The study of organelle DNA variability
in alloplasmic barley lines in the NGS era

M.G. Siniauskaya ✉, A.M. Makarevich, I.M. Goloenko, V.S. Pankratov, A.D. Liaudanski, N.G. Danilenko, N.V. Lukhanina, A.M. Shimkevich, O.G. Davydenko

Institute of Genetics and Cytology of the National Academy of Sciences of Belarus, Minsk, Belarus
✉ e-mail: cytoplasmic@mail.ru

Abstract. Alloplasmic lines are a suitable model for studying molecular coevolution and interrelations between genetic systems of plant cells. Whole chloroplast (cp) and mitochondrial (mt) genome sequences were obtained by the MiSeq System (Illumina). Organelle DNA samples were prepared from a set of 12 alloplasmic barley lines with different cytoplasms of *Hordeum vulgare* ssp. *spontaneum* and *H. vulgare* ssp. *vulgare*, as well as from their paternal varieties. A bioinformatic approach for analysis of NGS data obtained on an organellar DNA mix has been developed and verified. A comparative study of *Hordeum* organelle genomes' variability and disposition of polymorphic loci was conducted. Eight types of chloroplast DNA and 5 types of mitochondrial DNA were distinguished for the barley sample set examined. These results were compared with the previous data of a restriction fragment length polymorphism (RFLP) study of organelle DNAs for the same material. Formerly established data about a field evaluation of alloplasmic barley lines were revised in the light of information about organelle genomes gained after NGS. Totally 17 polymorphic loci were found at exons of chloroplast genomes. Seven of the SNPs were located in the genes of the Ndh complex. The nonsynonymous changes of nucleotides were detected in the *matK*, *rpoC1*, *ndhK*, *ndhG* and *infA* genes. Some of the SNPs detected are very similar in codon position and in the type of amino acid substitution to the places where RNA editing can occur. Thus, these results outline new perspectives for the future study of nuclear-cytoplasmic interactions in alloplasmic lines.

Key words: barley; alloplasmic lines; chloroplast DNA; mitochondrial DNA; next generation sequencing (NGS).

Изучение изменчивости ДНК органелл
аллоплазматических линий ячменя
в эпоху высокопроизводительного секвенирования

М.Г. Синявская ✉, А.М. Макаревич, И.М. Голоенко, В.С. Панкратов, О.Д. Левданский, Н.Г. Даниленко, Н.В. Луханина, А.М. Шимкевич, О.Г. Давыденко

Институт генетики и цитологии Национальной академии наук Беларуси, Минск, Беларусь
✉ e-mail: cytoplasmic@mail.ru

Аннотоция. Аллоплазматические линии являются подходящей моделью для изучения молекулярной ко-эволюции и взаимодействий между генетическими системами растительной клетки. С использованием MiSeq System (Illumina) были определены полногеномные последовательности ДНК органелл клетки – хлоропластов и митохондрий. ДНК органелл выделена из 12 образцов коллекции аллоплазматических линий ячменя с цитоплазмами *Hordeum vulgare* ssp. *spontaneum* (*H. spontaneum*) и *H. vulgare* ssp. *vulgare* (*H. vulgare*), а также из сортов ячменя доноров ядра. Разработан и верифицирован подход к анализу результатов NGS смесей хлоропластной и митохондриальной ДНК для сборки новых полных сиквенсов пластидных и митохондриальных геномов *H. vulgare* и *H. spontaneum*. Проведено сравнительное изучение изменчивости геномов органелл, локализованы полиморфные участки. Выделено восемь типов хпДНК и пять типов мтДНК. Полученная информация сопоставлена с результатами предыдущих исследований этих же линий методом полиморфизма длин рестрикционных фрагментов ДНК органелл. На основании сравнения полногеномных последовательностей хпДНК и мтДНК аллоплазматических линий и сортов доноров ядерных геномов пересмотрены полученные для них ранее данные по формированию признаков, связанных с продуктивностью. Семнадцать полиморфных локусов обнаружено в экзонах пластидных генов. Семь из них расположены в генах Ndh комплекса. Несинонимические замены нуклеотидов идентифицированы в генах *matK*, *rpoC1*, *ndhK*, *ndhG*, *infA*. Вероятно, некоторые SNP являются точками, где происходит эдитинг,

М.Г. Синявская, А.М. Макаревич, И.М. Голоенко …
Н.В. Луханина, А.М. Шимкевич, О.Г. Давыденко

Изучение изменчивости ДНК органелл
аллоплазматических линий ячменя

2020
24 • 1

о чем свидетельствуют позиции замены в кодоне и тип аминокислотной замены. Проведенное исследование открывает новые перспективы для изучения ядерно-цитоплазматических взаимодействий на примере аллоплазматических линий.

Ключевые слова: ячмень; аллоплазматические линии; хлоропластная ДНК; митохондриальная ДНК; высокопроизводительное секвенирование.

Introduction

Barley is one of the most important cereals in the world, after wheat and rice. The history of its cultivation stretches from ancient times (Pankin, von Korf, 2017). Peculiar distribution of different nuclear and cytoplasmic gene loci of barley across localities occurred together with the process of migration of humans (Saisho, Purugganan, 2007). In parallel with the process of barley spreading from the centers of origin to different climatic zones, definite changes in nuclear and cytoplasmic genes have happened. This led to change in the interactions of the nucleus and cytoplasm, the establishment of new nuclear-cytoplasmic relations.

Chloroplasts and mitochondria are essential organelles in plant cells and play an important role in sustaining life. The genomes of organelles have a number of properties that make them indispensable for studying in various areas of modern biology: a large number of copies per cell; relatively conservative sequences; lack of recombination; maternal inheritance; coding the proteins for photosynthesis and respiration; endosymbiotic origin (Danilenko, Davydenko, 2003; Daniell et al., 2016).

