Plant Molecular Biology **22**: 195–205, 1993.

Structure of the amplified 5-enolpyruvylshikimate-3-phosphate synthase gene in glyphosate-resistant carrot cells

Hyang Suh[1], Angus G. Hepburn[2], Alan L. Kriz[3] and Jack M. Widholm*
*Department of Agronomy, University of Illinois, Plant and Animal Biotechnology Laboratory, 1201 W. Gregory Ave., Urbana, Il 61801, USA (*author for correspondence); current addresses: [1] Agricultural Biotechnology Institute, Rural Development Administration, Suweon, 441-701, Korea; [2] Profile Diagnostic Sciences, Inc., 510 E. 73rd St., New York, NY 10021, USA; [3] DEKALB Plant Genetics, 62 Maritime Drive, Mystic, CT 06355, USA*

Received 21 October 1992; accepted 16 February 1993

Key words: carrot, cell culture, 5-enolpyruvylshikimate-3-phosphate synthase, gene amplification

Abstract

The structure of amplified 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS) DNA of carrot suspension-cultured cell lines selected for glyphosate resistance was analyzed to determine the mechanism of gene amplification in this plant system. Southern hybridization of the amplified DNA digested with several restriction enzymes probed with a petunia EPSPS cDNA clone showed that there were differences in fragment sizes in the amplified DNA from one highly resistant cell line in comparison with the parental line. Cloning of the EPSPS gene and 5′ flanking sequences was carried out and two different DNA structures were revealed. A 13 kb clone contained only one copy of the EPSPS gene while a 16 kb clone contained an inverted duplication of the gene. Southern blot analysis with a carrot DNA probe showed that only the uninverted repeated DNA structure was present in all of the cell lines during the selection process and the inverted repeat (IR) was present only in highly amplified DNA. The two structures were present in about equal amounts in the highly amplified line, TC 35G, where the EPSPS gene was amplified about 25-fold. The presence of the inverted repeat (IR) was further verified by resistance to S1 nuclease hydrolysis after denaturation and rapid renaturation, showing foldback DNA with the IR length being 9.5 kb. The junction was also sequenced. Mapping of the clones showed that the size of the amplified carrot EPSPS gene itself is about 3.5 kb. This is the first report of an IR in amplified DNA of a target enzyme gene in selected plant cells.

Introduction

The enzyme 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS) catalyzes the conversion of shikimate-3-phosphate (S-3-P) and phosphoenolpyruvate (PEP) to EPSP. EPSP is utilized to make chorismate, which is the branch point substrate in the aromatic amino acid biosynthetic pathway. Glyphosate (*N*-phosphonomethylglycine) is a competitive inhibitor with respect to PEP and an uncompetitive inhibitor with regards to the other substrate, S-3-P, in the EPSPS reaction [28].

Glyphosate, which is a herbicide lacking spec-

ificity between weeds and crops, has been used as a selective agent for microorganisms and higher-plant cells. These biological systems use two different mechanisms to overcome EPSPS inhibition by glyphosate. One is to produce glyphosate-tolerant, mutant EPSPS enzyme which has a different affinity for one or both substrates as well as for glyphosate. Selected lines from *Salmonella typhimurium*, *Aerobacter aerogenes* and *Escherichia coli* are examples (reviewed by Kishore and Shah [14]). The second mechanism is to produce elevated levels of EPSPS enzyme activity as has been reported for *A. aerogenes* and plant cells. The molecular basis of the EPSPS activity increase can be either by an increase in gene expression (*Corydalis sempervirens* [12]) or by gene amplification (petunia [24], carrot [11, 25, 26, 32] and tobacco [2, 9, 30]). Although the amplification of the EPSPS gene has been well documented, the actual mechanism has not been determined.

Only two other cases of gene amplification in plants have been reported and these were also in response to selection pressure using herbicides. In one of these cases the glutamine synthetase (GS) inhibitor phosphinothricin was used. Stepwise selection of alfalfa cells with phosphinothricin led to the isolation of cell lines which had elevated levels of GS activity as well as an increased number of GS genes [1]. The amplified alfalfa GS gene was cloned and the 5′ and 3′ flanking sequences were used to show that the size of the amplified units was larger than 35 kb [29].

Very recently, a tobacco cell line was selected with gradually increasing levels of sulfonylurea herbicides [10]. The line contained 6 to 7 times the normal activity of the target enzyme, acetolactate synthase. One of the two acetolactate synthase genes found in tobacco was amplified about 20-fold. The gene that was amplified also contained a base change which rendered the enzyme more resistant to inhibition by the herbicides.

The plant gene amplification systems studied thus far appear to be different from the intensively studied mammalian cancer cell cultures [27] since amplification occurs more slowly and to much lower levels in plant cells. There is also no evidence for extrachromosomal element involvement since the amplification is relatively stable even at early selection stages [25, 26] and in protoplast fusion hybrids [11, 32] and double minute chromosomes have not been observed. The amplified units may also be much smaller as demonstrated by the results presented here since in animal cells the units are usually from a few hundred to as many as 10000 kb in length.

To date the studies of gene amplification in plant systems have not investigated the structure of the amplified units to indicate any possible mechanism. This paper reports studies of the amplified EPSPS gene structure of selected carrot cells as an initial step to determine the gene amplification mechanism. An understanding of this genetic instability might provide insight into such phenomena as somaclonal variation and the formation of gene families in plants.

Materials and methods

Cell lines

The carrot (*Daucus carota* L. cv. Danvers Half Long) cell lines used were the wild-type cell line TC (usually denoted C1 in other publications) and the glyphosate-resistant cell lines TC xG (where x denotes the glyphosate concentration in mM used in the last selection step; for example, TC 35G is a TC-derived cell line which was selected stepwise with increasing concentrations up to 35 mM glyphosate in MX [18] liquid medium). The glyphosate concentrations used for selection were from 0.35 to 35 mM. PR is a glyphosate-resistant cell line selected against 25 mM glyphosate from TC cells in 1984 by Nafziger *et al.* [18].

