Plant Molecular Biology **18**: 83–95, 1992.

Structural organization of the chloroplast genome of the chromophytic alga *Vaucheria bursata*

Karl-Heinz Linne von Berg [1] and Klaus V. Kowallik*
*Institut für Botanik der Heinrich-Heine-Universität, Universitätsstrasse 1, D-4000 Düsseldorf, FRG (*author for correspondence); [1]present address: Botanisches Institut der Universität Köln, Gyrhofstrasse 15, D-5000 Köln 41, FRG*

Received 13 May 1991; accepted in revised form 9 September 1991

Key words: chloroplast genome, gene mapping, inverted repeat, flip-flop recombination, Chromophyta, *Vaucheria bursata*

Abstract

The chloroplast genome of the chromophytic alga *Vaucheria bursata* has been characterized by restriction site and gene mapping analysis. It is represented by a circular molecule 124.6 kb in size. An inverted sequence duplication (IR) not larger than 5.85 kb carries the rRNA genes and separates two single-copy regions of 64.6 kb and 48.3 kb from one another. The *Vaucheria* plastid genome exists in two equimolar isomers which is due to intramolecular flip-flop recombination within the IR sequences. The coding sites for 21 structural and soluble proteins have been mapped on both single-copy regions using heterologous gene sequences as probes. Although the overall gene order is found to be rearranged when compared with other chromophytic algal and land plant chloroplast genomes, most of the transcriptional units of cyanobacteria and land plant chloroplast genomes appear to be conserved. The phylogenetic implications of these findings are further discussed.

Introduction

The genus *Vaucheria* comprises some fifty species of world-wide distribution, which mostly inhabit freshwater and brackish water habitats [38]. Biochemical characters including the presence of chlorophyll *a* and *c* render this genus a member of the heterokont chromophytes. Among these, *Vaucheria* is placed within the Xanthophyceae. It differs, however, from all other members of this algal class by the ability to form distinct vegetative and sexual reproductive organs. This, together with certain cytological characters not found in other xanthophycean algae, is responsible for the peculiar position of this genus among the heterokont chromophytes.

Chromophytic algae have gained special interest because of their complex internal organization. Since it was already shown for certain chromophytes that their plastids may be interpreted as part of a widely reduced eukaryotic endosymbiont [11, 47], it is now generally accepted that probably all chromophytic plastids may have resulted from eukaryotic/eukaryotic endocytoses. However, both the origin and nature of the eukaryotic endosymbiont remain obscure, as in most chromophytes this endosymbiont is extremely reduced except for its plastids.

In order to elucidate possible phylogenetic relationships between the putative endosymbiont and any other prokaryotic or eukaryotic photosynthetic organism, it is necessary to gain detailed information about the genetic background of chromophytic plastids. Knowledge of chromophytic plastid genomes available to date is sparse and does not allow detailed comparisons with the well known chloroplast genomes of land plants. However, besides certain similarities such as structural organization and gene composition of plastid genomes from both chlorophyll $a+b$ and $a+c$ plants, specific characters of chromophytic plastid genomes now become apparent. One of the most striking features is the highly rearranged gene order which even among closely related genera exceeds the degree of reorientation within the chloroplast genomes of land plants [17]. Even within a single *Vaucheria* species, strains may exhibit differences in their restriction fragment patterns that are mainly caused by single-base mutations [22]. Since it appears difficult at present to explain this high degree of plastid genome variation within the chlorophyll $a+c$ lineage, it remains speculative as to whether these characters reflect the influence of the second eukaryotic host, or a high degree of phylogenetic distance, or simply specific intramolecular properties. We therefore report here on the restriction and gene map of the plastid genome of the xanthophycean alga *Vaucheria bursata* in order to obtain more information about the organization and evolution of chromophytic plastid genomes.

Material and methods

Cultivation of Vaucheria

Vaucheria bursata was obtained from the Culture Collection of Algae, University of Göttingen. Stock cultures were maintained in a 14/10 h light/dark cycle at a light intensity of 120 μE PAR at 15 °C in 100 ml Erlenmeyer flasks containing Waris medium [43] enriched with vitamins [41]. Mass cultures were established using dissected stock material and grown in aerated 25 l flasks at an optimum growth temperature of 18 °C. The algae were harvested 4–6 weeks after inoculation by filtration through nylon gauze, followed by gentle shaking in demineralized water overnight.

Isolation of DNA, restriction site and gene mapping

The procedures used to isolate morphologically intact plastids and to extract and purify the DNA were the same as mentioned previously [22]. Restriction enzyme analysis using various endonucleases followed standard protocols. The restriction fragments generated by *Sal* I, *Bam* HI, *Pvu* II, and *Eco* RI were mapped on the circular plastid chromosome by reciprocal digestions of primary fragments isolated from LGT agarose tube gels [20].

The positions of individual genes were determined by Southern hybridizations with gene probes from spinach, pea, tobacco, *Chlamydomonas* and *Dictyota* (Table 1). Contaminating vector DNA was removed by electro-elution. The probes were labelled by nick translation using ^{32}P-α-dATP and hybridized to restricted plastid DNA bound to nitrocellulose filters (Schleicher and Schuell) overnight at 22 °C in hybridization buffer [24] containing 50% or 35% formamide. The filters were washed twice (5 × SSC, 50% or 35% formamide) for 30 min at room temperature, followed by a short rinse in 2 × SSC, and exposed to Kodak XAR 5 film for 3 h to several days.

Cloning of restriction fragments

About 1 μg of plastid DNA was digested with *Eco* RI in the presence of 0.5 μg pEMBL 8$^-$ vector DNA [8], heated to 72 °C, and precipitated with ethanol. The DNA was religated in a total volume of 30 μl and subsequently used to transform *Escherichia coli* strain RR1ΔM15. Recombinant clones were preselected by digesting minilysates with *Eco* RI, and checked by Southern hybridizations against *Eco* RI-restricted plastid DNA. The cloned restriction fragments were

Table 1. Gene probes used for Southern hybridization experiments, encoding subunits of PS I (*psaA* and *psaB*), PS II (*psbA–psbF*); ATPase (*atpA, atpB, atpE, atpF, atpH, atpI*), photosynthetic electron transport (*petA, petB, petD*), ribosomal RNAs (*rrnS* and *rrnL*), ribulose 1,5-bisphosphate carboxylase (large subunit, *rbcL*), elongation factor of translation (*tufA*), RNA polymerase (*rpoA–rpoC*).

