computational Modeling, Siberian Branch, Russian Academy of Sciences, 660036 Krasnoyarsk, Russia. <sup>8</sup>Department of Plant Physiology, UPSC, Umeå University, S-90187 Umeå, Sweden. <sup>9</sup>Department of Forest Genetics and Forest Tree Breeding, Georg-August University of Göttingen, 37077 Göttingen, Germany.

**Goal:** Most plants have two types of organelles with their own genome: mitochondria and chloroplasts; that former is inherited over the maternal lines and is rather complex, in plants. Currently, only for four gymnosperms have published mitochondrial genomes. We isolated, sequenced and assembled *de novo* mitochondrial genome of *Pinus sibirica*, one of the most widespread gymnosperm species at Siberia.

**Methods:** Isolation of intact mitochondria was carried out at the Laboratory of Plant Genetic Engineering, SIFiBR SB RAS. In the forest genomics laboratory of Siberian Federal University,

Table 1: Repeats

DNA transposons	5
LTRs	8
LINEs	2
SINEs	1
Unclear	6