Curr Genet (1986) 10:931–941

Current Genetics

The location and possible evolutionary significance of small dispersed repeats in wheat ctDNA

Catherine M. Bowman and Tristan A. Dyer

Plant Breeding Institute, Maris Lane, Trumpington, Cambridge CB2 2LQ, UK

Summary. Low-stringency hybridisation between recombinant plasmids representing the complete *T. aestivum* chloroplast genome has revealed small repeated DNA segments dispersed through the molecule. Thirty-two repeated DNA segments were detected, and they could be divided into 12 unrelated sets; no repeat was detected as multiple copies. The longest of the small repeats mapped just within the large inverted repeat in spinach and mung-bean ctDNAs. It was found to have been duplicated after the divergence of a cereal progenitor to generate a third, dispensible copy, 0.2 kbp downstream of *rbc*L. In maize at least, this copy has also become integrated, with *rbc*L, in the mitochondrial genome. Another of the repeats is thought to have mediated a chloroplast DNA inversion (Howe 1985). Thus the diverse collection of small repeats probably represents some consequences and causes of past recombination events as well as a mechanism for further intramolecular ctDNA recombination. Their possible significance in the restructuring and evolution of chloroplast genomes is discussed.

Key words: *T. aestivum* – Chloroplast DNA – Repeat DNA – Evolution

Introduction

Gross alterations involving DNA deletion, insertion and inversion are common contributors to the variation seen among higher plant chloroplast DNAs (ctDNAs) (for review see Palmer 1985). The variant ctDNAs of the closely-related wheats and goatgrasses have accumulated a typical spectrum of alterations during their evolution (Bowman et al. 1983). In contrast to the seemingly gradual accumulation of wheat ctDNA alterations during evolutionary time, many different and gross rearrangements are also found in the chloroplast genomes of wheat (and barley) albino seedlings recovered from anther culture (Day and Ellis 1984, 1985). This implies that plastid DNA rearrangement can occur during the development and/or culture of the anthers, and that sequences allowing that rearrangement are present in hexaploid wheat ctDNA.

Recombination is a major cause of gross DNA alteration, and repeated DNA sequences are a feature common to many recombination systems. Homologous recombination (reciprocal and non-reciprocal) between repeated sequences is proposed as an important mechanism for the restructuring and evolution of nuclear and mitochondrial genomes (see Flavell 1986; Dover et al. 1982; Scherer and Davis 1980; Jeffreys et al. 1985; Lonsdale et al. 1984). Repeat DNA is also an integral part of transposable elements and insertion sequences IS's), which are known to mediate DNA insertion, deletion and inversion by non-homologous (and sometimes homologous) recombination mechanisms (e.g. see Iida et al. 1983). DNA segments inverted by site-specific recombination also carry terminal inverted repeats (Watson 1984; Broach 1982).

Recombination has been demonstrated in chloroplast genomes. In the lower eukaryote, *Euglena gracilis*, homologous recombination between the tandemly-repeated chloroplast ribosomal RNA operons generates deletion and insertion mutants (Nicolas et al. 1985). In higher plants, the multiplicity of the chloroplast genome and its predominantly maternal inheritance means that ctDNA recombination is rarely observed. However, in a recent and elegant experiment, protoplasts from ctDNA mutants of two *Nicotiana* species were fused, and a regenerant with a recombinant ctDNA phenotype was selected.

Offprint requests to: C. M. Bowman

Examination of its ctDNA showed that there had been extensive exchange of segments between the two parental ctDNAs by intermolecular homologous recombination (Medgyesy et al. 1985).

It was therefore decided to examine the ctDNA of *T. aestivum* for small repeated sequences which might indicate the kinds of mechanism responsible for the DNA rearrangements observed in *T. aestivum* ctDNA following anther culture, and in the variant ctDNAs of the divergent wheats and goatgrasses. This was done by low-stringency hybridisation of recombinant clones representing the complete *T. aestivum* chloroplast genome. The more interesting small repeats were then mapped more precisely. The possible association of one of the repeats with known ctDNA alterations was studied by hybridisation to ctDNAs containing defined deletions and inversions, and also by hybridisation to segments of "promiscuous" ctDNA from maize mitochondria (Stern and Lonsdale 1982).

Materials and methods

Isolation of DNA and restriction endonuclease analysis. Chloroplast DNA was isolated from non-aqueous chloroplasts of the following: cultivated wheat *(Triticum aestivum)* var. "Mardler", the wild diploid wheats and goatgrasses *T. boeoticum, T. monococcum, Aegilops bicornis* and *Ae. squarrosa,* also mung bean *(Vigna radiata),* pea *(pisum sativum)* var. "Feltham First", and spinach *(Spinacea oleracea)* var "Hybrid 102", as described elsewhere (see Bowman et al. 1983).

Plasmid DNA was isolated by the method of Birnboim and Doly (1979).

Zea mays ctDNA and DNA from cosmids containing *Z. mays* mitochondrial DNA (mtDNA) were a gift from Dr. D. Lonsdale.

Samples of DNA were digested with restriction endonucleases as described by the suppliers of the enzymes (Bethesda Research Laboratories). Digests were fractionated by electrophoresis in horizontal slab gels of 0.85% or 1% agarose, or in 5% polyacrylamide gels.

Transfer of DNA to nitrocellulose and hybridisation to ^{32}P-labelled probes. DNA digests, fractionated by agarose gel electrophoresis, were transferred to Schleicher and Schull BA85 nitrocellulose, essentially as described by Wahl et al. (1979).

Probes were made from complete recombinant plasmid DNA, or from cloned wheat ctDNA inserts recovered from plasmid DNA digests by electroelution from agarose gels. The DNA (0.1 to 0.5 μg) was labelled with ^{32}P by nick-translation (Rigby et al. 1977).

In experiments to detect small repeated sequences in ctDNA, low stringency hybridisation conditions were used to allow stable hybrids to form between sequences which were short or of incomplete homology. The prehybridisation and hybridisation medium was: 0.6 M NaCl, 4 mM EDTA, 100 μg/ml carrier DNA, 2% SDS, 10X Denhardt's solution, 20 mM PIPES pH 6.8. Hybridisation was overnight at 55 °C. After rinsing in 4 x SSC, unhybridised material was removed from the nitrocellulose by three