The efficiency of photosynthesis (the cooperative work of chloroplast and nuclear genes) and the functioning of the mitochondrial respiratory chain complexes (the joint work of different subunits of mitochondrial complexes coded by nDNA and mtDNA) basically determine the productivity of plants.

Among the suitable models for study of nuclear-cytoplasmic interactions are allo- and isoplasmic plant lines. Alloplasmic lines are usually created via hybridization when the nucleus from one plant (species or subspecies) are placed into the cytoplasm of another plant through substitution backcrossing. Evolutionarily established relations between nucleus and cytoplasm are disrupted at this case, and various consequencies could arise as the result. There is the big set of data about different effects of nuclear-cytoplasmic interactions. Wheat alloplasmic lines with the nucleus from *Triticum aestivum* and cytoplasm from different *Aegilops* and *Triticum* species were studied (Kihara, 1951; Fukasawa, 1959; Mukai et al., 1978; Maan, 1979; Tsunewaki, 1980, 1993). Barley alloplasmic lines were also created and thoroughly studied (Batura et al., 1989; Krepak et al., 1996). Different manifestations of nuclear-cytoplasmic interactions were investigated with the help of the alloplasmic lines collections: the influence of plasmon on morphogenesis, photosynthesis and respiration, fertility, different stress conditions, transmission and recombination of nuclear genes of plants (Palilova, Sylkova, 1987; Nakamura et al., 1991; Sychjova et al., 1998; Goloenko et al., 2002; Tsunewaki et al., 2002, 2019).

NGS of the whole chloroplast (cp) and mitochondrial (mt) genomes from different matrices (either total cell DNA, or a mixture of organelle DNA, or pure plastid and mitochondrial DNA) allows to explore large number of samples simultaneously, to obtain qualitatively new comparative data about the variability of cytoplasmic genomes (Nock et al., 2011; Twy-ford, Ness, 2017). Especially the sequencing of pure organelle DNA templates – chloroplast or mitochondrial (or mixture of organelle DNAs), gained after organelle lysis appears to be the most promising for organelle genome studies. In this case numerous genomic and organelle DNA-like sequences from nucleus are drawn away from analysis facilitating the process of organelle genomes assembly.

Complete chloroplast and mitochondrial genome sequences are essential for realizing and reshaping the phylogenetic relationships between closely related taxa and for improving our understanding of the evolution of plant species (Gornicki et al., 2014; Givnish et al., 2018).

The comparative study of the plastid genomes variability in alloplasmic barley lines (differing in origin of the cytoplasm donor) and their euplasmic analogues were performed.

Materials and methods

Study material. Three barley varieties Vezha, Roland, Visit as well as twelve alloplasmic lines created and maintained in the Lab of Cytoplasmic Inheritance (Institute of Genetics and Cytology, NAS Belarus) were used for organelle DNA isolation (Table 1).

We explored alloplasmic barley lines with cytoplasm of *H. spontaneum* and *H. vulgare*, as unfortunately we lack the original seeds of *H. spontaneum* and *H. vulgare* donors of cytoplasm (maternal parent). This plant material was created in 1990th and then maintained by self pollinating.

The maternal inheritance of organelle genomes is well documented in angiosperms, and considered as one of the laws of cytoplasmic inheritance (Birky, 2008), though very seldom occasional paternal transmission can occur (Reboud, Zeyl, 1994). We considered the plasmon of definite alloplasmic line is equivalent to the original plasmon of corresponding wild *H. spontaneum*, *H. vulgare*. Moreover, one of the recent studies of alloplasmic lines in *Triticum-Aegilops* complex demonstrated the stability of organellar DNA characteristics between native euplasmic plants and alloplasmic line, indicating generally the usefulness of organellar genome in tracing the maternal lineage of species (Tsunewaki et al., 2019).

The whole sequences of *H. vulgare* chloroplast (NC_008590, Saski et al., 2007) and mitochondrial (AP017301, Hisano et al., 2016) genomes accessible in NCBI nucleotide database (GenBank) were used as references for assembly of new barley organelle genomes.

DNA extraction. Organelles were extracted from 7 days seedlings. The fraction of organelles was obtained by differential centrifugation (Triboush et al., 1998). Chloroplast and mitochondrial DNA were obtained by lysis of isolated organelles with subsequent phenol-chloroform deproteinization. This approach allowed simultaneous sequencing of both genomes from each sample. The quality and concentration of DNA was evaluated after 0.8 % agarose gel electrophoresis and at NanoDrop 8000 spectophotometer (Thermo Fisher Scientific).

Table 1. Sources of nucleus and cytoplasm at the alloplasmic lines of barley

| Alloplasmic line | The nuclear donor (*H. vulgare* ssp. *vulgare*) | The cytoplasmic source | | | |
		Subspecies	Working code	CPI No.	Origin
Vezha (W3) BC$_7$	Vezha	*H. vulgare* ssp. *spontaneum*	W3	77129	Israel, Atlit-37
Roland (W3) BC$_7$	Roland				
Visit (W3) BC$_7$	Vizit				
Roland (W4) BC$_7$	Roland	*H. vulgare* ssp. *spontaneum*	W4	77129	Israel, Atlit- 55
Visit (W4) BC$_7$	Vizit				
Vezha (W8) BC$_7$	Vezha	*H. vulgare* ssp. *spontaneum*	W8	77135	Israel, Wadi Quilt-23
Roland (W8) BC$_7$	Roland				
Visit (W8) BC$_7$	Vizit				
Roland (W1) BC$_7$	Roland	*H. vulgare* ssp. *spontaneum*	W1	77133	Israel, Hermon-9
Roland (W9) BC$_7$			W9	77141	Israel, Sede-Boker-21
Roland (Atlas) BC$_7$		*H. vulgare* ssp. *vulgare*	Atlas	77151	California cultivar
Roland (Himalaya) BC$_7$			Himalaya	94435	Nepal cultivar

NGS. The NGS was perfomed on the MiSeq System (Illumina Inc., SanDiego, CA, USA), library preparation kit NexteraXT, MiSeq Reagent Kit v3, read lengths 300 bp.