DNA preparation and Southern hybridization analysis

Genomic DNA was isolated from carrot suspension cultures using the method of R.M. Hauptmann as reported by Menancio *et al.* [17]. Briefly,

cells were collected and ground with liquid nitrogen in a mortar and pestle. The ground cells were thawed on ice in the presence of 0.5–1.0 ml of TESE buffer per g fresh weight of cells (TESE is 50 mM Tris-HCl pH 8.0, 50 mM EDTA, 50 mM NaCl, 2% *N*-lauroyl sarcosine and 400 μg/ml ethidium bromide). Debris was removed by centrifugation and the supernatant was used for CsCl density gradient centrifugation. Following removal of ethidium bromide by isopropanol saturated with 20 × SSC (1 × SSC is 0.15 M sodium chloride, 0.015 M sodium citrate), DNA (5 or 10 μg) was digested with various restriction enzymes (BRL), separated by electrophoresis on agarose gels, and transferred to positively charged nylon membranes (Micron Separations). Southern transfer and hybridization were performed as described by Sambrook *et al.* [22]. The DNA was crosslinked covalently to the membrane by using a Stratalinker 1800 apparatus (Stratagene). Prehybridization was done at 65 °C for 2–6 h in 6 × SSC, 5 × Denhardt's solution (1 × Denhardt's solution is 0.02% Ficoll, 0.02% polyvinylpyrrolidone, 0.02% bovine serum albumin), 0.1% SDS, 5% dextran sulfate and 100 μg/ml sheared, denatured salmon sperm DNA. Hybridization was done overnight under the same conditions as the prehybridization except that ^{32}P-labeled probe DNA was added to a concentration of 1–2 × 10^6 cpm/ml. The petunia EPSPS cDNA sequence used as a probe was the 900 bp *Eco* RI fragment of pMON 9543 [24] obtained from Dr D. Shah of Monsanto Company. This sequence includes most of the mature protein coding region excluding the 3′ end, so contains about 2/3 of the entire EPSPS gene. Carrot DNAs used as probes included the 2.0 kb *Bgl* II fragment (map position 6.8 kb to 8.8 kb in Fig. 2) and the 1.05 kb *Bgl* II/*Eco* RI fragment (map position 7.7 kb to 8.8 kb in Fig. 2) of the 16 kb clone described here. The radioactive probes were prepared by the method of Feinberg and Vogelstein [3, 4] using a random-primed DNA labeling kit (U.S. Biochemicals). After hybridization, membranes were washed twice at room temperature for 15 min in 2 × SSC and 0.1% SDS and twice at 55 °C for 30 min in 2 × SSC and 0.1% SDS for the petunia probe or twice at 65 °C for 30 min in 0.2 × SSC and 0.1% SDS for the carrot probes. The sizes of the fragments that hybridized to the probes were determined using lambda DNA digested with *Hin*d III as standards.

Genomic library construction

TC 35G DNA was digested with *Bam* HI, subjected to electrophoresis in a 0.5% low-melting-point agarose gel, and 10–20 kb fragments were eluted from the gel using phenol extraction by the method of Wieslander [31] as described by Sambrook *et al.* [22]. The purified DNA was ligated to the Lambda DashII/*Bam* HI vector (Stratagene) and packaged *in vitro* with Gigapack gold II packaging extract (Stratagene) as described by the manufacturer.

Identification of the EPSPS clones

Approximately 3000 phage were screened with the petunia EPSPS cDNA probe from pMON 9543. Prehybridization and hybridization conditions were the same as those of genomic Southern hybridization except that 5 × 10^5 cpm/ml of the probe was used. The frequency of positive phage was 13 out of 3000, i.e. approximately one per two hundred. Of the thirteen positive phage plaques, six were purified and bulk DNA prepared by the method of Reddy *et al.* [20] using DEAE-cellulose column chromatography. The purified DNA was cleaved with *Bam* HI and three of the six phage were shown to contain a 13 kb insert while the remaining three contained 16 kb inserts. The identity of the two insert families was verified by restriction endonuclease digestion patterns.

Gene mapping

On each of the 13 kb and 16 kb clones was mapped as described in the Gene Mapping Kit instruction manual (1989, Stratagene).

Analysis of inverted duplications

The presence of inverted duplications in the cellular DNA was demonstrated by the method of Ford *et al.* [5] and Ford and Fried [6] with some modifications. Fresh preparations of CsCl-purified genomic DNA (100 μg/ml) were denatured by adding 10 M NaOH to a final concentration of 50 mM, renatured by addition of HCl to a final concentration of 50 mM and 50 mM Tris-HCl pH 7.5, and were digested with S1 nuclease. S1-resistant DNA was recovered by precipitation with ethanol and analyzed following electrophoresis on an agarose gel with or without restriction digestion by Southern transfer and hybridization with the pMON 9543 probe.

DNA sequencing

Appropriate restriction fragments of the 13 kb or 16 kb *Bam* HI fragments were subcloned into pUC 19. The dideoxy termination reaction was performed by the method of Sanger *et al.* [23] following the instruction manual of Sequenase version 2.0 (U.S. Biochemical Corporation). The reaction products were analyzed on 6% acrylamide gels containing 8 M urea in TBE (90 mM Tris-borate pH 8.2, 1 mM EDTA). Computer analysis of DNA sequences was performed on an Apple PC Macintosh IIcx with MacVector 3.5 software.

Results and discussion

As a first step in characterizing the structure of the amplified DNA during the gene amplification process, we determined if there were differences in the flanking sequences and EPSPS gene structure of highly glyphosate-resistant carrot cells in comparison with the wild-type cells. Southern hybridization with wild type, TC, and the most resistant selected line, TC 35G, DNA digested with five different enzymes using the petunia EPSPS cDNA clone pMON 9543 as probe showed that there were additional fragments evident with *Bam* HI, *Bgl* II and *Eco* RI digests of TC 35G DNA relative to those of TC (Fig. 1). These changes might indicate rearrangements or deletions of the proximal flanking sequences or changes in the EPSPS gene itself during the selection process.