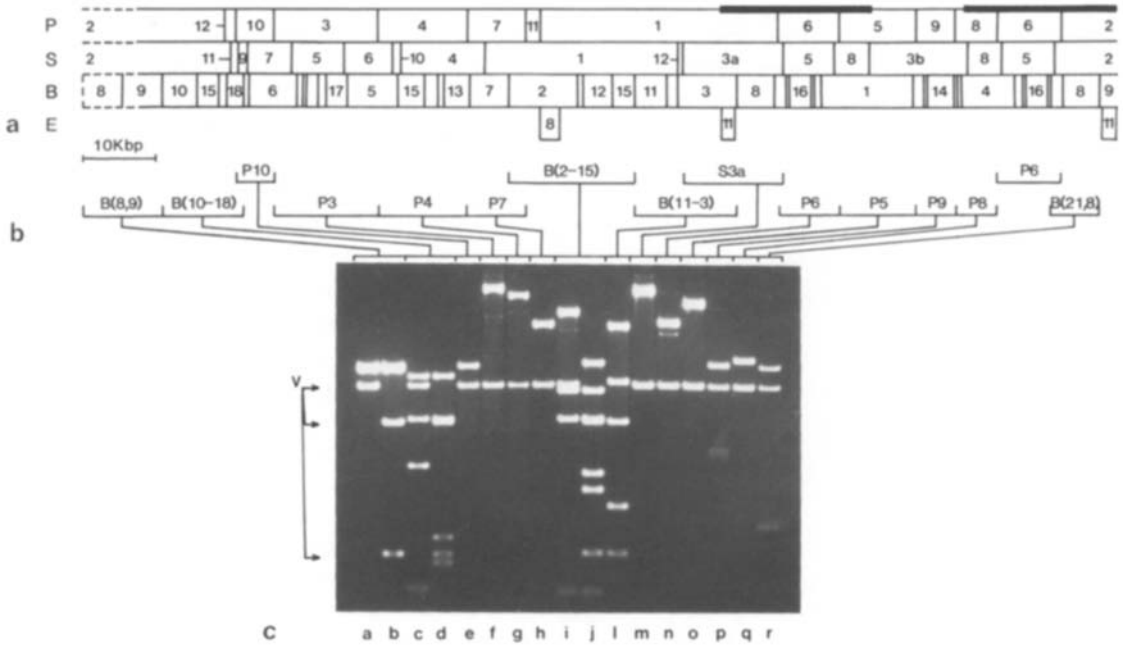

Fig. 1a–c. Physical map of wheat ctDNA aligned with recombinant plasmids representing the genome. **a** Location of target sites for the restriction endonucleases PstI (P), SalGI (S) and BamHI (B) and the location of two relevant EcoRI fragments (E). Although circular, the map is drawn in linear form with the large IR *(thickened upper line)* to the right. Fragments drawn in *dotted lines* have been added on the left of the map for convenient alignment of the plasmids. **b** Segments of wheat ctDNA contained by individual pTac recombinant plasmids chosen to represent the genome (for nomenclature see Materials and methods). **c** Fractionation by agarose gel electrophoresis, of the pTAc- plasmid DNAs digested (a–r) as described in the legend to Fig. 2. Vector (pBR322) DNA bands are *arrowed*

30 min washes at the same stringency as the hybridisation, i.e. at 55 °C in 4 x SSC (0.6 M NaCl. 0.06 M Na citrate) containing 0.1% SDS. The wet nitrocellulose sheets were wrapped in Clingfilm and exposed to X-ray film using intensifying screens at −70 °C. To detect the faint signals due to hybrids between short repeat sequences it was necessary to over-expose the film with respect to the signal due to hybridisation betwen long homologous sequences (e.g. Fig. 2). The stability of the weak hybrids was tested by subsequently giving the nitrocellulose two 30 min washes at successively higher stringencies (the stringency was increased by reducing the salt concentration). Those hybrids no longer detectable after washing in 0.2 x SSC at 55 °C were considered unstable.

For the detection of hybrids between long homologous sequences, filters were hybridised overnight at 65 °C using the medium described above, and washed at 65 °C in 2 x SSC.

Cloning of wheat ctDNA fragments and mapping of Bam HI target sites. Recombinant plasmids containing wheat ctDNA PstI fragments were obtained by "shotgun" cloning a PstI digest of the DNA using the vector pBR322 and the RecA$^-$ host NEM 256. Clones containing the largest PstI fragments P1 (33.0 kbp) and P2 (20.0 kbp) were not recovered. Recombinant plasmids containing some SalGI fragments, cloned using the vector pBR322 and the RecA$^+$ host, C600, were a gift from L. Hanley-Bowdoin.

In order to complete a bank of overlapping clones of the wheat chloroplast genome, and to order the BamHI fragments, to allow the detection and mapping of the short repeats, wheat ctDNA was partially digested with BamHI and the partial fragments were cloned using the vector pBR322 and the RecA$^-$ host ED8767. Approximately 500 recombinant (Amp^R Tet^S) clones were obtained. Two hundred of these were analysed directly by BamHI digestion of minipreparations of the plasmid DNA. Appropriate BamHI fragments were isolated from those clones that contained underrepresented regions of the genome. These fragments were used as probes to screen the remaining clones by colony hybridisation (Grunstein and Hogness 1975) to recover all those recombinants containing ctDNA from this region.

The resulting Bam-partial recombinant plasmids contained between 2 and 5 BamHI fragments. Because they were not obtained from size-fractionated ctDNA digests, their validity was thoroughly tested and eighty Bam-partial recombinant clones were finally stored to complete the bank of overlapping clones of the wheat chloroplast genome. They contained sufficient incremental sets of BamHI fragments to establish a map of the BamHI restriction sites as shown in Fig. 1. Thirty-eight differently-sized fragments were detected, but only BamHI fragments B1 to B18 inclusive are numbered on the map. Not all the target sites are represented at this resolution.

The recombinant clones are named according to the wheat ctDNA restriction fragments they contain. In the case of the clones obtained by partial BamHI digestion, the ctDNA insert fragment is defined by the two BamHI fragments that border the inserted ctDNA and are therefore adjacent to the vector. For example, pTacB(8,9) contains Bamfragments 8 and 9, whereas pTacB(2–15) contains Bam fragments B2, B26, B12 and B15 (see Fig. 1). The clone nomenclature is therefore to be used with reference to the map. (Further information concerning either the wheat ctDNA BamHI map or the recombinant clones is available on request).

Results

The detection of small repeated segments in wheat ctDNA

A set of fourteen recombinant plasmids containing contiguous BamHI fragments or individual PstI or SalGI fragments was chosen from the bank to represent the entire *T. aestivum* chloroplast genome. The DNA from each clone was digested with appropriate restriction enzymes such that, when fractionated by electrophoresis in 0.85% agarose, no ctDNA band migrated too close to a vector DNA band and all the fragments were contained by the gel. An example of such a gel is shown in Fig. 1. The segment of wheat ctDNA represented by each plasmid is indicated. Several such gels were run, and the fractionated DNA was transferred to nitrocellulose. Each filter was then hybridised with a ^{32}P-labelled DNA probe made from one of the set of recombinant plasmids, or in the case of pTacB(2–15), a portion of it. The following recombinant clones were used as probes to represent the wheat ctDNA region containing Bam fragments 2–15: pTacB2, pTacE8 and pTacB(26-15). Each ctDNA fragment was therefore probed with all the ctDNA in the remainder of the genome. Hybridisation was carried out at low stringency to preserve hybrids between short or incomplete homologies, and the filters were then autoradiographed and washed at successively higher stringencies, as described in Materials and methods, to compare the stability of different hybrids.

Figure 2 shows examples of autoradiographs from 3 filters probed with recombinant plasmid DNA from pTacP7, pTacP10 or pTacE8. The faint bands, marked with white dots on the autoradiographs, and coinciding with ctDNA bands on the original gels (e.g. in Fig. 1), are due to authentic hybrids between ctDNA sequences present in the probe fragment and also repeated elsewhere in the wheat chloroplast genome. The intensity of the band gives a rough indication of the length and/or degree of homology in the hybridising sequence, and the identity of the hybridising fragment locates the repeat approximately on the wheat chloroplast genome.

