ISSN 0006-2979, Biochemistry (Moscow), 2017, Vol. 82, No. 6, pp. 678-691. © Pleiades Publishing, Ltd., 2017.
Original Russian Text © N. P. Yurina, L. S. Sharapova, M. S. Odintsova, 2017, published in Biokhimiya, 2017, Vol. 82, No. 6, pp. 900-915.

REVIEW

Structure of Plastid Genomes of Photosynthetic Eukaryotes

N. P. Yurina*, L. S. Sharapova, and M. S. Odintsova

Bach Institute of Biochemistry, Research Center of Biotechnology, Russian Academy of Sciences, 119071 Moscow, Russia; E-mail: nyurina@inbi.ras.ru, NadezhdaYurina@hotmail.com

Received February 6, 2017
Revision received March 10, 2017

Abstract—This review presents current views on the plastid genomes of higher plants and summarizes data on the size, structural organization, gene content, and other features of plastid DNAs. Special emphasis is placed on the properties of organization of land plant plastid genomes (nucleoids) that distinguish them from bacterial genomes. The prospects of genetic engineering of chloroplast genomes are discussed.

DOI: 10.1134/S0006297917060049

Keywords: plastids, photosynthetic eukaryotes, chloroplast genome

The endosymbiotic origin of plastids in plant and algal cells is now commonly accepted. Plastids originated over a billion years ago as a result of symbiosis between nonphotosynthetic eukaryotic cells and photosynthetic cyanobacteria [1]. The emergence of plastids and, eventually, plants and algae, has had an enormous effect on the evolution of life on Earth. Plastids of plants and green algae (chloroplasts), red algae (rhodoplasts), and glaucophytes (cyanelles, or cyanoplasts) originated from the primary symbiosis.

Plants from several photosynthetic eukaryotic taxa contain plasmids that emerged from secondary or tertiary symbiosis, i.e. symbiosis of nonphotosynthetic eukaryotes with free-living photosynthetic eukaryotes (e.g. red or green algae) [2]. Secondary plastids are typical for cryptophytes, chlorarachniophytes, haptophytes, euglenophytes, most dinoflagellates, and diatomic algae. Diatomic algae, as well as stramenopiles (heterokonts), were formed by symbiosis between green and red algae [3-5]. Tertiary plasmids originated when photosynthetic eukaryotes lost secondary plasmids and replaced them with plastids from other symbionts. Tertiary plastids have been found in dinoflagellates [3].

Unlike primary plastids that emerged from the symbiosis between eukaryotes and prokaryotes and are surrounded by two membranes, secondary and tertiary plastids are bound by more than two membranes and typically lack the nucleus in photosynthetic eukaryotes [6]. Plants from some taxa with secondary plasmids, such as chlorarachniophytes and cryptophytes, retain a relic eukaryotic nucleus – the nucleomorph – located between the inner two and outer chloroplast membranes [4, 6]. Nucleomorphs of the single-cell cryptophyte *Guillardia theta* and chlorarachniophyte *Bigelowiella natans* contain three miniature chromosomes bearing 487 and 331 protein-coding genes, respectively [4, 7].

During evolution, the plastid genome (plastome) of originally genetically autonomous endosymbiont was considerably reduced in size due primarily to gene loss or transfer to the nuclear genome of the host organism [2, 8]. It was suggested that most chloroplast genes had been lost during the evolutionary stage between original endosymbiotic cyanobacterium (~2000 protein-coding genes) and the common ancestor of all existing chloroplast genomes (~210 protein-coding genes) [9]. The nuclear genome of the model plant *Arabidopsis thaliana* contains at least 2000 genes of cyanobacterial origin [10]. The plastome size in chloroplasts of land plants is 150 kb, while the cyanobacterial genome contains 3000 kb [11]. Since chloroplasts are composed of many more proteins (2100-3600) than could be encoded by the chloroplast genome [10, 12, 13], most of these proteins (>95%) are encoded in the nucleus, synthesized in the cytoplasm, and then transported to the chloroplast. The fact that many proteins encoded by the nuclear-located plastid genes are directed to the chloroplasts suggests active intracellular gene migration during evolution [4, 14].

Abbreviations: IR, inverted repeat; kb, thousand base pairs; LSC, large single copy region; nupDNA, nuclear-localized plastid DNA; SSC, small single copy region.

* To whom correspondence should be addressed.

Plant nuclear genomes contain multiple insertions of plastid DNA (nuclear localized plastid DNAs, nupDNAs). The abundance of nupDNAs correlates positively with nuclear genome size: the relatively large genome of the soybean plant *Glycine max* contains 1718 nupDNAs, whereas the small *Arabidopsis thaliana* genome contains only 85 nupDNAs. The nupDNAs include genes involved in energy metabolism, photosynthesis, fatty acid metabolism, cofactor biosynthesis, metabolite transport, and other cell processes. The number of nupDNAs also depends on the number of plastids in the cell. Genome polyploidy is typical for all chloroplasts. In vascular plants, the number of plastome copies varies from less than a hundred in meristematic cells to several thousands in fully developed diploid cells of leaf parenchyma [15]. Analysis of nupDNAs from 23 plant species confirmed considerable interspecific differences in the size and gene content of these nupDNAs [16]. The nupDNAs are not evenly distributed over the chromosomes: approximately 22.5% of all insertions in the rice nuclear genome are located on chromosome 10. In contrast, chromosome 11, which is larger than chromosome 10, only contains 0.9% of the nupDNAs [17]. It is believed that nupDNAs were initially inserted close to centromeres and then fragmented and distributed along the entire genome by transposable elements [18].

Along with plastome gene fragments, nuclear genomes contain multiple full-size copies of protein-coding plastid genes, some of which remain functional and could be transcribed by the nuclear transcription apparatus [4]. However, it appears that many of the putative promoter sequences of plastid genes were translocated or eliminated following plastid DNA insertion into the nuclear genome. For example, 55.93% of photosynthesis-related and 70.67% of energy metabolism-related plastid genes in the rice *Oryza sativa* nuclear genome lack promoter sequences. In the *Medicago truncatula* genome, 37.5% of promoter sequences for the photosynthesis-related genes and 96.08% of promoter sequences for the energy-metabolism related genes are eliminated [16]. It was suggested that transcription of nuclear-located plastid genes requires the presence of nuclear promoters. However, it was demonstrated recently in a model plant system that the plastid gene promoter *16S rrn* can drive the reporter gene transcription immediately after transfer to the nucleus without involvement of the nuclear sequences. Although this promoter might be sufficient for gene transcription, the resulting mRNAs might lack signals required for translation [16, 19]. The transfer of plastid DNA into nuclear genomes happened in the early stages of organelle evolution and is supposedly still ongoing in many eukaryotes. Since nupDNAs might retain the features of ancient plastomes, such as unique insertions and deletions (indels) or specific DNA sequences, their comparison to plastomes of modern plants provide new insights into the evolution of eukaryotic cells [18].

In this review, we present currents concepts on the structural organization of plastid genomes (plastomes) and their properties in different taxa of photosynthetic eukaryotes with special emphasis on chloroplasts. Although the major function of chloroplasts is photosynthesis, they are also involved in other processes essential for plant growth and development, such as biosynthesis of fatty acids, pigments, starch, and amino acids [1, 2, 9]. Here, we review data on the genetic content and properties of plastid nucleoids and briefly described the prospects of genetic engineering of plastids. The problems of gene expression in plastids and plastid origin and evolution require separate discussion and have been omitted from this review.

PLASTID NUCLEOIDS

Plastid DNA is organized into high molecular weight protein- and RNA-containing complexes that are attached to the inner plastid membranes. These structures are similar to bacterial nucleoids (hence the name) and sometimes are called plastid nuclei [2, 20]. Nucleoids are considered the be the major form of plastome organization in organelles [15].

Plastid nucleoids of higher plant plastids are highly dynamic structures: their number, morphology, structural organization, and protein content depend on the environmental conditions and alter considerably during chloroplast development [13, 21]. Plastid nucleoids are located in the envelope membrane of immature proplastids and then relocate to the thylakoids in mature chloroplasts, where they decrease in size and become more compact and abundant. Mature chloroplasts contain many small nucleoids attached to the thylakoids [11, 22]. It was suggested that nucleoids can differ structurally and functionally within the same chloroplast. On average, plastid nucleoids contain 10-20 copies of the plastid genome [23, 24].

Plastid nucleoids have unique composition and structure – they resemble both prokaryotic nucleoids and eukaryotic chromatin [2, 20]. The fundamental difference between genome organization in plastids and bacteria is that plastids have multiple nucleoids with a varying number of genome copies, whereas bacteria have only a single nucleoid containing a variable number of DNA molecules that is different in different bacterial species. The compact structure of DNA in plastid nucleoids has been compared with chromatin of the eukaryotic cell nucleus. The central body of the nucleoid with dense packaging (core) can correspond to eukaryotic heterochromatin, while more disperse peripheral regions (areas of active transcription) resemble eukaryotic euchromatin [20, 24]. In the nucleus, the availability of DNA for transcription is regulated mainly by posttranslational modifications, whereas in bacteria regulation of transcription involves the exchange

of DNA-binding proteins and changes in the compaction of the nucleoid. It seems that in plastids, the structural organization of the nucleoids is regulated by both kinds of mechanisms [20, 24].

Proteomic analysis of nucleoids identified not only all basic enzymes required for DNA replication, repair, and transcription, but also several proteins involved in posttranscriptional events, such as mRNA processing and editing and ribosome assembly [25]. Interestingly, most of the proteins involved in posttranscriptional processes in plastid nucleoids are not related to bacterial proteins [20, 22].