Gene probe	Fragment size (bp)	Origin	Reference
psaA	1098	*Spinacia*	[16]
psaB	1680	*Spinacia*	[16]
psbA	670	*Spinacia*	[49]
psbB	1160	*Spinacia*	[28]
psbC	837	*Spinacia*	[2]
psbD	989	*Spinacia*	[2]
psbE	194	*Spinacia*	[15]
atpA	780	*Spinacia*	[45]
atpB	1260	*Spinacia*	[50]
atpE	435	*Spinacia*	[50]
atpF	260	*Spinacia*	[14]
atpH	183	*Spinacia*	[14]
atpI	480	*Spinacia*	[14]
petA	478	*Spinacia*	[1]
petB	309	*Spinacia*	[13]
petD	275	*Spinacia*	[13]
rrnS	1400	*Dictyota*	[19]
rrnL	5500	*Dictyota*	[19]
rbcL	1750	*Spinacia*	[48]
tufA	374	*Chlamydomonas*	[44]
rpoA	1040	*Spinacia*	[40]
rpoB	1063	*Nicotiana*	[29]
rpoC	3100	*Pisum*	[6]

used to confirm the validity of the physical map of the plastid genome. They were also used to construct fine restriction maps in those cases where several genes hybridized to the same primary or secondary restriction fragment.

Results

Restriction site mapping and flip-flop recombination

Single digestions using the restriction endonucleases *Sal* I, *Bam* HI, *Pvu* II, and *Eco* RI yielded between 6 and 28 fragments. In double digestions the number of detectable DNA bands was equivalent to the sum of the corresponding primary fragments when separated in agarose gels ranging from 0.25 to 1.4% (Fig. 1). We were able to detect restriction fragments of less than 400 bp when appropriate gel strength and sufficient amount of DNA was provided. The size of fragments exceeding 15 kb was calculated from their corresponding secondary fragments. The stoichiometry of bands from single and double digestions comprising more than one fragment (e.g. S/E 18–20, B/E 17–20) was determined according to the fragment patterns obtained from redigested primary fragments (Table 2). Thus, a total mean size of the plastid genome of *Vaucheria bursata* of 124.6 kb was calculated (Table 3).

The strategy to align the restriction fragments generated by the four enzymes used resulted in the construction of a circular map of the plastid genome (Fig. 2). Out of a total of 60 restriction sites, 56 were located on the circular map. Only the positions of the *Eco* RI fragments E 10, E 16, E 27 and E 28, which are internal to *Pvu* II fragment P 3, and of E 13, E 21 and E 24, which belong to P 1, remained unresolved. The order of the *Pvu* fragments P 12, 14, 15, 16 and 17, which are internal to E 3, was clarified by a fine restriction analysis of the cloned fragment E 3. The validity of the restriction map was corroborated by Southern hybridizations using cloned *Eco* RI fragments and heterologous gene probes.

The physical map shows two pairs of neighbouring fragments, E 22/BE 17 and E 23/BE 18, respectively, which are identical in size. Because of their inverse orientation, the presence of an inverted repeat sequence (IR) of at least 4.45 kb may be assumed. The existence of an IR sequence is further established by analysing the restriction patterns obtained from *Pvu* II and *Sal* I/*Pvu* II digestions. These enzymes generate two pairs of understoichiometric fragments (*Pvu* II: P 1, 3 and P 1a, 3a; *Sal* I/*Pvu* II: SP 1, 2 and SP 1a, 2a; Fig. 1c, f). The sizes of these corresponding pairs of fragments are identical when added up. Since the inversely duplicated fragments E 22/BE 17 and E 23/BE 18 are internal regions of S 1 or P 1/3, respectively, a recombination process or flip-flop mechanism acting within the IR sequences becomes obvious.

The maximum extension of the IR sequences

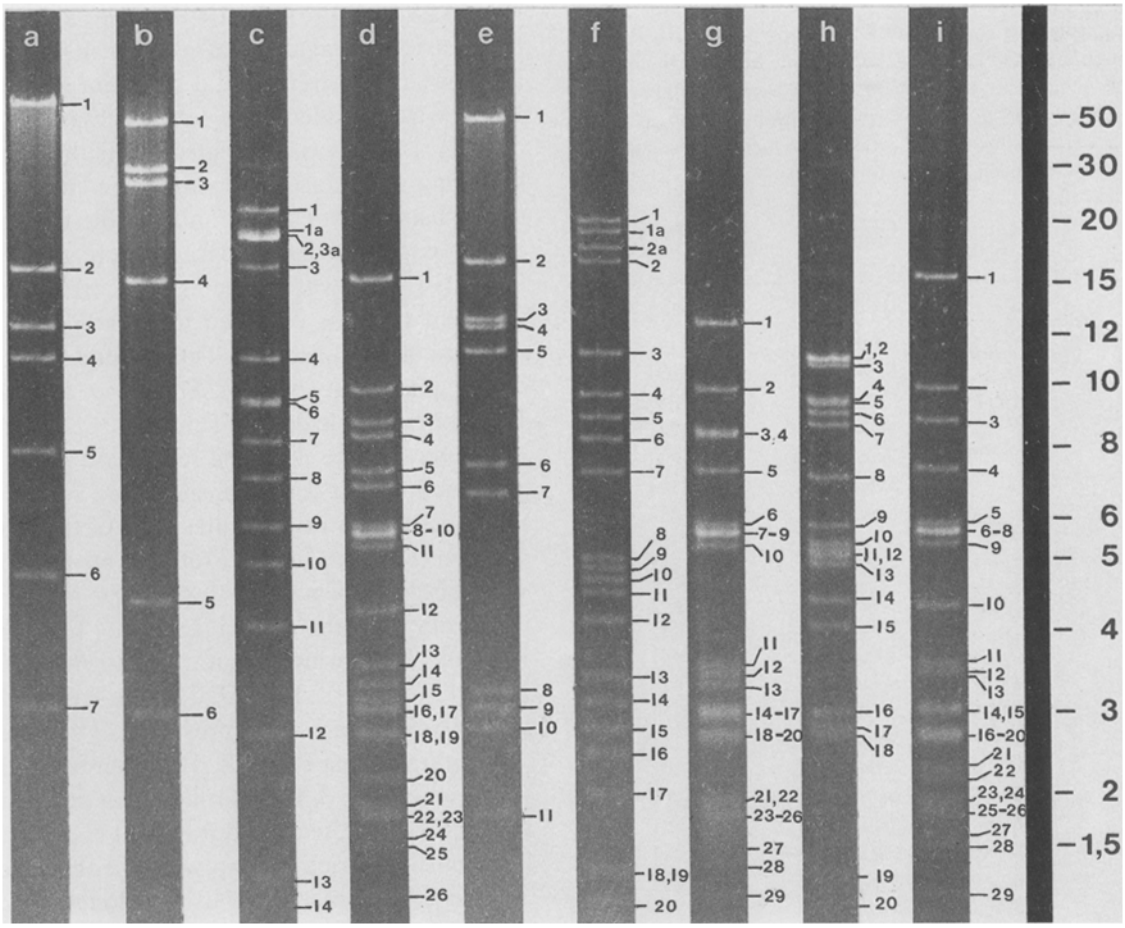

Fig. 1. Restriction patterns of *Vaucheria bursata* plastid DNA following digestion with *Sal* I (a), *Bam* HI (b), *Pvu* II (c), *Eco* RI (d), *Sal* I + *Bam* HI (e), *Sal* I + *Pvu* II (f), *Sal* I + *Eco* RI (g), *Bam* HI + *Pvu* II (h), and *Bam* HI + *Eco* RI (i). The restriction fragments were electrophoresed in 0,4% high-gelling agarose. A size scale in kb is shown at the right-hand side, indicating that restriction fragments smaller than 1,2 kb have migrated out of the gel. In two digestions (c, f) half stoichiometric pairs of fragments resulting from a flipping mechanism (see text) are indicated as 1, 3/1a, 3a (lane c), and 1, 2/1a, 2a (lane f), respectively.