In order to isolate the EPSPS gene and its flanking sequences, we constructed a genomic library with TC 35G DNA and screened the library with pMON9543. The *Bam* HI site of Lambda Dash II was used for cloning since there were two large *Bam* HI fragments in the

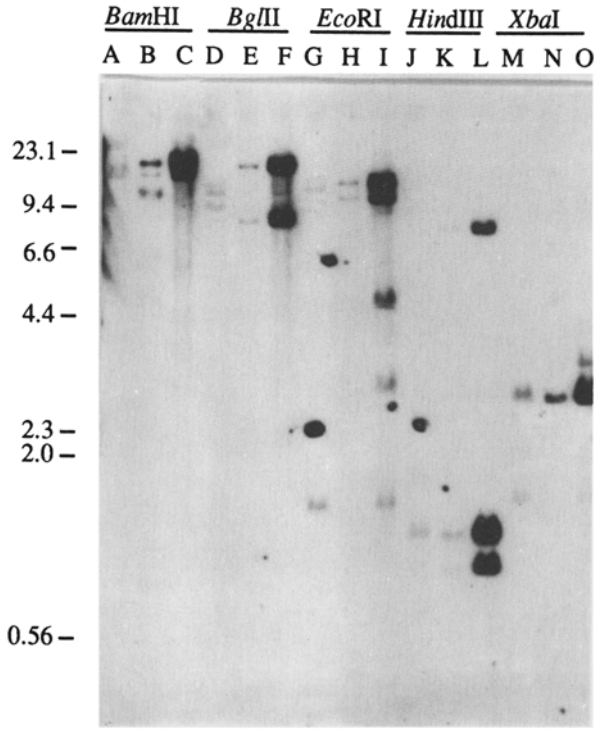

Fig. 1. Hybridization pattern of genomic DNA of wild type (TC) and a selected line (TC 35G). Genomic DNA was digested with *Bam* HI (lanes A–C), *Bgl* II (lanes D–F), *Eco* RI (lane G–I), *Hin*d III (lane J–L) and *Xba* I (lane M–O). The cut DNA was loaded onto an agarose gel in a fashion that the first lane contains 10 μg wild-type DNA, the second lane contains 0.4 μg of the selected line DNA which gives a comparable signal strength to the wild type and the third lane contains 10 μg seleced line DNA. The digested DNA was separated on a 0.6% agarose gel by electrophoresis and transferred onto a charged nylon membrane. The filter was hybridized with the ^{32}P-labeled insert of pMON 9543, which consisted of most of the mature protein coding regions of petunia EPSPS cDNA.

TC 35G DNA which were identified using the pMON 9543 probe (Fig. 1). Only two discrete insert sizes were observed with the six clones examined. One group of three clones contained a 13 kb insert while the remaining three contained 16 kb inserts. These two classes differed with respect to restriction maps as well as insert size but were identical within each group. However, Southern analysis of the two cloned DNAs probed with pMON 9543 gave identical hybridization patterns, indicating that the EPSPS gene itself was not altered but the sequences flanking the gene were different in the two clones. From restriction maps of the clones we concluded that the 13 kb clone contained only one copy of the EPSPS gene while the 16 kb clone had two copies of the EPSPS genes oriented in opposite directions (Fig. 2). Southern hybridization of restriction digests of cloned DNA probed with 5′ (400 bp *Eco* RI/*Bam* HI), middle (200 bp *Bam* HI/*Hin*d III), or 3′ (300 bp *Hin*d III/*Eco* RI) fragments of pMON 9543 (petunia EPSPS cDNA clone) was done to determine the orientation of the gene as shown in Fig. 2. This experiment also showed that the clones contained 5′ flanking sequences of the EPSPS gene and most of the carrot EPSPS coding sequence, because 3′ sequences of the petunia EPSPS cDNA hybridized to the 1.3 kb *Bam* HI/*Hin*d III end fragment of the clones.

About 250 bases from the left end of the 3.4 kb *Bam* HI/*Eco* RI fragment were sequenced. The first 120 bp of the sequence was at least 65% homologous to the cDNA sequences of petunia within about 100 bp of the 3′ end of the petunia EPSPS gene. Thus the 13 kb and 16 kb *Bam* HI fragments include almost the entire EPSPS gene except for about 100 bp of the protein coding sequences at the 3′ end in the last exon (data not shown). Therefore the genomic sequence of the carrot EPSPS gene is at least 3.4 kb in length (Fig. 2). The size of the EPSPS gene varies among plant species due to differences in the intron sizes, being about 2.3 kb in *Arabidopsis* [15], 2.8 kb in *Brassica napus* [8] and 6.9 kb in petunia [7].

The IR structure found in the 16 kb clone has a unique junction point as shown in Fig. 2. To determine if the IR existed in the parent or was an apparent product of the amplification process, the fragment containing this junction from the 16 kb clone was used as a probe in Southern analysis of the cloned 13 kb and 16 kb *Bam* HI fragments by using appropriate restriction endonucleases to differentiate between them. This