Several enzymes involved in genome replication, including DNA polymerase, DNA primase, DNA helicase, DNA topoisomerase, DNA ligase and several DNA recombination-associated enzymes, have been identified in plastids. These enzymes are encoded in the nucleus and then transported to the organelles. The regulatory mechanisms controlling the plastid genome remain obscure. It is believed that in studied green and red algae, DNA replication in plastids is similar to that in cyanobacteria. Land plants and multicellular plants have more complex regulatory mechanisms controlling the replication of organellar genomes. In land plants, replication of organellar genomes is restricted to meristematic tissues and not associated with the cell cycle or organellar division [26, 27].

Only few of identified nucleoid proteins have been characterized functionally [2, 28, 29]. Recently, proteomic analysis revealed two new DNA-binding proteins (plastid nucleoid-associated proteins, ptNAPs) that play an important regulatory role in the expression of plastid genes and chloroplast development [28, 30].

In prokaryotic cells (*E. coli*), the structure of bacterial DNA affects gene expression and is regulated by nucleoid proteins HU, FIS, and H-NS. One of these proteins, HU, is responsible for chromosome packaging. Homologs of bacterial HU proteins, namely HU-like proteins, have been found in the nuclei of most dinoflagellates and some algae, where they play the role of histones. HU-like proteins have been found in cyanobacteria, the primitive red alga *Cyanidioschyzon merolae*, and the green alga *Chlamydomonas*. However, in land plants, genes for the HU-like proteins have neither been found in any of the sequenced plastid genomes, nor in any of the sequenced nuclear genomes. Land plants, including mosses and flowering plants, have lost not only the HU genes, but also genes for other prokaryotic DNA-binding proteins [11, 24].

Eukaryotic sequence motifs (e.g. nuclear transcription factor motif) have been found in some of recently identified ptNAPs [20, 29, 31]. Thus, *Arabidopsis thaliana* contains 20 proteins with eukaryotic SWIB domains. At least four of these SWIB domain proteins are targeted to chloroplasts. These low molecular weight proteins were found to have a high Lys content and high isoelectric points; they might act as functional analogs of bacterial histone-like DNA-binding HU proteins. One of the chloroplast SWIB domain proteins, SWIB4, which has been localized to both nucleoids and cell nuclei, has a histone H1 motif next to the SWIB domain. Moreover, recombinant SWIB4 was shown to induce compaction and condensation of nucleoids. It is interesting that SWIB domain proteins have also been found in the bacterium *Chlamydophila felis*. This bacterium contains histone H1-like protein (Hc1) and a low molecular weight stand-alone SWIB domain protein that is the only type of SWIB protein found in Chlamydiae, a group of bacteria that live as endosymbionts or parasites in other bacteria or in eukaryotic cells. Although no Chlamydiae species, which are mostly animal and human pathogens, have been found in plant cells, many chlamydial genes share similarity with plant genes. Phylogenetic analysis suggests that an ancestral member of the Chlamydiae group facilitated the establishment of the primary endosymbiosis between cyanobacteria and an early eukaryote and that Chlamydiae have contributed at least 55 genes to plant genomes. Genes coding members of this subgroup of SWIB domain proteins have been found in the sequenced genomes of all land plants, but not in those of algae. The homology is very high among the sequences found in angiosperms, gymnosperms, mosses, and clubmosses (Lycopodiaceae) [24, 32].

It has been suggested that eukaryotic-type DNA-binding ptNAPs originated from eukaryotic nuclear proteins involved in DNA replication, repair, and transcription that have changed their subcellular localization during evolution and became plastid or dually targeted proteins. A considerable number of these proteins with a nuclear localization sequence also have the *N*-terminal plastid targeting sequence [20]. Several plastid DNA-binding proteins have motifs that are typically found in nuclear transcription factors. Examples are PEND, CND41, PD1, and PD3 [20], as well as Whirly1 (WHY1) [33].

Whirly proteins play an important role in nucleotide metabolism. They are located in mitochondrial and plastid nucleoids and can bind single-stranded RNA and DNA. WHY1 is involved in posttranscriptional regulation in chloroplast nucleoids and the coupling of transcription and splicing. In addition to its role in nucleoids, AtWHY1 has been reported to act as a nuclear transcription factor regulating the salicylic-acid dependent defense system [20, 22].

DNA-binding proteins PEND and MFP1 (MAR-binding filament-like protein) are anchor proteins that provide nucleoid attachment to chloroplast membranes. PEND is composed of cbZIP (leucine zipper) and C-terminal hydrophobic domains. The cbZIP domain is involved in PEND dimerization and sequence-specific DNA binding, whereas the C-terminal hydrophobic domain is required for targeting PEND to the chloroplast

envelope membrane, where DNA replication takes place. MFP1 attaches nucleoids to the thylakoid membranes, and its C-terminal domain exhibits DNA-binding activity. Both PEND and MFP1 are believed to be involved in nucleoid translocation from the chloroplast envelope membrane to the thylakoid membranes [20, 22, 24].

Sulfite reductase (SiR) is a 70-kDa soluble enzyme that catalyzes sulfite reduction into sulfide. SiR is one of the most abundant proteins in the plastid nucleoid. When bound to DNA, SiR induces compaction of plastid DNA and efficiently represses chloroplast DNA replication and transcription *in vitro*. Depletion of SiR leads to a significant downregulation of expression of some chloroplast genes, whereas other genes are unaffected or even upregulated. This provides important evidence that DNA compaction and gene expression are linked and that the level of compaction is not homogenous [20]. SiR interacts with thioredoxin Z (TrxZ) and two nucleoid superoxide dismutases, FSD2 and FSD3, which in turn bind the plastid-encoded RNA polymerase (PEP) [34]. The sulfite reductase activity of SiR is regulated by photoreduced ferredoxin. DNA binding to SiR does not influence enzyme activity and is not sequence-specific. However, DNA binding to SiR (similarly to MFP1 and SWIB4) is regulated by protein phosphorylation – phosphorylated SiR does not bind DNA. Phosphorylation plays an important role in posttranslational protein modification in eukaryotic cells. In chloroplasts, protein phosphorylation regulates photosynthesis, metabolic functions, and transcription [34]. A surprising feature of some DNA-binding plastid nucleoid proteins, such as SiR, FSD2/FSD3, CND41, WHY1, and MSH1, is their multifunctionality [20, 29].

SVR4 (suppressor of variegation 4, also known as MRL7) and its homolog SVR4-like (MRL7-like) are two proteins that were originally identified in *Arabidopsis* nucleoids. These proteins have orthologs in all monocotyledonous and dicotyledonous plants with sequenced genomes. In lower land plants, moss *Physcomitrella patens* and lycophyte *Selaginella moellendorffii*, only one protein with sequence similarities to both SVR4 and SVR4-like was found. In the absence of either SVR4 or SVR4-like, the nucleoid organization is disturbed. Fewer and larger nucleoids with the tendency to form ring-like structures are detected in the mutants [29]. Both proteins are located in chloroplasts and expressed at the early stages of chloroplast development. The primary amino acid sequences of SVR4 and SVR4-like contain 20% negatively charged glutamate and aspartate residues. Both proteins were found to be essential for normal functioning of the plastid transcriptional apparatus. It has been proposed that SVR4 and SVR4-like act as molecular chaperones involved in the establishment of proper DNA–protein interaction in chloroplasts [2, 24, 29]. It is possible that the two proteins might have similar functions, but at various stages of chloroplast development [29]. Immunological analysis revealed that SVR4 is a component of transcriptionally active chromosome (TAC) in chloroplasts.

Electron microscopy of isolated plastid TACs revealed chromatin-like beaded structures with several protruding DNA loops, suggesting that TAC is a subdomain of the chloroplast nucleoid. TAC forms a central core of the plastid nucleoid and serves as a "transcription factory" [11]. The protein composition of TAC is poorly studied. The proteome of spinach TAC contains CP29B and CP31 ribonucleoproteins and pentatricopeptide repeat (PPR)-containing proteins with molecular masses of 97 (pTAC2) [35], 98, and 67 kDa [32]. Since these proteins participate in organellar RNA processing and translation, it is reasonable to suggest that posttranscriptional mechanisms of gene expression regulation (e.g. RNA stability and exchange) at least partly involve nucleoid TACs. Proteomic analysis of TAC proteins revealed 18 novel nonbacterial proteins (pTAC1-pTAC18) that are important for the activity of cyanobacterial plastid RNA polymerase (PEP). At least 10 TAC proteins have been found to form complexes with PEP; some of these proteins interact directly with PEP. For example, pTAC3 associates with the PEP complex at all three stages of the transcription cycle (initiation, elongation, and termination), and its defect results in decreased transcriptional activity of PEP. All genes coding for PEP-interacting proteins (except *TrxZ*) have been found in land plants, but not in the green alga *Chlamydomonas*, which suggests that during evolution land plants acquired new non-cyanobacterial protein components of the PEP complex for the regulation of plastid transcriptional activity [28]. Therefore, PEP-catalyzed transcription is mediated by several pTAC proteins. Another protein involved in plastid transcription is nuclear-encoded plastid RNA polymerase (NEP). So far, NEP has not been identified in any of the studied TAC or nucleotide proteomes [20, 22]. Immunological analysis of the TAC proteome revealed TCP34 (molecular mass, 34 kDa) that (similarly to MFP1) anchors DNA to the thylakoid membrane. Protein pTAC16 might also perform the same anchoring function. It was suggested that pTAC16 phosphorylation is regulated by the redox state of the photosynthetic apparatus and controls the ability of nucleoids to bind to the membranes [29, 30]. WHY1/pTAC1 is another TAC protein involved in nucleoid attachment to thylakoid membranes. Barley and maize nucleoids contain WHY1 protein only [29]. The TAC proteome of *Arabidopsis* chloroplasts contains AtWHY1 and AtWHY3 proteins (pTAC1 and pTAC11, respectively). In *Arabidopsis*, both these proteins function in chloroplasts as anti-recombination proteins and act as safeguards against plastid genome instability [20]. It was proposed that WHY1 forms 24-subunit homooligomers that can be destabilized by changes in the redox state of the photosynthetic apparatus. The resulting monomers are translocated to the

nucleus, where they might regulate expression of nuclear genes [29]. Beside Whirly proteins, TAC also includes other single-stranded DNA-binding proteins, OSB1 and OSB2 (pTAC9), that control recombination and provide DNA integrity and DNA stability and control recombination [30]. The pTAC12/HEMERA protein is localized to both nuclei and chloroplasts and is involved in phytochrome light signaling. pTAC13 presumably participates in the regulation of plastid elongation. It is believed that TAC contains most chloroplast nucleoid proteins [11, 20, 22].