NGS data processing. The raw data were exported for the primary analysis. The algorithm of sequencing data processing included the following steps: trimming of raw reads (Trimmomatic; Bolger et al., 2014); aligning reads to the "double" reference, containing full sequences of chloroplast and mitochondrial barley genomes (Bowtie2, http://bowtie-bio.sourceforge.net/bowtie2/index.shtml); obtaining mapping statistics (bash scripts, BCFtools, https://samtools.github.io/bcftools/bcftools.html); alignment visualization (Tablet; Milne et al., 2013); generating VCF files (Samtools; Li et al., 2009); filtering VCF files (VCFlib, https://bio.tools/vcflib). The algorithm was tested on artificial Illumina reads synthesized using the ART program. Ultimately, VCF files containing all polymorphic loci of the chloroplast and mitochondrial genomes were obtained (Makarevich et al., 2018).

Comparative analysis of cp genomes. The VCF files for cp genomes of all studied samples as well as accessible complete cp genome sequences of *H. vulgare* ssp. *vulgare* (NC_008590) and *H. vulgare* ssp. *spontaneum* (KC912688, KC912689) taken from the NCBI GenBank database were involved in whole-genome comparison analysis. To visually display the phylogenetic relationships between the studied samples and divide them into groups by the types of chloroplast genomes, the maximum parsimony-cladogram was constructed on the base of all founded SNPs and INDELs. Complete chloroplast genome sequences of *H. jubatum* (KM974741) and *H. bulbosum* (KY636105) available in NCBI GenBank acted as outgroups. The cladogram was built in Excel (Version 14.0.6112.5000) with searching for homologous regions in the outgroup genomes using the SnapGene V.4.3.10.

Results and discussion

15 chloroplast and 15 mitochondrial genome sequences were obtained after NGS (Illumina, MiSeq) of organelle DNA samples (plastid + mitochondrial fraction). Bioinformatics' approaches have been optimized for the processing of the "raw" data after sequencing the mixture of cp and mtDNA.

The comparative analysis of the obtained sequences as well as available in NCBI nucleotide database (GenBank) was carried out, that promoted the assessment of the total level of sequence variation between samples of the same species, defined the regions, where changes more often occur.

Generally, 103 polymorphic loci of cpDNA were revealed after comparison analysis of 15 obtained new full organelle sequences and 4 sequences from GenBank (see Materials and methods). Among them 78 differences for the alloplasmic lines and barley paternal varieties were detected: 56 SNP (39 of them at introns of various genes, 17 – at exons (see Table 4)), 14 SSR loci and 8 indels.

The cladogram for the complete chloroplast genomes of barley was constructed on the base of polymorphic loci found, except for the SSR regions (see the Figure). It displays the diversity of chloroplast genomes of *H. vulgare* and *H. spontaneum*, which allows to subdivide the plasmotypes in the study set of barley (see below Table 2).
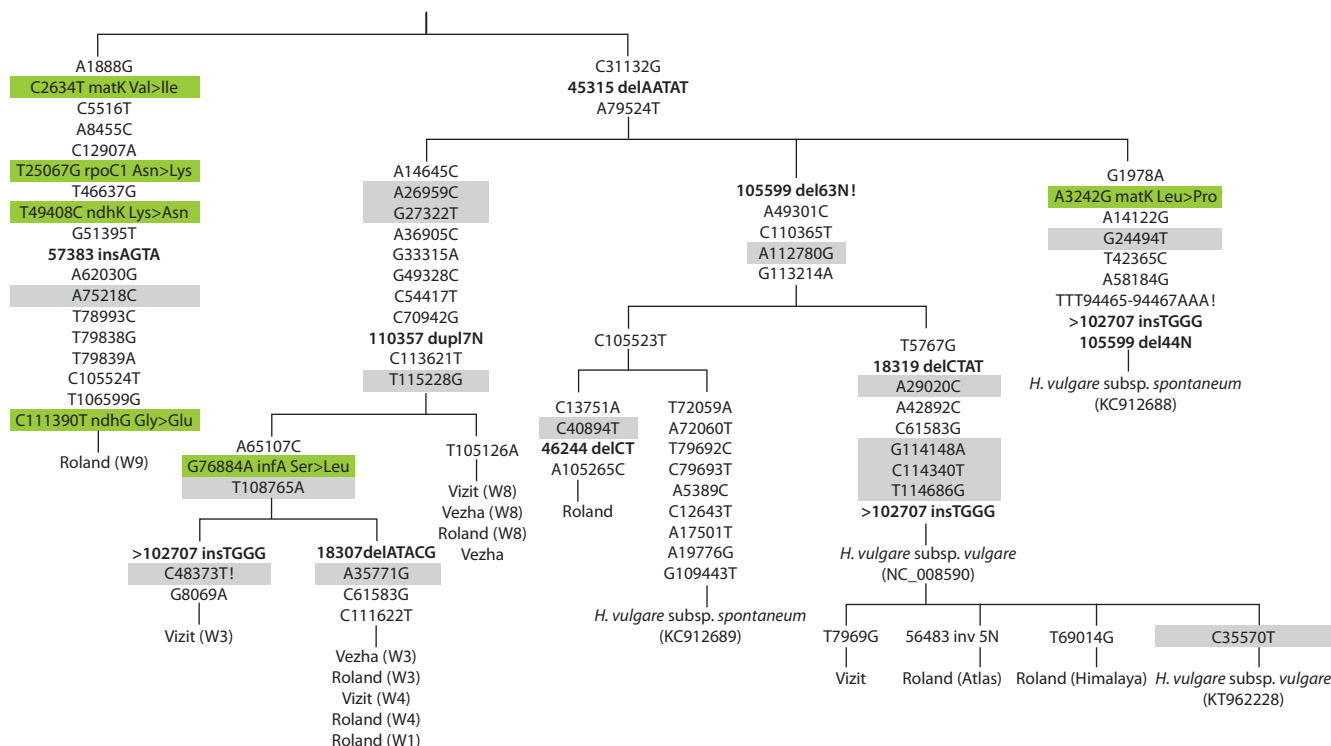
20 loci of mtDNA variability were identified. The level of sequence variation was much lower compared to chloroplast DNA. Only two types of changes were detected – point nucleotide substitutions (19) and insertion (1). All the indels and the most significant SNP were checked by Sanger sequencing to verify the assembly of both genomes.