can be determined by considering the two singular restriction sites next to the duplicated 4.45 kb sequence. The restriction site closest to E 23 is a *Pvu* II site yielding an overlapping fragment EP 38 of 0.8 kb. Since this *Pvu* II site is not present next to E 22, the IR sequence terminates within EP 38. Distal to EP 38 the subfragment BE 34 of 0.6 kb overlapping B 1 and E 14 contains the second border of the IR. Consequently, the IR sequences do not exceed 5.85 kb and thus separate a large single copy (LSC) region of 64.6 kb from a small single copy (SSC) region of 48.3 kb.

In order to verify the presence of a flipping mechanism acting on the plastid genome of *Vaucheria* we isolated the four understoichiometric P 1/3 and P 1a/3a fragments from LGT agarose gels and subjected them to *Eco* RI digestion. Both pairs contained the *Eco* RI fragments E 4, 10, 13, 14, 16, 21, 22, 23, 24, 27, 28 (not shown), thus indicating their sequence identity. Hence it can be concluded that the circular plastid chromosome of *Vaucheria* exists in two equimolar isomers that differ from one another by the relative orientation of their single copy regions (Fig. 3).

Table 2. Corresponding primary and secondary restriction fragments obtained by reciprocal redigestions. Marginal subfragments are marked with an asterisk. Cloned *Eco* RI fragments are underlined.

Primary fragment	Secondary fragments generated with			
	Sal I	*Bam* HI	*Pvu* II	*Eco* RI
S1		1, 4*, 7*, 9	1*, 2, 3, 7, 8, 12, 16*, 18	2, 4, 6, 7, 8, 9, 11, 13, 14, 18, 21*, 22, 23, 24, 25*, 26, 27, 32, 33
S2		2	5*, 6*	5, 10, 12*, 35*
S3		3	4, 17*, 19*	1
S4		5	10*, 15, 20, 21, 23, 24*, 25	3*, 15*
S5		6*, 13*	11*, 13*	16, 19, 28*, 31*
S6		8*, 11*	9	17*, 29, 34*
S7		10*, 14*	14	20
S8		12	22	30
B1	1		1*, 2, 3*, 8, 13, 19	2, 5, 7, 8, 11, 13, 19, 20*, 23, 27, 28, 32, 33, 34*
B2	2, 5, 12, 13*, 14*		4, 6*, 11*, 18, 20, 21, 22, 23	3, 4, 9, 10*, 12*
B3	3, 6*, 7*		5, 7*, 9, 17*	1, 15, 16, 17*, 25, 31*
B4	4*, 8*		10*, 12*, 15, 24	6, 18*, 22, 26, 29, 30*
B5	10*, 11*		14	21*, 24*
B6	9		16	14
P1	1*, 17*	3*, 7*, 16		2, 11*, 14, 23, 24, 28, 37*
P2	6*, 9, 14, 16*, 22	6*, 10*, 14		3, 8*, 16, 21, 32, 44*
P3	2	1*, 12*		5, 15, 17, 25, 38*, 41, 42, 46*
P4	3	2		4, 13*, 29*
P5	4	5		1
P6	5*, 24*	4		6, 18*, 36*
P7	10*, 13*	11*, 17*		9, 26*, 27*
P8	7	8		10*, 30, 33*
P9	11*, 19*	9		20, 22*, 35*
P10	8	13		7
P11	12	15		12
P12	15	18		19
P13	18	19		31
P14	20	20		34
P15	21	21		39
P16	24	22		40
P17	25	23		43
P18	26	24		45
E1	1, 21*, 31*	1	1*, 11, 22*	
E2	2	2	7, 13*, 37*	
E3	3*, 35*	3	19, 26*, 34*, 36, 39, 40, 43	
$\overline{\text{E4}}$	4	14, 17*, 20*	2	
E5	5	4	8*, 18*	
$\overline{\text{E6}}$	12*, 20*, 30	10*, 21*	3	
$\overline{\text{E7}}$	6	5	10*, 31, 46*	
E8	7	6	12, 38*, 44*, 45	
E9	8	7	4	
E10	9	8	5	
$\overline{\text{E11}}$	10	9	6	
E12	15*, 28*	12*, 31*	9	
$\overline{\text{E13}}$	11	11	14	
E14	13	18*, 34*	15	
E15	17*, 36*	24*, 30*	16	
$\overline{\text{E16}}$	14	13	17	
E17	16	15	27*, 35*	
$\overline{\text{E18}}$	18	16	30*, 33*	
$\overline{\text{E19}}$	19	19	20	
E20	25*, 34*	22	21	
E21	22	23	23	
E22	23	25	24	
$\overline{\text{E23}}$	24	26	25	
$\overline{\text{E24}}$	26	27	28	
E25	27	28	29	
E26	29	29	32	
E27	32	32	41	
E28	33	33	42	

Table 3. Sizes (kb) of restriction fragments following single and double digestions. Restriction fragments from double digestions which are identical to primary fragments are marked with an asterisk.

Fragment	*Sal* I	*Bam* HI	*Pvu* II	*Eco* RI	S/B	S/Py	S/E	B/Py	B/E
1	67.9	46.6	22.3	15.6	46.6*	20.4	12.6*	11.4	15.6*
2	16.3	29.4	18.4	9.7	16.3*	16.3*	9.7*	11.2*	9.7*
3	12.6	26.4	16.3	8.6	12.6*	11.2*	8.3	10.6	8.6*
4	11.2	15.1	11.2	8.2	12.2	9.5*	8.2*	9.5*	7.3*
5	7.9	4.4	9.5	7.3	11.2*	8.6	7.3*	9.2*	6.1*
6	4.9	2.85	9.2	6.7	7.2	7.6	6.1*	8.8	5.9*
7	3.15		7.9	6.1	6.5	7.0*	5.9*	8.5	5.8*
8	0.7		7.0	5.9	3.1	5.0*	5.8*	7.0*	5.7*
9			5.8	5.8	2.85*	4.9*	5.7*	5.8*	5.6*
10			5.0	5.7	2.65	4.8	5.6*	5.3	4.3
11			4.1	5.6	1.7	4.5	3.6*	5.3	3.6*
12			2.7	4.4	0.7*	4.1*	3.4	5.3	3.55
13			1.4	3.6	0.65	3.2	3.1*	5.0*	3.0*
14			1.2	3.2	0.45	3.15*	3.0	4.4*	2.85*
15			0.75	3.1		2.7*	2.95	4.1*	2.85*
16			0.7	2.9		2.4	2.85*	2.85*	2.75*
17			0.55	2.85		2.0	2.8	2.7	2.7
18			0.5	2.7		1.4*	2.7*	2.7*	2.7
19				2.7		1.4	2.7*	1.4*	2.7*
20				2.4		1.2*	2.7*	1.2*	2.7
21				2.0		0.75*	2.1	0.75*	2.4
22				1.85		0.7*	2.0*	0.7*	2.4*
23				1.85		0.65	1.85*	0.55*	2.0*
24				1.7		0.6*	1.85*	0.5*	1.9
25				1.5		0.55*	1.8		1.85*
26				1.15		0.5*	1.8*		1.85*
27				0.68			1.5*		1.8*
28				0.65			1.45		1.5*
29							1.15*		1.15*
30							0.7		1.1
31							0.7		0.8
32							0.7*		0.7*
33							0.65*		0.65*
34							0.6		0.6
35							0.3		
36							0.3		
Total	124.6	124.8	124.5	124.5	124.7	125.0	124.5	124.8	124.7

