Restriction mapping of *Cucurbitae pepo* L. chloroplast DNA *Pst*I 9.8 kb fragment and detection of ORF 2280 transcripts

Silas D. Obukosia¹*, Charles D. Boyer², Christina M. Richards³

¹Department of Crop Science, University of Nairobi, P. O. Box 29053, Nairobi, Kenya.
²Department of Horticulture, 4017 Agriculture and Life Sciences Building Oregon State University, Corvallis, OR 97331-7304, USA.
³Department of Horticulture, Pennsylvania State University, 215, Tyson PA 16802, USA.

Accepted 20 July, 2004

Using tomato heterologous probe, the ORF 2280 was located on the *Pst*I 9.8 kb fragment in the inverted repeat A (IRA) and on the *Pst*I 12.1 kb fragment in the inverted repeat B (IRB) on the physical map of squash chloroplast DNA. A fine map of the *Pst*I 9.8 kb fragment was developed using *Sal*I, *BamH*I and *EcoR*I restriction enzymes. Two transcripts of 1.4 and 2.6 kbs of ORF 2280 of squash plastid DNA were detected in fruits but not in the leaves of Early Prolific cultivar using a heterologous *Pst*I 1.2 kb tomato plastid DNA probe internal to ORF 2280. In Bicolor gourd fruits, 1.4 and 2.6 kb transcripts of ORF 2280 were detected both in the fruits and in leaves.

Key words: *Cucurbitae pepo*, restriction mapping, *Pst*I 9.8 kb fragment, ORF 2280 transcripts.

INTRODUCTION

Previous studies on plastid gene expression have mainly focused on plastid genes of known functions such as *rbcL* gene encoding the large subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase and related rubisco enzymatic activity (Obukosia et al., 2002; Berry et al., 1986, 1988, 1990; Edmondson et al., 1990; Gutteridge and Julien, 1989; Inamine et al., 1985) and *psbA* gene encoding the 32 kd photosystem II quinone binding protein (Gruissem and Zurawski, 1985a, b). In addition to these genes, the plastids contain many unidentified open reading frames (ORFs) (Hiratsuka et al. 1989; Ohyama et al., 1986; Shinozaki et al., 1986). A couple of recent studies have focused on the large ORF, capable of encoding a protein of 2100-2280 amino acids. This large ORF has been identified in several plant species including the tobacco ORF 2280 (Palmer, 1991; Shinozaki et al., 1986), spinach ORF 2132 (Zhou et al., 1988), liverwort ORF 2136 (Glick and Sears, 1993) and broad bean ORFx (Herdenberg et al., 1988). The rice chloroplast genome is an exception in that it lacks the large ORF (Hiratsuka et al., 1989; Shinozaki et al., 1986) and the recently sequenced chloroplast genome of unicellular green alga *Chlorella vulgaris* has no inverted repeat and the equivalent of ORF 2280 (Wakasugi et al., 1997).

*Corresponding author. E-Mail: citrustc@wananchi.com
Fax 254-2-020-632121, Tel: 254-2-630340, 631340/44
Abbreviations: ORF- open reading frame; IRA- inverted repeat A; IRB- inverted repeat B of cpDNA; cpDNA- chloroplast DNA.
In addition to the prevalence of the large ORF among plant species, there is also sequence homology between the large ORFs of different species. For example, the partial sequence of the tomato large ORF matched the tobacco ORF 2280 at 96 out of 97 positions with an exception of 9 nucleotide deletion in tomato or insertion in tobacco (Richards et al., 1991). ORFx of broad bean had 80% amino acid homology with the tobacco ORF 2280, 50% homology with liverwort ORF 2135 and also high homology with the spinach ORF 2132 (Herdenberg et al., 1988). High amino acid homology has also been reported between tobacco and liverwort ORFs (80-90%), while spinach and liverwort ORFs contain a region of high homology (71-73%) and low homology (35-50%). The prevalence of the large ORF 2280 in several plant species including the parasitic beechdrops (Epiphasgus virginiana) and the high sequence homology among plant species, led to the suggestion that the ORF 2280 is a functional gene (Glick and Sears, 1993). The initial steps to unravel the function of the large ORF involved a search for ORF transcripts. Five ORF 2280 transcripts of sizes 1.7, 2.4, 2.7 and 8.3 kbs have been detected in tomato fruits and leaves (Richards et al., 1991). The amounts of the 8.3 kb transcript increased with fruit ripening in tomato cultivar Count II.

