



The low level of variability is depicted in plastomes of early soybean varieties

Aleksandrovich Valeryia, Siniauskaya Maryna, Liaudanski Aleh, Mishuk Yahor, Shatarnova Tatjana, Danilenko Nina, Davydenko Oleg

Institute of Genetics and Cytology, NAS, Minsk, Belarus
Laboratory of cytoplasmic inheritance

valeria.alexandrovich@gmail.com
cytoplasmic@mail.ru

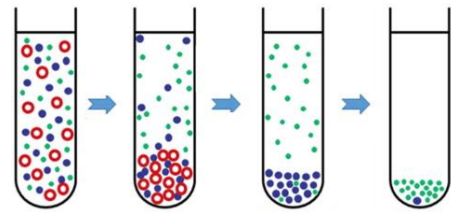
Motivation and aim

The chloroplast genome of land plants is usually conservative in structure, but variable in sequence. Intraspecific diversity of organelle possibly plays an important role in successful breeding and creation of the new varieties in plants.

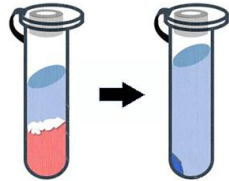
The **aim of our study** was to evaluate the level of chloroplast genomes diversity in specific group of *Glycine max* varieties. The study sample contained **24 early-maturing** *G. max* cultivars from our collection. All studied soybean varieties have different geographic origin.



Methods



1) Isolation of the chloroplast fraction by differential centrifugation from the 7-day-old soybean seedlings*



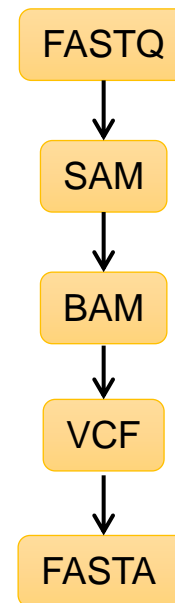
2) Phenol-chloroform DNA extraction



3) DNA sequencing on Illumina MiSeq using Illumina DNA Prep library preparation kit and MiSeq Reagent Kit v3 (600-cycle)



4) The NGS data processing algorithm included:



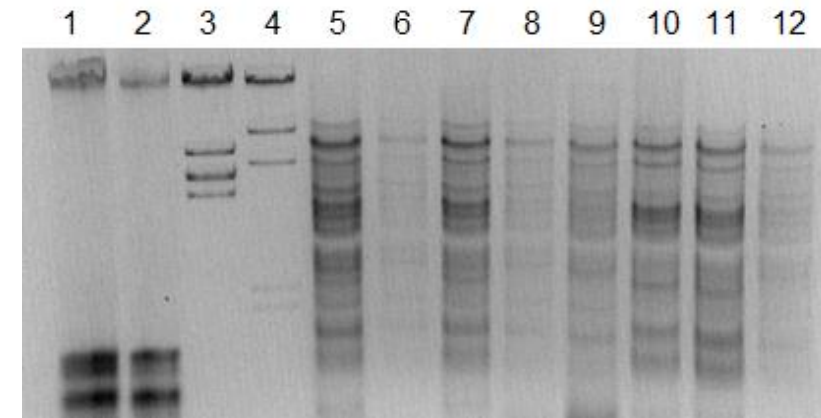
Reads alignment on reference sequence (MW357264.1) (Burrows-Wheeler Alignment Tool)

Converting to .bam format and sorting (Samtools)

Variant calling (HaplotypeCaller) followed by filtering (Bcftools)

VCF files contained all polymorphic loci of chloroplast genomes of each sample. Complete FASTA sequences of the soybean chloroplast genomes were generated based on these VCF files.

