Chloroplast genetic system of higher plants: Chromosome replication, chloroplast division and elements of the transcriptional apparatus

Régis Mache* and Silva Lerbs-Mache

Université de Grenoble and Centre National de la Recherche Scientifique, Laboratoire de Génétique Moléculaire des Plantes, BP 53, F-38041 Grenoble cedex 9, France

DNA replication, transcription of the circular plastid genome and plastid division represent fundamental processes that take place in all types of the multifunctional, multiple-form plastids. The components of the plastid genetic system are highly expressed during early phases of chloroplast development corresponding to the establishment of 'housekeeping' functions before the onset of photosynthesis. A survey of fundamental and recent work is presented concerning plastid division, the organization of the plastid genome and the characterization of enzymes implicated in replication and transcription.

AN essential function, directly or indirectly, for all forms of present-day living organisms is photosynthesis. This function is maintained in chloroplasts which are present in all plants, with the exception of the few parasitic plants which have lost autotrophy. To maintain a high photosynthetic capacity the number of chloroplasts per cell has tremendously increased during evolution. In higher plants, chloroplasts are derived from proplastids, from pre-existing chloroplasts or from other forms of plastids. They differentiate into reversible plastid forms in relation with tissue differentiation. They are named leucoplasts and amyloplasts in roots, chloroplasts in leaves, chromoplasts in several organs such as petals and fruits.

Originating from meristematic cells, proplastids develop according to the tissue in which they are located. When seeds are germinating in the soil, i.e. without light, the cotyledon cells contain etioplasts which are characterized by the presence of a pseudocrystalline structure, the prolamellar body and many ribosomes. In the presence of light, thylakoids are formed, emerging from the crystalline body. The presence of ribosomes in etioplasts suggests that the building of the plastid genetic system in plastids precedes the setting of the photosynthetic apparatus. This assumption has been verified in spinach by showing that very early after the beginning of seed imbibition, nuclear genes encoding plastid ribosomal proteins are activated. The expression of these genes is dependent on the activation of specific promoter elements and precedes the expression of nuclear genes encoding photosynthetic products by two days[1,2]. Thus, the determination of the components of the chloroplast genetic system and the analysis of their regulation is essential in at least two ways: (1) to decipher the steps occurring in early chloroplast differentiation, and (2) to understand the functioning of chloroplast gene expression. We report here a survey of fundamental works concerning the chloroplast genetic system in higher plants, with the exception of the translational apparatus. Organization of the chloroplast genome in an evolutionary background will be considered first, then some problems of plastid division, and finally, recent discoveries concerning elements of the transcriptional apparatus will be described.

The plastid chromosome: Organization and evolution

Origin and gene transfer during evolution

The plastid chromosome has been highly conserved during evolution since the earliest green algae arose. *Mesostigma viride* represents a unicellular flagellate that belongs to the earliest diverging green plant lineage discovered to date[3]. Its chloroplast genome has the same overall structure as that of the higher plant plastid chromosome and a similar size. The cyanobacterial origin of the plastid genome is well documented now and the endosymbiotic hypothesis has been verified by several molecular data.

The cyanobacterial genome contains more than 3000 potential protein genes, whereas the present-day chloroplast genome contains only about 75 protein genes. Thus, most of the genes of the endosymbiotic ancestor were either lost or transferred to the nucleus. Indeed, the sequencing of the *Arabidopsis* chromosome 4 shows that a large number of transit peptide-containing, nucleus-encoded, proteins are highly similar to those in

*For correspondence. (e-mail: regis.mache@ujf-grenoble.fr)

Synechocystis sp., a cyanobacterium[4]. The similarity of the gene content of the plastid genomes of evolutionary distant plants indicates that most genes have been transferred from the endosymbiotic ancestor to the nucleus very early in plant evolution. Although, the transfer process is not yet complete, but it occurs rarely. An indication of a relatively recent transfer has been obtained for the *rpl*22 gene encoding the ribosomal protein L22 (ref. 5). In many legumes, the gene has been relocated into the nucleus. An additional protein coding region was added after the gene transfer which has a plastid targeting activity. In this particular case, the transit peptide which allows chloroplast targeting, results from an exon shuffling into the 5' of the gene, separated by an intron from the exon sequence for the mature protein. In one case, the origin of the 5' shuffled exon has been identified[6]. The fast evolution of presequences after shuffling suggests a rapid adaptive evolution. In conclusion, a large number of genes were transferred to the nucleus which thus controls the biogenesis of chloroplasts.