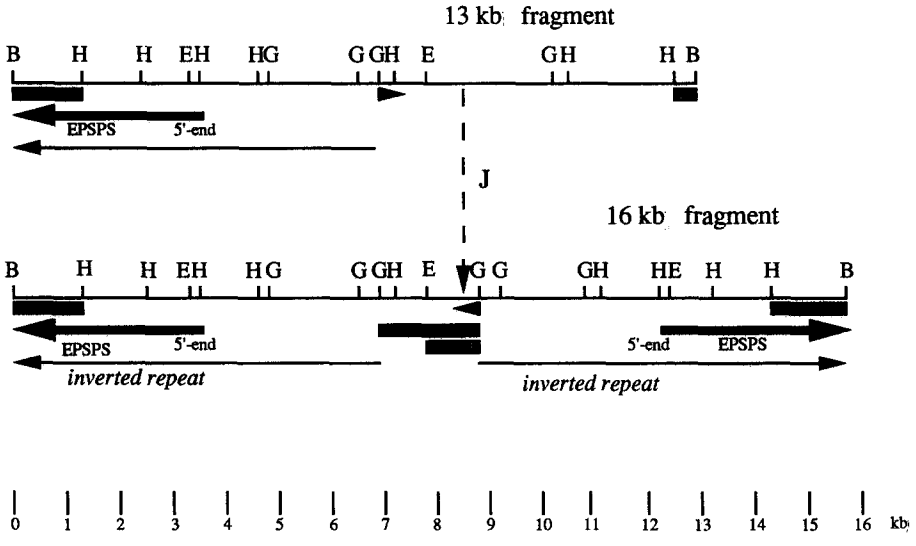

Fig. 2. Restriction endonuclease map of 13 kb and 16 kb *Bam* HI fragments of a selected carrot line, TC 35G. B, *Bam* HI; E, *Eco* RI; G, *Bgl* II; H, *Hin*d III. The position of the junction is indicated as J. The blocks indicate the probes used in the experiments (see the text). The filled triangles indicate the regions and direction of the sequencing as shown in Fig. 6.

method was also applied to genomic DNA isolated from TC and TC 35G. The results of Southern hybridization using a 1.05 kb *Bgl* II/*Eco* RI fragment from the 16 kb clone (position 7.7 kb to 8.8 kb of the restriction map in Fig. 2 which spans the junction) as a probe are shown in Fig. 3. While the *Bam* HI/*Eco* RI digests (lanes A–F) of wild type DNA (lane A) yielded the 5.2 kb fragment, the TC 35G DNA (lane B) had both the 4.6 kb and 5.2 kb fragments, which are diagnostic for the 13 kb and 16 kb structures, respectively. *Bam* HI/*Eco* RI digests of the 16 kb clone (lanes C–D) and 13 kb clone (lanes E–F) resulted in 4.6 kb or 5.2 kb fragments, respectively. Therefore, we could conclude that wild-type DNA contained only the 13 kb *Bam* HI fragment and TC 35G DNA had both the 13 kb and 16 kb *Bam* HI fragments. This conclusion is confirmed in additional experiments with *Bam* HI/*Hin*d III digests (lanes G–L). While the 13 kb clone (lanes K–L) and 16 kb clone (lanes I–J) had 3.4 kb and 4.0 kb fragments, respectively, the 3.4 kb fragment was detected in the wild type (lane G) and both the 3.4 kb and 4.0 kb fragments were found in TC 35G DNA (lane H). The intensities of the band indicate that TC 35G cells have about equal concentrations of each fragment.

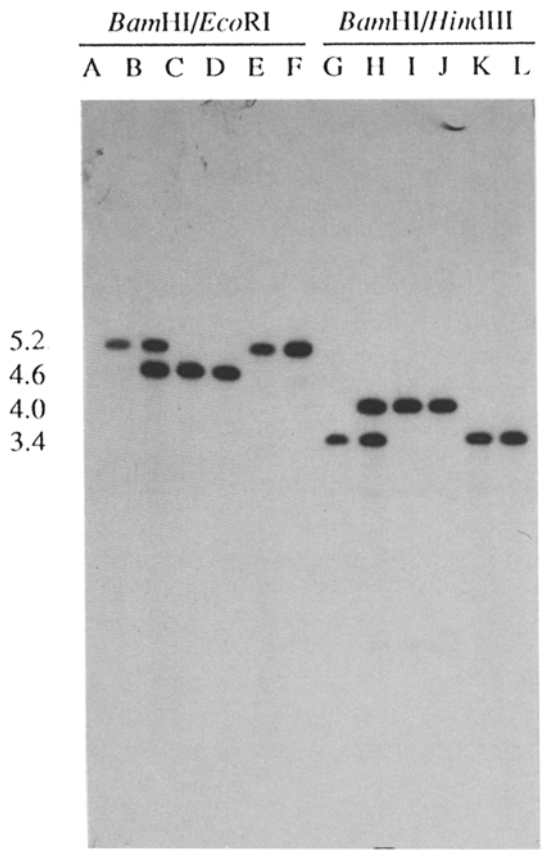

Fig. 3. Southern hybridization of genomic DNA and genomic clones showing that the wild type had only the 13 kb *Bam* HI fragment and the selected line TC 35G had both the 13 kb and 16 kb *Bam* HI fragments. Genomic DNA (5 μg) or phage DNA (4 ng) was digested with *Bam* HI/*Eco* RI or with *Bam* HI/*Hin*d III. The fragments were separated on a 0.7% agarose gel and transferred to a charged nylon membrane. The filter was hybridized with the ^{32}P-labeled carrot probe which was the 1.05 kb *Eco* RI/*Bgl* II fragment of the 16 kb clone (position 7.7–8.8 kb in the restriction map shown in Fig. 2). Lanes A–F, *Bam* HI/*Eco* RI-digested DNA of TC (lane A), TC 35G (lane B), two 16 kb clones (lanes C and D) and two 13 kb clones (lanes E and F). Lane G–L, *Bam* HI/*Hin*d III-digested DNA of TC (lane G), TC 35G (lane H), two 16 kb clones (lanes I and J) and two 13 kb clones (lanes K and L).

To determine when the inverted repeat structure of the 16 kb *Bam* HI fragment was formed during the glyphosate selection process, several selected cell lines were examined at different steps during one selection experiment. When *Bgl* II digested genomic DNA was probed with the 2.0 kb *Bgl* II fragment of the 16 kb clone (part of this fragment is also contained in the 3.5 kb *Bgl* II fragment in the 13 kb clone as shown in Fig. 2), only the 3.5 kb fragment was present in the wild type up to TC 22G lines (lanes A–D of Fig. 4). However, both the 3.5 kb and 2.0 kb fragments were present in TC 35G and the intensity of each band was approximately equal (lane E of Fig. 4). This indicates that only the 13 kb *Bam* HI fragment was present in the selected lines up to 22 mM glyphosate, and both the 13 kb and 16 kb *Bam* HI fragments were present in equal amounts in the 35 mM glyphosate selected line. Therefore, the inverted repeat was formed or was enriched enough to become detectable between the 22 mM and 35 mM glyphosate selection levels.