Gene mapping

Twenty-three genes have been mapped on the chloroplast genome of *Vaucheria bursata*. These include genes coding for thylakoid polypeptides (photosystems I and II, electron transport chain, ATP synthase) as well as those coding for soluble proteins (RNA polymerase, elongation factor Tu of translation, Rubisco). In addition, the coding sites of the ribosomal RNAs have been determined.

The extension of the IR sequences as inferred from restriction site analysis was confirmed by Southern hybridizations using cloned fragments containing the rRNA genes of *Dictyota dichotoma* [19]. Since a *Bam* HI site separates the 16S from the 23S rRNA in *Dictyota*, and the sizes of the cloned fragments do not exceed the extension of

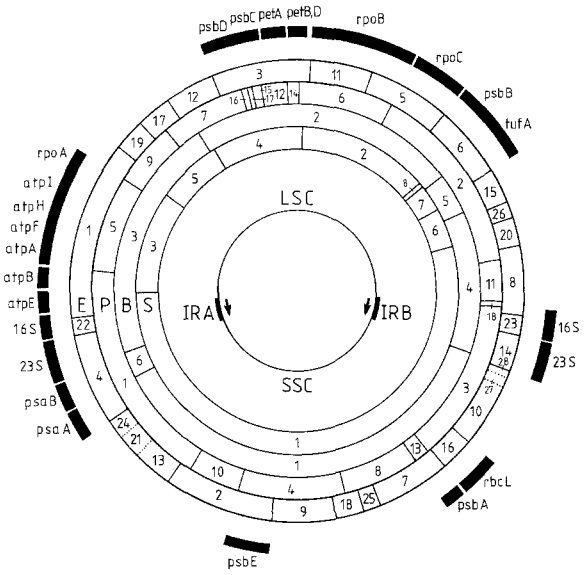

Fig. 2. Physical map of the plastid genome of *Vaucheria bursata* showing the restriction sites for *Sal* I (S), *Bam* HI (B), *Pvu* II (P), and *Eco* RI (E). The borders of fragments whose positions relative to one another remain undetermined, are dotted. Gene positions are indicated according to the smallest hybridizing fragments. The maximum extension of the IR sequences, as well as the large (LSC) and small (SSC) single-copy regions are shown on the innermost circle. Arrows mark the putative direction of transcription of the rRNA genes.

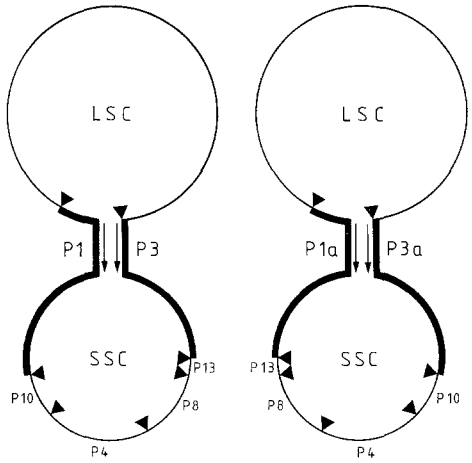

Fig. 3. Stem-loop illustration of the two isometric forms of the plastid genome of *Vaucheria bursata* following digestion with *Pvu* II. Arrows along the IR sequences indicate the inferred direction of transcription of the rRNA genes towards the SSC region. *Pvu* II restriction sites of the SSC region are marked by triangles, illustrating the inverse orientation of the SSC region relative to the LSC region. The maximum extension of the IR sequences towards the LSC region is indicated by the border of P3.

either the 16S or the 23S rRNA genes, the location of these genes on the plastid genome of *Vaucheria* is unequivocal. In particular, the 16S rRNA gene probe strongly hybridizes with E 22/23, but a faint signal with E 1 and E 8 becomes apparent after prolonged exposure (Fig. 4i). Correspondingly, the probe pDdBS10 containing the *Dictyota* 23S rRNA chloroplast gene hybridized extensively with E 4 and E 14, but also showed weak signals with E 22/23 (Fig. 4g). The existence of two isomeric populations of molecules has been verified additionally by the observation that the four half-stoichiometric fragments P 1/3 and P 1a/3a hybridized with the 23S gene probe (Fig. 4h). As expected, also the 16S gene probe hybridized with these *Pvu* II fragments involved in the high-frequency recombination process (not illustrated).

Four ATPase genes (*atpI, H, F, A*) that constitute a transcription unit in land plants hybridize with P 5 which is internal to SB 3 and E 1 (Fig. 4a–d). The hybridizations therefore cannot elucidate the serial order of these genes within P 5. However, as the plastid genomes of three other *Vaucheria* species (*V. aversa, V. racemosa, V. hercyniana*), which are collinear with the plastid genome of *V. bursata* with respect to gene order, clearly distinguish the coding sites for the ATPase genes (unpublished results), one may deduce that in *V. bursata* the genes *atpI*, *atpH*, *atpF* and *atpA* map in the given order, the gene *atpA* being proximal and *atpI* distal to the IR A sequence.

Next to *atpA* maps the gene *atpB* within SP 17 which overlaps S 3 and P 1 (Fig. 4e). Since the epsilon gene probe binds to E 1 (Fig. 4f) and SP 1 (not shown), the orientation of the *atpB* and *atpE* genes relative to one another is unambiguous. Occasionally, a second signal resulting from the *atpE* gene probe is encountered with P 14. By analogy with other chromophytic plastid genomes (unpublished data) we consider P 14 to be a site that contains some sequence similarity but not the functional gene.