Nucleoids contain nuclear-encoded MSH1 protein (MutS protein homolog 1) that functions in plastids and mitochondria and influences organellar genome behavior and plant growth patterns. The plastid-localized form of this protein interacts with the plastid genome and influences genome stability and plastid development, these developmental changes including altered patterns of nuclear gene expression. Disruption of the MHS1 protein leads to variegation of the plant. Variegated tissues show changes in their redox state together with enhanced plant survival and protection under photooxidative light conditions. It was suggested that the triggered plastids are an adaptive response of affected plants to naturally occurring light stress [36].

The roles of some other nucleoid proteins have been determined. Immunological analysis using specific antibodies identified DNA-binding PD1 and PD3 proteins that contain specific motifs for binding AT-rich DNA sequences (AT-hooks) and attach nucleoid to the chloroplast envelope membranes. Both PD1 and PD3 are absent from the chloroplast TAC proteome. Chloroplast nucleoid protein CND41 (molecular mass, 41 kDa) nonspecifically binds plastid DNA. CND41 is believed to be a negative transcription regulator, because its high content in various tissues causes downregulation of certain plastid mRNAs. The amino acid sequence of CND41 resembles the sequences of aspartate proteases. YLMG1 protein is involved in proper nucleoid distribution. Overexpression of *ylmg1* causes formation of an irregular network of chloroplast nucleoids. In contrast, knockout of *ylmg1* results in concentration of the nucleoids into a few large structures. The protein localizes on thylakoids and probably associates with nucleoids. Further molecular characterization of plastid nucleoid proteins will elucidate complex regulatory mechanisms of plastid gene expression that take place during plastid differentiation or under different environmental conditions [20].

STRUCTURAL ORGANIZATION AND GENE CONTENT OF PLASTID GENOMES

Understanding of plastid genome organization has been considerably advanced by the development of whole-genome sequencing methods. The first sequenced chloroplast genomes were from tobacco (1986) and Marchantia (1986); the rice chloroplast genome was sequenced later [37, 38].

Plastomes from green algae and angiosperm, seed, and land plants have been extensively studied, while data on the plastome properties from other taxa are scarcer [1]. The size of the plastid genome in land plants and photosynthetic algae varies from 120 to 190 kb [34, 39] (for example, the size of the *Arabidopsis thaliana* plastome is 154 kb [14]).

Although plastid genomes are relatively small, coding sequences comprise only 50% of land plant plastomes. The rest of the plastome is represented by introns, regulatory sequences, and intergenic spacers. The most compact plastome among photosynthetic land plants (66% of coding sequences) was found in *Welwitschia mirabilis* (family Welwitschiaceae), a plant from the gnetophyte lineage of gymnosperms [40]. In photosynthetic algae, the content of coding sequences in the plastome varies from 50% (green alga *Chlamydomonas reinhardtii*) to 93.5% (red alga *Cyanidioschyzon merolae*) [14].

Plastomes are also highly AT-rich (60-70%); overall GC content is typically 30-40%, although in some regions that do not encode proteins, AT content exceeds 80%. The proportion of GC, which is higher in the protein-coding sequences, varies across plastomes. For example, genes coding photosynthetic proteins have the highest GC content, while the NAD(P)H dehydrogenase genes have the lowest [27]. An extremely high AT content was found in the remnant plastid (apicoplast) genomes of parasitic Apicomplexa, such as *Plasmodium* and *Toxoplasma* [41].

The gene content and general structure of chloroplast genomes in land plants are very conserved [9]. The plastome of tobacco *Nicotiana tabacum* (155,939 bp) is considered an etalon of ancient plastome organization. The location of genes in this plastome is typical for angiosperms, whose plastomes have undergone no significant rearrangements during evolution.

Due to progress in whole-genome sequencing technologies, complete nucleotide sequences for hundreds of plastomes from plants of different taxa have been determined. However, most of the sequenced plastomes belong to only nine families: Asteraceae, Brassicaceae, Fabaceae, Magnoliaceae, Malvaceae, Myrtaceae, Pinaceae, Poaceae, and Theaceae [42].

The chloroplast genome (plastome) is composed of many circular and, probably identical, double-stranded DNA molecules. Each circular DNA molecule in higher plants contains 100-150 genes (<5% of the typical cyanobacterial genome) [14]; in terrestrial plants and photosynthesizing algae, the number of genes in one circular DNA molecule varies from 100 to 200 [10]. In addition to circular DNAs, plastomes can include alternative DNA forms such as multimeric circles and linear and branched DNA molecules [27, 43].

The typical plastome is quadripartite, with large single copy (LSC, ~80 kb) and small single copy (SSC, ~20 kb) regions separated by two identical inverted repeats (IRs, ~25 kb) that are required for stabilization of the plastid genome structure. In *Arabidopsis*, the LSC and SSC comprise 54 and 12% of the plastid genome, respectively. The IRs contain identical genes that are oriented in the opposite direction (i.e. the plastid genome contains two copies of IR genes). The differences in plastome sizes in higher plants are in most cases related to differences in IR length (7 to 76 kb). For example, the unusually large plastome of the geranium *Pelargonium hortorum* (217 kb) contains extremely long IRs (~76 kb) [37]. Domain organization of the plastid genome resembles the macrodomain structure of the bacterial genome; however, it remains unclear whether plastome domains are associated with specific ptNAPs, as demonstrated for bacterial genomes [29].

IRs are the most conserved elements of the plastome. In land plants, the IRs usually contain a core set of genes for four rRNAs (4.5S, 5S, 16S, 23S) and five tRNAs (trnA-UGC, trnI-GAU, trnN-GUU, trnR-ACG, trnV-GAC). In addition to this core rRNA/tRNA cluster, the IRs of many land plants, particularly vascular plants, also contain a variety of other genes. For example, in the neotropical liana *Tanaecium tetragonolobum* (family Bignoniaceae), 10 out of 86 protein-coding genes are located within the IR region and thus fully duplicated in the plastome [42]. The boundaries of IRs have shifted during evolution due to gene transfer from single copy regions to IRs and vice versa, causing considerable expansions/reductions of IRs, respectively. In closely related species, these IR boundary shifts tend to be relatively minor, resulting in the gain or loss of a small number of genes [44, 45]. However, large-scale expansions (over several kilobases) were reported for a few lineages, such as *Pelargonium*, *Psilotum*, and Trochodendraceae, which transferred numerous genes from the single copy regions into the IRs [44]. In land plants, the last full-length IR gene at the IR/SSC boundary is trnN-GUU. Excluding seed plants (i.e. angiosperms and gymnosperms), the IR generally terminates at the trnV-GAC gene at the IR/LSC boundary. However, in some representative of the genus *Lamiales* (family Bignoniaceae), the boundary between the LSC and IR regions is located within the *rps19* gene, resulting in the loss of the functional gene and formation of the *rps19* pseudogene. The boundary of the SSC/IR junction in *Lamiales* chloroplast genomes is located within the *ycf1* gene, also resulting in the formation of the *ycf1* pseudogene. The maximal length of the *rps19* pseudogene in the studied *Lamiales* species is 100 bp; the *ycf1* pseudogene varies in length from 816 to 1301 bp. Therefore, the IR expansion/contraction in *Lamiales* has led to changes in the structure of the chloroplast genome, contributing to the formation of pseudogenes [42].

The presence of the IR has a major impact on the rate of plastome sequence evolution. It was shown that in land plants the rates of synonymous and nonsynonymous substitutions are, on average, 3.7 times lower in IRs than in the single copy regions. Similar low rates of synonymous and nonsynonymous substitutions were also found for the IRs of insectivorous plants and cycads (Cycadophyta), which are the second in size group of gymnosperms after conifers. The nucleotide substitution rate in cycad IRs is approximately two times slower than in the single copy regions. It is possible that low evolutionary rate is a hallmark feature of the IR in the plastome [44].

Despite their presence in most angiosperm plastomes, IRs do not seem to be essential. One IR copy has been lost in some eudicots (at least twice in rosids and twice in asterids). The Fabaceae family includes a clade (including *Trifolium*, *Pisum*, *Cicer*, and *Medicago*) that lacks both IRs. The loss of IRs was observed in *Erodium texanum* of the Geraniaceae family. IRs are absent in plastomes of some nonphotosynthetic parasitic plants of the Orobanchaceae family (*Conopholis americana*, *Striga asiatica*, *Phelipanche ramosa*). The close relative of *Phelipanche ramosa*, *Phelipanche purpurea*, was found to have a small IR, perhaps accounting for the 1182-bp difference in the sizes of their plastomes [27]. Partly because of independent losses of one IR, the sizes of plastomes in the Orobanchaceae family vary 3.5-fold, with American squawroot (*Conopholis americana*) having the smallest plastome (45 kb) reported for land plants [41].