The 16 kb cloned fragment appears to comprise a large inverted duplication which extends beyond the *Bam* HI ends of the fragment. To confirm the presence of the inverted duplication in the glyphosate selected cells and to determine

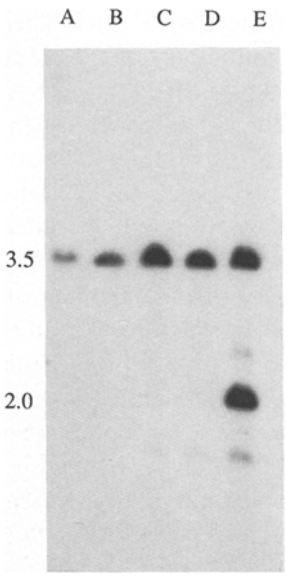

Fig. 4. Southern hybridization of stepwise selected lines showing the presence of only the 13 kb *Bam* HI fragment in the wild type and lines selected with up to 22 mM glyphosate (TC 22G) and the presence of both the 13 kb and 16 kb *Bam* HI fragments in the 35 mM glyphosate selected line (TC 35G). Genomic DNA (10 μg) was digested with *Bgl* II. The cut DNA was separated on a 0.6% agarose gel and transferred to a charged nylon membrane. The filter was hybridized with the ^{32}P-labeled 2.0 kb *Bgl* II fragment of the 16 kb clone. Lane A, wild type DNA; lane B, TC 4.4G DNA; lane C, TC 12G DNA; lane D, TC 22G DNA; lane E, TC 35G DNA.

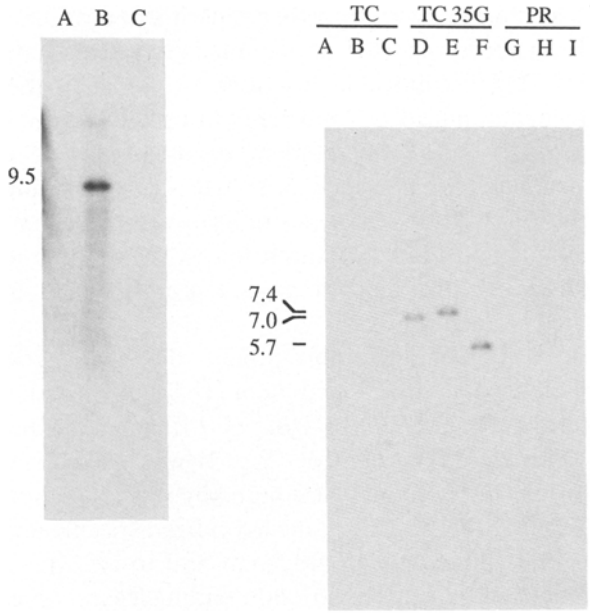

Fig. 5. Southern hybridization of S1-resistant TC 35G DNA showing the presence of inverted duplications in the cell. Fresh preparations of CsCl purified genomic DNA were denatured with NaOH, neutralized with HCl to anneal the inverted repeats and then treated with S1 nuclease. The S1-resistant DNA was precipitated and digested with restriction enzymes if necessary. The S1-treated (and digested) DNA was separated on a 0.6% agarose gel and transferred onto a charged nylon membrane. The filter was hybridized with ^{32}P-labeled insert of pMON 9543. 10 μg of only S1-treated TC (lane A), TC 35G (lane B) and PR (lane C) DNA. B. 30 μg of S1-treated and restriction-digested TC (lanes A–C), TC 35G (lanes D–F) and PR (lanes G–I) DNA. The first lane of each cell line contained *Bam* HI-digested DNA, the second lane *Bgl* II-digested DNA, and the third lane *Eco* RI-digested DNA.

the full length of the inverted repeat (IR) DNA unit, S1 nuclease resistance experiments were carried out as described by Ford *et al.* [5] and Ford and Fried [6]. Following denaturation and rapid renaturation, only IR DNA anneals fast enough to form double strands resistant to the single strand specific endonuclease S1. Because the two copies of the gene appear to be present as a palindrome on the 16 kb fragment we would expect to find an S1 nuclease fragment equal in size to the stem of the palindromic region. When S1-resistant TC 35G DNA was analyzed by Southern hybridization without any restriction digestion, a 9.5 kb fragment hybridized to the pMON 9543 probe (lane B of Fig. 5A). When this S1-resistant DNA was further digested with *Bam* HI, a 7.0 kb fragment was found (lane D of Fig. 5B), which is the size predicted from the gene map in Fig. 2. *Bgl* II or *Eco* RI digests of S1 resistant TC 35G DNA showed 7.4 kb or 5.7 kb fragments, respectively (lanes E–F of Fig. 5B), which are the expected sizes if the length of the IR stem is 9.5 kb. Therefore, we could conclude that the length of the IR stem region is about 9.5 kb, and extends ca. 2.5 kb past the *Bam* HI site ends of the 16 kb clone.

Wild-type TC and another line, PR, which was selected from TC cells in 1984 [18] for growth on 25 mM glyphosate and selected recently for growth on 35 mM glyphosate, did not show any bands in the same experiments (Fig. 5A and 5B) indicating that the IR is present only in TC 35G,

which agrees with the other experimental results. It has been reported in mammalian systems that inverted duplications are often observed in the highly amplified DNA where during the early selection stages only tandem head-to-tail repeat structures are present [13, 15]. It is possible that the 13 kb clone represents one copy of these tandem head-to-tail structures, but as yet we do not know how the 13 kb fragments are organized in the genomic DNA.