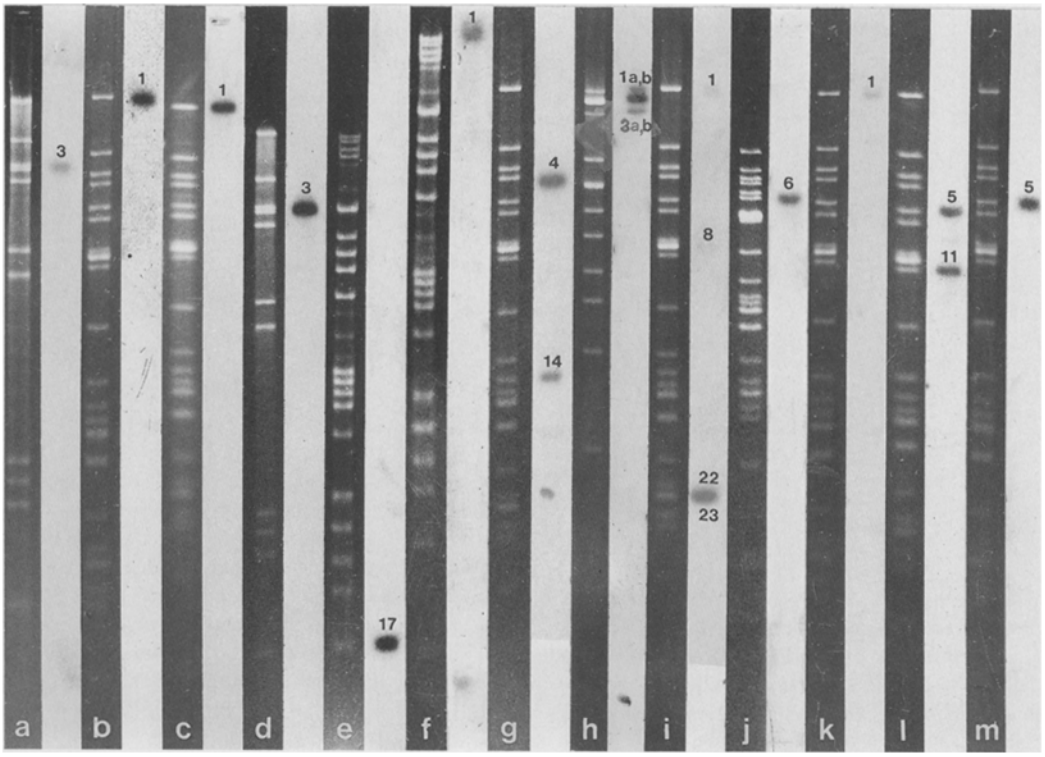

Fig. 4. Restriction patterns of *Vaucheria bursata* plastid DNA showing hybridization signals from heterologous gene probes. a, *atpI* (*Sal* I + *Bam* HI); b, *atpH* (*Eco* RI); c, *atpF* (*Eco* RI); d, *atpA* (*Sal* I + *Bam* HI); e, *atpB* (*Sal* I + *Pvu* II); f, *atpE* (*Eco* RI); g, *rrnL* (*Eco* RI); h, *rrnL* (*Pvu* II); i, *rrnS* (*Eco* RI); j, *tufA* (*Eco* RI); k, *rpoA* (*Eco* RI); l, *rpoB* (*Eco* RI); m, *rpoC* (*Eco* RI).

The genes *psaA* and *psaB* coding for the two P700 chorophyll *a* binding polypeptides of photosystem I (PS I) map adjacent to each other on the fragments SB 1 (*psaA*) and SB 9 corresponding to B 6 (*psaB*) within the SSC region (Fig. 5a, b). In addition, the *psaA* gene probe also hybridizes with B 6, whereas the *psaB* gene probe faintly binds to SB 1. This cross-reactivity which is known from the corresponding land plant genes [10, 16] also indicates some sequence homology for these genes in *Vaucheria*.

Five genes coding for photosystem II (PS II) polypeptides were localized at four distinct regions on the *Vaucheria* plastid genome. Two of them (*psbA*, Fig. 5c; *psbE*, Fig. 5g) map within the SSC region. Only the genes *psbC* and *psbD* hybridized to the same restriction fragment (SB 5, equivalent to S 4; Fig. 5e, f). The orientation of these genes relative to each other was evaluated by a fine restriction map of E 3 and by nucleotide sequencing of the *psbDC* operon (unpublished results). Some 20 kb apart from the *psbDC* operon is the coding site of the 51 kDa chlorophyll *a* apoprotein (*psbB*) within E 6 (Fig. 5d).

The positions of three genes for polypeptides of the photosynthetic electron transport chain that are plastid-encoded in land plants (*petA*, *petB*, *petD*) were determined within E 3, adjacent to the *psbDC* operon (Fig. 5h–j). Unlike the genes *petB* and *petD*, which indicate a high degree of sequence homology with the corresponding spinach gene probes, the signal resulting from the *petA* gene probe usually remained weak. Even under reduced stringency conditions, hybridizations required a strongly labelled probe.

As with the structural proteins, soluble polypeptides are also encoded in both single-copy regions. The Rubisco large subunit gene (*rbcL*) maps within E 16 (Fig. 5k) and P3/P3a (data not shown). It is not yet possible, however, to deter-

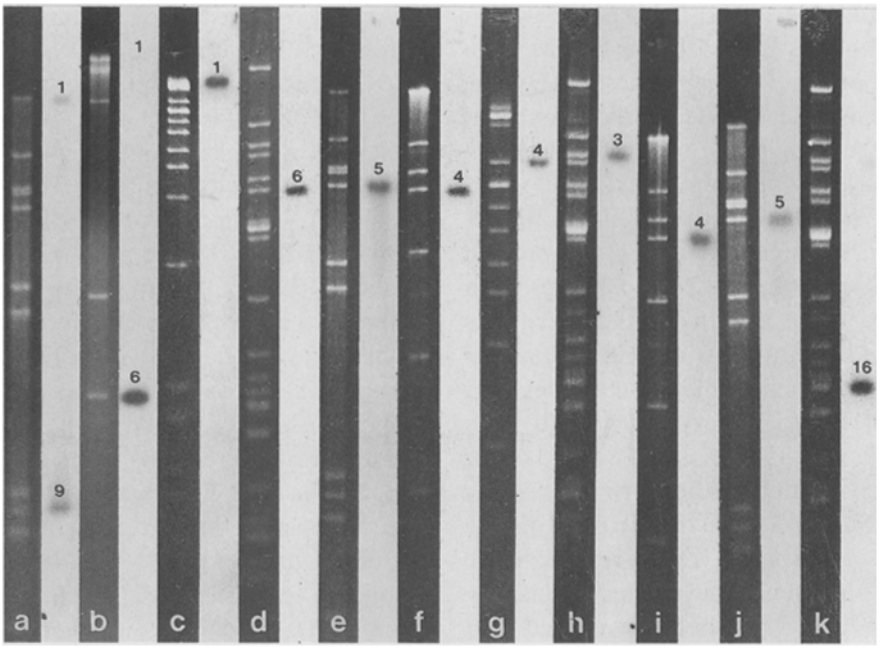

Fig. 5. Restriction patterns of *Vaucheria bursata* plastid DNA, together with hybridization signals obtained from heterologous gene probes. a, *psaA* (*Sal* I + *Bam* HI); b, *psaB* (*Bam* HI); c, *psbA* (*Pvu* II); d, *psbB* (*Eco* RI); e, *psbC* (*Sal* I + *Bam* HI); f, *psbD* (*Sal* I); g, *psbE* (*Pvu* II); h, *petA* (*Eco* RI); i, *petB* (*Sal* I); j, *petD* (*Sal* I + *Bam* HI); k, *rbcL* (*Eco* RI).

mine the position of this gene within the physical map more precisely, as the serial order of four *Eco* RI fragments, including E 16, which are internal to P 3, remains open.