The amounts of this transcript also increased in Traveller 76 cultivar, while the levels of the 1.7 and to 3.2 kb transcripts increased only slightly in both cultivars. The 8.3 transcript was present in the in the leaf RNA but at lower levels compared to small transcripts (Richards et al., 1991). Similarly, a 10 kb ORF transcript has been detected in broad bean (Herdenberg et al., 1988).

Further steps to elucidate the function of ORF involved a search for ORF translation products. Potentially the ORF 2280 is capable of encoding a protein of about 300 Kda. However, immunological study of ORF translation products detected proteins with apparent molecular weights ranging from 170-180 Kda in tobacco, spinach and Oenothera chloroplasts (Glick and Sears, 1993) and in tomato (Richards et al., 1994). The discrepancy between the observed ORF proteins (170-180 Kda) and the size of molecular weight of the protein the ORF is potentially capable of encoding (about 300 Kda), has been attributed to post-transcriptional and post-translational processing (Glick and Sears, 1993).

Whereas the function of the ORF remains an enigma, it is unlikely that the ORF encodes a protein that is functional in the photosynthetic apparatus, since no protein of that size has been detected in the electron transport chain of the thylakoid membrane (Anderson, 1986; Glick and Sears, 1993; Kaplan and Amtzen, 1982). Based on the observation that ORF transcripts were higher in the chromoplasts than in the chloroplasts of tomato, Richards et al. (1991) suggested that ORF of tomato might encode a protein that is mainly functional in the chromoplasts. In the plastid DNA of beechdrops, a non-photosynthetic angiosperm, most of the photosynthetic genes and other ORFs are absent but genes for 3 rRNAs, 7 tRNAs and region homologous to ORF 2280 are present (dePamphilis and Palmer, 1990; Morden et al., 1991) suggesting that the ORF protein may not be a component of the photosynthetic apparatus (Glick and Sears, 1993). Similarities between predicted amino acids sequence of ORF 2280 and the β subunit of the RNA polymerase point to the possibility that the ORF protein may be a partial product of the secondary ORF transcript in the chloroplasts (Glick and Sears, 1993).

The presence of the large ORF in several plant species, the detection of ORF transcripts and translation products, the sequence homology of ORF between species and the homology of ORF sequence with the β subunit of RNA polymerase are strong evidence that ORF 2280 is a functional gene. Although the specific function is still unknown, further study is needed in aspects such as search for transcripts and translation products from ORF 2280 in other plant species and characterization of ORF by DNA sequencing. The restriction and the genetic maps of squash chloroplast DNA have been developed (Lim et al., 1991). The objectives of this study were to (i) locate the ORF on the PstI cpDNA restriction map, (ii) fine map the restriction fragment containing the ORF 2280, and (iii) determine whether ORF 2280 transcripts are present.

Materials and Methods

Sources of ORF 2280 probes

Heterologous ORF 2280 specific probes were obtained from tomato cpDNA. The DNA sequence at the terminal ends of the PstI 7.7 kb fragment of tomato perfectly aligned with the tobacco ORF sequence covering the 147,176 to 147,281 and 154,674 to 155,788 regions, except for a 6 basepair deletion which did not alter the reading frame (Richards et al., 1991) (Figure 1). Using the published tobacco ORF sequence (Shinozaki et al., 1986), the BamHI restriction sites were located in the tobacco ORF 2280 region that corresponded to the tomato PstI 7.7 fragment. The 7 predicted restriction fragments after digestion with BamHI were 0.2, 0.3, 0.7, 1.05, 1.1, 1.19 and 2.93 kbs (with the 0.7 and 1.05 kb flanking the fragment), approximated to 0.2, 0.3, 0.7, 1.0, 1.1, 1.2 and 3 kbs, respectively. It was presumed that digestion of the tomato PstI 7.7 fragment would produce similar fragments as predicted from tobacco sequence. The tomato cpDNA PstI 7.7 fragment was digested with BamHI and produced the predicted 7 fragments of which the 1.19 (approximated 1.2 kb) and the 0.7 kbs fragments were isolated and cloned in pUC18 and Bluescript SK+, respectively. The 12 kb fragment was later used as a heterologous internal probe for ORF 2280 and for the detection of ORF 2280 cpDNA transcripts.