Using this approach, faint bands that are due to contamination or partial digestion, are less likely to mislead because such faint bands are only considered valid if they coincide with ctDNA insert bands on the gel.

The autoradiograph in Fig. 2 (panel 1) shows hybridisation of pTacP7 to ctDNA fragments B15 (track c) and P10 (track e). The hybrids were not stable in 0.2 x SSC at 55 °C (panel 2). This type of result was typical of most probes. Hybridisation to P8, detected on all filters, was not reciprocated when P8 was used as a probe and was therefore considered spurious.

The most interesting hybridisation patterns were observed when two particular recombinant plasmids were used as probes: pTacP10 (see autoradiograph panels 3

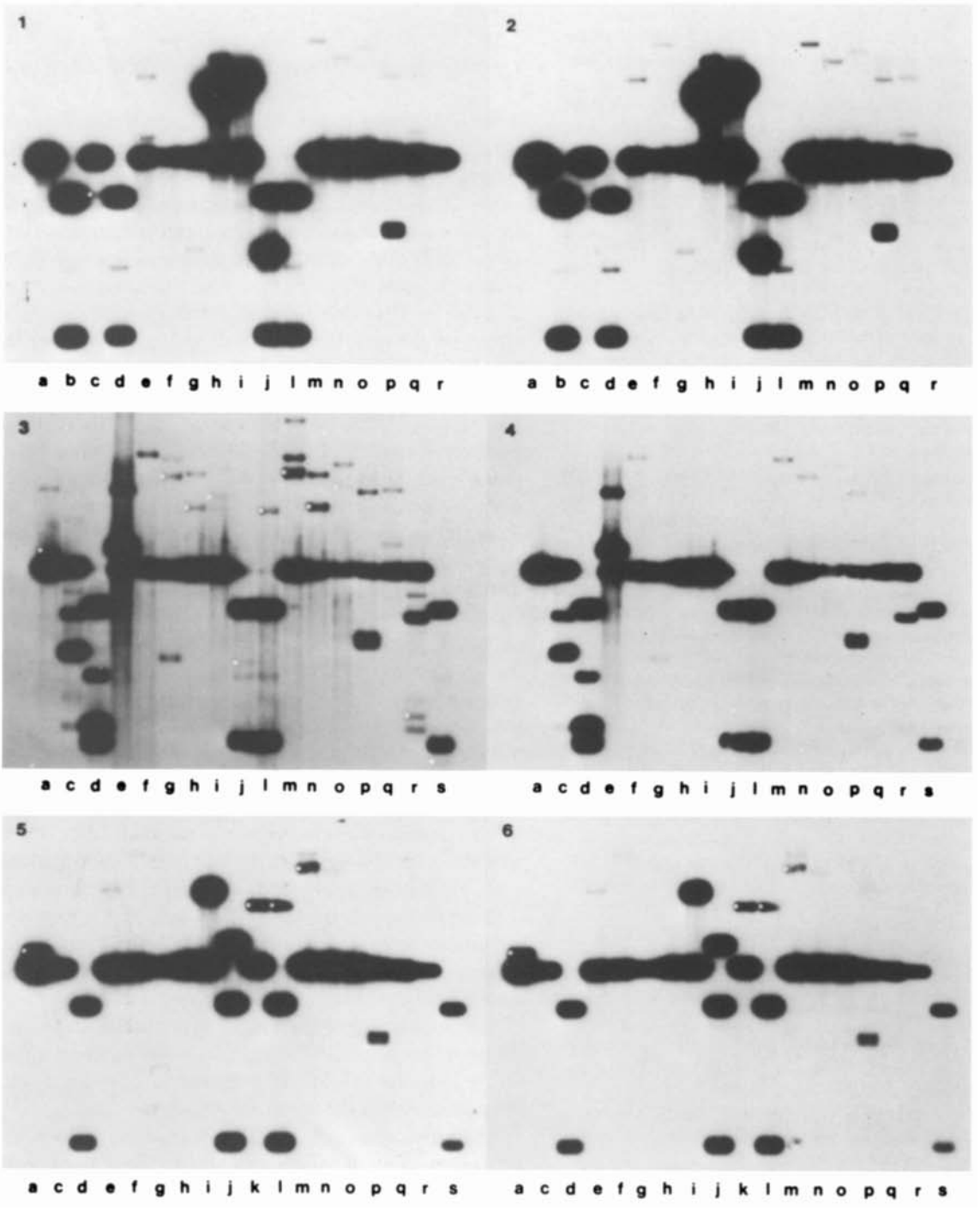

Fig. 2. Detection of small repeated sequences in wheat ctDNA. Tracks in all panels represent the pTac- plasmid DNAs shown in Fig. 1, digested as follows: *a, b,* pTacB(8.9) digested with *a* BamHI and *b* BamHI/PstI; *c, d,* pTacB(10–18) digested with *c* BamHI and *d* BamHI/PstI; *e–h,* PstI digest of *e* pTacP10, *f* pTacP3, *g* pTacP4 and *h* pTacP7; *i, j,* pTacB(2–15) digested with *i* BamHI and *j* BamHI/PstI; *k, l,* pTacB(11–13) digested with *k* BamHI and *l* BamHI/PstI; *m* pTacS3a digested with SalGI; *n–q* PstI digests of *n* pTacP6, *o* pTacP5, *p* pTacP9 and *q* pTacP8; *r, s,* pTacB(21,8) digested with *r* BamHI and *s* BamHI/PstI. Lettering is not consecutive on all autoradiographs. Hybridisation probes were pTacP7 (panels 1 and 2), pTacP10 (panels 3 and 4) and pTacE8 (panels 5 and 6). Filter washing was at 55 °C in either 4 x SSC (panels 1, 3 and 5) or 0.2 x SSC (panels 2, 4 and 6). Two sets of bands on all autoradiographs are due to hybridisation between vector DNA sequences. These are the very intense bands coinciding with major vector bands on the gels (e.g. see Fig. 1), and the faint bands that coincide with very faint bands on the gels due to partial digestion of the plasmid DNA. Very intense bands also appear on each pair of autoradiographs after hybridisation to an individual probe. These bands coincide on the original gels (see Fig. 1) with ctDNA bands in the plasmid from which the probe was made, or with bands comprising ctDNA that overlaps the probe fragment(s) in the chloroplast genome. Bands representing hybrids between authentic ctDNA repeats are marked on their left with a *white dot*