The gene *tufA* can be localized close to the gene *psbB* within E 6, to which it strongly hybridizes even under high stringency conditions (Fig. 4j). Two genes, which code for subunits of the RNA polymerase (*rpoB*, *rpoC*), map adjacent to this site. Since *rpoB* hybridizes both with E 5 and E 11 (Fig. 4l) and *rpoC* exclusively with E 5 (Fig. 4m), the linear arrangement of these genes appears unequivocal. In comparison to these two genes, however, the presence and precise location of the *rpoA* gene encoding the α-subunit of the RNA polymerase appears questionable. Even under reduced stringency conditions only faint hybridization signals are obtained with E 1 (Fig. 4k).

Discussion

Restriction site and partial gene maps of chromophytic plastid genomes have now been reported for three diatoms [3, 17], two brown algae [18, 19, 23], and flagellates belonging to the Cryptophyceae [9], Chrysophyceae [5], and Rhaphidophyceae [37]. All these species exhibit features including circularity and segmentation that are in principle similar to those known from most chlorophytic plants. Noticeable differences include small IR sequences that do not largely extend the rRNA operon, a reduced genome size of about 120–125 kb, and a highly scrambled gene order [17]. Among the above mentioned chromophytes, only the plastid genome of the raphidophyte *Heterosigma* (formerly called *Olisthodiscus*) resembles that of land plants with respect to both its overall sequence complexity and the sizes of the IR sequences, but differs in gene arrangement and gene composition [37].

Restriction site analysis and electron microscopy [21] revealed that the plastid genome of *Vaucheria bursata* is unicircular. Furthermore, we have shown that this plastid genome exists in two equimolar isomers resulting from intramolecular recombination processes within the IR sequences,

similar to the plastid genomes of *Chlamydomonas* [33] and land plants [30]. Obviously, the small IR sequences of the *Vaucheria* plastid genome, which alone may explain the reduction in size of about 20% when compared to land plant chloroplast genomes, do not impair the flip-flop recombination. We consider such small IR sequences that are typical for most chromophytic plastid genomes [9, 17, 19, 23] to represent a primitive rather than a derived character. This view is substantiated by the observation that during the evolution of land plants IR sequences tend to increase in length by expanding into both single-copy regions [35].

A second major difference between the *Vaucheria* plastid genome and those of land plants is the presence of an enlarged SSC region relative to the LSC region. Consequently, genes that exclusively map within the LSC region of the conserved land plant chloroplast genomes are scattered over both single-copy regions in *Vaucheria*. There are even genes which reside in the LSC region of certain chromophytes [17], but which map within the SSC region in *Vaucheria* (*psaA*, *psaB*, *psbE*). If only inversions would have been responsible for this feature, one has to take into account a loss of the IR sequences, at least temporarily, which then would have changed into direct repeats. Such direct rDNA repeats which are dispersed within the circular molecule are, however, unknown for chloroplast genomes. They are prominent features of plant mitochondrial genomes and are considered to be responsible for the generation of subgenomic circles from a genome-sized master chromosome [34]. A shift of genes between the two single-copy regions in chromophytic algae is, therefore, more easily explained as the result of transposition events that were shown to cause rearrangements in subclover chloroplast DNA [27].

It appears that extensive gene scrambling is typical for chlorophyll $a+c$-containing algae. Nevertheless, there are certain gene clusters preserved in both chlorophyll $a+b$ and $a+c$ lineages. These include, besides the rRNA cistrons, genes which encode polypeptides of the PS I (*psaA*, *B*) and PS II (*psbD*, *C*) reaction centres. Among the clustered genes are also those coding for subunits of the electron transport chain (*petB*, *D*), of the RNA polymerase (*rpoB*, *C*) as well as of the F_0F_1 ATP synthase (*atpB*, *E*; *atpI*, *H*, *F*, *A*). Several such gene clusters, if not all, may constitute transcription units in *Vaucheria* (data not shown), similar to cyanobacteria [12], and chloroplasts from other chromophytes [5] and land plants [39, 42]. It remains to be shown whether the *atpA* gene cluster in *Vaucheria* also contains the genes *atpG* and *atpD*, as found in other chlorophyll $a+c$-containing algae (P.G. Pancic *et al.*, unpublished results; M.G. Kuhsel *et al.*, unpublished results).

On the other hand, genes including those of the *psbB* operon that are transcribed into a single mRNA in land plants [46] map at different positions in *Vaucheria*, similar to other chromophytes [17]. Thus, differences in gene order and mode of transcription are less significant in elucidating phylogenetic relationships among major algal lineages. Such relationships are rather pronounced in the primary structure of conserved genes. As an example, the two chlorophyll *a*-binding polypeptides of the PSI reaction centre of both cyanobacteria [4] and land plant chloroplasts [10, 16] share common sequences, suggesting a putative ancestral gene for *psaA* and *psaB*. Southern hybridizations show that such sequence similarities also exist for these two genes in *Vaucheria*, both of which cross-hybridize with either *psaA* or *psaB* from spinach. This feature strongly suggests a close phylogenetic relationship between cyanobacterial and probably all eukaryotic P 700 chlorophyll *a*-binding proteins.

Similarities in gene structure and gene assembly between plastid genomes of chlorophytes and chromophytes contrast strikingly, however, with specific characters of chromophytic plastid genomes. It is these genomes that will become increasingly important for evolutionary considerations. Among these characters are the number and order of ATP synthase genes, six of which (*atpA*, *B*, *E*, *F*, *H*, *I*) reside in the plastids of land plants, whereas three (*atpC*, *D*, *G*) are nuclear genes. It is suggested that these latter genes migrated from the ancestral prokaryotic genome to

the nucleus during the course of plastid evolution in chlorophyll $a+b$ plants [14, 31]. Unexpectedly, however, both in the brown alga *Dictyota dichotoma* (M.G. Kuhsel *et al.*, in preparation) and the diatom *Odontella sinensis* [36], the gene *atpD* encoding the delta subunit of CF_1 turned out to reside in the plastid genome at the same position as in cyanobacteria [7] and the cyanelles of *Cyanophora paradoxa* (D.A. Bryant, V.L. Stirewalt and M.B. Annarella, unpublished results). This again may be indicative of a much closer relationship between cyanobacteria and chromophytic plastids as compared to chlorophyll $a+b$ plastids. In addition, the *psbD* and *psbC* genes that overlap by 53 bp in land plant chloroplast genomes share 17 nucleotides in *Vaucheria* (unpublished results), as has been found in cyanobacteria [12]. There is thus no reason to argue for prokaryotes of the *Heliobacterium* type as the putative ancestors of chromophytic plastids [25].