Qualitatively new data about the variability of *Hordeum* organelle genomes and disposition of polymorphic loci were obtained after NGS. It turned out to be higher than described in earlier reports based on RFLP data or SSR analysis of relatively few loci (Neale et al., 1988; Provan et al., 1999; Russell et al., 2003; Lukhanina et al., 2006; Sipahi et al., 2013).

Differentiation of the alloplasmic lines collection according various molecular methods

As a result of NGS study, eight types of chloroplast DNA and five types of mitochondrial DNA were distinguished for the examined barley sample set (Table 2). We were able to highlight eight plasmotypes totally.

Earlier in 1984 the same set of cytoplasms as well as some others: W1, W3, W4, W5, W7, W8, W9, W10, Atlas, Himalaya, L1, L2 – were studied via restriction of cpDNA with

М.Г. Синявская, А.М. Макаревич, И.М. Голоенко …
Н.В. Луханина, А.М. Шимкевич, О.Г. Давыденко

Изучение изменчивости ДНК органелл
аллоплазматических линий ячменя

2020
24 • 1

The cladogram for the complete chloroplast genomes of barley (based on variability regions).

Maximum parsimony-cladogram was constructed on the base of founded SNPs and INDELs. Chloroplast genome sequences of *H. jubatum* (KM974741) and *H. bulbosum* (KY636105) available in NCBI GenBank were used as outgroups. The polymorphic loci are given according to reference *H. vulgare* (NC_008590). Highlighted in green – synonymous nucleotide substitutions, grey – nonsynonymous substitutions. Indels are written in bold. Exclamation mark (!) – reverse replacements , ">" – parallel mutations.

Table 2. The specific combinations of chloroplast and mitochondrial DNA types in the collection of alloplasmic barley lines (NGS data)

Cp DNA type	Mt DNA type				
	1	2	3	4	5
1	I Roland (W9)				
2				II Vezha (W3), Vizit (W4), Roland (W3), Roland (W4), Roland (W1)	
3			III Vizit (W3)		
4					IV Vizit (W8), Vezha (W8), Roland (W8), Vezha
5	V Roland				
6			VI Vizit		
7		VII Roland (Atlas)			
8		VIII Roland (Himalaya)			

4 endonucleases *BamHI*, *BclI*, *EcoRI*, *HindIII* (Clegg et al., 1984). Five different chloroplast RFLP patterns were detected. Afterwards mtDNA endonuclease digestion was performed for 12 alloplasmic lines with the above mentioned cytoplasms and nucleus from barley variety Vezha. Four groups of mtDNA types were detected (Sychjova et al., 1998). Combining these results, 9 specific plasmotypes were distinguished (Table 3).

NGS data for chloroplast genome (this work) and earlier obtained RFLP data (Clegg et al., 1984) were compared. RFLP analysis having quite low resolution was not able to show all the differences between samples. So, we expected that sequencing of full chloroplast and mitochondrial genomes of alloplasmic lines will shed light on the organelle genome variability of this collection.

Specific motifs for *BamHI*, *BclI*, *EcoRI*, *HindIII* endonucleases were determined in obtained full chloroplast genomes (SnapGene V.4.3.10). Despite many variability positions between samples detected by NGS (see the Figure), they mainly didn't affect the enzyme cutting sites (c. s.). All samples bear the same endonuclease digestion points for *BamHI* (54 c. s.), *BclI* (53 c. s.), *EcoRI* (97 c. s.), *HindIII* (40 c. s.). Just only G33315A substitution led to appearance the new cutting site for *BclI* in W1, W3, W4, W8 and Vezha cytoplasm (*BclI* is *Dam* methylation-sensitive enzyme, so it could influence the RFLP results). All together they bear 54 *BclI* sites.

The Roland (W9) appeared to be especially unique, because only this line had four new SNP: A1888G (created the new digestion site for *HindII*), T25067C, T79838G and T79839A

Table 3. The specific combinations of chloroplast
and mitochondrial DNA types in the collection
of alloplasmic barley lines (according to previously
published data of RFLP analysis)

Cp DNA type	Mt DNA type			
	A	C	D	E
1	I W10, Atlas, Himalaya			
2	II L1, L2			
3	III W7, W4	V W8	VI W9	
4	VII W1			VIII W5
5	IX W3			

(that promote the emergence of two new *EcoRI* cutting points).
Thus Roland (W9) bears 41 *HindIII* and 99 *EcoRI* cleavage
points.