The plastid chromosome

The plastid DNA, also named plastid chromosome or plastome, is a circular double-stranded molecule of 120 to 180 kilobase pairs (kbp). Each plastid contains tens to hundreds of copies of the molecule, organized into several nucleoids. These nucleoids are readily observed in chloroplasts stained by DAPI (4',6-diamidino-2-phenylindole) using fluorescent microscopy. These observations confirm the pioneering work of Kowallik and Herrmann[7] identifying a series of discrete areas spread throughout the plastid. Nucleoids appear interconnected in young and mature chloroplasts[8]. A small number of nucleoids is present in proplastids but this number increases in young plastids. At this stage of development, plastids are actively dividing and nucleoids are associated with the inner membrane of the plastid envelope. A 130 kDa DNA binding protein which recognizes specific sequences of plastid DNA has been identified in pea[9]. This protein named PEND (for plastid envelope DNA binding) is certainly involved in the binding of nucleoids to the envelope of young plastids and might be involved in DNA replication. Interestingly, the PEND protein is not detected in mature chloroplasts in which nucleoids are more likely attached to thylakoids. In later stages of development the nucleoids are dispersed and associated with grana thylakoids to which they are attached[10].

The plastid chromosome exists as a negatively supercoiled molecule. The analysis of DNA conformation by pulse-field electrophoresis showed that molecules are present as monomers, dimers, trimers and tetramers in a relative amount of 1, 1/3, 1/9 and 1/27, respectively[11]. The high percentage of oligomeric forms is surprising. Bacterial or yeast plasmid DNA are also existing as dimers, but at a much lower abundance.

The complete plastid DNA sequences from eight higher plant genera are presently known, including gymnosperms and angiosperms[12]. From all these data, and the sequence data from algae, it can be concluded that the chromosome organization is highly conserved. The plastid chromosomes from plants, with some exceptions in the Fabaceae species, contain two inverted repeat (IR) regions separating a large and a small single-copy (LSC and SSC, respectively) region. The observed difference in size between the different plastid chromosomes is mainly due to the length of the IR and not to additional genes. A large inversion in the LSC region has occurred during evolution after the bryophytes[13]. Smaller rearrangements in the plastome organization have occurred mainly in some monocotyledonous species.

The A+T content is not evenly distributed in the plastome. It is higher in non-coding regions and is lower in regions coding for tRNA and for the rRNAs. The plastome of higher plants contains 4 ribosomal RNA genes, 30 tRNA genes, more than 72 genes encoding polypeptides and several conserved reading frames (ycf) coding for proteins of yet unknown function. In contrast to prokaryotic tRNA genes, no plastid tRNA gene codes for its 3'-CCA end, although in several cases the first C is present. This element is added post-transcriptionally. The set of 30 tRNAs is sufficient to read all amino acid codons[14]. It is important to state that no RNA, even of small size, is imported into chloroplasts. Besides its role in translation, one of the three tRNA–Glu is involved in the synthesis of γ-aminolevulinic acid, the precursor of chlorophyll[15]. The plastid genes coding for polypeptides can be classified into several categories: genes coding for the prokaryotic RNA polymerase core-enzyme; genes coding for proteins of the translational apparatus; for the photosynthetic apparatus and genes encoding subunits of the NADH-dehydrogenase (*ndh*). One gene (*clpP*) codes for a protease subunit and, in dicotyledons, the gene *acc*D encodes an acetyl-CoA-carboxylase subunit. Some divergence in the presence of genes exists between monocotyledons and dicotyledons. The chromosomes of rice and maize do not contain the *acc*D gene nor the full *ycf*2 open reading frame. Exceptionally, genes are present as pseudo-genes[16,17]. A number of sites have to be edited on the mRNA. These sites have been detected in several different plastomes in all families of land plants[18].

Many genes are interrupted by intronic sequences. In contrast to algae[19], plastid genes of higher plants contain single introns, with the exception of the *rpl*12 gene which contains two introns and requires *trans* splicing. The *trn*K intron contains an ORF presumably encoding a maturase related to mitochondrial maturases[20]. Introns

are also present in protein encoding genes. Based on conserved secondary structures, the introns have been classified into two groups, group I and II[19]. Intron sites of group II introns have been determined experimentally in some cases[20,21]. The *trn*L(UAA) intron can be folded in a group I intron structure which should be auto spliceable. But this has not yet been shown experimentally for land plants. Other tRNA gene containing introns belong to group II[14].

Chloroplast division

Control of the number of plastids per cell

All terrestrial plants, including the first land plants such as the liverworts which belong to early Bryophytes, have chloroplasts of similar shape that contain stacked thylakoids. It is worth noting that many of the Bryophytes and Pteridophytes possess one single plastid per cell, as it is the case for *Anthoceros* sp. for example, in evolutionary continuity with a number of green algae such as *Chlamydomonas* sp. Interestingly, *Isoetes*, an evolved fern, possesses one chloroplast per meristematic cell and several chloroplasts in mature cells[22]. This species might represent a link with plants which originated later in evolution and possess several plastids per cell. In angiosperms, the dark-green spinach leaves contain more than 200 chloroplasts per mesophyll cell and *Arabidopsis* contains more than 100 plastids per mesophyll cell[23]. An explanation for this enhancement of chloroplast number per cell during evolution is to allow the leaf to catch the maximum of light. An optimal size of chloroplasts facilitates their mobility within the cell, to expose the maximum of chloroplast surface to light. Thus a high number of optimally sized chloroplasts would correspond to an adaptation of plants to the environment in order to increase the photosynthetic capacity of leaves.

