The Female-Specific Cs-ACS1G Gene of Cucumber. A Case of Gene Duplication and Recombination between the Non-Sex-Specific 1-Aminocyclopropane-1-Carboxylate Synthase Gene and a Branched-Chain Amino Acid Transaminase Gene

Ronit Rimon Knopf and Tova Trebitsh *
Department of Life Sciences, Ben-Gurion University of the Negev, POB 653, Beer-Sheva 85105, Israel

Cucumber (Cucumis sativus L.) is a monocoeic plant in which female sex expression (gynoecy) is controlled by the Female (F) locus that can be modified by other sex-determining genes as well as by environmental and hormonal factors. As in many other cucurbits, ethylene is the major plant hormone regulating female sex expression. Previously we isolated the Cs-ACS1 (ACS, 1-aminocyclopropane-1-carboxylate synthase) gene that encodes the rate-limiting enzyme in the ethylene biosynthetic pathway. We proposed that Cs-ACS1 is present in a single copy in monoecious (ffMM) plants whereas gynoecious plants (FFMM) contain an additional copy Cs-ACS1G that was mapped to the F locus. To study the origin of Cs-ACS1G, we cloned and analyzed both the gynoecious-specific Cs-ACS1G gene and the non-sex-specific Cs-ACS1 gene. Our results indicate that Cs-ACS1G is the result of a relatively recent gene duplication and recombination, between Cs-ACS1 and a branched-chain amino acid transaminase (BCAT) gene. Taking into consideration that the Cs-ACS1G gene was mapped to the F locus, we propose that this duplication event gave rise to the F locus and to gynoecious cucumber plants. Computer analysis of the 1 kb region upstream of the transcription initiation site revealed several putative cis-acting regulatory elements that can potentially confer the responsiveness of Cs-ACS1G to developmental and hormonal factors and thereby control female sex determination in cucumber. These findings lead us to a model explaining the action of Cs-ACS1 and Cs-ACS1G in cucumber floral sex determination.

Keywords: 1-Aminocyclopropane-1-carboxylate synthase — Branched-chain amino acid transaminase — Cucumis sativus — Ethylene — Gynoecy — Sex determination.

Abbreviations: ACC, 1-aminocyclopropane-1-carboxylate; ACS, ACC synthase; BCAT, branched-chain amino acid transaminase; CDS, coding sequences; GM, gynoecious monocoeic; G-GM, gynoecious-GM; IPCR, inverse PCR; M-GM, monocoeic-GM; UTR, untranslated region.

Sequence data have been submitted to the GenBank/EMBL databases under accession numbers DQ839406, DQ839407, DQ839408, DQ839409 and DQ839410.

*Corresponding author: E-mail, trebitsh@bgu.ac.il; Fax, +972-8-646-1710.

Introduction

Floral sex determination is a developmentally regulated process culminating in plants bearing unisexual flowers. Sex determination genes control sex expression, although plant hormones as well as environmental factors (such as day length and temperature) can modulate expression of the sex-determining genes, leading in many cases to complete floral sex reversal (reviewed in Dellaporta and Calderon-Urrea 1993, Lebel-Hardenack and Grant 1997, Khryanin 2002, Tanurdzic and Banks 2004). This phenomenon led to the suggestion that the role of sex-determining genes is to trigger the switch between alternative developmental programs by modulating the level and/or ratios of endogenous hormones (Lebel-Hardenack and Grant 1997). Studies in Mercurialis annua, maize, cucumber and other plant species have implicated sex-determining genes in the production of plant hormones (Durand and Durand 1991, Yin and Quinn 1995, Lebel-Hardenack and Grant 1997). In the majority of plant species studied to date, unisexual flower development is achieved through selective abortion or developmental arrest of pre-formed sex organs of the initially bisexual flower primordia (with some exceptions such as hemp, Mercurialis or spinach) (Dellaporta and Calderon-Urrea 1993, Lebel-Hardenack and Grant 1997). A breakthrough in understanding hermaphrodite flower development was achieved with the realization that floral organ position and identity are controlled by a combinatorial action of homeotic genes (belonging to class A, B and C), in three overlapping regions of the floral primordia (Coen and Meyerowitz 1991). The current thought is that the function of floral homeotic genes precedes that of the sex determination cascade (Lebel-Hardenack and Grant 1997).

Sex expression in cucumber is a combinatorial interaction between genetic, environmental and hormonal factors (Mälepszy and Niemirioicz-Szczytzyt 1991, Perl-Treves 1999 and references therein). The extensive physiological and genetic studies conducted in cucumber, as well as a wide range of sex phenotypes and genotypes, designated...
cucumber as a model plant for the study of sex determination in monoecious plants. In monoecious wild-type cucumber, flowers are produced in a pre-set, developmental sequence along the main stem, with a first phase of staminate flowers, followed by a mixed phase of staminate and pistillate flowers, and terminated by a pistillate flowers phase (Shifriss 1961). The nature and length of each phase is regulated by sex-determining genes: a dominant Female (F) locus, and two other recessive loci andromonoecious (m, staminate and hermaphrodite flowers) and androecious (a, staminate flowers). Female sex expression (gynoecy) is determined by the F locus that interacts with the other two loci and can be modified by additional genes (In-F, gy) to enhance femaleness (Pierce and Wehner 1990 and references therein). In the FF genotype, all flowers contain pistils whereas in heterozygous plants (Ff) the pistillate flower phase can be preceded by a short mixed phase (2-4 nodes) of staminate and pistillate flowers (Shifriss 1961). Similar to other unisexual flowering plants, cucumber floral primordia are originally hermaphroditic, containing both male and female reproductive organs (reviewed in Dellaporta and Calderon-Urrea 1993, Tanurdzic and Banks 2004). Cucumber unisexuality is achieved by the developmental arrest of pre-formed reproductive organs, and floral organ identity genes, belonging to class B and C, have been implicated in the developmental arrest of male organs (Kater et al. 2001, Bai et al. 2004).