We tried to collate the types of cytoplasms obtained by
two different methodological approaches – RFLP and NGS
sequencing (see the Figure and Table 2, 3).

Roland (Atlas) and Roland (Himalaya) refer to the same
plasmotype according to RFLP data, but NGS differentiated it.

According to Clegg (Clegg et al., 1984), W3 and W4 belong
to different RFLP because of different cpDNA types and one
common mitochondrial. However, the NGS study combined
them in one group, with one exeption – Vizit (W3). Note an
interesting fact – W3 and W4 lines originated from one locality
Atlit at Israel and they are most likely closer genetically than
other lines studied. Sequencing data put alloplasmic lines with
W3 and W4 cytoplasm at one subcluster. So, Vezha (W3),
Roland (W3), Roland (W4), Vezha (W4) are identical for all
polymorphic loci and Vizit (W3) is very similar to other al-
loplasmic lines with W3 cytoplasm, bearing variability only
at 7 polymorphic loci.

Alloplasmic lines with W8 cytoplasm formed separate
group according to both methods.

Roland (W9) line was placed into the special RFLP group;
according to NGS data, this line has more than 20 specific
SNP, thus it appears to be remote from all other lines studied
(see the Figure).

SNP detected at chloroplast genome sequences, their possible importance for plants

Only single nucleotide substitutions were revealed at exons of
chloroplast genes, 12 of them were synonymous: *rpoA*, *rpoC1*,
rpoC2, *ndhA*, *ndhD*, *ndhG*, *ndhH*, *ndhK*, *atpA*, *psaA* genes.
The 5 nonsynonymous changes of nucleotides were detected
in *matK*, *rpoC1*, *ndhK*, *ndhG* and *infA* genes (Table 4).

Some of SNP detected are very similar (position in codon
and also the type of amino acid substitution) to the places

Table 4. List of SNP detected at exons of cpDNA after comparative analysis of obtained whole genomes

No.	Location at cpDNA	Gene	Gene function	Nucleotide substitution	Amino acid replacement	Codon position	Samples in which SNP was detected
1	2634	*matK*	Maturase within trnK intron	C > T	Val370Ile	1	Roland (W9)
2	25067	*rpoC1*	RNA polymerase beta' subunit-1	T > G	Asn571Lys	3	
3	26959	*rpoC2*	RNA polymerase beta' subunit-2	A > C	Ser449=	3	Lines with W1, W3, W4, W8 cytoplasms, Vezha
4	27322			G > T	Gly570=	3	
5	29020			A > C	Ser1136=	3	Vizit, Roland (Atlas), Roland (Himalaya)
6	35771	*atpA*	ATPase alpha subunit	A > G	Gly312=	3	Vezha (W3), Roland (W3), Vizit (W4), Roland (W4), Roland (W1)
7	40894	*psaA*	PSI p700 apoprotein A1	C > T	Pro329=	3	Roland
8	49408	*ndhK*	NADH-dehydrogenase 27kD subunit	T > C	Lys233Asn	3	Roland (W9)
9	75218	*rpoA*	RNA polymerase alpha subunit	A > C	Leu227=	3	Roland (W9)
10	76884	*infA*	Initiation factor 1	G > A	Ser87Leu	2	All lines with W1, W3, W4 cytoplasms
11	108765	*ndhD*	NADH-dehydrogenase ND4	T > A	Gly257=	3	All lines with W1, W3, W4 cytoplasms
12	111390	*ndhG*	NADH-dehydrogenase ND6	C > T	Gly16Glu	2	Roland (W9)
13	112780	*ndhA* exon2	NADH-dehydrogenase ND1	A > G	Gly213=	3	Vizit, Roland
14	114148	*ndhA* exon1	NADH-dehydrogenase ND1	G > A	Pro101=	3	Vizit, Roland (Atlas), Roland (Himalaya)
15	114340			C > T	Leu37=	3	
16	114686	*ndhH*	NADH-dehydrogenase 49kD subunit	T > G	Ser316=	3	Vizit, Roland (Atlas), Roland (Himalaya)
17	115228			T > G	Arg136=	1	All lines with W1, W3, W4, W8 cytoplasms, Vezha

N o t e: The grey filling in the table – marked SNP, where nonsynonymous changes of nucleotides occurred.

М.Г. Синявская, А.М. Макаревич, И.М. Голоенко …
Н.В. Луханина, А.М. Шимкевич, О.Г. Давыденко

Изучение изменчивости ДНК органелл
аллоплазматических линий ячменя

2020
24·1

where RNA editing can occur. Editing usually appeared to be C-U replacement at the mRNA level, it corrects the "wrong" nucleotide at the DNA level. It leads to restoration of phylogenetically conservative amino acid sequences of the proteins (Maier et al., 1996; Danilenko, Davydenko, 2003). Editing is the important posttranscriptional control of the gene expression. It is necessary for plastids functioning and plants survival (Takenaka et al., 2013). Nucleotide substitutions in *matK* 2634 C>T (p.Val370Ile), *ndhK* 40894 C>T (p.Lys233Asn), *infA* 76884 G>A (p.Ser87Leu), *ndhG* 111390 C>T (p.Gly16Glu), *ndhA* exon2 112780 A>G (p.Gly213=), *ndhA* exon1 114148 G>A (p.Pro101=), *ndhA* exon1 114340 C>T (p.Leu37=) possibly mark positions where editing could take place.