Among gymnosperms, species of the *Gnetales*, *Cycadales*, and *Ginkgoales* genera have retained a canonical IR ranging from 17.3 to 25.1 kb in size [46]. Plastomes of conifers from the Pinaceae family contain only one highly reduced IR copy. For example, black pine *Pinus thunbergii* and *Cedrus deodara* have exceptional small IR copy (IR_b) of only 495 and 424 bp, respectively. Highly reduced IR copies in Pinaceae plastomes include only the full trnI-CAU gene and, in most species, a portion of the *psbA* gene. Plastomes of another group of conifers, cupressophytes, have also lost one IR copy (IR_a) during evolution, but they have acquired one or more short novel IRs. In *Cephalotaxus oliveri*, the size of novel IR is 544 bp (this novel IR contains duplicated gene trnQ-UUG). In *Cryptomeria japonica* and *Taiwania cryptomerioides*, this novel IR is smaller – only 280 bp. It is believed that the independent losses of IR copies in the evolution of coniferous plants happened twice: once in Pinaceae and once in Cupressophyta [1, 46]. Sequencing of plastomes from the relict dawn redwood *Metasequoia glyptostroboides* (family Cupressaceae), sister species *Cryptomeria japonica*, and *Taiwania cryptomerioides* revealed that the *M. glyptostroboides* plastome lacks one IR copy, namely IR_a [46]. Plastomes of three green algal species of the *Chlorella* genus (*Chlorella vulgaris* C-27, *Chlorella variabilis* NC64A, *Chlorella* sp. ArMOO29B**)** have also lost one IR copy during evolution [1, 47].

Plastid genes can be divided into three functional groups: genes for photosynthetic apparatus components, genes for components of the genetic system, and genes for proteins involved in other cell processes (biosynthesis of amino acids, fatty acids, pigments, etc.) [37, 38, 48]. Typical plastid genomes of higher plants encode all types of rRNAs (23S, 16S, 5S, 4.5S), 27 to 31 tRNA genes, and a set of proteins (e.g. ~85 photosynthetic proteins) [29].

A core of about 45 genes has been found in plastomes of almost all photosynthetic organisms [49]. Plastids also contain numerous noncoding RNAs (including many antisense RNAs) [50]. Chloroplasts of dinoflagellates have a very unusual plastome: these algae have retained about a dozen genes in small (2.5-3.0 kb) plasmids termed minicircles. The retained genes are a subset of those present in other plastids; they encode core subunits of four photosynthetic membrane complexes and rRNAs. All other genes typical for plastomes have been transferred to the dinoflagellate nucleus or lost during evolution [49].

Structurally unique plastid genomes have been found in three families of angiosperm plants: Campanulaceae, Cyphiaceae, and Lobeliaceae. Plastomes in the clade that includes these families exhibit more inversions than plastomes of all angiosperms taken together and are characterized by the presence of dozens of large foreign DNA inserts containing protein-coding genes of nuclear origin. It is known that most plastome evolutionary changes result from point mutations, tandem duplications, deletions, and small-scale rearrangements, such as hairpin inversions or intron loss. Although loss of cyanobacterial ancestral genes during the evolution of plastids has been commonly accepted, acquisition of novel genes by the plastomes of a small clade was found for the first time. Detailed study of the unique evolutionary potential of Campanulaceae plastids to acquire new DNA may help bioengineers to incorporate genes into plastids of other plants [51].

Extremely reduced plastomes have been found in nonphotosynthetic organisms. In these plants, photosynthesis-associated genes have either become pseudogenes or were deleted, resulting in functional and physical reduction of the plastid genome [41]. Despite the lack of pigment, plastids of nonphotosynthetic organisms perform basic metabolic functions and, similarly to their photosynthetic counterparts, contain a genome that is composed mostly of genes involved in the expression of plastid DNA. An example of a nonphotosynthetic organism is the colorless heterotrophic flagellate protist *Astasia longa* closely related to the photoautotrophic *Euglena gracilis*. The circular plastome of *A. longa* is 73 kb, which is about half the size of the *E. gracilis* plastome (143 kb). All photosynthesis-related genes are completely absent from the *A. longa* plastome except for *rbcL* coding for the ribulose-1,5-bisphosphate carboxylase large subunit. Most of the identified genes encode components of the plastid transcriptional and translational machinery. Among these are genes for three subunits of chloroplast RNA polymerase, 20 chloroplast ribosomal protein genes, the gene for plastid Tu elongation factor, plastid tRNA genes, and three tandemly arranged repeats of 16S, 23S, and 5S rDNAs [52]. *Astasia* and *Euglena* contain several protein-coding genes that are absent in chloroplast genomes of angiosperms, but they lack some homologs of gymnosperm genes (*clpP*, *accD*, *orf1738*, *orf2216*) [53].

Many nonphotosynthetic plastid-bearing eukaryotic species are parasites, such as the malaria parasite *Plasmodium falciparum*, green alga *Helicosporidium* sp. (insect gut parasite), and its close relative *Prototheca wickerhamii*, which causes protothecosis in humans [54]. One of the most extensively studied lineages of parasitic plants is the genus *Cuscuta* (family Convolvulaceae). Whole plastome sequencing of four *Cuscuta* species demonstrated changes to both plastid gene content and structure. Two clades in the subgenus *Grammica* were characterized by considerable loss of plastid genes, including otherwise highly conserved genes for proteins of the small and large ribosomal subunits, which presumably was a consequence of the complete loss of photosynthesis. Two clades in the subgenus *Monogynella* retained much of their plastid genomes (121-125 kb), with losses being restricted primarily to the chlororespiratory (*ndh*) genes and noncoding regions, such as intergenic spacers and introns. Clades of the *Grammica* subgenus have substantially smaller plastomes (85-87 kb). In addition to losses shared with species of the subgenus *Monogynella*, they also lack the RNA polymerase gene (*rpo*) and some housekeeping genes [55]. Despite all these changes, plastomes from the four *Cuscuta* species retain many plastid genes required for photosynthesis, such as *rbcL*, *psa*, *psb*, *pet*, and *atp*. For comparison, the plastome of the nonphotosynthetic parasitic flowering plant *Epifagus virginiana* (family Orobanchaceae) contains only 42 genes, at least 38 of which are required for the expression of plastid genes [53].

It has been found during the last few years that some nonphotosynthetic species, both parasitic and free-living, have no plastid genomes. The parasitic protist *Cryptosporidium parvum* of the type Apicomplexa lacks not only the plastome, but plastids as well. It was suggested that a common ancestor of all Apicomplexa had plastids that were lost at an early evolutionary stage in the lineage leading to *C. parvum*. Interestingly, this parasite also lacks the mitochondrial genome [49]. Another species that contains no plastid genome is the nonphotosynthetic free-living fresh-water single-cell green alga *Polytomella* sp. that is closely related to the model organism *Chlamydomonas reinhardtii*. *Polytomella* sp., which acquired plastids via primary endosymbiosis, can perform metabolic functions similar to those observed in the plastids of other nonphotosynthetic algae, but it lacks genes

coding proteins required for expression, replication, and repair of plastid DNA. The genes for plastid ribosomal proteins and proteins involved in the regulation of transcription and translation in the plastids are absent as well. The genus *Polytomella* is the first example of a primary plastid-bearing lineage without a plastid genome. It is puzzling why *Polytomella* sp. lost its plastome when so many other heterotrophic lineages have kept it. It is possible that *Polytomella* sp. has a propensity for organelle genome reduction, since its mitochondrial genome is the smallest and has the most reduced gene content among all green plants [54].

Some parasitic plants do not have plastomes. About 1% of all known angiosperm species are parasitic plants. Plant parasitism is an evolutionary adaptation arising independently at least 12-13 times in flowering plants. The genus *Rafflesia* (family Rafflesiaceae, order Malpighiales) that is widespread in the Philippines is one of eight known genera of plant obligate parasites. No substantial portions of plastid genomes were found in *Rafflesia lagascae*, which is endemic to the Philippines, except short chloroplast genome fragments including photosynthesis and energy metabolism genes (*atp*, *ndh*, *pet*, *psa*, *psb*, *rbcl*), rRNA genes (*rrn16*, *rrn23*), ribosomal protein genes (*rps7*, *rps11*, *rps16*), tRNA genes, as well as *matK*, *accD*, *ycf2*, and multiple nongenic regions from the IRs. *Rafflesia lagascae* parasitizes the tropical vine *Tetrastigma* (family Vitaceae). It is possible that identified fragments of *R. lagascae* chloroplast genome are located in the *Tetrastigma* (host) nuclear genome similarly to plastid gene inserts (nupDNAs) in nuclear genomes of other plant species. Although all attempts to find an intact plastid genome in *R. lagascae* have failed, this plant was found to contain a large mitochondrial genome that has been partially sequenced. The plastid genome is supposedly absent in another *Rafflesia* species, *R. leonardi*. Parasitic flowering species of the *Rafflesia* genus are the first known group of primary plastid-bearing plants that lost their plastid genome [56].

Most data on plastid genomes and genes have been obtained in seed plants. Information on the properties of plastomes in other taxonomic groups is scarce.