Detection of ORF 2280 on cpDNA restriction map

Chloroplast DNA was extracted from leaves of three-week old squash grown in the greenhouse using a procedure previously described (Gounaris et al., 1986, Lim 1990). Four restriction
enzymes, namely \textit{Psfl}, \textit{Sall}, \textit{PvuII} and \textit{SacI}, that had been used in the development of squash cpDNA physical map (Lim et al., 1991) (Figure 2), as recommended by the supplier (New England Biolabs, Beverly MA), were used in the study. The restricted products were fractionated on 0.7% agarose gel, southern-botted and DNA immobilized onto nitrocellulose membrane and eventually probed with the 1.2 kb (ORF 2280 specific probe from tomato cpDNA) in order to locate the position of ORF 2280 on the squash cpDNA. The 9.8 kb fragment (Figure 2) to which the ORF 2280 was located was isolated and cloned in pUC18 as described below.

Cloning and fine mapping of the \textit{Psfl} 9.8 cpDNA fragment:

Chloroplast DNA was digested with \textit{Psfl} and fractionated on 0.7% agarose and the 9.8 kb band dissected out of the gel. The DNA band was extracted by electoelution into aqueous solution of 0.5 X TBE (0.089 M Tris-HCl, 0.089 M borate, 0.001 M EDTA). DNA was purified in 3 steps: extraction with phenol, phenol/chloroform/isoamyl alcohol (25:25:1), and once with chloroform/butanol (4:1). The plasmid vector (pUC18) was digested with the \textit{Psfl} restriction enzyme at 37°C. The cpDNA/pUC18 fragments were ligated in a water bath at a temperature range of 6 to 14.5°C for 16 h using T4 ligase. The \textit{E. coli} cells strain JM83 were made competent by treatment with CaCl₂ solution. Transformed clones were selected using the antibiotic, ampicillin, and screened for cpDNA insert using 5-bromo-4-chloro-3-indolyl-

In order to characterize ORF of squash cpDNA, the \textit{Psfl} 9.8 fragment was fine mapped by digestion with \textit{Sall} (S), \textit{BamH} (B), \textit{EcoRI} (E), double digested with \textit{S} x \textit{B}, \textit{E} x \textit{B}, and triple digestion with \textit{P} x \textit{E} x \textit{B}, \textit{P} x \textit{B} x \textit{S}. The 4.7, 3.2 and 1.9 fragments were cloned into pUC18 and sequenced (unpublished).

Extraction of total RNA:

Early Prolific cultivar of squash was grown in the greenhouse while the Bicolor cultivar was grown in the field. The squash fruits were harvested at 3, 7, 14 and 21 days post-pollination in early Prolific varieties, and at 14 and 21 days in Bicolor fruits and frozen in liquid nitrogen and stored at -70°C. Total RNA was isolated by selective LiCl precipitation (Piechulla et al., 1986). The frozen tissue was first ground in a pre-cooled mortar and pestle under liquid nitrogen, and resuspended in 50 ml homogenization buffer (0.35 M sorbitol, 50 mM Tris-Hcl pH 8.0, 25 mM EDTA, 15 mM 2-mercaptoethanol, 2 mM dithiothreitol, 0.1% polyvinylpyrrolidone, SmM aurintricarboxylic acid). One-tenth lysis buffer (5% sodium lauroylsarcosinate, 50 mM Tris-Hcl pH 8.0, 25 mM EDTA) was added and the homogenate incubated for 5 - 15 min at room temperature. After lysis, an equal volume of phenol:chloroform:isoamyl alcohol (25:24:1) was added to the homogenate and mixed thoroughly using a pasteur pipette. The mixed homogenate was centrifuged at 10,000 X g for 15 min to remove the protein. The phenol/chloroform extraction was repeated 3-4 times depending on the protein content in the tissue. Finally, the top aqueous phase was recovered and 1/10 volume of 5 M ammonium acetate (pH 5.2) and 2 volumes absolute ethanol were added and stored at -70°C for 1 h or at -20°C overnight, to
precipitate the total nucleic acids. It was observed that when RNA was precipitated from the nucleic acids directly, the resultant RNA product was hard to dissolve. This problem was circumvented by redissolving the total nucleic acid pellet into TE buffer (10 mM Tris, 1 mM EDTA) and adding 1/4 volumes of 10 M ammonium acetate, followed by centrifugation at 8,000 g for 30 min. The supernatant was saved and the white pellet discarded (Richards et al., 1991). 10 M LiCl was added to the collected supernatant (to give a final concentration of 2.5M LiCl), to selectively precipitate high molecular weight RNA. The precipitated RNA was pelleted by centrifugation at 10,000 x g for 1 h. The RNA pellet was vacuum dried, resuspended in appropriate volume of diethylpyrocarbonate (DEPC) treated water and stored in aliquots at -70°C.