According to Table 4, SNP at cpDNA 76884 nucleotide position resulted in amino acid substitution p.Ser87Leu at *infA* gene for lines *Hordeum vulgare* subsp. *spontaneum* with W3 and W4 cytoplasm. *infA* is the factor 1 of translation initiation (IF1) encoded in chloroplast. It is highly conserved and universal in all living organisms (Roll-Mecak et al., 2001). Remarkably we found polymorphic loci at *infA* gene, position 76884 G>A of cpDNA. This SNP occurred at the second position of codon, usual position for editing (87 % of all edited sites – the second nucleotide positions of triplets encoding amino acids). The barley with cytoplasms W3, W4, and W1 bear A (or T at opposite DNA strand) and they needn't correction C>T at the RNA level (editing), but all the other samples studied possess C and it can be corrected by editing machinery. This substitution will lead to Ser>Leu change – most common conversion during editing (Tsudzuki et al., 2001). Some further work to support this idea and to determine functional peculiarities of this event is needed.

About half of SNP (7 out of 17) are located at the genes of *NADH* complex. Chloroplast DNA of barley encodes 11 *Ndh* genes. Of note that these genes are quite variable in different taxons and still remain between the relatively few number of genes retained at cpDNA of higher plants in comparison with genome of prokaryotic *Cyanobacteria* ancestor. They have been lost in *Gnetales*, pines, *Erodium* species, some parasitic plants, endowing their characteristic cpDNA features (Peredo et al., 2013; Sabater, 2018). Nevertheless, most land photosynthetic plants contain these genes. It is considered, that they are important for adaptation of plants to photosynthesis (Martin, Sabater, 2010; Shikanai, 2016). *Ndh* complex function is necessary to optimize photophosphorylation rates under different stress conditions (Rumeau et al., 2007; Martin, Sabater, 2010).

Nearly 50 % of editing sites of flowering plants concentrate at the *ndh* group of genes (Martin, Sabater, 2010). We found loci of variability at five of them. Possibly nucleotide substitution C>T at 111390 position of cpDNA resulted in amino acid change Gly16Glu, can be the place of editing event which needs future verification.

The productivity characters of substituted barley lines with five different nuclear genomes and six plasmotypes were thoroughly tested in field conditions for several years (Goloenko et al., 2000). The direct correlation between productivity and differences in the structure of organelle genomes was not found. The effects of cytoplasm substitution on the economically valuable traits of plants varied greatly depending on nuclear genome, significant impact for various traits was revealed for definite nuclear-cytoplasmic combinations. Nevertheless, two groups of cytoplasms with minimal – W4, W8, and highest influence on productivity – A, L1, W3, W5 were defined on the basis of comparative analysis.

Conclusion

We have got detailed information about sequences of barley organelle genomes of set of alloplasmic lines. At present we cannot precisely foresee which cytoplasm type gives definite phenotypic effect. The results prove that interaction of definite nucleus with specific organelle genome in every case can be unique. Nevertheless, each line was marked and its peculiar features were defined. It represents the new type of starting material for prospective studies in breeding and molecular genetics in Poaceae.

Furthermore, alloplasmic lines appeared to be the promising material for the study of molecular coevolution between genetic systems of plant cell. If editing really occurs at some loci of organelle genes of definite barley cytoplasms it needs special transfactors, at least PLS-type PPR proteins (encoded in nucleus). These proteins are highly specific for their RNA targets. If we combine in alloplasmic lines cpDNA which needs editing and nucleus where transfactor for this event is not encoded, how could plants overcome this inconsistency? Another case – the cytoplasm loci need no editing, but transfactor is encoded in nucleus. Amazingly, phylogenetic comparison of editing sites and corresponding PPR proteins leads to the very interesting fact: if conversion C to T at the DNA level takes place, no need in editing exists and PPR protein can gradually disappear from the nucleus of this taxon. Moreover, as revealed for DOT4 PPR protein – the loss of editing target site through C to T conversion allowed DOT4 in Poaceae to adapt for new function. Also some cases were described where the target sequences for editing exist, but no corresponding PPR is encoded by nucleus (Hein et al., 2019).

Thus, the field for further studies of nuclear-cytoplasmic cooperation is quite large. Possibly editing (PLS-PPRs) or other processes and their players both from nucleus and cytoplasm are involved, that bring us to discovery of fine mechanisms of different cell genetic systems interaction.

References

Batura F.N., Davydenko O.G., Kadyrov M.A. The substitution of cytoplasm in barley varieties and it breeding impact. Doklady AN BSSR = Reports of the Academy of Sciences of the BSSR. 1989; 33(7):657-659. (in Russian)

Birky C.W. Uniparental inheritance of organelle genes. Curr. Biol. 2008;18(16):R692-R695. DOI 10.1016/j.cub.2008.06.049.

Bolger A.M., Lohse M., Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30(15):2114-2120. DOI 10.1093/bioinformatics/btu170.

Clegg M.T., Brown A.N.D., Whitfeld P.R. Chloroplast DNA diversity in wild and cultivated barley: implication for genetic conservation. Genet. Res. 1984;4:339-343. DOI 10.1017/S0016672300026112.

Daniell H., Lin C.S., Yu M., Chang W.J. Chloroplast genomes: diversity, evolution, and applications in genetic engineering. Genome Biol. 2016;17(1):134. DOI 10.1186/s13059-016-1004-2.

Danilenko N.G., Davydenko O.G. Worlds of Organelle Genomes, Minsk: Tekhnalogiya Publ., 2003. (in Russian)

Fukasawa H. Nucleus substitution and restoration by means of successive backcrosses in wheat and its related genus *Aegilops*. Jpn. J. Bot. 1959;17:55-91.