To understand more about the sequences flanking the two cloned *Bam* HI fragments, the terminal 0.4 kb *Hin*d III/*Bam* HI fragment of the 13 kb *Bam* HI fragment (Fig. 2) was used as a probe in genomic Southern hybridization of TC 35G DNA. This probe hybridized specifically to the 13 kb *Bam* HI fragment and to two fragments when other restriction endnucleases were used. This indicates that there are two different flanking sequences next to the 13 kb *Bam* HI fragment. When the 1.3 kb *Bam* HI/*Hin*d III fragment, which is found at the other end of the 13 kb clone, and also at both ends of the 16 kb clone, was used as a probe in Southern hybridization, two fragments were revealed with each enzyme, instead of the three fragments which could be expected from the three ends. We know that the IR extends only about 2 kb past the *Bam* HI sites of the 16 kb *Bam* HI fragment from the S1 nuclease resistance experiments, so the flanking sequences past that point are not the same. It is probable that the extension from one end of the 16 kb *Bam* HI fragment is the same as that of the 13 kb *Bam* HI fragment while the sequences flanking the other end are specific to the 16 kb *Bam* HI fragment as indicated by the S1 nuclease resistance experiments and Southern hybridization with the 1.3 kb *Bam* HI/*Hin*d III end fragment as a probe.

This is the first report of the involvement of IR in gene amplification in plants. Stepwise selection of alfalfa cells with phosphinothricin produced cultures which had increased numbers of GS genes. There were no changes in the structure of the GS genes and the flanking sequences in the amplified DNA compared to that of the wild type [29]; however, the alfalfa cell line contained only a 5- to 10-fold increase in the GS gene copy number [29]. Densitometry of the Southern hybridization signals of the TC 35G DNA showed that there was about a 25-fold increase in EPSPS gene copy number [25, 26]. In glyphosate-resistant tobacco lines, rearrangement of the EPSPS gene was detected by Southern hybridization [9, 30] but detailed analysis was not reported.

The presence of IRs as shown by the resistance to S1 nuclease digestion [5, 6] and by mapping of the cloned genes [13, 19] has been mechanistically linked to mammalian gene amplification. The structure of the IR-containing fragment in TC 35G might therefore be able to give clues to the gene amplification mechanism in carrot. Ruiz and Wahl [21] showed that a 9.2 kb dihydrofolate reductase (DHFR) gene construct in a plasmid vector transformed into DHFR gene-deficient mammalian cells formed a submicroscopic molecule containing IRs spanning more than 70 kb comprising the DHFR gene and host DNA. These extrachromosomal elements were at least 750 kb in size and were readily amplified under the selection pressure, but were lost in the absence of drug selection. In this case, the formation of an IR was proposed to be an initial step in gene amplification. In contrast, our selected lines were relatively stable, maintaining half of the amplified copies of EPSPS genes without selection pressure for two years [25, 26]. The possible involvement of extrachromosomal elements in the formation of the IR during the EPSPS gene amplification in carrot cells therefore seems unlikely.

In our studies, comparison of the 13 kb and 16 kb clone restriction maps suggests that approximately 1.5 kb (map position 7.0–8.5 kb in Fig. 2) of the 13 kb *Bam* HI fragment is not duplicated in the second copy of the inverted repeat of the 16 kb *Bam* HI fragment and a junction (J) was found between the *Bgl* II and *Eco* RI sites (see arrowheads in Fig. 2) in the 16 kb clone. Similar IR structures have been observed in mammalian gene amplification [13, 14, 19, 21]. In such cases, the two palindromic arms, which may extend for many hundreds of kb, flank a much smaller nonpalindromic central loop of variable size (200–1000 bp). Several explanations for such

structures have been proposed. From the structure and sequence of the inversion joint, Nalbantoglu and Meuth [19] suggested that, in a rare event, replication switches strands and proceeds around the replication fork rather than in a bidirectional manner producing an IR from the two arms of the fork. Hyrien *et al.* [13] proposed a chromosomal spiral model based on the multicopy appearance of a large inverted duplication and the sequence at the inversion joint. Nalbantoglu and Meuth [19] also suggested that the unduplicated region could be produced by loopout of a single strand in advance of the replication complex before the strand switching event. It is also possible, however, that the IR structure with the unduplicated loop region and junction is formed by recombination events after unscheduled replication as in the onion skin model [27].

Joints linking two amplified units together have been reported to be formed during mammalian gene amplification. Break points lying in the loops of stable stem-loop structures were found at each point where breaks had occurred by Hyrien *et al.* [13] but no sequence peculiarity was found by Nalbantoglu and Meuth [19]. There has been no absolute requirement for a particular type of sequence at the points where the breaks occurred; however, A + T-rich elements seem to be involved frequently in mammalian systems [27]. When we compared the sequence of the 237 nucleotides starting from the *Bgl* II site at the 8.8 kb map position (see Fig. 2 for map) at the junction in the 16 kb *Bam* HI fragment with the 259 nucleotides starting from the *Bgl* II site at the 6.8 kb map position (see Fig. 2 for map) in the 13 kb *Bam* HI fragment, they were identical for the first 88 bases going in the direction of the arrows (Fig. 6). The end of the matching sequence represents the putative break point in the 13 kb clone and the putative junction point (or joint) in the 16 kb clone (Fig. 6). Our sequencing data shows that the break point is at the end of a very A + T-rich region as has been reported in the mammalian systems.