Formaldehyde gel electrophoresis of RNA and Northern Blotting

Equal amounts of total RNA (5 g) were denatured and fractionated on 1.2% agarose-formaldehyde gels according to the procedure described by Sambrook et al., (1990). This protocol, which was adapted from those of Lehrach et al., (1977), Goldberg (1990) and Seed (1982), was used to separate RNA species according to their molecular weights. The loading of the RNA was adjusted such that the 28S cytoplasmic RNA bands, as scanned by densitometry on photographic negatives, were equal for each stage. The RNA samples were fractionated for 12 - 16 h at 17 mA, 20 V h, under a hood. The buffer was mixed regularly to maintain constant pH.

At the end of the run, the gel was stained with ethidium bromide (0.5 µg/ml in 0.1 M ammonium acetate) for 45 - 60 min, and later destained in DEPC water for at least 2 h. A transparent ruler was aligned with the gel and photographed by ultraviolet illumination.

The photograph was used to measure the distance from the loading well to each of the bands of the RNA. The log₁₀ of the size of the RNA fragments were plotted against the distance migrated. The resulting curve was used to calculate the sizes of RNA species detected by hybridization after blotting.

The fractionated RNAs were transferred to nylon membrane by capillary method as described by Sambrook et al, (1990). The blotted RNA was immobilized by UV irradiation for 30 s at 6 W/m² (Church and Gilbert, 1984).

Labelling of probes and hybridization

The 1.2 kb ORF 2280 specific probe isolated from the tomato cpDNA was used to detect ORF 2280 transcripts. 32P probes were prepared by the random primer labelling procedure developed by Feinberg and Vogelstein (1983). Typically, 50 ng of DNA probes were labelled to specific activities of 10⁵ to 10⁶ cpm/µg DNA. Unincorporated precursor dNTPs were removed from the radiolabelled probe by DNA Quick Spin Columns (G-50 sephadex).

Blots were prehybridized at 55°C (for heterologous probes) or at 65°C (for homologous probes) in 0.5 M NaHPO₄, 7% sodium dodecyl sulphate (SDS) for at least 40 min. Prior to hybridization, the purified probes were denatured in a boiling water bath for 5 min and immediately chilled on ice. Hybridization was performed overnight under similar conditions as prehybridization. The hybridized blots were washed once in preheated 2.5% SEN (2.5% SDS, 1mM EDTA and 40mM Sodium phosphate) for 10 minutes, and twice in preheated 0.5% SEN (0.5% SDS, 1 mM EDTA and 40 mM Sodium-phosphate) for 20 min each turn. The blots were sealed in plastic bags and auto-radiographed by exposing them to X-ray film (Kodak XAR-2) at 70°C with intensifying screens.

**Figure 2.** Restriction and genetic map of squash chloroplast DNA showing the *Pst*I 9.8 and *Pst*I 12.1, in IRA and IRB respectively. Haktae et al. (1991).
RESULTS AND DISCUSSION

Isolation of ORF 2280 specific probe from tomato cpDNA

The predicted fragments of sizes: 0.2, 0.3, 0.7, 1.0, 1.1, 1.2, 2.6 (pUC18) and 3.0 kbs, were obtained from the digestion PstI 7.7 fragment of tomato, with PstI and BamHI (Figure 3). Because of very close sizes of the fragments they were resolved by electrophoresis at three agarose gel concentrations of 1.2%, 1.5% and 2.0% (Figure 3). The 1.2 kb fragment was cloned in pUC18, while the 0.7 kb fragment was cloned in Bluescript SK+

The chloroplast DNA was digested by the following restriction enzymes: PstI, PvuII, BglII and SacII. Probe of the Southern blot with BamHI-1.2 kb fragment internal to tomato ORF 2280) localized the squash ORF in the inverted repeat on the following fragments: 12.1 and 9.8 kb fragments of PstI (lane 1); 56.7 and 16.7 kbs of PvuII (lane 2), 31.0 and 20.3 kb fragments of BglII lane 3; 27.5 kb of SacII lane 4 (Figure 4). The cpDNA of squash contains two 27.5 Kb SacII DNA fragments, one located in each inverted repeat, IRA and IRB, as shown in Figure 2. This was why the signal for SacII digest is both very strong and thick (Figure 4, panel B lane 4). The Southern blot band signal therefore runs ahead and behind the 31.0 BglII. Except for the 27.5 kb fragments of SacII, the rest of the fragments namely, the 12.1 kb and the 9.8 kb fragments of PstI; the 31.0 and 20.3 kb fragments of BglII; the 56.7 kb and the 16.7 kbs of PvuII are located at appropriate relative positions and in accord with the published cpDNA restriction map (Figure 2) (Lim, 1991).
Fine restriction mapping of the 9.8 kb \textit{PstI} fragment