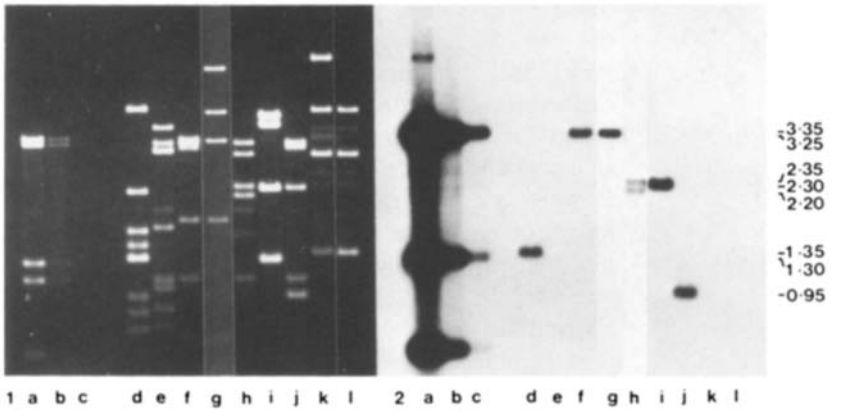

Fig. 3. Mapping of the "P10" small unstable repeats in wheat ctDNA. *Tracks a–c*, pTacP10 digested with BamHI/PstI and loaded at *a* 1-fold, *b* 10-fold and *c* 100-fold dilution; *d*, pTacB(8,9) digested with EcoRI/HindIII; *e*, pTacP3 digested with BamHI/PstI; *f, g*, pTacP4 digested with *f* BamHI/PstI and *g* BamHI; *h*, pTacP7 digested with BamHI/PstI/SalGI; *i*, pTacB3 digested with BamHI/EcoRI; *j*, pTacP6 digested with BamHI/PstI; *k, l*, SalGI/Pst digest of *k* pTacP5 and *l* pTacP8. *Panel 1*, fractionation of the digests by agarose gel electorphoresis. *Panel 2*, autoradiography after hybridisation with the isolated wheat ctDNA fragment P10

and 4) and pTacE8 (panels 5 and 6). Fragment P10 hybridised with 9 other ctDNA fragments (see panel 3): B(8,9), P4 and P7, the segment of B2 that overlaps P7, and also B3, S3a, P6, P5 (faintly) and B21. All of these hybrids were unstable in 0.2 x SSC (panel 4). The repeated segments contained by P10 were therefore called small unstable repeats. Fragment E8 hybridised with fragments B(8,9), B3, and S3a (see panel 5), and all these hybrids were stable in 0.2 x SSC (panel 6). The repeated DNA segment(s) contained by E8 was therefore called the "small stable repeat".

More precise mapping of small repeated sequences

After the initial screening, plasmids containing ctDNA that hybridised with pTacP10 and pTacE8 were analysed further where it was considered necessary to locate the repeated DNA more precisely. To simplify the interpretation of the autoradiographs, the P10 and E8 fragments were isolated from the plasmid DNA for use as probes.

The results of hybridisation of P10 to mapped restriction fragments from pTacB(8,9), pTacP3, pTacP4, pTacP7, pTacB3, pTacP6, pTacP5 and pTacP8 DNA are shown in Fig. 3. Fragment P3 DNA was included as a control, and serial dilutions of digests of cloned P10 DNA were included for comparison of hybridisation signal strength. As expected, the probe did not hybridise with P3 (track e). Also, the initial hybridisation with fragments P5 (track k) and P8 (track l) was not reproducible. Hybridisation of P10 to the mapped subfragments of the other plasmid DNAs did confirm the initial results, and located the repeats on the wheat ctDNA as summarized in Fig. 5. Hybridisation of P10 with a 2.3 kbp Eco fragment (E11) of B3 (track i) and a 1.35 kbp HindIII/Eco subfragment of E11 in B(8,9) (track d) maps this repeat just inside the large IR (repeat 3 in Fig. 5). Hybridisation to the 3.25 kbp Bam fragment (B15) in Bam/Pst (track f) and Pst (track g) digests of P4, diagnoses repeat No. 2 in Fig. 5. In track h, hybridisation is shown to a 2.35 kbp fragment and a 2.2 kbp Bam/Pst fragment of P7. These hybrids locate repeats 4 and 5 in Fig. 5. Finally, hybridisation to a 0.95 kbp Sal/Pst fragment of P6 (track j) maps repeat no. 6 in Fig. 5. Hybridisation of P10 with the majority of these fragments (Fig. 3, tracks d–j) gives signals comparable in intensity with those observed when P10 hybridises with subfragments of itself, but at approximately 100-fold dilution (Fig. 3, track c). Fragment P10 is 5.2 kbp long. Most of the repeats having a segment in P10 are therefore of the order of 50 bp long, or are longer and mismatched. Repeats no. 4 and 5 (Fig. 5) found in P7 (Fig. 3, track h) are exceptions.

Hybridisation of the isolated E8 fragment to mapped restriction fragments from pTacB(8,9), pTacB3, and pTacS3a DNA (not shown) located the small stable repeat to the same 2.3 kbp Eco fragment E11, found in the large IR, that hybridised with P10. The E11 fragment was isolated and used as a probe to locate the small stable repeat more accurately on the fragment E8. The autoradiograph in Fig. 4 (panel 2) shows hybridisation to the common 2.3 kbp E11 fragment itself in digests of cloned B3 and B(8,9) DNA (tracks a and b respectively). Hybridisation is shown to the 2.7 kbp fragment E8 (track c) and to a 0.35 kbp Xba fragment of E8 (track d) that is 0.2 kbp downstream of the *rbc*L coding region (sequence to be presented elsewhere). The location of the small stable repeat on the wheat ctDNA map is shown as repeat 9 in Fig. 5. Given that the entire repeat is contained by E8 (track c), the intensity of hybridisation to the 0.35 kbp Xba fragment (track d) suggests that this small fragment contains most of the small stable repeat.

Figure 5 summarizes the results from all the hybridisation experiments. The experimental design limited the detection of small repeats to those detectable by routine autoradiography (>20 bp) and those present in different restriction enzyme fragments. By cross-referencing all

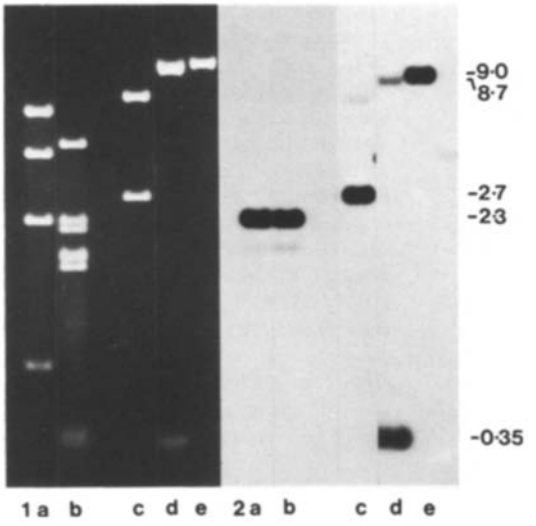

Fig. 4. Mapping of the "E8" small stable repeat in wheat ctDNA. *Panels1* and *2: track a,* pTacB3 digested with EcoRI; *b,* pTacB-(8,9) digested with BamHI/EcoRI; *c–e,* pTacE8 digested with *c* EcoRI, *d* XbaI and *e* HindIII. *Panel 1,* fractionation of the digests by agarose gel electrophoresis. *Panel 2,* autoradiography after hybridisation with the isolated wheat ctDNA fragment E11

the hybrids, it was possible to deduce the smallest ctDNA fragment within which a given repeat segment was found. These fragments are boxed in Fig. 5. Occasionally a probe hybridised, apparently convincingly, with a given fragment but the hybridisation was not reciprocated when that fragment was used as a probe. (For example, the probe P9 hybridised with P10, but not the reverse, and the probe B2 hybridised with P3 and P4 but not the reverse). Such discrepancies may have been due to artifactual hybridisation, or to differences in the specific radioactivity of the different probes. Such hybrids are not included in Fig. 5.