During the last decade, our knowledge concerning the control of plastid number has made a great deal of progress. Genetic analyses have shown that this control is made by nuclear genes and comprises interactive regulation of plastid number and size. The discovery in *Arabidopsis* of *arc* (accumulation and replication of chloroplasts) mutants has been an important step in unravelling the genetic basis for chloroplast division[24,25]. The *arc* mutants have the same phenotype as wild type plants, but mesophyll cells contain an abnormal number of chloroplasts. The mutants *arc3*, *arc5*, *arc6* and *arc12* have a reduced number of chloroplasts per mesophyll cell by comparison with the wild type, whereas the mutants *arc1* and *arc7* have a higher number of chloroplasts. An important observation made after the analysis of these mutants is that an inverse relationship exists between the number of chloroplasts and size. The mesophyll cell of the *arc6* mutant contains two chloroplasts, but they are much larger[26]. Conversely, the mesophyll cell of the *arc7* mutants contains more but smaller chloroplasts than does the wild type. Another important observation has been made for several dicotyledonous species that correlates the total surface of chloroplasts to cell size. The number of chloroplasts increases when the cell enlarges. When a blockage in chloroplast division is introduced, by *arc* mutation for example, the growth of chloroplast continues to compensate for the lack of total chloroplast surface. Thus the ratio of cell volume to total chloroplast area appears to be the controlling factor. The sensor genes for this factor are unknown but in wild type mesophyll cells the final target genes are those implicated in chloroplast division and secondarily, the genes implicated in chloroplast growth.

Plastid division per se

Many problems are still unsolved concerning plastid division. When the number of plastids per cell is low, plastid division is probably controlled by the cell cycle. But detailed studies are missing concerning this problem. When cells contain a large number of plastids, they do not divide synchronously. The regulatory pathway that determines when a plastid enters the division cycle is also unknown. One determinant could be the size of a given plastid: division would occur only when the plastid volume has reached a defined value. The number of nucleoids present in the plastid is probably another factor[27]. Also, it has been recently discovered that division of plastids, besides the overall control by the cell, has conserved prokaryotic-like mechanisms.

In bacteria, a ring made of the protein FtsZ is formed at the division site within the cell, at the surface of the cytoplasmic membrane. The separation of the two daughter cells is made by the invagination of the cytoplasmic membrane, driven by the contracting FtsZ ring. Recent structural studies have shown similarities between the structure of FtsZ and that of tubulins. An homologue of FtsZ has been found in *Arabidopsis*, that is imported into chloroplasts[28]. FtsZ plays a role in plastid division as shown by plants transformed with an antisense construction of the *Fts*Z gene. In these plants, plastid division is modified in mesophyll cells and in meristematic cells[25]. Chloroplast division is also altered in the mutant moss *Physcomitrella patens* with a null gene for *Fts*Z[29]. Thus, the function of FtsZ is conserved. This protein could start the division process by pulling the envelope towards the point where the plastids can be separated.

Other bacterial genes like the *min*C, *min*D and *min*E genes, are important for the correct positioning of the inner ring in the middle of the cell. An homologue of

one of these genes has been recently found in *Arabidopsis*[25]. Also, light and electron microscopy have shown that chloroplast division is initiated by a constriction in the middle of the plastid. This type of control is surely important in order to assure an equilibrated distribution of essential components of the chloroplast into the two daughter chloroplasts. It can be expected that the position of the FtsZ ring within the chloroplast is under the control of genes similar to those functioning in bacteria.

Replication of plastid DNA

The synthesis of DNA is a prerequisite for plastid division. Plastid DNA synthesis has been first demonstrated by the incorporation of labelled ^{3}H-thymidine into chloroplast DNA of dividing tissues. The label distributed equally in the daughter cells[30], showing that all plastids undergo DNA replication. It was also shown that all the plastid chromosomes (about 10,000) in a cultured cell of tobacco replicate in one cell cycle[31]. Thus, the high degree of polyploidy of the plastid genome is conserved in all cells. Also, a relationship of plastid DNA replication to the cell cycle events seems to exist. It has been generally assumed that DNA synthesis occurs in nucleoids. However, a recent contribution[32] has shown that a 68 kDa DNA compacting nucleoid protein inhibits DNA synthesis *in vitro*. This result suggests that DNA synthesis occurs outside of compact nucleoids. It is therefore possible that partial release from the nucleoid constraints is necessary for replication elongation.

Detailed studies of chloroplast DNA content and synthesis in dividing spinach leaf cells have been made[23,33]. In young spinach cells, chloroplast DNA synthesis keeps pace with chloroplast division and the amount of DNA per chloroplast is relatively low. In elongating cells, efficient chloroplast DNA synthesis occurs in less dividing chloroplasts. In adult cells, DNA synthesis is stopped. The DNA content per plastid decreases because chloroplasts are still dividing. In barley, a monocotyledonous plant, development of chloroplasts can be followed in parallel with leaf development. Using this system, Baumgartner and Mullet[34] have shown that high rates of DNA synthesis occur in the leaf basal meristem of dark-grown leaves. The basal cells contain proplastids. DNA synthesis activity decreases in later stages of chloroplast development.