Givnish T., Zuluaga A., Spalink D., Soto Gomez M., Lam V.K.Y., Saarela J.M., Sass C., Iles W.J.D., de Sousa D.J.L., Leebens-Mack J., Chris Pires J., Zomlefer W.B., Gandolfo M.A., Davis J.I., Stevenson D.W., dePamphilis C., Specht C.D., Graham S.W., Barrett C.F, Ané C. Monocot plastid phylogenomics, timeline, net rates of species diversification, the power of multi-gene analyses, and a functional model for the origin of monocots. Am. J. Bot. 2018;105(11):1888-1910. DOI 10.1002/ajb2.1178.

Goloenko I.M., Davydenko O.G., Shimkevich A.M. The disturbance of splitting by nuclear marker genes in allo- and isoplasmic barley lines. Genetika = Genetics (Moscow). 2002;38(7):944-949. (in Russian)

Goloenko I.M., Teljatnicova A.A., Davydenko O.G. Some nuclei cytoplasmic combinations of barley substituted lines collection change the productivity characteristics. Barley Genet. Newsl. 2000;30:28-31.

Gornicki P., Zhu H., Wang J., Challa G., Zhang Z., Gill B., Li W. The chloroplast view of the evolution of polyploid wheat. New Phytologist. 2014;204(3):704-714. DOI 10.1111/nph.12931.

Hein A., Brenner S., Knoop V. Multifarious evolutionary pathways of a nuclear RNA editing factor: disjunctions in coevolution of DOT4 and its chloroplast target rpoC1eU488SL. Genome Biol. Evol. 2019; 11(3):798-813. DOI 10.1093/gbe/evz032.

Hisano H., Tsujimura M., Yoshida H., Terachi T., Sato K. Mitochondrial genome sequences from wild and cultivated barley (*Hordeum vulgare*). BMC Genomics. 2016;17(1):824. DOI 10.1186/s12864-016-3159-3.

Kihara H. Substitution of nucleus and its effects on genome manifestations. Cytologia. 1951;16:177-193. DOI 10.1508/cytologia.16.177.

Krepak I.M., Davydenko O.G., Triboush S.O., Danilenko N.G. The creation of allo- and isoplasmic barley lines. In: Molecular-Genetic Markers in Plants: Abstracts of Int. Conf. Yalta, Nov. 11–15, 1996. Kiev: Agrarna Nauka Publ., 1996;74. (in Russian)

Li H., Handsaker B., Wysoker A., Fennell T., Ruan J., Homer N., Marth G., Abecasis G., Durbin R. 1000 Genome Project Data Processing Subgroup. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009;25(16):2078-2079. DOI 10.1093/bioinformatics/btp352.

Lukhanina N.V., Siniauskaya M.G., Goloenko I.M., Davydenko O.G. Chloroplast microsatellites in barley: the reduction of variability spectrum in cultivated forms. Ekologicheskaya Genetika = Ecological Genetics. 2006;IV(1):17-21. (in Russian)

Maan S.S. Specificity of nucleo-cytoplasmic interactions in *Triticum* and *Aegilops* species. Wheat Inform. Service. 1979;50:71-79.

Maier R., Zeltz P., Kössel H., Bonnard G., Gualberto J., Grienenberger J. RNA editing in plant mitochondria and chloroplasts. Plant Mol. Biol. 1996;32(1-2):343-365. DOI 10.1007/BF00039390.

Makarevich A., Pankratov O., Sinyavskaya M., Lukhanina N., Shymkevich A., Liaudansky A., Goloenko I., Danilenko N., Davydenko O. NGS data processing method for the mixture of chloroplast and mitochondrial DNA of barley. In: Systems Biology and Bioinformatics (SBB-2018): The Tenth International Young Scientists School (27–31 Aug. 2018, Novosibirsk, Russia): Abstracts. Novosibirsk, 2018;29. DOI 10.18699/SBB-2018-23.

Martin M., Sabater B. Plastid *ndh* genes in plant evolution. Plant Physiol. Biochem. 2010;48(8):636-645. DOI 10.1016/j.plaphy.2010.04.009.

Milne I., Stephen G., Bayer M., Cock P.J.A., Pritchard L., Cardle L., Shaw P.D., Marshall D. Using Tablet for visual exploration of second-generation sequencing data. Brief. Bioinformatics. 2013;14(2):193-202. DOI 10.1093/bib/bbs012.

Mukai Y., Maan S.S., Panayotov I., Tsunewaki K. Comparative studies of the nucleus-cytoplasm hybrids of wheat produced by three research groups. In: Proc. 5th Int. Wheat Genet. Symp. 1978;1:282-292.

Nakamura Ch., Yamakawa S., Suzuki T. Recovery of normal photosynthesis and respiration in common wheat with *Agropyron* cytoplasms by telocentric *Agropyron* chromosomes. Theor. Appl. Genet. 1991;81:514-518. DOI 10.1007/BF00219442.

Neale D.B., Saghai-Maroof M.A., Allard R.W., Zhang Q., Jorgensen R.A. Chloroplast DNA diversity in populations of wild and cultivated barley. Genetics. 1988;120(4):1105-1110.

Nock C.J., Waters D.L.E., Edwards M.A., Bowen S.G., Rice N., Cordeiro G.M., Henry R.J. Chloroplast genome sequences from total DNA for plant identification. Plant Biotechnol. J. 2011;9:328-333. DOI 10.1111/j.1467-7652.2010.00558.x.

Palilova A.N., Sylkova T.A. Formation of productivity in the new series of alloplasmic wheat lines under the influence of alien cytoplasm. Selskokhozyaistvennaya Biologiya = Agricultural Biology. 1987; 12:3-5. (in Russian)

Pankin A., von Korff M. Co-evolution of methods and thoughts in cereal domestication studies: a tale of barley (*Hordeum vulgare*). Curr. Opin. Plant Biol. 2017;36:15-21. DOI 10.1016/j.pbi.2016.12.001.