Our results do not explain whether there was a single prokaryotic/eukaryotic endocytosis leading to the evolution of both the chlorophyll $a+b$ and $a+c$ lineages or whether several such events occurred, involving different types of cyanobacteria. The present data on chromophytic plastid genomes are meagre and contrast to the vast amount of information about chlorophytic plastid genomes. Nevertheless, differences in gene content between the two major photosynthetic lineages are indicative of specific evolutionary events that occurred within one lineage but not in the other. It may be possible that, due to secondary eukaryotic/eukaryotic endocytoses which gave rise to the now living chlorophyll $a+c$ algae, chromophytic plastid genomes evolved at a much faster rate than those of land plants. Different evolutionary rates are known, for instance, to occur in plant and animal mitochondrial genomes [32]. In a similar way, restriction site polymorphisms among strains of *Vaucheria bursata* [22], which are extremely rare in land plant chloroplast genomes [26], could be interpreted as a result of faster substitution rates in chromophytic plastid genome evolution.

Acknowledgements

We thank Dr R.G. Herrmann, University of Munich, Dr J.D. Palmer and Dr M.G. Kuhsel, University of Indiana, Bloomington, for providing gene probes and cloned restriction fragments from spinach and *Dictyota* chloroplast DNAs. We also thank Dr P.L. Beech, University of Cologne and Mrs N. Shukla for critical reading of the manuscript.

References

1. Alt J, Herrmann RG: Nucleotide sequence of the gene for pre-apocytochrome *f* in the spinach plastid chromosome. Curr Genet 8: 551–557 (1984).
2. Alt J, Morris J, Westhoff P, Herrmann RG: Nucleotide sequence of the clustered genes for the 44 kd chlorophyll *a* apoprotein and the '32 kd'-like protein of the photosystem II reaction center in the plastid chromosome. Curr Genet 8: 597–606 (1984).
3. Bourne CM, Stoermer EF, Palmer JD: Structure of the chloroplast genome in closely related species of *Cyclotella* (Bacillariophyceae). J Phycol 25 (Suppl): 5 (1989).
4. Cantrell A, Bryant DA: Molecular cloning and nucleotide sequence of the *psaA* and *psaB* genes of the cyanobacterium *Synechococcus* sp. PCC 7002. Plant Mol Biol 9: 453–468 (1987).
5. Cattolico RA, Loiseaux-de Goer S: Analysis of chloroplast evolution and phylogeny: a molecular approach. In: Green JC, Leadbeater BSC, Diver WL (eds) The Chromophyte Algae. Problems and Perspectives, vol 38, pp. 85–100. The Systematics Association/Clarendon Press, Oxford (1989).
6. Cozens AL, Walker JE: Pea chloroplast DNA encodes homologues of *Escherichia coli* ribosomal subunit S2 and the β' subunit of RNA polymerase. Biochem J 236: 453–460 (1986).
7. Cozens AL, Walker JE: The organization and sequence of the genes for ATP synthase subunits in the cyanobacterium *Synechococcus* 6301. Support for an endosymbiotic origin of chloroplasts. J Mol Biol 194: 359–383 (1987).
8. Dente L, Cesareni G, Cortese R: pEMBL: a new family of single stranded plasmids. Nucl Acids Res 11: 1645–1655 (1983).
9. Douglas SE: Physical mapping of the plastid genome from the chlorophyll *c*-containing alga *Cryptomonas* Φ. Curr Genet 14: 591–598 (1988).
10. Fish LE, Kück U, Bogorad L: Two partially homologous adjacent light-inducible maize chloroplast genes encoding polypeptides of the P_{700} chlorophyll *a*-protein complex of photosystem I. J Biol Chem 260: 1413–1421 (1985).