To understand the mechanisms governing the replication of plastid DNA, electron microscopic observations were made with plastids isolated from several plant species. In a pioneering work, Kolodner and Tewari[35] had shown that replication of plastid DNA from maize and pea is initiated with the formation of two displacement loops (D-loops), located 7 kbp apart on opposite strands of the chromosome and expanding towards each other to form a Cairns-type of replicative intermediate. The plastid DNA contains also a rolling circle replicative intermediate resulting from the continuation of the Cairns rounds of replication. The origins of replication have been mapped by electron microscopic analysis and by two-dimensional agarose gel electrophoresis. In all dicotyledonous plants analysed, two origins were identified as D-loops. In pea, one origin (*ori*A) is located in the spacer region between the 16S and the 23S rRNA genes, and the second one (*ori*B) is located downstream the 23S rRNA gene[36]. In tobacco and *Oenothera*, the initiation sites *ori*A and *ori*B are also located in the flanking regions of the 23S rRNA gene[37,38]. In maize, the replication origins have been identified by selection of plastid DNA fragments efficiently used as template for DNA synthesis in the presence of chloroplast extracts. The putative replication origins are located in regions of the chromosome other than for dicotyledonous plants reported above in the large single copy region, in the 5' end of the *rpl*16 gene for one origin, and in the 3' end of the *psb*A gene for the second origin[39]. Albino mutants of wheat and barley contain plastid DNA with large deletions, yet they are able to replicate[40]. A common region in these deletion mutants contains the *psb*A gene. Other data are necessary to know whether monocotyledonous species in general possess different replication origins than dicotyledonous species.

A chloroplast DNA polymerase, named γ-DNA polymerase, has been purified from spinach and pea chloroplasts[41,42]. The enzyme differs from the α class of DNA polymerase in several properties: it is smaller, resistant to aphidicolin and is inhibited by ethidium bromide. A 43 kDa protein that interacts with the chloroplast DNA polymerase and increase the processivity of replication has also been discovered[43]. Another protein participating in the assembly of the replication complex has been identified as a 120 kDa primase. It is required for the initiation of the synthesis of Okazaki fragments on the lagging strand during the rolling circle replication. The enzyme is capable of priming DNA replication on a single-stranded template. It synthesizes a short complementary RNA and requires the four rNTP for activity[44].

As a circular double-stranded DNA molecule, the chloroplast genome raises topological problems concerning the opening of the two strands that is necessary for replication initiation. The prokaryotic topoisomerase I relaxes the torsion created by the opening of the replication fork. At the end of replication, topoisomerase II liberates the two concatenated daughter duplexes. These relaxing enzymes, including the eubacterial gyrase (a topoisomerase II), are also necessary for local transcription and for recombination. Helicases represent another kind of enzymes which intervene in the production of positive supercoils, interact with topoisomerases and

are ATP dependent. They play a role in replication, DNA repair, transcription and recombination processes. Several of these enzymes have been identified in chloroplasts. Surprisingly, two different topoisomerases I have been purified from intact pea chloroplasts[45]. One, of 112 kDa size, is related to the Mg^{++}-dependent prokaryotic type. The second one, unexpectedly, is an active monomer of 69 kDa, and is clearly of the eukaryotic type. The function of this eukaryotic type topoisomerase *in vivo* is unknown. An accessory protein in the replication complex has also been characterized, which interacts with the pea topoisomerase I and affects its DNA relaxation activity[46]. Topoisomerase II activities have been detected in chloroplasts of higher plants and maximal level was found in chloroplasts at the time of chloroplast DNA replication[47]. Also, a gyrase activity has been detected in pea chloroplasts[48]. A pea DNA helicase of 78 kDa has been purified to homogeneity. The enzyme contains a DNA-dependent ATPase activity and unwinds DNA fork-like replication structures[49].

Chloroplast DNA is not efficiently protected against UV irradiation by the surrounding cell structures. Therefore, we can expect UV-induced damages and correlatively, repair processes. Surprisingly, the chloroplast genome is highly conserved and the mutation rate is low by comparison with that of the plant nuclear genomes[50]. This means that DNA repair mechanisms are very efficient. Photoproducts of UV-light could be removed by photoreactivation and excision repair, as in bacteria. The RecA protein is involved in recombination and in DNA repair. A cDNA clone encoding a RecA-like pre-protein addressed to chloroplast has been identified in several plants, providing evidence for the presence of a recombination/repair system in chloroplasts. The protein is homologous to the cyanobacterial RecA[51].