DNA samples isolated from TC 35G on several occasions over a period of time involving

```
                    BglII
                    ------  10         20         30         40         50
1.05 kb Bg/RI     1 AGATCTTTCA GTAGTATTGT AGTAATATAC CTTAAAATTA GCCTTTCAAC    50
0.9  kb Bg/RI     1 ********** ********** *********G ********** T*********    50

                                                          junction point
                                                                ↙
                          60         70         80         90         100
1.05 kb Bg/RI    51 AGCAGAATGG GACACAGAAA GAAACTAAAG AAAGAAAACC GTGTGACGAC   100
0.9  kb Bg/RI    51 ********** ********** ********** ********AG AATACG*TGA   100

                                                                 break point
                         110        120        130        140        150
1.05 kb Bg/RI   101 TTGTCAGACA TACACTCCTG AGGTCCCTGC CGATGCCACC AGAAATCTCT   150
0.9  kb Bg/RI   101 AAT*TT***C ATA*G**T*T TAA*G*AAT* *AGCCAA*A* *CTCTCA*TA   150

                         160        170        180        190        200
1.05 kb Bg/RI   151 CTCAGGCATG TCTAAATGCC CTTGCACCAG CAGTTCCTGG CCTAATCGGT   200
0.9  kb Bg/RI   151 *AACAA*CCA *ACC**ACTT TCAA*TTT*A G*ACAGAAAC AAA*GGATA*   200

                         210        220        230        240        250
1.05 kb Bg/RI   201 GGTAGTGCTG ACTTTGCTTC CTCCAAACAC ACTTGAA... .........   250
0.9  kb Bg/RI   201 TT*CAGCAAC *A*CCAAAC* TCAGGGT*CA TTCCTTCACT TTCAATTGGT   250

                         260        270        280        290        300
1.05 kb Bg/RI   251 ......... ..........  .........  .........  .........   300
0.9  kb Bg/RI   251 TGTATGTGC. ..........  .........  .........  .........   300
```

Fig. 6. DNA sequences showing the putative break point in the 0.9 kb *Bgl* II/*Eco* RI (Bg/RI) fragment in the 13 kb *Bam* HI fragment and the putative junction point in the 1.05 kb *Bgl* II/*Eco* RI (Bg/RI) fragment of the 16 kb *Bam* HI fragment. The position and direction of the sequencing starting from *Bgl* II sites are indicated by the filled triangles in Fig. 2. Matched sequences are shown as bars and unmatched sequences are shown as lower case letters from the comparison of the sequences of 237 bp of the 1.05 kb *Bgl* II/*Eco* RI fragment and 259 bp of the 0.9 kb *Bgl* II/*Eco* RI fragments.

subculturing showed the same relative concentrations of the 13 and 16 kb fragments. It is unlikely therefore that TC 35G contains a mixed population of cells since this would require that the relative cell densities were constant despite cell growth during subculturing. We therefore believe that both molecular structures are present in the same cell and that the IR probably derives from the linear structure. This would suggest that either a single event generated the IR which was subsequently rapidly amplified or that the event happened many times in the population of amplified linear structures, possibly during the resolution of an instability generated in the amplification step from 22 to 35 mM glyphosate. Study of larger DNA fragments to determine whether there is any physical relationship between the IR and linear forms may indicate whether single or multiple events were involved in the amplification. We do not know why the IR structure is found only in the highly amplified line, TC 35G. It is possible that a population of cells having the IR structure became enriched in the later selection stages, which made detection of the IR by Southern hybridization possible. However, Southern hybridization of the *Bgl* II digested DNA from several lines obtained from colonies selected on 1, 2, and 10 mM glyphosate-containing medium also failed to detect the IR using the junction fragment of the 16 kb *Bam* HI fragment as a probe, indicating that the IR structure was not present even in small populations of the cells at the early selection stages (S. Caretto and J.M. Widholm, unpublished data). To verify if the IR structure and non-IR structure are present in the same cell, cloning of the TC 35G cells is required.

In conclusion, two different arrangements of the highly amplified EPSPS genes of the glyphosate-resistant carrot cells were identified from two genomic clones. One of the clones had only one copy of the EPSPS gene, which may represent the unit fragment of a tandem head-to-tail repeat structure. This structure is present in both wild type and selected cells. The other structure which was present only in the highly amplified line, TC 35G, contained a 9.5 kb inverted duplication of the EPSPS gene. The presence of IRs in the later stages of selection and of junctions has been reported in several amplified mammalian cell systems. The size of the IRs in the case of carrot are much smaller than those found in animal cells. Our results suggest that similar gene amplification mechanisms may occur in mammalian and plant systems even though plant gene amplification occurs more slowly and to a lower level and does not appear to involve extrachromosomal elements as determined by relative stability even at early stages of selection [25, 26].

This carrot system is ideal for gene amplification (genome instability) studies due to the ease and repeatability of the selection. Further study should yield information about how plant genomes can evolve to produce phenomena such as repeated DNA and gene families.

Acknowledgements

We would like to thank Dr Faith Belanger for advice on Southern hybridization, gene cloning and plaque lifts and Y. Joanne Shyr for DNA from some selected cell lines. This work was supported by funds from the Illinois Agricultural Experiment Station.

References

1. Donn G, Tischer E, Smith JA, Goodman HM: Herbicide-resistant alfalfa cells: an example of gene amplification in plants. J Mol Appl Genet 2: 621–635 (1984).
2. Dyer WE, Weller SC, Bressan RA, Herrmann KM: Glyphosate tolerance in tobacco (*Nicotiana tabacum* L.). Plant Physiol 88: 661–666 (1988).
3. Feinberg AP, Vogelstein B: A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal Biochem 132: 6–13 (1983).
4. Feinberg AP, Vogelstein B: Addendum to 'A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity'. Anal Biochem 137: 266–267 (1984).
5. Ford M, Davies B, Griffiths M, Wilson J, Fried M: Isolation of a gene enhancer within an amplified inverted duplication after 'expression selection'. Proc Natl Acad Sci USA 82: 3370–3374 (1985).
6. Ford M, Fried M: Large inverted duplications are associated with gene amplification. Cell 45: 425–430 (1986).