PLASTOMES OF FERNS AND BRYOPHYTES. RNA EDITING

Ferns (Monilophyta) are an ancient lineage of land plants that have been an important component of the Earth's flora. It is the second in size group of vascular plants that includes over 11,000 species and occupies diverse ecological niches (especially, in tropical zones). Due to their unique phylogenetic position, ferns are very important for elucidating the evolution of land plants (Embryophyta) and understanding the origin and diversification of traits found in seed and model plants (e.g. rice and *Arabidopsis*) [57, 58]. Table 1 compares some fern and seed plant plastomes. The number of genes in plastomes varies (99 to 122) in the plants shown in Table 1. This difference between relatively small gene sets of organellar genomes might be a result of the evolutionary loss or transfer of the endosymbiont genes to the nuclear genome. Unfortunately, only few fern plastomes have been completely determined, compared to hundreds of fully sequenced plastomes of seed plants. It was found that unlike seed plants, whose plastomes are structurally conserved, most modern fern species (~90%) have considerably reorganized plastomes. The plastome of the eusporangiate fern *Angiopteris* (family Marattiaceae) retained an ancient gene order that is almost identical to that in the etalon seed plant *Nicotiana tabacum*, even though their lineages diverged in the Devonian more than 350 million years ago [57]. The unusual plastome organization in most modern ferns is a result of a series of genomic inversions. Plastomes of the studied fern species contain five inversions. One inversion (~3.3 kb) is present in the LSCs in all fern species. Two large (18 and 21 kb) and overlapping inversions are in the IR, with rRNA

Table 1. Characteristics of plastomes from fern and seed plant species [44]

Characteristics	Ferns		Gymnosperms			Angiosperms		
	Ehye	Aang	Gbil	Ggne	Pstr	Agra	Pmar	Pmed
Genome size, bp	131,760	153,596	157,002	115,022	115,576	152,849	158,358	164,130
IR, bp	10,093	21,676	17,733	20,051	472	25,822	33,735	38,398
SC, bp	111,574	110,244	121,536	74,920	114,632	101,205	90,888	87,334
GC content, %	33.7	35.5	39.6	38.2	38.8	38.7	38.6	38
Number of unique genes	120	122	118	99	108	112	113	113
Protein-coding genes in IRs	0	3	3	3	0	6	11	15
rRNA genes in IRs	4	4	4	4	0	4	4	4
tRNA genes in IRs	5	8	6	8	1	7	7	7
Repeats (not including IRs), %	1.8	1.6	1.2	1.4	4.3	0.47	2.5	2.7

Note: Ehye, *Equisetum hyemale*; Aang, *Angiopteris angustifolia*; Gbil, *Ginkgo biloba*; Ggne, *Gnetum gnemon*; Pstr, *Pinus strobus*; Agra, *Acorus gramineus*; Pmar, *Plantago maritima*; Pmed, *Plantago media*.

genes occurring in the reverse order compared to all other plants. The LSCs also contain two small overlapping inversions. Ferns and seed plants differ in the rate of synonymous and nonsynonymous substitutions in plastomes – in most cases the substitution rate is higher in ferns. Although the rate of nucleotide substitutions is an important indication of the gene evolutionary rate, the conclusion that fern plastid genes evolve at a higher rate than plastid genes in seed plants needs to be confirmed by sequencing more fern plastomes. Fern plastomes are also characterized by active RNA editing (posttranscriptional mRNA modification).

RNA editing is a significant process in gene expression in angiosperm organelles, where it might occur in the form of C-to-U editing, i.e. substitution of cytidine (C) by uridine (U) in mRNA, or in rare cases, U-to-C editing [59, 60]. No examples of RNA editing have been found in transcripts of algal plastid genes and cyanobacterial genes. The number of RNA editing sites in plastomes of angiosperm plants is low and, according to various sources, varies from 25 to 30 [59] or from 20 to 37 [9]. Most seed plant plastomes have about 30-40 RNA editing sites [57]; the number of RNA editing sites in mitochondrial genomes is much higher (>1000). The plastome of the homosporous bracken fern *Pteridium aquilinum* contains over 1000 putative RNA editing sites, the highest number yet detected for a plastome. Unlike angiosperm plastomes, many U-to-C RNA editing events were detected in addition to C-to-U RNA editing events in plastid genomes of ferns and bryophytes (mosses, liverworts, and hornworts). RNA editing in organelles has been lost during evolution in some liverworts of the order Marchantiales (e.g. *Marchantia polymorpha*) [59, 60]. Table 2 shows rates and types of plastome RNA editing in representatives from three main lineages: ferns, hornworts, and seed plants. As seen from Table 2, RNA editing occurs in all types of plastid genes, including protein-coding, tRNA, and rRNA genes, as well as intergenic and intron regions. Most RNA editing sites occur in protein-coding sequences, with some RNA editing events modifying the start or stop codons of the resulting mRNA sequences. It appears that RNA editing rates are the highest in hornworts. Fern plastomes contain fewer RNA editing sites than liverworts, but more than seed plants. However, the rates of RNA editing might considerably vary within each of the clades. Studies of RNA editing led to the suggestion that it first appeared in bryophytes, which are generally believed to be the earliest land plants. Bryophyte plastomes are very conserved and resemble most land plant plastomes. Structural alterations in moss plastomes, such as a 71-kb inversion in the LSC of mosses from the order *Funariales* (*Physcomitrella patens*) or loss of the *rpoA* gene in arthrodont mosses, are extremely rare. Plastomes of some moss species (*Takakia lepidozioides*, *Sphagnum palustre*, and *Andreaea nivalis*) have been studied for their properties, mostly for the presence of pseudogenes and gene loss. It was found that plastomes of two peristomate mosses, *Tortula ruralis* and *Tetraphis pellucida*, have lost the *petN* gene, which makes these two species the only land plants that lack this gene in their chloroplasts. Overall GC content is moss plastomes is 29.4%, which is similar to other known bryophyte chloroplast genomes (28-33%) and significantly less than the 34-40% found in seed plants. Despite identical or near identical gene content, the *Tetraphis pellucida* chloroplast genome (127,489 bp) is approximately 5 kb longer than those of *Tortula ruralis* (122,530 bp) and *Physcomitrella patens* (122,890 bp), the difference being accounted for by an increased total length of intragenic spacer regions in the LSC [64]. Gene loss (e.g. *cysA* and *cysT*) was also observed in plastomes of liverwort species [65].

The levels of organellar RNA editing differ in moss clades: the number of RNA editing sites is higher in early diverging clades than in later diverging clades, with numbers of both chloroplast and mitochondrial sites being higher in peristomate mosses than in arthrodont mosses and numbers of chloroplast sites very much lower than those found in the early diverging moss lineage *Takakia*,

Table 2. Number of RNA editing sites in plastomes of plants from different taxonomic groups [57]

Parameter	Fern *Adiantum* [61]	Hornwort *Anthoceros* [62]	Seed plant *Arabidopsis* [63]
Plastome size, bp	150,568	161,162	154,478
Number of RNA editing sites	350	979	34
Number of C-to-U substitutions	311	533	34
Number of U-to-C substitutions	39	446	0
Number of RNA editing sites in:			
protein-coding genes	344	972	32
tRNA genes	1	1	0
rRNA genes	0	0	0
introns	0	3	1

Note: Plastomes of all three species contain RNA editing sites in start codons. Plastomes of *Adiantum* and *Anthoceros* (but not of *Arabidopsis*) contain RNA editing sites in stop codons.

as well as in hornworts. The chloroplast genome of *Takakia lepidozioides* contains 302 putative RNA editing sites, while chloroplast genomes of *Physcomitrella patens*, *Tortula ruralis*, and *Tetraphis pellucida* have only 2, 3, and 15 sites, respectively [64].

The chloroplast genome of the hornwort *Anthoceros angustus* (earlier erroneously called *Anthoceros formosae*) is one of the largest among land plants (161,162 bp). It consists of double-stranded circular DNA molecules and contains genes for 76 proteins, 32 tRNAs, 4 rRNAs, and 10 open reading frames (ORFs) [66]. The *Anthoceros* plastome is larger than any bryophyte plastome (161,162 vs. 108,007-123,500 bp, respectively) due mostly to the larger IR that contains duplicated *ndhB* and *rps7* genes and the 3′-exon of the *rps12* gene. The genes *matK* and *rps15*, commonly found in the chloroplast genomes of land plants, are pseudogenes. Another feature of the *Anthoceros* plastome is the presence of a type I intron in the 23S rRNA gene (*rrn23*) that is highly similar (~50%) to introns of the green algae *Chlorella* and *Chlamydomonas* [66]. More than half of the *Anthoceros* protein-coding genes have nonsense codons (UGA, UAA, UAG), which are converted into sense codons (CGA, CAA, CAG) by RNA editing. The plastome of the hornwort *Nothoceros aenigmaticus* is structurally similar to those of most bryophytes, but it differs from the *Anthoceros* plastome by several gene regions located within the IR in *Anthoceros* being in the LSC in *Nothoceros*, by the *rrn23* gene lacking an intron, and by *rpl2* being a pseudogene. Structural differences between hornwort plastomes are more significant than between plastomes of mosses or liverworts. However, it should be noted that by 2014 only a few bryophyte plastomes had been fully sequenced (two liverworts, three mosses, and two hornworts) [64, 67].