When one fragment (e.g. P4 or P10) hybridised with several others, it was usually possible to deduce whether this fragment contained several unrelated repeats or a single segment that was also repeated many times elsewhere in the genome. In Fig. 5, the fragments containing releated repeats are drawn on the same line and linked. Unrelated repeats are drawn on different lines.

Using this approach it was possible to describe 12 different sets of small repeats in wheat ctDNA, which together contributed a total of 32 repeated segments. Fig. 5 illustrates their dispersion through the genome. There was no evidence of a single segment being repeated many times. The hybridisation of P10 with several different ctDNA fragments (Figs. 2, 3) rather reflects the fact that P10 contains at least 5 unrelated repeats (repeats 2–6 inclusive).

While it is recognised that these hybridisation experiments do not reveal all the small repeats in wheat ctDNA, it is probably still significant that the detectable repeats are not scattered randomly through the genome (see Fig. 5). Of the 12 sets of repeats, 5 have one copy in the region represented by the overlapping fragments P10

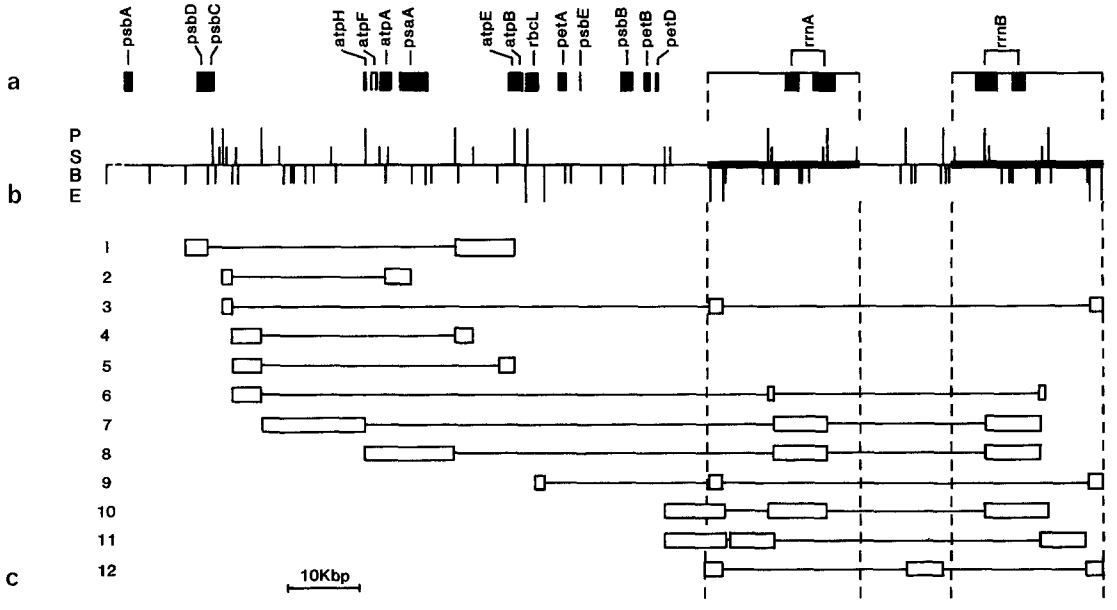

Fig. 5a–c. Location of some small repeated sequences in the wheat chloroplast genome. **a** Position of mapped genes. For original mapping references see Courtice et al. (1985) Genes are named according to Hallick and Bottomley (1984). **b** Physical map showing target sites for restriction endonucleases PstI (P), SalGI (S) and BamHI (B) and the position of two EcoRI fragments (E). The map is drawn as in Fig. 1, the *thickened line* representing the large IR. **c** Location of the detectable small repeats. The extent of the large IR is indicated by the *vertical dotted lines.* The *boxes* represent the smallest fragment within which a given repeat is found. Fragments containing related repeats are drawn on the same line and linked. Repeats drawn on different lines are unrelated

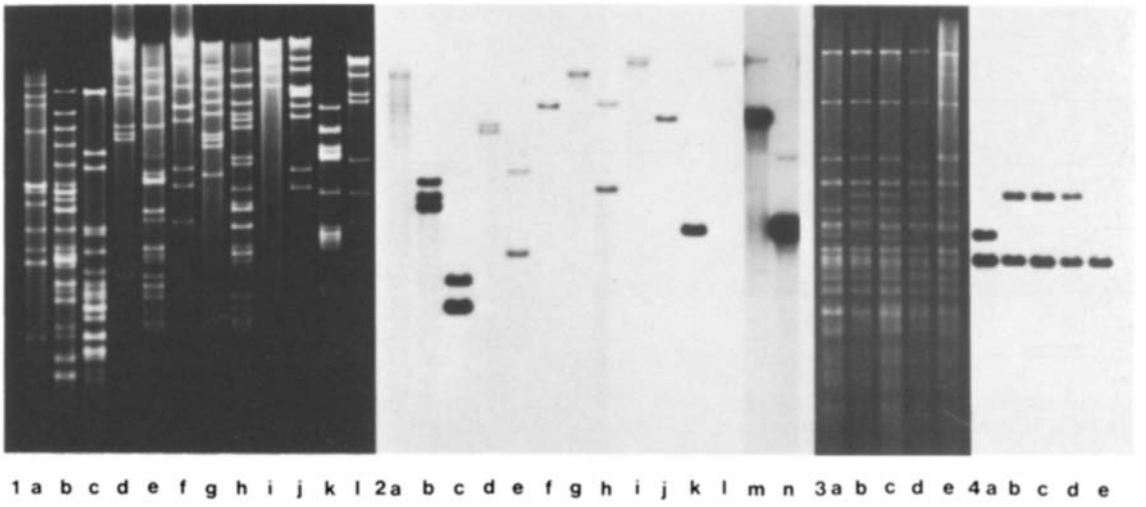

Fig. 6. Detection of sequences related to the wheat small stable repeat in some altered ctDNAs. *Panels 1* and *2: tracks a–c,* maize ctDNA digested with *a* PstI, *b* BamHI and *c* EcoRI; *d–f,* spinach ctDNA digested with *d* PstI, *e* BamHI and *f* SalGI; *g–i,* mung-bean ctDNA digested with *g* PstI, *h* BamHI and *i* SalGI; *j–n,* pea ctDNA digested with *j, m* PstI, *k, n* BamHI and *l* KpnI. *Panels 3* and *4:* EcoRI digests of ctDNA from *a T. aestivum, b Ae. bicornis, c T. monococcum, d T. boeoticum* and *e Ae. squarrosa. Panels 1* and *3:* fractionation of the digests by agarose gel electrophoresis. *Panels 2* and *4:* autoradiography after hybridisation with the isolated *T. aestivum* ctDNA 0.35 kbp XbaI fragment

and B18. This region represents only 6% of the single-copy DNA but contains 30% of the total number of small repeats detected in the single-copy regions of the genome. It may also be significant that 8 of the 12 repeat sets have a segment in one copy of the large IR.