7. Gasser CS, Winter JA, Hironaka CM, Shah DM: Structure, expression, and evolution of the 5-enolpyruvylshikimate-3-phosphate synthase genes of petunia and tomato. J Biol Chem 263: 4280–4289 (1988).
8. Gasser GC, Klee HJ: A *Brassica napus* gene encoding 5-enolpyruvylshikimate-3-phosphate synthase. Nucl Acids Res 18: 2821 (1990).
9. Goldsbrough PB, Hatch EM, Huang B, Kosinki WG, Dyer WG, Herrman KG, Weller SC: Gene amplification in glyphosate tolerant tobacco cells. Plant Science 72: 53–62 (1990).
10. Harms CT, Armour SL, DiMaio JJ, Middlesteadt LA, Murray D, Negrotto DV, Thompson-Taylor H, Weymann K, Montoya AL, Shillito RD, Jen GC: Herbicide resistance due to amplification of a mutant acetohydroxy-acid synthase gene. Mol Gen Genet 233: 427–435 (1992).
11. Hauptmann RM, della Cioppa G, Smith AG, Kishore GM, Widholm JM: Expression of glyphosate resistance in carrot somatic hybrid cells through the transfer of an amplified 5-enolpyruvylshikimate-3-phosphate synthase gene. Mol Gen Genet 211: 357–363 (1988).
12. Hollander-Czytko H, Johanning D, Meyer HE, Amrhein N: Molecular basis for the overproduction of 5-enolpyruvylshikimate-3-phosphate synthase in a glyphosate tolerant cell suspension culture of *Corydalis sempervirens*. Plant Mol Biol 11: 215–220 (1988).
13. Hyrien O, Debatisse M, Buttin G, Robert de Saint Vincent B: The multicopy appearance of a large inverted duplication and the sequence at the inversion joint suggest a new model for gene amplification. EMBO J 7: 407–417 (1988).
14. Kishore GM, Shah DM: Amino acid biosynthesis inhibitors as herbicides. Annu Rev Biochem 57: 627–663 (1988).
15. Klee HJ, Muskopf YM, Gasser CS: Cloning of an *Arabidopsis thaliana* gene encoding 5-enolpyruvylshikimate-3-phosphate synthase: sequence analysis and manipulation to obtain glyphosate-tolerant plants. Mol Gen Genet 210: 437–442 (1987).
16. Ma C, Loney JE, Leu T-H, Hamlin JL: Organization and genesis of dihydrofolate reductase amplicons in the genome of a methotrexate-resistant Chinese hamster ovary cell line. Mol Cell Biol 8: 2316–2327 (1988).
17. Menancio DI, Hepburn AG, Hymowitz T: Restriction fragment length polymorphism (RFLP) of wild perennial relatives of soybean. Theor Appl Genet 79: 235–240 (1990).
18. Nafziger ED, Widholm JM, Steinrucken HC, Killmer JL: Selection and characterization of a carrot cell line tolerant to glyphosate. Plant Physiol 76: 571–574 (1984).
19. Nalbantoglu J, Meuth M: DNA amplification-deletion in a spontaneous mutation of the hamster *n* locus: structure and sequence of the novel joint. Nucl Acids Res 14: 8361–8371 (1986).
20. Reddy KJ, Kuwabara T, Sherman LA: A simple and efficient procedure for the isolation of high quality phage lambda DNA using a DEAE-cellulose column. Anal Biochem 168: 324–331 (1988).
21. Ruiz JC, Wahl GM: Formation of an inverted duplication can be an initial step in gene amplification. Mol Cell Biol 8: 4302–4313 (1988).
22. Sambrook J, Fritsch EF, Maniatis T: Molecular Cloning: A Laboratory Manual, 2nd ed., pp. 9.31–9.58. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1989).
23. Sanger F, Niklen S, Coulson AR: DNA sequencing with chain terminating inhibitors. Proc Natl Acad Sci USA 74: 5463–5467 (1977).
24. Shah DM, Horsch RB, Klee HJ, Kishore GM, Winter JA, Tumer NE, Horonaka CM, Sanders PR, Gasser CS, Aykent S, Siegel NR, Rogers SG, Fraley RT: Engineering herbicide tolerance in transgenic plants. Science 233: 478–481 (1986).
25. Shyr YYJ, Hepburn AG, Widholm JM: Glyphosate selected amplification of the 5-enolpyruvylshikimate-3-phosphate synthase gene in cultured carrot cells. Mol Gen Genet 232: 377–382 (1992).
26. Shyr Y, Widholm JM: Glyphosate resistance and gene amplification in selected *Daucus carota* suspension cultures. In: Nijkamp HJJ, van der Plas LHW, van Aartrijk J (eds) Progress in Plant Cellular and Molecular Biology, Current Plant Science and Biotechnology, pp. 148–152. Kluwer Academic Publishers, Boston (1990).
27. Stark GR, Debatisse M, Giulotto E, Wahl GM: Recent progress in understanding mechanisms of mammalian DNA amplification. Cell 57: 901–908 (1989).
28. Steinrucken HC, Amrhein N: The herbicide glyphosate is a potent inhibitor of 5-enolpyruvylshikimate-3-phosphate synthase. Biochem Biophys Res Commun 94: 1207–1212 (1980).
29. Tischer E, DasSarma S, Goodman HM: Nucleotide sequence of an alfalfa glutamine synthetase gene. Mol Gen Genet 203: 221–229 (1986).
30. Wang Y, Jones JD, Weller SC, Goldsbrough PB: Expression and stability of amplified genes encoding 5-enolpyruvylshikimate-3-phosphate synthase in glyphosate-tolerant tobacco cells. Plant Mol Biol 17: 1127–1138 (1991).
31. Weislander L: A simple method to recover intact high M.W. RNA and DNA after electrophoretic separation in low gelling temperature agarose gels. Anal Biochem 98: 305 (1979).
32. Ye J, Hauptmann RM, Smith AG, Widholm JM: Selection of a *Nicotiana plumbaginifolia* universal hybridizer and its use in intergeneric somatic hybrid formation. Mol Gen Genet 208: 474–480 (1987).