In the last 20 years, chloroplast genomes (plastomes) have been widely used for genetic engineering of plants. The plastome is an ideal "platform" for biosynthesis of target proteins in large amounts. The method involves plastome transformation by incorporation of vectors bearing foreign DNA genes resulting in the generation of cells and plants with transgenic plastid genomes (transplastomic plants) (for detailed description of plastome transformation methods and vectors, see [66-75]). Stable chloroplast transformation has been performed in the single cell green alga *Chlamydomonas reinhardtii*, tobacco *Nicotiana tabacum*, *Arabidopsis thaliana*, tomato *Lycopersicon esculentum*, potato *Solanum tuberosum*, eggplant *Solanum melongena*, rapeseed *Brassica napus*, lettuce *Lactuca sativa*, soybean *Glycine max*, cabbage *Brassica oleracea*, carrot *Daucus carota*, and some other plants [71, 72]. Several methods for chloroplast transformation have been developed that could be used in fundamental and applied studies [76]. However, stable plastid transformation is not a routine procedure yet. Despite that fact that over 800 plastomes have been sequenced, transplastome approaches have been successfully used for only a relatively small number of plants [75]. Thus, all the attempts to transform plastomes of monocotyledons, some of which are important food crops, have failed. Plastid genome transformation has also been inefficient in non-green tissues containing different types of plastids (chromoplasts, amyloplasts, leucoplasts). Since chloroplasts are absent in fruits, tubers, and seeds, expression of photosynthesis-associated genes in these organelles is suppressed [77].

Plant cells have three genomes: nuclear, plastid, and mitochondrial. Both nuclear (~30,000 genes) and chloroplast (~120 genes) genomes can be genetically modified; however, no successful transformation of the mitochondrial genome (~60 genes) has been achieved. Transgene expression in the plastid genome poses certain advantages over nuclear expression. The first and the major advantage is many plastids and therefore, plastid genome copies (~100) in a cell, which would allow achieving high levels of foreign gene expression (tens to hundreds of times higher than when expressed in the nuclear genome) [78]. Theoretically, if all chloroplasts are transformed in each cell of every leaf in a plant, this makes very high expression of a foreign protein possible. In fact, very high expression has been achieved, reaching up to 72% of total leaf protein and 70% of total soluble protein. Nuclear transformation, which usually results in the incorporation of 1-2 transgene copies into the nuclear genome, provides considerably lower expression levels (on average, 0.01-0.4% of total soluble protein) [79]. Second, transgene insertion into the plastid genome occurs by homologous recombination (i.e. the insert position in all the resulting transformants is the same), which allows choice of the insertion site for the optimal gene expression and avoids undesirable positional effects. This makes plastid genome transformation a high-precision technique of plant genetic engineering [71, 78]. Nuclear genome transformation is usually a result of nonhomologous recombination, which means that transformants differ in the sites of transgene incorporation. Third, due to the endosymbiotic origin of plastids, their genomes have no silent genes and other epigenetic mechanisms that would prevent stable transgene expression. The nuclear genome is regulated epigenetically, which affects the yield of expressed foreign proteins. In general, plastid transformation is more stable and predictable than transformation of the nuclear genome. The fourth advantage is that similarly to bacterial genes, most plastid genes are organized in operons. This allows grouping multiple transgenes into artificial operons that could be expressed under control of the same promoter as polycistronic transcripts that would then be processed into monocistronic mRNAs. The last advantage is the maternal type of plastid inheritance in most species of angiosperms, which considerably decreases the probability of transgene loss and excludes the possibility of crossbreeding between transplastomic

plants and any wild-growing species, thereby providing biosafety of plant genetic engineering [75, 76, 78-81].

Genetic engineering of plastid genomes in algae and land plants (Embryophyta) attests to enormous potential of chloroplast transformation in key areas of biotechnology [76]. Chloroplast transformation is widely used for the integration of genes into plastid genomes to confer herbicide and pest resistance, express recombinant proteins (e.g. vaccines), enhance agronomic traits of agricultural plants (e.g. by photosynthesis modification), and study cell signaling mechanisms and metabolic processes [77]. For example, expression of antioxidant enzymes could increase plant tolerance to abiotic stress [77]. Plastid transformation with the *dao* gene coding D-amino acid oxidase from *Schizosaccharomyces pombe* provides plant tolerance against potential D-alanine-based herbicides and allows the use of inexpensive and widely available D-amino acids, which are relatively nontoxic to animals and microbes [81]. During the last five years, many therapeutic proteins (biopharmaceuticals) have been expressed in transgenic plastids, some of them (e.g. bacteriophage endolysins) in very large amounts. Considerable progress has been achieved in the expression of antibody fragments, blood coagulation factors for hemophilia treatment, and transforming growth factor β (TGF-β), a cytokine protein that promotes wound healing and diminishes scarring. Recombinant glucocerebrosidase synthesized in carrot cells has been approved for replacement therapy in Gaucher disease (hereditary glucocerebrosidosis) [75, 79]. Of particular interest is the use of plastid genomes for antibiotic and vaccine production. The rapid rise in the resistance of pathogenic bacteria to most traditional antibiotics generated the need for a new generation of antibiotics, some of which could be of bacteriophage origin. Hydrolytic enzymes of bacteriophages (endolysins) cleave peptidoglycans of bacterial cell wall. When expressed in plasmids, endolysins will be stable in these organelles, because the plasmid proteolytic apparatus strongly resembles that of eubacteria, to which bacteriophages have developed resistance. Antimicrobial peptides (AMPs) are another type of new generation antibiotics that have significant advantages over traditional antibiotics. AMPs display antimicrobial activity toward a broad range of known pathogens; they are efficient against many microbes resistant to currently used drugs. Moreover, pathogens develop no resistance against AMPs. The major obstacle for the wide clinical application and marketing of AMPs is their high cost. However, if AMP genes are introduced into chloroplast genomes, the resulting transplastomic plants could be used as "green biofactories" for the production of AMP-based antibiotics against human pathogens. In addition, chloroplast transformation with the AMP genes might enhance plant resistance as well. Two AMPs, retrocyclin 101 and protegrin 1, have been expressed in chloroplasts; both AMPs were active against bacterial and viral pathogens [75]. Many infectious diseases might be prevented and cured with vaccines. Since traditional methods of vaccine production are costly, and plastid biotechnology is a promising approach for creation of inexpensive, stable, safe, and efficient vaccines in large quantities. Tobacco chloroplasts have been used to express antigens for the generation of vaccines against widespread dangerous infectious human diseases such as tuberculosis, anthrax, human papilloma virus, poliomyelitis, plague, cholera, and malaria [75, 77]. Plant-produced vaccine against Ebola virus has been successfully used for the treatment of infected patients in West Africa. More than 40 biopharmaceutics and vaccine antigens have been expressed in chloroplast genomes [77]. Recently achieved superexpression of biopharmaceuticals in plastid genomes in leaves of edible plants promoted the development of methods for their oral delivery and considerably decreased (by ~90%) production costs, which is extremely important in developing countries. The efficiency of oral delivery of biopharmaceuticals has been proven in the treatment of Alzheimer's disease, diabetes mellitus, pulmonary hypertension, hemophilia, and retinopathy [75]. Lettuce is now the only consistent reproducible transplastomic system for the production of biopharmaceuticals and vaccines for oral administration. Beside other advantages, plant-produced vaccines are bioencapsulated and, therefore, protected from gastric juices. Only after entering the intestine, plant cell walls are broken by the gut microflora, and released antigens become available for the immune system [75]. The first vaccine antigen against human disease was obtained in tobacco *Nicotiana tabacum* chloroplasts over 15 years ago. Although the methods for vaccine antigen production have significantly improved since then and now can achieve very high protein expression levels, chloroplast expression is still performed in tobacco plants in 80-90% cases. Studies of chloroplast expression in other plant species are relatively few. High levels of protein expression and large biomass make tobacco an ideal plant for production of large quantities of vaccines. Lettuce is emerging as an alternative system for vaccine production due to the large biomass and broad leaves. The efficiency and high immunogenicity of plant vaccines have been proven in many animal models; however, the safety and immunogenicity of these vaccines in humans remain to be studied. No plant-produced vaccine against human disease has entered the market yet [79].

Production of enzymes for industrial use is a new and rapidly developing field of chloroplast biotechnology. Recent studies showed that most enzymes required for cellulose biomass conversion into fermented sugars (cellulases, xylanases, glucosidases, pectate lyases, and cutinases) could be expressed in the plastid genome. What is especially important is that since the sources for some of the enzyme genes were thermophilic organisms, the encoded proteins are thermostable and can be easily used

for industrial treatment of cellulose biomass. Interest in transplastomic plants as biofactories for the production of "green chemicals", i.e. raw materials and building blocks for chemical industry, in on the rise. There is reason to hope that further development of plant biotechnology will solve many significant problems of human society.

Acknowledgments

This work was supported by the Molecular and Cell Biology Program of the Presidium of the Russian Academy of Sciences and by the Russian Foundation for Basic Research (project No. 16-04-01626).