For further study of the ORF of squash cpDNA, a fine restriction map of the \textit{PstI} 9.8 kb fragment of squash cpDNA was developed. The fragment was digested with \textit{SalI} (S), \textit{BamHI} (B), \textit{EcoRI} (E) and double digested \textit{S} x \textit{B}, \textit{E} x \textit{B}. A double digestion of the \textit{PstI} 9.8 kb fragment with \textit{PstI} (P) x \textit{B} gave 3 fragments of 4.6, 3.3 and 1.9 kbs. In order to align the 3.2, 1.9 and 4.7 kb, the \textit{PstI} fragment 9.8 kb fragment was double and triple digested with restriction enzymes. The purpose of including the \textit{P} in the digestion was to remove the insert from the vector, so it can be disregarded in the following analysis. The pUC18 vector (2.69 kb) was left intact and can also be disregarded. Triple digestion with \textit{P} x \textit{E} x \textit{B} produced 3.3 kb and 1.9 kb fragments common to the \textit{P} x \textit{B} digestion (Figure 5). However, the 4.6 kb fragment was absent; instead it was replaced by two new fragments of 0.7 and 3.9 kbs (Figure 5). It can be inferred from these results that 4.6 kb fragment is located between 1.9 and 3.2 kb fragments (Figure 6). Other double and triple digestion with \textit{P}, \textit{S} and \textit{B} confirmed the above alignment of fragments. A triple digestion with \textit{P} x \textit{B} x \textit{S} gave 1.9 and one 3.3 kb fragments and the 4.6 kb fragment was again...
absent (Figure 5, panel D). Instead two fragments of sizes 1.3 and another 3.3 kbs appeared (Figure 5). Based on these results a linear map of the *PstI* fragment with four restriction enzymes was developed (Figure 6).

**Figure 6.** The restriction map of the 9.8 kb fragment of squash cpDNA constructed using 4 restriction enzymes. The *psbA* gene was located on the 1.9 *PstI/BamHI* fragment; part of the ORF 2280 was located in the 4.7 kb *BamHI* fragment.

**Figure 7.** The 2.6 and 1.4 kb transcripts detected in the Early Prolific cultivars fruits harvested at 3, 7, 14 and 21 days post-pollination. The transcripts were detected by the 1.2 kb probe internal to tomato ORF 2280. The X-ray film was exposed for 6 days with a screen.

**Figure 8.** The 2.6 and 1.4 kb transcripts detected in the leaf and fruit pericarp of Bicolor gourd cultivar of squash harvested at 14 and 21 days post-pollination. Y refers to the pericarp of the yellow portion, while the G to the green portion of the same fruit. The X-ray film was exposed for 6 days with a screen.

Transcripts from the ORF 2280

Total plastid RNA from 3 squash cultivars was probed with heterologous 1.2 kb probe internal to ORF 2280 of tomato. Two transcripts of 2.6 and 1.4 kbs were detected in Early Prolific fruits with a heterologous 1.2 kb probe (Figure 7). The X-ray film was exposed for 6 days with a screen due to weakness of the signal. Similarly, a 2.6 and 1.4 kb fragments were detected in the chromoplasts and chloroplasts of Bicolor fruits and in the leaves (Figure 8). The amounts of transcripts detected in the leaves exceeded the levels detected in the fruits of Bicolor gourds. Two transcripts of sizes 2.6 and 1.2 kbs were also detected in the EP fruits using a homologous cpDNA *PstI* 9.8 kb probe. However, the signal was rather weak and the X-ray film had to be exposed for 7 days with a screen. The 2.6 kb fragment corresponded to the 2.7 kb mRNA detected in the tomatoes using 1.2 kb ORF 2280
specific internal probe. The 1.4 kb transcript detected in squash corresponds to the 1.7 kb transcript detected in tomato fruits (Richards et al., 1991). However, in the current study there was no corresponding 8.3 kb ORF 2280 mRNA as detected in tomato by Richards et al. (1991). Our study provides initial evidence that the ORF 2280 of squash chloroplast DNA is transcribed giving a 2.6 kb mRNA transcript and a smaller 1.4 kb fragment.

References