Transcriptional apparatus

Regulation of chloroplast gene expression occurs at several levels, including transcription. It was clearly demonstrated that transcription regulation exists for several sets of genes[52]. Transcription activity largely varies from one gene to another. In addition, gene transcription is modulated during chloroplast development or according to environmental changes. For instance, genes encoding the genetic system are preferentially expressed during the first steps of chloroplast differentiation[53]; the *psb*A gene is highly activated under light[54], but such genes are weakly expressed in differentiated chromoplasts or amyloplasts[55]. Regulation of transcription is adapted to the different needs and is controlled at several levels which have been partially unveiled recently. First, several RNA polymerases, PEP and NEP, are present in the chloroplasts having defined functions[56,57]. Second, sigma-like transcription initiation factors are controlling the activity of the plastid encoded plastid RNA polymerase (PEP)[58]. Lastly, transcription factors can interact with the two types of RNA polymerase and thus regulate the choice of the transcriptional system[59]. Why are so many regulatory elements necessary to transcribe the small plastid chromosome? This question is still unanswered, but some hypotheses could be developed. These different points will be considered in more detail.

Plastid-encoded plastid RNA polymerases

The plastid chromosome contains genes encoding polypeptides similar to the RNA polymerase subunits of cyanobacteria, representative of the endosymbiotic ancestor. Several procedures have been used to purify nearly homogenous plastid RNA polymerase, from maize, pea and spinach. The enzyme is referred to as 'prokaryotic', because of its genetic origin and is now named PEP for plastid-encoded plastid RNA polymerase. However, the direct proof that the plastid-encoded *rpo*A, *rpo*B, *rpo*C1 and *rpo*C2 genes correspond really to polypeptides present in a highly purified RNA polymerase preparation was obtained only in maize so far[60,61]. Other highly-purified enzymes have been described for spinach and pea[62,63], but their polypeptides have not been microsequenced. The pea or spinach purified fractions contain additional polypeptides. They could represent degradation products or polypeptides of yet unknown function. Recently, β and β' subunits have been assigned to the glycerol gradient purified RNA polymerases from mustard[64].

The presence, in higher plant chloroplasts, of transcription initiation factors of the sigma-70 type has been revealed recently[56,58,65]. The translated products of six different cDNAs show strong similarities with the prokaryotic sigma-70 factors. For three of them, it has been shown that they are transported into chloroplasts[65,66]. The proteins deduced from the genes would have a size ranging from 60 to 70 kDa. Five different cDNAs coding for sigma-like proteins have also been cloned and analysed from maize[67,68]. In mustard and spinach, functional assays have been performed that show transcription initiation-specific functions of polypeptides that correspond to sigma gene product[69,70]. An interesting model was proposed to explain the regulation of the transcriptional activity of gene expression based on studies using the *psb*A gene encoding an essential element of the PSII photosystem. Light-induced regulation of gene expression could involve phosphorylation–dephosphorylation of sigma-like factors[71]. The holoenzyme present in etioplasts would transcribe prokaryotic-type promoters after dephosphorylation of a sigma-like factor by a light-induced phosphatase.

Recently, the differences between three plant sigma factors with respect to promoter recognition have been determined by *in vitro* transcription assays[70]. It has been shown that the N-terminal parts of the plant factors have different functions than the N-terminal part of the sigma-70 factor from *E. coli*. The C-terminal part which is responsible for the DNA promoter recognition, is functionally conserved between prokaryotes and plastids. One of the plastid factors, SIG1, is the most prokaryotic-like of the six plant sigma-like factors and recognizes all essential *E. coli* promoters. It recognizes specifically the plant prokaryotic-type *rbc*L promoter. In contrast, SIG2 recognizes specifically the less-conserved prokaryotic-type P1 promoter of the *rrn* operon encoding the rRNA species. The function of SIG3 is less specific. SIG3 recognizes all plastid prokaryotic promoters that have been analysed.

Several factors binding plastid DNA have been identified, that play a role in transcription regulation. The AGF factor binds to a specific promoter element named AAG box necessary for the transcription of the blue-light activated *psb*D operon in barley[72–74]. Another factor, CDF2, has been identified in spinach. CDF2 binds specifically to the promoter region of the *rrn* operon and regulates expression of rRNA in plastids[75–77].

Nuclear-encoded plastid RNA polymerases

Plastid RNA polymerases whose subunits are encoded in the nuclear genome are named NEP, for nucleus encoded plastid RNA polymerase. Nuclear genes encoding organellar destined RNA polymerases have been recently identified[56,78]. They encode RNA polymerases resembling the bacteriophage T7 RNA polymerase. One of them has been shown to be located in plastids. Indeed, the biochemical identification in chloroplast of a 110 kDa monomeric RNA polymerase, recognizing a T7 promoter was made several years ago[79]. The question, which plastid genes are expressed by NEP, could be solved by using transformed tobacco plants with the deletion of one *rpo* gene encoding a subunit of the PEP. Using this method, it has been shown that transformed plants are white coloured and are not able to perform photosynthesis[80–82]. It has been shown that NEP transcribes the genes encoding elements of the genetic system, rather than the photosynthesis genes[83]. A series of NEP promoters have been also mapped using mutant plants or photosynthetically inactive plant cell cultures and plant organs[56,84].