REFERENCES

1. Xu, J. H., Liu, Q., Hu, W., Wang, T., Xue, Q., and Messing, J. (2015) Dynamics of chloroplast genomes in green plants, *Genomics*, **106**, 221-231.
2. Jensen, P. E., and Leister, D. (2014) Chloroplast evolution, structure and functions, *F1000Prime Rep*., **6**, 40.
3. Howe, C. J., Barbrook, A. C., Nisbet, R. E., Lockhart, P. J., and Larkum, A. W. (2008) The origin of plastids, *Phil. Trans. R. Soc. B*, **363**, 2675-2685.
4. Green, B. R. (2011) Chloroplast genomes of photosynthetic eukaryotes, *Plant J*., **66**, 34-44.
5. Ruck, E. C., Nakov, T., Jansen, R. K., Theriot, E. C., and Alverson, A. J. (2014) Serial gene losses and foreign DNA underlie size and sequence variation in the plastid genomes of diatoms, *Gen. Biol. Evol*., **6**, 644-654.
6. Barbrook, A. C., Howe, C. J., Kurniawan, D. P., and Tarr, S. J. (2010) Organization and expression of organellar genomes, *Phil. Trans. R. Soc. B*, **365**, 785-797.
7. Curtis, B. A., Tanifuji, G., Burki, F., Gruber, A., Irimia, M., Maruyama, S., Arias, M. C., Ball, S. G., Gile, G. H., Hirakawa, Y., Hopkins, J. F., Kuo, A., Rensing, S. A., Schmutz, J., Symeonidi, A., Elias, M., Eveleigh, R. J., Herman, E. K., Klute, M. J., Nakayama, T., Obornik, M., Reyes-Prieto, A., Armbrust, E. V., Aves, S. J., Beiko, R. G., Coutinho, P., Dacks, J. B., Durnford, D. G., Fast, N. M., Green, B. R., Grisdale, C. J., Hempel, F., Henrissat, B., Höppner, M. P., Ishida, K., Kim, E., Koreny, L., Kroth, P. G., Liu, Y., Malik, S. B., Maier, U. G., McRose, D., Mock, T., Neilson, J. A., Onodera, N. T., Poole, A. M., Pritham, E. J., Richards, T. A., Rocap, G., Roy, S. W., Sarai, C., Schaack, S., Shirato, S., Slamovits, C. H., Spencer, D. F., Suzuki, S., Worden, A. Z., Zauner, S., Barry, K., Bell, C., Bharti, A. K., Crow, J. A., Grimwood, J., Kramer, R., Lindquist, E., Lucas, S., Salamov, A., McFadden, G. I., Lane, C. E., Keeling, P. J., Gray, M. W., Grigoriev, I. V., and Archibald, J. M. (2012) Algal genomes reveal evolutionary mosaicism and the fate of nucleomorphs, *Nature*, **492**, 59-65.
8. Lloyd, A. H., and Timmis, J. N. (2011) The origin and characterization of new nuclear genes originating from a cytoplasmic organellar genome, *Mol. Biol. Evol*., **28**, 2019-2028.
9. Raman, G., and Park, S. (2015) Analysis of the complete chloroplast genome of a medicinal plant, *Dianthus superbus* var. *longicalyncinus*, from a comparative genomics perspective, *PLoS One*, **10**, e0141329.
10. Rochaix, J.-D., and Ramundo, S. (2015) Conditional repression of essential chloroplast genes: evidence for new plastid signaling pathways, *Biochim. Biophys. Acta*, **1847**, 986-992.
11. Yagi, Y., and Shiina, T. (2014) Recent advances in the study of chloroplast gene expression and its evolution, *Front. Plant Sci*., **5**, 61.
12. Leister, D. (2003) Chloroplast research in the genomic age, *Trends Genet*., **19**, 47-57.
13. Woodson, J. D., and Chory, J. (2012) Organelle signaling: how stressed chloroplasts communicate with the nucleus, *Curr. Biol*., **22**, R690-692.
14. Ortelt, J., and Link, G. (2014) Plastid gene transcription: promoters and RNA polymerases, *Methods Mol. Biol*., **1132**, 47-72.
15. Golczyk, H., Greiner, S., Wanner, G., Weihe, A., Bock, R., Borner, T., and Herrmann, R. G. (2014) Chloroplast DNA in mature and senescing leaves: a reappraisal, *Plant Cell*, **26**, 847-854.
16. Chen, H., Yu, Y., Chen, X., Zhang, Z., Gong, C., Li, J., and Wang, A. (2015) Plastid DNA insertions in plant nuclear genomes: the sites, abundance and ages, and a predicted promoter analysis, *Funct. Integr. Genomics*, **15**, 131-139.
17. Cullis, C. A., Vorster, B. J., Van Der Vyver, C., and Kunert, K. J. (2009) Transfer of genetic material between the chloroplast and nucleus: how is it related to stress in plants? *Ann. Bot*., **103**, 625-633.
18. Hsu, C. Y., Wu, C. S., and Chaw, S. M. (2014) Ancient nuclear plastid DNA in the yew family (Taxaceae), *Genome Biol. Evol*., **6**, 2111-2121.
19. Wang, D., Qu, Z., Adelson, D. L., Zhu, J. K., and Timmis, J. N. (2014) Transcription of nuclear organellar DNA in a model plant system, *Genome Biol. Evol*., **6**, 1327-1334.
20. Krupinska, K., Melonek, J., and Krause, K. (2013) New insights into plastid nucleoid structure and functionality, *Planta*, **237**, 653-664.
21. Pfalz, J., and Pfannschmidt, T. (2015) Plastid nucleoids: evolutionary reconstruction of a DNA/protein structure with prokaryotic ancestry, *Front. Plant Sci*., **6**, 220.
22. Yagi, Y., and Shiina, T. (2012) Evolutionary aspects of plastid proteins involved in transcription: the transcription of a tiny genome is mediated by a complicated machinery, *Transcription*, **3**, 290-294.
23. Melonek, J., Mulisch, M., Schmitz-Linneweber, C., Grabowski, E., Hensel, G., and Krupinska, K. (2010) Whirly1 in chloroplasts associates with intron containing RNAs and rarely co-localizes with nucleoids, *Planta*, **232**, 471-481.
24. Powikrowska, M., Oetke, S., Jensen, P. E., and Krupinska, K. (2014) Dynamic composition, shaping and organization of plastid nucleoids, *Front. Plant Sci*., **5**, 424.
25. Bohne, A. V. (2014) The nucleoid as a site of rRNA processing and ribosome assembly, *Front. Plant Sci*., **5**, 257.
26. Moriyama, T., and Sato, N. (2014) Enzymes involved in organellar DNA replication in photosynthetic eukaryotes, *Front. Plant Sci*., **5**, 480.
27. Ruhlman, T. A., and Jansen, R. K. (2014) The plastid genomes of flowering plants, *Methods Mol. Biol*., **1132**, 3-38.