The small stable repeat and ctDNA alterations

The small stable repeat maps within 0.2 kbp of the gene *rbc*L (see Fig. 5), and the position of *rbc*L has been used to diagnose many ctDNA alterations. If spinach ctDNA is considered archetypal, *rbc*L lies near one end-point of two (probably independent but similar) inversions that have taken place in mung bean (Palmer and Thompson 1981) and *Oenothera* (Herrmann et al. 1983) ctDNAs respectively. The position of *rbc*L is also changed in the much-rearranged pea ctDNA (Palmer and Thompson 1981) and it borders two different deletions in *Triticum* and *Aegilops* ctDNA (Bowman et al. 1983). Coding sequence of *rbc*L has also been detected in two segments of "promiscuous" ctDNA found in maize mitochondria (see Lonsdale et al. 1983).

To see whether the small stable repeat has been involved in the ctDAN alterations described above, the 0.35 kbp Xba fragment that contains most of the repeat, but no *rbc*L coding sequence, was isolated and used as a probe to locate related sequences in some of the altered ctDNAs. Hybridisations to the mung-bean, pea, maize and spinach ctDNAs were done at low stringency because the extent to which the sequences may have diverged was not known.

The results shown in Fig. 6 demonstrate that a sequence related to the *T. aestivum* small stable repeat was detectable in all the ctDNAs tested. Digests of ctDNA from maize, spinach, mung-bean and pea are shown in panel 1 and autoradiography of the hybridising fragments is shown in panel 2. The 0.35 kbp XbaI fragment hybridised as follows: in maize ctDNA to PstI fragments of 16.9 kbp and 15.3 kbp (tracks (a) and by deduction), BamHI fragments of 5.3 kbp, 4.6 kbp and 4.3 kbp (track b) and EcoRI fragments of 2.7 kbp and 2.2 kbp (track c); in spinach ctDNA to PstI fragments of 8.9 kbp and 8.1 kbp (track d), BamHI fragments of 5.9 kbp and 3.0 kbp (track e) and a double SalGI fragment of 10.6 kbp (track f); in mung-bean ctDNA to PstI fragments of 17.2 kbp and 16.2 kbp (track g), BamHI fragments of 10.8 kbp and 5.1 kbp (track h) and SalGI fragments of 24.4 kbp and 20.5 kbp (Track i) and lastly, in pea ctDNA to PstI fragments of 10.3 kbp (tracks j and m) and 21.6 kbp (track m), BamHI fragments of 4.8 kbp (tracks k and n) and 6.4 kbp (track n) and a KpnI fragment of 21.5 kbp (track l).

Figure 6, panel 3 shows EcoRI digests of ctDNA from five closely related species of wheat and goatgrass that each have a different ctDNA type (Bowman et al. 1983). The ctDNAs of *T. aestivum* itself (track a) and of *Ae. squarrosa* (track e) have suffered different deletions within 2 kbp 3′ of *rbc*L (see Fig. 5), the other three types (tracks b–d) are normal in this respect. The autoradiograph in panel 4 shows that, as expected, the 0.35 kbp XbaI fragment hybridises with the 2.3 kbp double Eco fragment E11 (see Fig. 5) in all these species. In

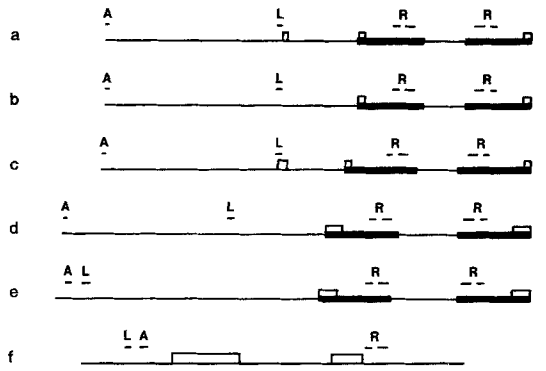

Fig. 7. Location of sequences related to the wheat small stable repeat in some altered ctDNAs. The maps are drawn as in Fig. 1 but the restriction enzyme sites are omitted. The *thickened line* represents the large IR. The smallest fragment containing sequences related to the small stable repeat is *boxed.* For orientation of the maps three genes are shown: psbA (A), rbcL (L) and rrn (R). The maps represent ctDNA from *a, T. aestivum, Ae. bicornis, T. monococcum,* and *T. boeoticum; b, Ae. squarrosa; c,* maize; *d,* spinach; *e,* mung-bean; *f,* pea

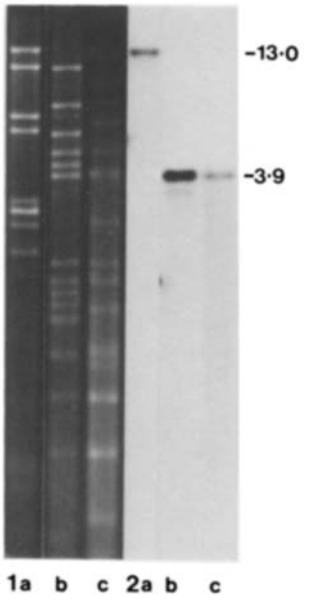

Fig. 8. Detection of sequences related to the wheat small stable repeat in the "promiscuous ctDNA" of maize mitochondria. *Panels 1* and *2:* tracks *a, b,* cosmid 2c87 DNA digested with *a* BamHI and *b* EcoRI; *c,* cosmid 2c7 DNA digested with EcoRI. *Panel 1,* fractionation of digests by agarose gel electrophoresis. *Panel 2,* autoradiography after hybridisation with the isolated *T. aestivum* ctDNA 0.35 kbp XbaI fragment

the "undeleted" forms of the ctDNA there is also hybridisation to a 3.4 kbp EcoRI fragment E4a (Bowman et al. 1983) (tracks b–d). In *T. aestivum* ctDNA, deletion has reduced the size of this hybridising fragment to 2.7 kbp (i.e. E8) (Fig. 5) (track a) and the small stable repeat is present. In *Ae. squarrosa* ctDNA however, a larger deletion has reduced the size of the homologous EcoRI fragment to 2.2 kbp (E13) and hybridisation to such a fragment is not detectable (track e).

These hybridisations allow sequences related to the small stable repeat to be located on the ctDNA maps for these species as shown in Fig. 7. As typified by *T. aestivum* ctDNA (Fig. 5), a homologous sequence is located close to both junctions of the large IR with the large single-copy region, in all those ctDNAs that do contain an inverted repeat (maps a to e inclusive). However, only in ctDNA from *T. aestivum,* and "undeleted" forms of wheat and goatgrass (map a) and from maize (map c) does the small stable repeat also map 3′ to *rbc*L. This repeated DNA segment has apparently been deleted from *Ae. squarrosa* ctDNA (map b) and may never have been present adjacent to *rbc*L in the ctDNA of spinach, mung-bean and pea (maps d, e and f respectively). Although pea ctDNA no longer contains an inverted repeat (Palmer and Thompson 1981), sequences homologous with the small stable repeat map to a fragment that extends, at most, 11.0 kbp away from the 16S rRNA gene (map f). This homology probably represents part of what remains of one copy of the large IR. There was also additional but extremely faint hybridisation to the 21.6 kbp fragment P1 and a 6.4 kbp BamHI fragment. It is not possible to deduce from the nearby protein-coding genes (see Courtice et al. 1985) whether this limited homology represents vestigial DNA from the "lost" copy of the large IR, or a rearranged segment of the copy that remains.