On the basis of plastid RNA expression studies obtained with mutant plants and leaf sections, a hypothesis has evolved envisaging a sequential action of the two enzymes during plant development. This hypothesis attributes specific functions to NEP in housekeeping gene expression during early phases of plant and plastid development, and to PEP in photosynthesis-related gene expression in later phases of plant and plastid development[79,85]. This hypothesis of sequential synthesis and action of the two enzymes is generally accepted to hold true for the differentiation of proplastids into chloroplasts, i. e. when meristematic cells differentiate into photosynthetically active tissues during plant growth and development. But what happens later in mature chloroplasts? The mechanism of transcription regulation by the two types of enzymes in relation to plant and plastid development is not yet clear. The fact that NEP is present in mature chloroplasts suggests division of labour between the two types of enzymes, NEP and PEP. For instance, it has been shown that two RNA polymerases coexist in leaf chloroplasts of very young spinach plants[79]. Under these conditions, plastid gene expression seems to be regulated by transcription factor-mediated modulation of the transcriptional activity of the two enzymes[76,77]. Additional work is needed to answer these questions in more detail.

A second NEP

The *rrn* operon contains the genes encoding the ribosomal RNA species (16S, 23S, 5S and 4,5S rRNAs) in addition to the genes encoding the tRNA species present in the 16S–23S spacer region and at the end of the operon. In the promoter region of the spinach plastid *rrn* operon, a sequence specific *cis*-element is present that serves to bind a transcription factor named CDF2. This factor plays an essential role in the transcription initiation of that operon. It binds to PEP and prevents transcription from two consensus prokaryotic promoters, P1 and P2 (ref. 76). An additional sequence-specific promoter, PC, is present and is exclusively used *in vivo*. This promoter is used by an enzyme named NEP2 which is not immunologically related to the previously characterized phage-type 110 kDa enzyme that was named NEP1 (ref. 77). NEP2 is recruited to the PC promoter by the CDF2 factor. Thus, a transcription factor (CDF2) controls transcription by interacting with two different RNA polymerases: it represses transcription by binding to PEP and activates transcription by binding to NEP2. The exact polypeptide composition of NEP2 is not yet known.

Conclusion

Since the first sequence of a plastid genome was obtained in 1986, many new elements which participate in the genetic system of plastids in higher plants have been characterized. Some of the findings were very new and unexpected. This was the case, in particular, for the presence of several plastid RNA polymerases and of several sigma-like initiation factors which have been

described here. These findings are surprising as the plastid genome is reduced by comparison with its cyanobacterial ancestor. These discoveries open new ways to understand how the plastid gene expression is controlled. Another point of interest resides in the fact that many key elements of the plastid genetic system are encoded in the nucleus. Systematic bioinformatic analysis of the *Arabidopsis* genomic and cDNA sequences should help to determine systematically all nuclear sequences coding for plastid-localized proteins. The analysis of their function and of the control of their synthesis should help to complete our understanding on the interrelationship between cell and plastid differentiation.