Peredo E.L., King U.M., Les D.H. The plastid genome of *Najas flexilis*: adaptation to submersed environments is accompanied by the complete loss of the NDH complex in an aquatic angiosperm. PLoS One. 2013;8(7):e68591. DOI 10.1371/journal.pone.0068591.

Provan J., Russell J.R., Booth A., Powell W. Polymorphic chloroplast simple sequence repeat primers for systematic and population studies in the genus *Hordeum*. Mol. Ecol. 1999;8(3):505-511. DOI 10.1046/j.1365-294X.1999.00545.x.

Reboud X., Zeyl C. Organelle inheritance in plants. Heredity. 1994;72: 132-140. DOI 10.1038/hdy.1994.19.

Roll-Mecak A., Shin B.S., Dever T.E., Burley S.K. Engaging the ribosome: universal IFs of translation. Trends Biochem. Sci. 2001;26: 705-709.

Rumeau D., Peltier G., Cournac L. Chlororespiration and cyclic electron flow around PSI during photosynthesis and plant stress response. Plant Cell Environ. 2007;30(9):1041-1051. DOI 10.1111/j.1365-3040.2007.01675.x.

Russell J.R., Booth A., Fuller J.D., Baum M., Ceccarelli S., Grando S., Powell W. Patterns of polymorphism detected in the chloroplast and nuclear genomes of barley landraces sampled from Syria and Iordan. Theor. Appl. Genet. 2003;107(3):413-421. DOI 10.1007/s00122-003-1261-9.

Sabater B. Evolution and function of the chloroplast. Current investigations and perspectives. Int. J. Mol. Sci. 2018;19(10):3095.

Saisho D., Purugganan M.D. Molecular phylogeography of domesticated barley traces expansion of agriculture in the Old World. Genetics. 2007;177(3):1765-1776. DOI 10.1534/genetics.107.079491.

Saski C., Lee S.B., Fjellheim S., Guda C., Jansen R.K., Luo H., Tomkins J., Rognli O.A., Daniell H., Clarke J.L. Complete chloroplast genome sequences of *Hordeum vulgare*, *Sorghum bicolor* and *Agrostis stolonifera*, and comparative analyses with other grass genomes. Theor. Appl. Genet. 2007;115(4):571-590. DOI 10.1007/s00122-007-0567-4.

Shikanai T. Chloroplast NDH: a different enzyme with a structure similar to that of respiratory NADH dehydrogenase. Biochim. Biophys. Acta. Bioenergetics. 2016;1857(7):1015-1022. DOI 10.1016/j.bbabio.2015.10.013.

Sipahi H., Meydan H., Özbek K. Genetic variation of barley germplasm from Turkey assessed by chloroplast microsatellite markers. Int. J. Biodivers. Conserv. 2013;5(11):775-781. DOI 10.5897/IJBC2013.0613.

Sychjova I.M., Aksjonova H.A., Davydenko O.G. The effect of intraspecific cytoplasmic substitution on the frequency of chiasmata and sister chromatid exchanges and on marker gene segregation. In: Lelley T. (Ed.). Current Topics in Plant Cytogenetics Related to Plant Improvement. Wien: WUV-Universitatsverlag, 1998;168-174.

Takenaka M., Zehrmann A., Verbitskiy D., Härtel B., Brennicke A. RNA editing in plants and its evolution. Annu. Rev. Genet. 2013; 47(13):335-352. DOI 10.1146/annurev-genet-111212-133519.

М.Г. Синявская, А.М. Макаревич, И.М. Голоенко …
Н.В. Луханина, А.М. Шимкевич, О.Г. Давыденко

Изучение изменчивости ДНК органелл
аллоплазматических линий ячменя

2020
24 • 1

Triboush S.O., Danilenko N.G., Davydenko O.G. Method for isolation of chloroplast DNA and mitochondrial DNA from sunflower. Plant Mol. Biol. Rep. 1998;16:183-189. DOI 10.1023/A:1007487806583.

Tsudzuki T., Wakasugi T., Sugiura M. Comparative analysis of RNA editing sites in higher plant chloroplasts. J. Mol. Evol. 2001;53: 327-332. pmid: 11675592.

Tsunewaki K. (Ed.). Genetic Diversity of the Cytoplasm in *Triticum* and *Aegilops*. Tokyo: Jpn. Soc. for the Promotion of Science, 1980.

Tsunewaki K. Genome-plasmon interactions in wheat. Jpn. J. Genet. 1993;68(1):1-34. DOI 10.1266/jjg.68.1.

Tsunewaki K., Mori N., Takumi S. Experimental evolutionary studies on the genetic autonomy of the cytoplasmic genome "plasmon" in the *Triticum* (wheat)-*Aegilops* complex. Proc. Natl. Acad. Sci. USA. 2019;116(8):3082-3090. DOI 10.1073/pnas.1817037116.

Tsunewaki K., Wang G.-Z., Matsuoka Y. Plasmon analysis of *Triticum* (wheat) and *Aegilops*. 2. Characterization and classification of 47 plasmons based on their effects on common wheat phenotype. Genes Genet. Syst. 2002;77(6):409-427. DOI 10.1266/ggs.77.409.

Twyford A.D., Ness R.W. Strategies for complete plastid genome sequencing. Mol. Ecol. Resour. 2017;17(5):858-868. DOI 10.1111/1755-0998.12626.

ORCID ID

M.G. Siniauskaya orcid.org/0000-0002-9800-4725