Hybridisation of the *T. aestivum* ctDNA 0.35 kbp XbaI fragment to mapped fragments from two maize mtDNA cosmids that contain promiscuous ctDNA (Stern and Lonsdale 1982) is demonstrated in the autoradiograph of Fig. 8. One of the cosmids (2c8) contains the gene *rbc*L, the other (2c7) overlaps 2c8 in the mitochondrial genome and contains *rbc*L plus a 12 kbp segment originating from the maize ctDNA large IR. The two promiscuous ctDNA segments are separated in the mitochondrial genome by "authentic" mtDNA. (For a map of the cosmids see Lonsdale et al. 1983.)

As shown in panel 2, the 0.35 kbp XbaI fragment hybridised with a 13.0 kbp BamHI fragment of 2c87 (track a). This is a vector/insert junction fragment. In tracks (b) and (c) hybridisation is to a 3.9 kbp EcoRI fragment common to both cosmids. In the mtDNA this EcoRI fragment maps in the DNA region overlapped by the two cosmids, and contains *rbc*L. Thus it appears that the small stable repeat maps near *rbc*L in the promiscuous ctDNA found in maize mitochondria. There was no evidence of any homologous sequences associated with the 12 kbp promiscuous segment from the ctDNA inverted repeat.

Discussion

The results described in this paper demonstrate that wheat ctDNA contains several small repeated sequences in addition to the large (21 kbp) IR. Twelve sets of repeats were identified, which together contribute 32 detectable small repeated segments to the wheat chloroplast genome. Do the characteristics of these repeats indicate the recombination mechanism(s), homologous or nonhomologous, that may have generated them, and do the repeats now represent the potential for further homologous recombination in wheat ctDNA?

There is of course a limitation in attempting to recognise possible recombination events from traces of participant repeats, particularly when studying evolutionary alterations to DNA. The approach is valid only when the sequence responsible for any recombination event that alters a region of DNA, is then retained in subsequent alterations to that same DNA region.

The low copy-number of the individual repeats shows that wheat ctDNA contains no multiple repeats of the type that can generate many different rearrangements in a single genome as a result of the many alternative copies of the repeat that can pair and recombine (see Flavell 1986; Jeffreys et al. 1985). Although it is not known how many copies of a particular repeat are represented by a given hybrid, all the repeats can be divided into 12 unrelated sets; none is present as multiple copies through the genome. The apparent size of most of the repeats is too small to be characteristic of "mobile" elements that are replicated during non-homologous mechanisms such as transposition. The small stable repeat is large enough to be a terminal repeat or insertion sequence (e.g. see Iida et al. 1983), but it is now known not to be so (data to be presented elsewhere). Small multiple dispersed repeats and larger repeated elements have, however, been found in some chloroplast genomes (see review of Palmer 1985): *Chlamydomonas reinhardtii* ctDNA contains 25–40 copies of small (0.1–0.3 kbp) inverted repeats scattered through the genome, and at least five copies of a larger (0.2–1.0 kbp) repeat element are apparently present in ctDNA from clover *(Trifolium subterraneum)*. Neither of these categories of repeat resembles those found in wheat ctDNA.

The small size of the wheat ctDNA repeats seems at first to argue against their possible involvement in homologous recombination. Rare intermolecular homologous recombination has recently been demonstrated between the chloroplast genomes of *Nicotiana* somatic hybrids causing mutual exchange of DNA segments (Medgyesy et al. 1985). Frequent intramolecular recombination between the 2 copies of the large IR is also the most likely, but not the only, explanation for the equimolar proportions of the two possible inversion isomers detected in ctDNA from several species (Palmer 1983; 1985). However, the mechanism of the inferred inversion is not known (see Palmer 1983). A site-specific mechanism is possible by analogy with DNA inversion in prokaryotes (Watson 1984) or yeast 2-micron DNA (see Broach 1982), but homologous recombination is equally likely. In these two examples, long stretches of DNA are available for homologous pairing and strand exchange, but the repeated sequences detected in the wheat ctDNA "single-copy" region are on average probably less than 100 bp. However, the assay of spontaneous deletions in *E. coli* with Rec^+ and Rec^- backgrounds has shown homologous recombination between sequences as short as 15 bp. The recombination frequency is also directly related to repeat length (Albertini et al. 1982).

Thus it is possible that the small repeats in wheat ctDNA do represent the potential for and/or results of homologous recombination involving the single-copy region of the genome. Their presence provides circumstantial evidence that at least some of the insertions, inversions and deletions observed in ctDNA have been generated by this mechanism.

This general interpretation is supported by a specific example, in which DNA sequencing has indicated the mediation of one of the "P10" repeats, in a particular ctDNA alteration (Howe 1985). This repeat (no. 2) is 70 bp long and has been mapped to the end-points of a $\simeq$ 20 kbp inversion that occurred in ctDNA at or after the divergence of the cereals (Palmer and Thompson 1981). The inversion was not simple and the repeats are now in a direct orientation. At this stage there is no reason to believe that repeat no. 2 is unique in this respect, so it is equally likely that some of the other repeats detected in the wheat ctDNA single-copy region may be/have been similarly involved in ctDNA restructuring. As well as containing an end-point of the inversion described above, the "variable" $\simeq$ 6 kbp P10–B18 region of wheat ctDNA has also sustained at least 3 deletions and an insertion during the more recent divergence of the closely-related wheats and goatgrasses (Bowman et al. 1983; Ogihara and Tsunewaki 1982). The location of copies of 5 of the repeat sets in this same DNA region (Fig. 5) may therefore be significant.

The size of the wheat small stable repeat allowed its further study by hybridisation to total ctDNA preparations from other plants. Several observations had suggested a possible role of this repeat in known ctDNA rearrangements: the location of 2 copies of the sequence within the large IR (a segment of ctDNA thought to recombine) in wheat, the proximity of the third copy of the sequence to *rbc*L and the proximity of *rbc*L in turn to several independent ctDNA alterations. It therefore seemed possible that the small stable repeat contained a recombination site common to several different ctDNA alterations.