11. Gibbs SP: Chloroplasts of some algae groups may have evolved from endosymbiotic eucaryotic algae. Ann N Y Acad Sci 361: 193–207 (1981).
12. Golden SS, Stearns GW: Nucleotide sequence and transcript analysis of three photosystem II genes from the cyanobacterium *Synechococcus* sp. PCC 7942. Gene 67: 85–96 (1988).
13. Heinemeyer W, Alt J, Herrmann RG: Nucleotide sequence of the clustered genes for apocytochrome *b6* and subunit 4 of the cytochrome *b/f* complex in the spinach plastid chromosome. Curr Genet 8: 543–549 (1984).
14. Hennig J, Herrmann RG: Chloroplast ATP synthase of spinach contains nine nonidentical subunit species, six of which are encoded by plastid chromosomes in two operons in a phylogenetically conserved arrangement. Mol Gen Genet 203: 117–128 (1986).
15. Herrmann RG, Alt J, Schiller B, Widger WR, Cramer WA: Nucleotide sequence of the gene for apocytochrome b-559 on the spinach plastid chromosome: Implications for the structure of the membrane protein. FEBS Lett 176: 239–244 (1984).
16. Kirsch W, Seyer P, Herrmann RG: Nucleotide sequence of the clustered genes for two P_{700} chlorophyll *a* apoproteins of the photosystem I reaction center and the ribosomal protein S14 of the spinach plastid chromosome. Curr Genet 10: 843–855 (1986).
17. Kowallik KV: Molecular aspects and phylogenetic implications of plastid genomes of certain chromophytes. In: Green JC, Leadbeater BSC, Diver WL (eds) The Chromophyte Algae, Problems and Perspectives, vol. 38, pp. 101–124. The Systematics Association/Clarendon Press, Oxford (1989).
18. Kuhsel M, Kowallik KV: The plastome of a brown alga, *Dictyota dichotoma* I. Physical properties and the *Bam* HI/*Sal* I/*Bgl* II cleavage site map. Plant Mol Biol 4: 365–376 (1985).
19. Kuhsel M, Kowallik KV: The plastome of a brown alga, *Dictyota dichotoma*. II. Localization of structural genes coding for ribosomal RNAs, the large subunit of ribulose-1,5-biphosphate carboxylase/oxygenase and for polypeptides of photosystems I and II. Mol Gen Genet 207: 361–368 (1987).
20. Lawn RM, Fritsch EF, Parker RC, Blake G, Maniatis T: The isolation and characterization of linked α- and β globin genes from a cloned library of human DNA. Cell 15: 1157–1174 (1978).
21. Linne von Berg K-H, Schmid M, Linne von Berg G, Sturm K, Hennig A, Kowallik KV: The chloroplast genome (plastome) from algae of different phylogenetic relationships. Br Phycol J 17: 235 (1982).
22. Linne von Berg K-H, Kowallik KV: Structural organization and evolution of the plastid genome of *Vaucheria sessilis* (Xanthophyceae). BioSystems 21: 239–247 (1988).
23. Loiseaux-de Goer S, Markowicz Y, Dalmon J, Audren H: Physical maps of the two circular plastid DNA modules of the brown alga *Pylaiella littoralis* (L.) Kjellm. Location of the rRNA genes and of several protein-coding regions on both molecules. Curr Genet 14: 155–162 (1988).
24. Maniatis T, Fritsch EF, Sambrook J: Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1982).
25. Margulis L, Obar R: *Heliobacterium* and the origin of chrysoplasts. BioSystems 17: 317–325 (1985).
26. Milligan BG: Differentiation of chloroplast DNA within populations of *Trifolium pratense*. Fourth International Congress of Systematics and Evolution Biology, College Park, Maryland, USA (1990).
27. Milligan BG, Hampton JN, Palmer JD: Dispersed repeats and structural reorganization in subclover chloroplast DNA. Mol Biol Evol 6: 355–368 (1989).
28. Morris J, Herrmann RG: Nucleotide sequence of the gene for the P_{680} chlorophyll *a* apoprotein of the photosystem II reaction center from spinach. Nucl Acids Res 12: 2837–2850 (1984).
29. Ohme M, Tanaka M, Chunwongse J, Shinozaki K, Sugiura M: A tobacco chloroplast DNA sequence possibly coding for a polypeptide similar to *E. coli* RNA polymerase β subunit. FEBS Lett 200: 87–90 (1986).
30. Palmer JD: Chloroplast DNA exists in two orientations. Nature 301: 92–93 (1983).
31. Palmer JD: Plastid chromosomes: Structure and Evolution. In: Bogorad L, Vasil IK (eds) The Molecular Biology of Plastids, vol 7. In: Cell Culture and Somatic Cell Genetics in Plants, in press.
32. Palmer JD: Comparison of chloroplast and mitochondrial genome evolution in plants. In: Herrmann RG (ed.) Plant Gene Research, vol. 6: Organelles. Springer-Verlag, Heidelberg/New York, in press.
33. Palmer JD, Boynton JE, Gillham NW, Harris EH: Evolution and recombination of the large inverted repeat in *Chlamydomonas* chloroplast DNA. In: Molecular Biology of the Photosynthetic Apparatus, pp. 269–278. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
34. Palmer JD, Herbon LA: Tricircular mitochondrial genomes of *Brassica* and *Raphanus*: Reversal of repeat configurations by inversion. Nucl Acids Res 14: 9755–9764 (1986).
35. Palmer JD, Nugent JM, Herbon LA: Unusual structure of *Geranium* chloroplast DNA: A triple-sized inverted repeat, extensive gene duplications, multiple inversions, and two repeat families. Proc Natl Acad Sci USA 84: 769–773 (1987).
36. Pancic PG, Strotmann H, Kowallik KV: The δ-subunit of the chloroplast ATPase is plastid-encoded in the diatom *Odontella sinensis*. FEBS Lett 280: 387–392 (1991).
37. Reith M, Cattolico RA: Inverted repeat of *Olisthodiscus luteus* chloroplast DNA contains genes for both subunits of ribulose-1,5-bisphosphate carboxylase and the 32,000-dalton Qb protein: phylogenetic implications. Proc Natl Acad Sci USA 83: 8599–8603 (1986).

38. Rieth A: Xanthophyceae. In: Ettl H, Gerloff J, Heynig H (eds) Süsswasserflora von Mitteleuropa, Band 4. Gustav Fischer Verlag, Stuttgart/New York (1980).
39. Shinozaki K, Ohme M, Tanaka M, Wakasugi T, Hayashida N, Matsubayashi T, Zaita N, Chunwongse J, Obokata J, Yamaguchi-Shinozaki K, Ohto C, Torazawa K, Meng BY, Sugita M, Deno H, Kamogashira T, Yamada K, Kusuda J, Takaiwa F, Kato A, Tohdoh N, Shimada H, Sugiura M: The complete nucleotide sequence of the tobacco chloroplast genome: its gene organization and expression. EMBO J 5: 2043–2049 (1986).
40. Sijben-Müller G, Hallick RB, Alt J, Westhoff P, Herrmann RG: Spinach plastid genes coding for initiation factor IF-1, ribosomal protein S11 and RNA polymerase α-subunit. Nucl Acids Res 14: 1029–1044 (1986).
41. Stosch HA von, Drebes G: Entwicklungsgeschichtliche Untersuchungen an zentrischen Diatomeen IV. Die Planktonalge *Stephanopyxis turris* – ihre Behandlung und Entwicklungsgeschichte. Helgol Wiss Meersunters 11: 209–257 (1964).
42. Tanaka M, Obokata J, Chunwongse J, Shinozaki K, Sugiura M: Rapid splicing and stepwise processing of a transcript from the *psbB* operon in tobacco chloroplasts: Determination of the intron sites in *petB* and *petD*. Mol Gen Genet 209: 427–431 (1987).
43. Waris H: The significance for algae of chelating substances in the nutrient solutions. Physiol Plant 6: 538–543 (1953).
44. Watson JC, Surzycki SJ: Extensive sequence homology in the DNA coding for elongation factor Tu from *Escherichia coli* and the *Chlamydomonas reinhardtii* chloroplast. Proc Natl Acad Sci USA 79: 2264–2267 (1982).
45. Westhoff P, Nelson N, Bünemann H, Herrmann RG: Localization of genes for coupling factor subunits on the spinach plastid chromosome. Curr Genet 4: 109–120 (1981).
46. Westhoff P, Herrmann RG: Complex RNA maturation in chloroplasts. The *psbB* operon from spinach. Eur J Biochem 171: 551–564 (1988).
47. Whatley JM, Whatley FR: Chloroplast evolution. New Phytol 87: 233–247 (1981).
48. Zurawski G, Perrot B, Bottomley W, Whitfeld PR: The structure of the gene for the large subunit of ribulose 1,5-bisphosphate carboxylase from spinach chloroplast DNA. Nucl Acids Res 9: 3251–3270 (1981).
49. Zurawski G, Bohnert HJ, Whitfeld PR, Bottomley W: Nucleotide sequence of the gene for the M_r 32,000 thylakoid membrane protein from *Spinacia oleracea* and *Nicotiana debneyi* predicts a totally conserved primary translation product of M_r 38,950. Proc Natl Acad Sci USA 79: 7699–7703 (1982).
50. Zurawski G, Bottomley W, Whitfeld PR: Structure of the genes for the β and ε subunits of spinach chloroplast ATPase indicate a dicistronic mRNA and an overlapping translation stop/start signal. Proc Natl Acad Sci USA 79: 6260–6264 (1982).