Quality of the isolated chloroplast DNA was checked by RFLP-analysis



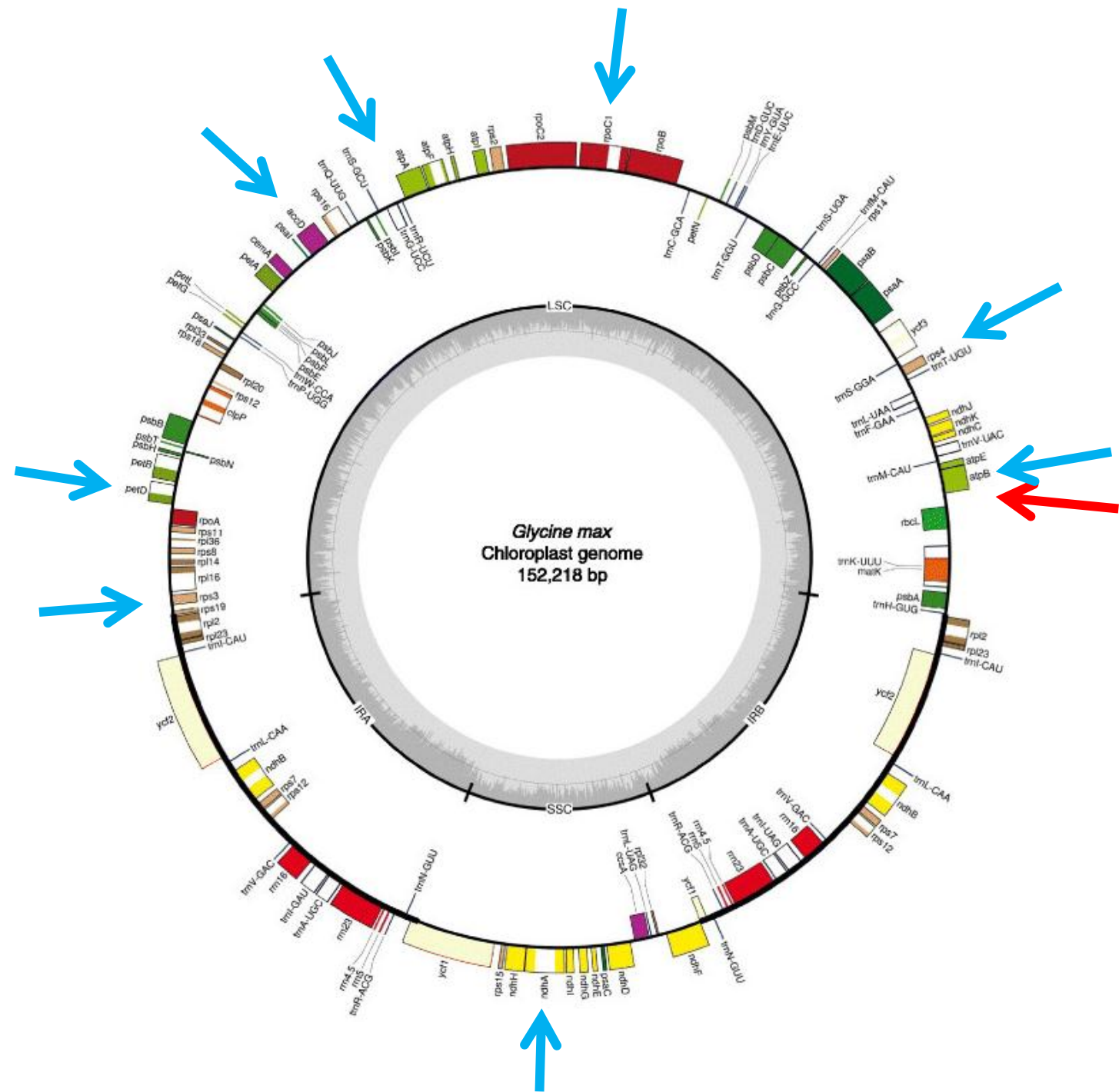
Ethidium bromide stained agarose gel of electrophoretically-separated soybean chloroplast DNA restriction fragments

1, 2 – undigested chloroplast DNA; 3 – Lambda DNA/EcoRI Marker; 4 – Lambda DNA/HindIII Marker; 5-12 – chloroplast DNA of different soybean varieties digested by EcoRI

* S.O. Triboush, N.G. Danilenko, O.G. Davydenko, "A method for isolation of chloroplast DNA and mitochondrial DNA from sunflower", *Plant Molecular Biology Reporter*, vol. 16, 1998, pp. 183- 189.

Results

Variability was detected in **8 positions** along the whole cpDNA molecule:

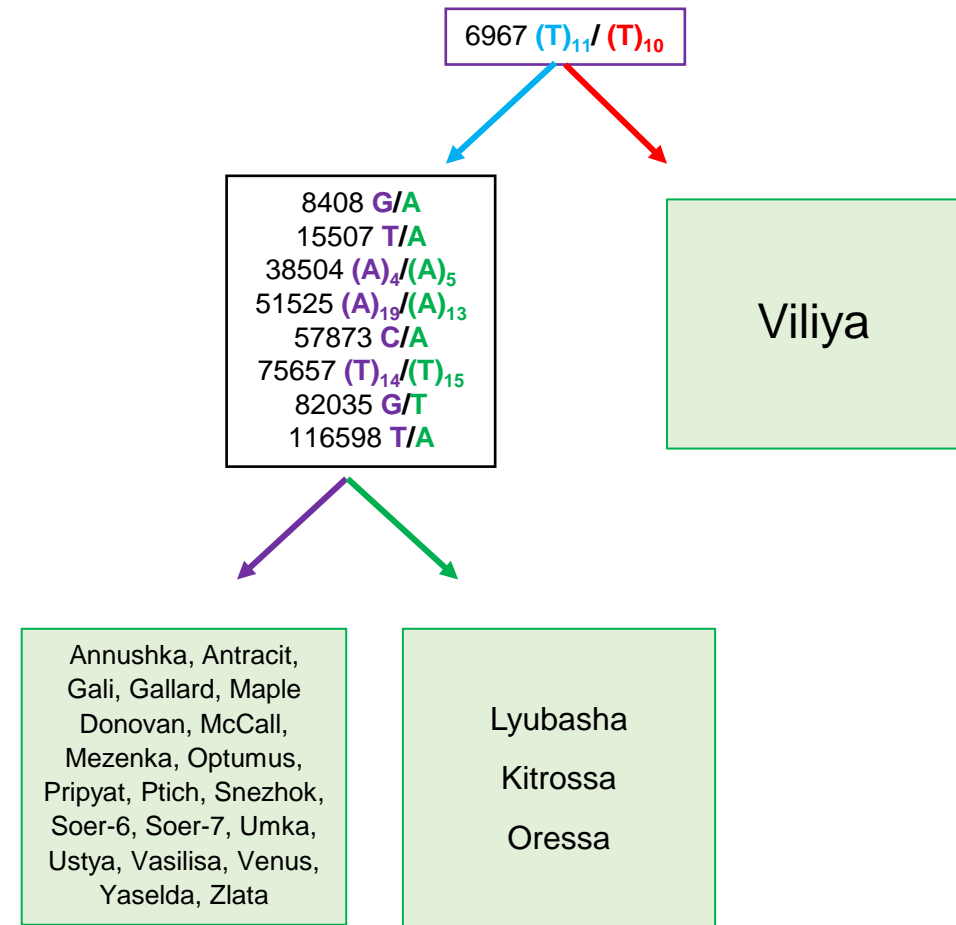


Comparative analysis of 24 complete soybean chloroplast genome sequences revealed 3 SSRs, 1 INDEL and 5 SNP between samples. 4 SNPs were located in coding sequences of *atpB*, *rps4*, *accD* and *rps3* genes. All of them were synonymic.

Discovered polymorphic sites of the soybean chloroplast genome

Location in the reference sequence (MW357264.1)	Reference allele	Alternative allele	Cultivars carrying the alternative allele
6967 rbcL – atpB intergenic region	(T) ₁₁	(T) ₁₀	Viliya
8408 atpB coding sequence	G	A	Lyubasha, Kitrossa, Oressa
15507 rps4 coding sequence	T	A	Lyubasha, Kitrossa, Oressa
38504 rpoC1 intron	(A) ₄	(A) ₅	Lyubasha, Kitrossa, Oressa
51525 atpA - trnR-UCU intergenic region	(A) ₁₉	(A) ₁₃	Lyubasha, Kitrossa, Oressa
57873 accD coding sequence	C	A	Lyubasha, Kitrossa, Oressa
75657 petD intron	(T) ₁₄	(T) ₁₅	Lyubasha, Kitrossa, Oressa
82035 rps3 coding sequence	G	T	Lyubasha, Kitrossa, Oressa
116598 ndhA intron	T	A	Lyubasha, Kitrossa, Oressa

On the basis of revealed diversity 3 haplotypes can be defined:



Conclusions

Analysis of 24 complete soybean chloroplast genome sequences demonstrated the drastically low level of genetic diversity in studied varieties of *Glycine max*. Possibly it is the result of the founder effect – evidently a narrow range of ancestors was used while this group of varieties was created.

Our results are consistent with earlier works, that revealed the low level of intraspecific diversity in soybean.

Quite another data was obtained earlier in our laboratory for Poaceae (barley): 9 INDELs, 19 SSRs and 79 SNPs. *

Three varieties (Lyubasha, Kitrosa and Oresa) were found to be extremely different from the other samples studied: 2 SSRs, 1 INDEL and 5 SNP. It is of special interest that maternal parents of these three varieties originated from China.

* A. Makarevich, M. Siniuskaya, I. Halayenka et al, "Complete sequencing of barley organellar genomes: new data for intraspecific differentiation", Book of abstracts of BGRS/SB-2020, Bioinformatics of Genome Regulation and Structure/Systems Biology: The Twelfth International Multiconference: 06–10 July 2020. – Novosibirsk: ICG SB RAS, 2020, p. 364.

The future plans:



To verificate of NGS data by Sanger



To obtain and compare full mitochondrial genomes of 24 *G. max* varieties listed above



To enlarge the range of *G. max* varieties for further study of their organelle genome diversity



To submit the obtained complete organelle sequences into NCBI GenBank database



To realize: Is there a link between successful hybrid performance and type of maternal cytoplasm?

Acknowledgments

The study was carried out as a part of the State Scientific Research Program «Biotechnologies» 2021–2023, Subprogram 2 «Structural and functional genomics», assignment 2.1.3