1. Harrak, H., Lagrange, T., Bisanz-Seyer, C., Lerbs-Mache, S. and Mache, R., *Plant Physiol.*, 1995, **108**, 685–692.
2. Lagrange, T., Franzetti, B., Axelos, M., Mache, R. and Lerbs-Mache, S., *Mol. Cell Biol.*, 1993, **13**, 2614–2622.
3. Lemieux, C., Otis, C. and Turmel, M., *Nature*, 2000, **403**, 649–652.
4. Mayer *et al.*, *Nature*, 1999, **402**, 769–777.
5. Gantt, J. S., Baldauf, S., Calie, P. J., Weeden, N. F. and Palmer, J. D., *EMBO J.*, 1991, **10**, 3073–3078.
6. Long, M., de Souza, S. J., Rosenberg, C. and Gilbert, W., *Proc. Natl. Acad. Sci. USA*, 1996, **93**, 7727–7731.
7. Kowallik, V. and Herrmann, R. G., *J. Cell Sci.*, 1972, **11**, 357–377.
8. Possingham, J., Chaly, N., Robertson, M. and Cain, P., *Biol. Cell*, 1983, **47**, 205–212.
9. Sato, N., Albrieux, C., Joyard, J., Douce, R. and Kuroiwa, T., *EMBO J.*, 1993, **12**, 555–561.
10. Rose, R. J., *J. Cell Sci.*, 1979, **36**, 169–183.
11. Deng, X.-W., Wing, R. A. and Gruissem, W., *Proc. Natl. Acad. Sci. USA*, 1989, **86**, 4156–4160.
12. Martin, W. and Herrmann, R. G., *Plant Physiol.*, 1998, **118**, 9–17.
13. Sugiura, M., *Essays Biochem.*, 1995, **30**, 49–57.
14. Maréchal-Drouard, L., Weil, J. H. and Dietrich, A., *Annu. Rev. Plant Physiol. Plant Mol. Biol.*, 1993, **44**, 13–52.
15. Shön, A., Krupp, G., Gough, S., Berry-Lowe, S., Kannangara, C. G. and Söll, D., *Nature*, 1988, **322**, 281–284.
16. Mache, R., *Plant Sci.*, 1990, **72**, 1–12.
17. Maier, R. M., Neckermann, K., Igloi, G. L. and Kössel, H., *J. Mol. Biol.*, 1995, **251**, 614–628.
18. Freyer, R., Kiefer-Meyer, M.-C. and Kössel, H., *Proc Natl. Acad. Sci. USA*, 1997, **94**, 6285–6290.
19. Copertino, D. W. and Hallick, R. B., *EMBO J.*, 1991, **10**, 433–442.
20. Zhou, D. X., Quigley, F., Massenet, O. and Mache, R., *Mol. Gen. Genet.*, 1989, **216**, 439–445.
21. Sugiura, M., *Plant Mol. Biol.*, 1992, **19**, 149–168.
22. Pettitt, J. M., *J. Cell Sci.*, 1976, **20**, 671–685.
23. Possingham, J. V., *Annu. Rev. Plant Physiol.*, 1980, **31**, 113–129.
24. Marrison, J. L., Rutherford, J. L., Robertson, E. J., Lister, C., Dean, C. and Leech, R. M., *Plant J.*, 1999, **18**, 651–662.
25. Pyke, K. A., *Plant Cell*, 1999, **11**, 549–556.
26. Robertson, E. J., Pyke, K. A. and Leech, R. M., *J. Cell Sci.*, 1995, **108**, 2937–2944.
27. Boffey, S. A. and Leech, R. M., *Plant Physiol.*, 1982, **69**, 1387–1391.
28. Osteryoung, K. W. and Vierling, E., *Nature*, 1995, **376**, 473–474.
29. Strepp, R., Schlotz, S., Kruse, S., Speth, V. and Reski, R., *Proc. Natl. Acad. Sci. USA*, 1998, **94**, 5967–5972.
30. Rose, R. J., Cran, D. G. and Possingham, J. V., *Nature*, 1974, **251**, 641–642.
31. Infante, D. and Weissbach, A., *Plant Mol. Biol.*, 1990, **14**, 891–897.
32. Cannon, G. C., Ward, L. N., Case, C. I. and Heinhorst, S., *Plant Mol. Biol.*, 1999, **39**, 835–845.
33. Lawrence, M. E. and Possingham, J. V., *Plant Physiol.*, 1986, **81**, 708–710.
34. Baumgartner, B. J. and Mullet, J. E., *J. Photochem. Photobiol.*, 1991, **11**, 203–218.
35. Kolodner, R. D. and Tewari, K. K., *Nature*, 1975, **256**, 708–711.
36. Meeker, R., Nielsen, B. and Tewari, K. K., *Mol. Cell Biol.*, 1988, **8**, 1216–1223.
37. Lu, Z., Kunnimalaiyaan, M. and Nielsen, B. L., *Plant Mol. Biol.*, 1996, **32**, 693–706.
38. Kunnimalaiyaan, M., Shi, F. and Nielsen, B. L., *J. Mol. Biol.*, 1997, **268**, 273–283.
39. Carillo, N. and Bogorad, L., *Nucleic Acids Res.*, 1988, **16**, 5603–5620.
40. Day, A. and Ellis, T. H. N., *Cell*, 1984, **39**, 359–368.
41. Sala, F., Amileni, A. R., Parisi, B. and Spadari, S., *Eur. J. Biochem.*, 1980, **112**, 211–217.
42. McKown, R. L. and Tewari, K. K., *Proc. Natl. Acad. Sci. USA*, 1984, **81**, 2354–2358.
43. Gaikwad, A., Tewari, K. K., Kumar, D., Chen, W. and Mukherjee, S. K., *Nucleic Acids Res.*, 1999, **27**, 3120–3129.
44. Nielsen, B. L., Rajasekhar, V. K. and Tewari, K. K., *Plant Mol. Biol.*, 1991, **16**, 1019– 1034.
45. Mukherjee, S. K., Reddy, M. K., Kumar, D. and Tewari, K. K., *J. Biol. Chem.*, 1994, **269**, 3793–3801.
46. Marrison, J. L. and Leech, R., *Plant J.*, 1992, **2**, 783–790.
47. Van Hop, D., Gaikwad, A., Yadav, B. S., Reddy, M. K., Sopory, S. and Mukherjee, S. K., *Plant J.*, 1999, **19**, 153–162.
48. Lam, E. and Chua, N.-H., *Plant Mol. Biol.*, 1987, **8**, 415–424.
49. Tuteja, N. and Phan, T.-N., *Plant Physiol.*, 1998, **118**, 1029–1039.
50. Wolfe, K. H., Li, W.-H. and Sharp, P., *Proc. Natl. Acad. Sci. USA*, 1987, **84**, 9054–9058.
51. Cerutti, H., Osman, M., Grandoni, P. and Jagendorf, A. T., *Proc. Natl. Acad. Sci. USA*, 1992, **89**, 8068–8072.
52. Igloi, G. L. and Kössel, H., *Crit. Rev. Plant Sci.*, 1992, **10**, 525–588.
53. Baumgartner, B. J., Rapp, J. C. and Mullet, J. E., *Plant Physiol.*, 1993, **101**, 781–791.
54. Klein, R. R. and Mullet, J. E., *J. Biol. Chem.*, 1990, **265**, 1895–1902.
55. Deng, X. W. and Gruissem, W., *EMBO J.*, 1988, **7**, 3301–3308.
56. Hess, W. R. and Börner, T., *Cytology*, 1999, **190**, 1–59.
57. Liere, K. and Maliga, P., in *Regulatory Aspects of Photosynthesis* (eds Anderson, B. and Aro, E. M.), Kluwer Academic Publishers, 2000, in press.
58. Allison, L. A., *Biochimie*, 2000, in press.
59. Lerbs-Mache, S., *Biochimie*, 2000, in press.
60. Hu, J. and Bogorad, L., *Proc. Natl. Acad. Sci. USA*, 1990, **87**, 1531–1535.
61. Hu, J., Troxler, R. F. and Bogorad, L., *Nucleic Acids Res.*, 1991, **19**, 3431–3435.
62. Lerbs, S., Brautingham, E. and Parthier, B., *EMBO J.*, 1985, **4**, 1661–1666.
63. Rajasekhar, V. K., Sun, E., Meeker, R., Wu, B.-W. and Tewari, K. K., *Eur. J. Biochem.*, 1991, **195**, 215–219.