However, the mapping of sequences related to the small stable repeat on examples of some of the altered DNAs (Figs. 6, 7) showed that this was not so. The location of a sequence homologous with the small stable repeat near both junctions of the large IR with the large single-copy region, in all the ctDNAs tested that contain a large inverted repeat, suggests that this is the original position of the sequence (Fig. 7). The absence of an extra copy in the large single-copy region of pea, spinach and mung-bean ctDNA (Fig. 7) further suggests that this segment of DNA was probably duplicated to its position 3′ of *rbc*L after the divergence of a cereal progenitor. This consequently implies that such a sequence was never adjacent to *rbc*L in mung-bean and pea progenitor ctDNAs to mediate the inversion and other rearrangements that have relocated the gene *rbc*L in these plants. In the *Triticum* and *Aegilops* ctDNAs, the "undeleted" forms of the genome do contain a third copy of the small stable repeat. The 0.7 kbp deletion sustained by *T. aestivum* ctDNA 3′ to *rbc*L has left the repeat in place (Fig. 7). However, in the 1.2 kbp deletion sustained in the same location by *Ae. squarrosa* ctDNA, the third copy of the small stable repeat has itself been deleted, arguing that this copy of the sequence is dispensible.

It thus appears that the small stable repeat is a result of ctDNA recombination. The "E8" region of the chloroplast genome between *rbc*L and *pet*A (see Fig. 5) must be either unstable and/or tolerant to change. During evolution it has sustained inversions, deletions and in this particular example, the insertion of a small segment of DNA from the large IR, which, becoming dispensible has subsequently been deleted. The presence of the small stable repeat in the promiscuous ctDNA of the maize mitochondrion can be interpreted in the same way. It is a dispensible piece of DNA from a variable region of maize ctDNA, which has been integrated into the mitochondrial DNA on the segment of DNA containing *rbc*L, presumably by the same mechanism that has integrated other ctDNA segments into the mitochondrial genome (for review, see Lonsdale 1985). The repeat itself is probably not part of that mechanism. By the same token, the significance of the clustering of repeats and DNA alterations in the B18–P10 region of the genome, may simply be that this is a region similarly tolerant to change.

To summarise: at this level of analysis the apparent size of the wheat ctDNA repeats indicates that they are large enough to undergo homologous recombination, but at present there is no evidence of their involvement in other recombination mechanisms. Two regions of the genome that have tolerated several DNA alterations do contain repeated sequences. One of these (repeat no. 2) is thought to have mediated a DNA inversion (Howe 1985) while another (repeat no. 9) was shown to have arisen in the single-copy region of the genome by duplication of a distant segment of the inverted repeat.

Thus the diverse collection of small repeats probably represents some causes and results of past recombination events as well as a mechanism for further DNA alteration by homologous recombination in wheat ctDNA.

In conclusion, what part might these small dispersed repeats play in the evolution of wheat chloroplast genomes? The generation of the repeats is in itself a facet of wheat ctDNA evolution that needs to be understood; perhaps mechanisms such as gene conversion are operating.

Once generated, what is the chance that small dispersed repeats will allow ctDNA alterations that then become fixed? The factors that facilitate recombination between sequences that can recombine, and the factors that affect the fixing and transmission of DNA alterations in ctDNA populations are so complex (eg. Birky 1983) that the fixing of neutral ctDNA alterations will not be discussed here. However, it is at least clear that alterations to DNA required for chloroplast function will be fixed only if they are not lethal. Functional constraints alone predict that future recombinations between several of the repeats described in this study are unlikely to contribute to the detectable rearrangement of ctDNA in a green wheat plant. For example, intramolecular homologous recombination between at least 2 copies of any of the 8 repeats that occur in the large IR and also in a single-copy region, will excise a large and necessary part of the genome. The same is true for repeat no. 2 which in its present orientation is a direct repeat (Howe 1985). In the special case of the albino plants recovered from wheat anther culture the situation is different; they can survive with at least photosynthetically non-functional chloroplasts, and their ctDNAs do contain many different large deletions (Day and Ellis 1984). These unusual mutants may therefore be interpreted as revealing the recombination potential present in wheat ctDNA, that can be realised if recombination is somehow facilitated, and functional constraints are removed. The results described in this paper indicate that at least some of that potential is provided by small dispersed repeats.

Acknowledgements. We thank Dr. R. B. Flavell for many helpful discussions.

References

Albertini A, Hofer M, Calos MP, Miller JH (1982) Cell 24:319–328

Birky CW Jr (1983) Science 222:468–475

Birnboim HL, Doly J (1979) Nucleic Acids Res 7:1513–1523

Bowman CM, Bonnard G, Dyer TA (1983) Theor Appl Genet 65: 247–262

Broach JR (1982) Cell 28:203–204

Courtice GRM, Bowman CM, Dyer TA, Gray JC (1985) Curr Genet 10:329–333
Day A, Ellis THN (1984) Cell 39:359–368
Day A, Ellis THN (1985) Curr Genet 9:671–678
Dover GA, Brown S, Coen E, Dallas J, Strachan T, Trick M (1982) The dynamics of genome evolution and species differentiation. In: Dover GA, Flavell RB (eds) Genome evolution. Academic Press, London New York, p 343
Flavell RB (1986) Phil Trans R Soc Lond B 312:227–242
Grunstein M, Hogness DS (1975) Proc Natl Acad Sci USA 72: 3961–3965
Hallick RB, Bottomley W (1984) Plant Mol Biol Rep 1:38–43
Herrmann RG, Westhoff P, Alt J, Winter P, Tittgen J, Bisanz C, Sears BB, Nelson N, Hurt E, Hauska G, Viebrock A, Sebald W (1983) Identification and characterisation of genes for polypeptides of the thylakoid membrane. In: Cifferi O, Dure L III (eds) Structure and function of plant genomes. Plenum Press, New York London, p 143
Howe CJ (1985) Curr Genet 10:139–145
Iida S, Meyer J, Arber W (1983) Prokaryotic IS elements. In: Shapiro J (ed) Mobile genetic elements. Academic Press, New York London, p 159
Jeffreys A, Wilson V, Thein SW (1985) Nature (London) 314: 67–73
Lonsdale DM, Hodge TP, Howe CJ, Stern DB (1983) Cell 34: 1007–1014
Lonsdale DM (1985) Chloroplast DNA sequences in the mitochondrial genome of maize. In: van Vloten-Doting L, Groot GSP, Hall TC (eds) Molecular form and function of the plant genome. Plenum Publishing Corporation, p 421
Lonsdale DM, Hodge TP, Fauron CM-R (1984) Nucleic Acids Res 12:9249–9261
Medgyesy P, Fejes E, Maliga P (1985) Proc Natl Acad Sci USA 82:6960–6964
Nicolas P, Ravel-Chapuis P, Heizmann P, Nigon V-M (1985) Curr Genet 10:229–238
Ogihara Y, Tsunewaki K (1982) Jpn J Genet 57:371–396
Palmer JD (1983) Nature (London) 301:92–93
Palmer JD (1985) Annu Rev Genet 19:325–354
Palmer JD, Thompson WF (1981) Proc Natl Acad Sci USA 78: 5533–5537
Rigby PWJ, Dieckmann M, Thodes C, Berg P (1977) J Mol Biol 133:273–251
Scherer S, Davis RW (1980) Science 209:1380–1384
Stern DB, Lonsdale DM (1982) Nature (London) 299:698–702
Wahl GM, Stern M, Stark GR (1979) Proc Natl Acad Sci USA 76:3683–3687
Watson M (1984) Trends in Biochem Sci 9:82–83

Communicated by C. J. Leaver

Received January 17 / March 17, 1986