28. Pfalz, J., and Pfannschmidt, T. (2013) Essential nucleoid proteins in early chloroplast development, *Trends Plant Sci.*, **18**, 186-194.
29. Powikrowska, M., Khrouchtchova, A., Martens, H. J., Zygadlo-Nielsen, A., Melonek, J., Schulz, A., Krupinska, K., Rodermel, S., and Jensen, P. E. (2014) SVR4 (suppressor of variegation 4) and SVR4-like: two proteins with a role in proper organization of the chloroplast genetic machinery, *Physiol. Plant.*, **150**, 477-492.
30. Majeran, W., Friso, G., Asakura, Y., Qu, X., Huang, M., Ponnala, L., Watkins, K. P., Barkan, A., and Van Wijk, K. J. (2012) Nucleoid-enriched proteomes in developing plastids and chloroplasts from maize leaves: a new conceptual framework for nucleoid functions, *Plant Physiol.*, **158**, 156-189.
31. Kodama, Y. (2007) Plastidic proteins containing motifs of nuclear transcription factors, *Plant Biotechnol.*, **24**, 165-170.
32. Melonek, J., Matros, A., Trosch, M., Mock, H. P., and Krupinska, K. (2012) The core of chloroplast nucleoids contains architectural SWIB domain proteins, *Plant Cell*, **24**, 3060-3073.
33. Grabowski, E., Miao, Y., Mulisch, M., and Krupinska, K. (2008) Single-stranded DNA-binding protein Whirly1 in barley leaves is located in plastids and the nucleus of the same cell, *Plant Physiol.*, **147**, 1800-1804.
34. Yu, Q. B., Huang, C., and Yang, Z. N. (2014) Nuclear-encoded factors associated with the chloroplast transcription machinery of higher plants, *Front. Plant Sci.*, **5**, 316.
35. Pfalz, J., Liere, K., Kandlbinder, A., Dietz, K. J., and Oelmuller, R. (2006) pTAC2, -6, and -12 are components of the transcriptionally active plastid chromosome that are required for plastid gene expression, *Plant Cell*, **18**, 176-197.
36. Xu, Y. Z., Arrieta-Montiel, M. P., Virdi, K. S., De Paula, W. B., Widhalm, J. R., Basset, G. J., Davila, J. I., Elthon, T. E., Elowsky, C. G., Sato, S. J., Clemente, T. E., and Mackenzie, S. A. (2011) MutS homolog1 is a nucleoid protein that alters mitochondrial and plastid properties and plant response to high light, *Plant Cell*, **23**, 3428-3441.
37. Odintsova, M. S., and Yurina, N. P. (2003) Plastidic genome of higher plants and algae: structure and function, *Mol. Biol. (Moscow)*, **37**, 1-16.
38. Odintsova, M. S., and Yurina, N. P. (2005) Genomics and evolution of cell organelles, *Genetika*, **41**, 1-12.
39. Wicke, S., Schneeweiss, G. M., De Pamphilis, C. W., Muller, K. F., and Quandt, D. (2011) The evolution of the plastid chromosome in land plants: gene content, gene order, gene function, *Plant Mol. Biol.*, **76**, 273-297.
40. McCoy, S. R., Kuehl, J. V., Boore, J. L., and Raubeson, L. A. (2008) The complete plastid genome sequence of *Welwitschia mirabilis*: an unusually compact plastome with accelerated divergence rates, *BMC Evol. Biol.*, **8**, 130.
41. Wicke, S., Muller, K. F., De Pamphilis, C. W., Quandt, D., Wickett, N. J., Zhang, Y., Renner, S. S., and Schneeweiss, G. M. (2013) Mechanisms of functional and physical genome reduction in photosynthetic and nonphotosynthetic parasitic plants of the broomrape family, *Plant Cell*, **25**, 3711-3725.
42. Nazareno, A. G., Carlsen, M., and Lohmann, L. G. (2015) Complete chloroplast genome of *Tanaecium tetragonolobum*: the first Bignoniaceae plastome, *PLoS One*, **10**, 129930.
43. Oldenburg, D. J., and Bendich, A. J. (2015) DNA maintenance in plastids and mitochondria of plants, *Front. Plant Sci.*, **6**, 883.
44. Zhu, A., Guo, W., Gupta, S., Fan, W., and Mower, J. P. (2016) Evolutionary dynamics of the plastid inverted repeat: the effects of expansion, contraction, and loss on substitution rates, *New Phytol.*, **209**, 1747-1756.
45. Wu, C. S., and Chaw, S. M. (2015) Evolutionary stasis in cycad plastomes and the first case of plastome GC-biased gene conversion, *Genome Biol. Evol.*, **7**, 2000-2009.
46. Guo, W., Grewe, F., Cobo-Clark, A., Fan, W., Duan. Z., Adams, R. P., Schwarzbach, A. E., and Mower, J. P. (2014) Predominant and substoichiometric isomers of the plastid genome coexist within *Juniperus* plants and have shifted multiple times during cupressophyte evolution, *Genome Biol. Evol.*, **6**, 580-590.
47. Jeong, H., Lim, J. M., Park, J., Sim, Y. M., Choi, H. G., Lee, J., and Jeong, W. J. (2014) Plastid and mitochondrion genomic sequences from Arctic *Chlorella* sp. ArM0029B, *BMC Genomics*, **15**, 286-300.
48. Tiller, N., and Bock, R. (2014) The translational apparatus of plastids and its role in plant development, *Mol. Plant*, **7**, 1105-1120.
49. Barbrook, A. C., Howe, C. J., and Purton, S. (2006) Why are plastid genomes retained in non-photosynthetic organisms? *Trends Plant Sci.*, **11**, 101-108.
50. Borner, T., Aleynikova, A. Y., Zubo, Y. O., and Kusnetsov, V. V. (2015) Chloroplast RNA polymerases: role in chloroplast biogenesis, *Biochim. Biophys. Acta*, **1847**, 761-769.
51. Knox, E. B. (2014) The dynamic history of plastid genomes in the *Campanulaceae sensu lato* is unique among angiosperms, *Proc. Natl. Acad. Sci. USA*, **111**, 11097-11102.
52. Gockel, G., and Hachtel, W. (2000) Complete gene map of the plastid genome of the nonphotosynthetic euglenoid flagellate *Astasia longa*, *Protist*, **151**, 347-351.
53. Wolfe, K. H., Morden, C. W., and Palmer, J. D. (1992) Function and evolution of a minimal plastid genome from a nonphotosynthetic parasitic plant, *Proc. Natl. Acad. Sci. USA*, **89**, 10648-10652.
54. Smith, D. R., and Lee, R. W. (2014) A plastid without a genome: evidence from the nonphotosynthetic green algal genus *Polytomella*, *Plant Physiol.*, **164**, 1812-1819.
55. Braukmann, T., Kuzmina, M., and Stefanovic, S. (2013) Plastid genome evolution across the genus *Cuscuta* (Convolvulaceae): two clades within subgenus *Grammica* exhibit extensive gene loss, *J. Exp. Bot.*, **64**, 977-989.
56. Molina, J., Hazzouri, K. M., Nickrent, D., Geisler, M., Meyer, R. S., Pentony, M. M., Flowers, J. M., Pelser, P., Barselona, J., Inovejas, S. A., Uy, I., Yuan, W., Wilkins, O., Michel, C. I., Locklear, S., and Purugganan, M. D., Concepcion, G. P., and Purugganan, M. D. (2014) Possible loss of the chloroplast genome in the parasitic flowering plant *Rafflesia lagascae* (Rafflesiaceae), *Mol. Biol. Evol.*, **31**, 793-803.
57. Wolf, P. G., Der, J. P., Duffy, A. M., Davidson, J. B., Grusz, A. L., and Pryer, K. M. (2011) The evolution of chloroplast genes and genomes in ferns, *Plant Mol. Biol.*, **76**, 251-261.
58. Sessa, E. B., Banks, J. A., Barker, M. S., Der, J. P., Duffy, A. M., Graham, S. W., Hasebe, M., Langdale, J., Li, F. W., Marchant, D. B., Pryer, K. M., Rothfels, C. J., Roux, S. J.,

Salmi, M. L., Sigel, E. M., Soltis, D. E., Soltis, P. S., Stevenson, D. W., and Wolf, P. G. (2014) Between two fern genomes, *Giga Sci.*, **3**, 15.
59. Odintsova, M. S., and Yurina, N. P. (2000) RNA editing in plant chloroplasts and mitochondria, *Fiziol. Rast.*, **47**, 307-320.
60. Yurina, N. P., and Odintsova, M. S. (2016) Mitochondrial genome structure of photosynthetic eukaryotes, *Biochemistry (Moscow)*, **81**, 101-113.
61. Wolf, P. G., Rowe, C. A., and Hasebe, M. (2004) High levels of RNA editing in a vascular plant chloroplast genome: analysis of transcripts from the fern *Adiantum capillus-veneris*, *Gene*, **339**, 89-97.
62. Kugita, M., Yamamoto, Y., Fujikawa, T., Matsumoto, T., and Yoshinaga, K. (2003) RNA editing in hornwort chloroplasts makes more than half the genes functional, *Nucleic Acids Res.*, **31**, 2417-2423.
63. Chateigner-Boutin, A. L., and Small, I. (2007) A rapid high-throughput method for the detection and quantification of RNA editing based on high-resolution melting of amplicons, *Nucleic Acids Res.*, **35**, 114.
64. Bell, N. E., Boore, J. L., Mishler, B. D., and Hyvonen, J. (2014) Organellar genomes of the four-toothed moss, *Tetraphis pellucida*, *BMC Genomics*, **15**, 383-394.
65. Wickett, N. J., Forrest, L. L., Budke, J. M., Shaw, B., and Goffinet, B. (2011) Frequent pseudogenization and loss of the plastid-encoded sulfate-transport gene *cysA* throughout the evolution of liverworts, *Am. J. Bot.*, **98**, 1263-1275.
66. Kugita, M., Kaneko, A., Yamamoto, Y., Takeya, Y., Matsumoto, T., and Yoshinaga, K. (2003) The complete nucleotide sequence of the hornwort (*Anthoceros formosae*) chloroplast genome: insight into the earliest land plants, *Nucleic Acids Res.*, **31**, 716-721.
67. Villarreal, J. C., Forrest, L. L., Wickett, N., and Goffinet, B. (2013) The plastid genome of the hornwort *Nothoceros aenigmaticus* (Dendrocerotaceae): phylogenetic signal in inverted repeat expansion, pseudogenization, and intron gain, *Am. J. Bot.*, **100**, 467-477.
68. Maliga, P. (2004) Plastid transformation in higher plants, *Annu. Rev. Plant Biol.*, **55**, 289-313.
69. Lutz, K. A., Azhagiri, A., Tungsuchat-Huang, T., and Maliga, P. (2007) A guide to choosing vectors for transformation of the plastid genome of higher plants, *Plant Physiol.*, **145**, 1201-1210.
70. Lutz, K. A., and Maliga, P. (2007) Construction of marker-free transplastomic plants, *Curr. Opin. Biotechnol.*, **18**, 107-114.
71. Maliga, P., and Bock, R. (2011) Plastid biotechnology: food, fuel and medicine for the 21st century, *Plant Physiol.*, **155**, 1501-1510.
72. Shchelkunov, S. N., Konstantinov, Yu. M., and Deineko, E. V. (2011) Transplastomic plants, *Vavilov. Zh. Genet. Selekt.*, **15**, 808-817.
73. Maliga, P. (2012) Plastid transformation in flowering plants, in *Genomics of Chloroplasts and Mitochondria* (Bock, R., and Knoop, V., eds.) Springer, pp. 393-414.
74. Hanson, M. R., Gray, B. N., and Ahner, B. A. (2013) Chloroplast transformation for engineering of photosynthesis, *J. Exp. Bot.*, **64**, 731-742.
75. Daniell, H., Chan, H.-T., and Pasoreck, E. K. (2016) Vaccination via chloroplast genetics: affordable protein drugs for the prevention and treatment of inherited or infectious human diseases, *Annu. Rev. Genet.*, **50**, 595-618.
76. Bock, R. (2015) Engineering plastid genomes: methods, tools, and applications in basic research and biotechnology, *Annu. Rev. Plant Biol.*, **66**, 211-241.
77. Jin, S., and Daniell, H. (2015) The engineered chloroplast genome just got smarter, *Trends Plant Sci.*, **20**, 622-640.
78. Bock, R. (2014) Genetic engineering of the chloroplast: novel tools and new applications, *Curr. Opin. Biotechnol.*, **26**, 7-13.
79. Waheed, M. T., Ismail, H., Gottschamel, J., Mirza, B., and Lossl, A. G. (2015) Plastids: the green frontiers for vaccine production, *Front. Plant Sci.*, **6**, 1005.
80. Bock, R. (2014) Engineering chloroplasts for high-level foreign protein expression, *Methods Mol. Biol.*, **1132**, 93-106.
81. Gisby, M. F., Mudd, E. A., and Day, A. (2012) Growth of transplastomics cells expressing D-amino acid oxidase in chloroplasts is tolerant to D-alanine and inhibited by D-valine, *Plant Physiol.*, **160**, 2219-2226.