64. Pfannschmidt, T., Ogrzewalla, K., Baginsky, S., Sickmann, A., Meyer, H. E. and Link, G., *Eur. J. Biochem.,* 2000, **267**, 1–10.
65. Isono, K., Shimizu, M., Yoshimoto, K., Niwa, Y., Satoh, K., Yokota, A. and Kobayashi, H., *Proc. Natl. Acad. Sci. USA*, 1997, **94**, 14948–14953.
66. Kanamaru, K. *et al.*, *Plant Cell Physiol.*, 1999, **40**, 832–842.
67. Lahiri, S. D., Yao, J., McCumbers, C. and Allison, L., *Mol. Cell Biol. Res. Comm.*, 1999, **1**, 14–20
68. Tan, S. and Troxler, R. F., *Proc. Natl. Acad. Sci. USA*, 1999, **96**, 5316–5321.
69. Kestermann, M., Neukirchen, S., Kloppstech, K. and Link, G., *Nucleic Acids Res.*, 1998, **26**, 2747–2753.
70. Hakimi, A. M., Privat, I., Valay, J. G. and Lerbs-Mache, S., *J. Biol. Chem.*, 2000, **275**, 9215–9221.
71. Link, G., *BioEssays*, 1996, **18**, 465–471.
72. Kim, J. K. and Mullet, J. E., *Plant Cell*, 1995, **7**, 1445–1457.
73. Kim, M., Christopher, D. A. and Mullet, J. E., *Plant Physiol.*, 1999, **119**, 663–670.
74. Kim, M., Thum, K. E., Morishige, D. T. and Mullet, J. E., *J. Biol. Chem.*, 1999, **274**, 4684–4692.
75. Baeza, L., Bertrand, A., Mache, R. and Lerbs-Mache, S., *Nucleic Acids Res.*, 1991, **19**, 3577–3581.
76. Iratni, R., Baeza, L., Andreeva, A., Mache, R. and Lerbs-Mache, S., *Genes Dev.*, 1994, **8**, 2928–2938.
77. Bligny, M., Courtois, F., Thaminy, S., Chang, C.-C., Lagrange, T., Baruah-Wolff, J., Stern, D. and Lerbs-Mache, S., *EMBO J.*, 2000, **19**, 1851–1860.
78. Hedke, B, Börner, T. and Weihe, A., *Science*, 1997, **277**, 809–811.
79. Lerbs-Mache, S., *Proc. Natl. Acad. Sci. USA*, 1993, **90**, 5509–5513.
80. Allison, L. A., Simon, L. D. and Maliga, P., *EMBO J.*, 1996, **15**, 2802–2809.
81. Serino, G. and Maliga, P., *Plant Physiol.,* 1998, **117**, 1165–1170.
82. De Santis-MacIossek, G., Kofer, W., Bock, A., Schoch, S., Maier, R. M., Wanner, G., Rudiger, W., Koop, H. U. and Herrmann, R. G., *Plant J.*, 1999, **18**, 477–489.
83. Hajdukiewicz, P. T. J., Allison, L. A. and Maliga, P., *EMBO J.*, 1997, **13**, 4041–4048.
84. Liere, K. and Maliga, P., *EMBO. J.*, 1999, **18**, 249–257.
85. Mullet, J. E., *Plant Physiol.*, 1993, **103**, 309–313.