In cucurbits such as cucumber and melon, long days, high temperature and gibberellins promote formation of staminate flowers, whereas short days, low temperature, ethylene and auxin promote the formation of pistillate flowers (reviewed in Frankel and Galun 1977, Malepszy and Niemirowicz-Szczytt 1991, Perl-Treves 1999). In cucumber, a high level of correlation exists between femaleness (F locus) and elevated endogenous levels of ethylene, whereas inhibitors of ethylene biosynthesis or perception suppress expression of the F locus (Atsmon and Tabbak 1979, Takahashi and Jaffe 1984). In addition, auxin has also been found to enhance female sex expression, possibly via induction of ethylene biosynthesis (Galun 1959, Takahashi and Jaffe 1984, Trebitsh et al. 1987). The involvement of ethylene in the development of gynoecium has also been shown in hermaphrodites such as Arabidopsis and tobacco (De Martinis and Mariani 1999, Tsuchisaka and Theologis 2004).

In recent years, the study of sex determination in cucumber has focused primarily on the elucidation of the molecular mechanisms by which ethylene controls gynoecy. A mechanistic model suggested by Yin and Quinn (1995) proposed that ethylene regulates both sexes in cucumber, with the F locus regulating ethylene levels and the M locus attenuating tissue sensitivity to ethylene. Characterization of ethylene receptor-related genes suggested that ethylene signals might indeed influence the product of the M locus and thus inhibit stamen development in cucumber (Perl-Treves 1999, Yamasaki et al. 2000, Yamasaki et al. 2001, Yamasaki et al. 2003a). Four family members of the gene family encoding the rate-limiting enzyme in the ethylene biosynthetic pathway (ACS, 1-aminocyclopropane-1-carboxylate synthase) have been isolated from cucumber. Two of these genes, Cs-ACS2 and Cs-ACS1G, were correlated with female flower production (Trebitsh et al. 1997, Mathooko et al. 1999, Yamasaki et al. 2000, Yamasaki et al. 2003a). In gynoecious plants and under female inductive conditions, Cs-ACS2 is expressed in pistil primordia, whereas in monoecious plants the transcript level is reduced and accumulates below the pistil primordia in the adaxial side of petals (Kamachi et al. 1997, Yamasaki et al. 2003b).

We previously reported the cloning of a partial genomic sequence of Cs-ACS1 from female (gynoecious FF) cucumber. We proposed that Cs-ACS1 is present as a single copy gene in monoecious (ff) cucumber plants and that gynoecious (FF) cucumbers possess two copies of the gene. The gynoecious-specific Cs-ACS1G was mapped to the F locus (Trebitsh et al. 1997). The presence of a second gene copy (Cs-ACS1G) in genotypes with a dominant F allele was recently confirmed (Mibus and Tatlioglu 2004). Interestingly, the 5' proximal flanking region isolated by Mibus and Tatlioglu (2004) was identical in the Cs-ACS1 and the Cs-ACS2 genes (Mibus and Tatlioglu 2004). However, vital information is still missing concerning the origin of the F allele and the homology between the two genes within the coding region and the distal promoter. This information is crucial since it has implications for gene regulation either at the transcriptional level or on enzyme activity by post-translational modification. In this report, we describe the isolation of the full-length genomic sequence of the two gene copies: the non-sex-specific Cs-ACS1 from monoecious (ff) and gynoecious (FF) plants, and the gynoecious-specific Cs-ACS1G from gynoecious plants. For each gene, we analyzed a 7.7 kb genomic sequence. We show that sequences corresponding to the proximal promoter, and the Cs-ACS1 transcript are identical in the two genes, thus precluding the possibility that enzyme activity is modulated by post-translational modifications. However, the distal promoter sequences of Cs-ACS1 and Cs-ACS1G completely diverge and gene-specific putative cis-regulatory elements identified in this region suggest the possibility of a differential gene regulation at the level of transcription. In addition, we show that the Cs-ACS1G gene originated from a gene duplication and recombination between Cs-ACS1 and a branched-chain amino acid transaminase (BCAT) gene. This unique gene duplication and the implications on the origin of the F allele and sex determination in cucumber are addressed herein.
Results

Cloning the full genomic sequence of Cs-ACS1 and Cs-ACS1G genes

We previously reported that the genome of gynoecious ( ff ) cucumber contains two copies of Cs-ACS1, whereas the genome of monoecious ( ff ) cucumber lines possesses a single copy of Cs-ACS1 (U59813, Trebitsh et al. 1997). The gynoecious-specific gene was designated Cs-ACS1G and mapped to the F locus (Trebitsh et al. 1997). To determine where in their sequence these genes differ, we probed genomic DNA from gynoecious and monoecious plants with a probe derived from the 5′ coding sequences (CDS) of the partial Cs-ACS1 that we previously isolated (Trebitsh et al. 1997). The probe recognized two different fragments in gynoecious genomic DNA digested with MspI: a fragment unique to the gynoecious genotypes (~3.5 kb) and a non-sex-specific fragment (~2.5 kb), common to both the sex genotypes (Fig. 1). This raised the possibility that the two genes diverge in their 5′-flanking region. To characterize this gene duplication, the full-length genomic sequence of the non-sex-specific Cs-ACS1 (from monoecious and gynoecious cucumber) and the sex-specific Cs-ACS1G (from gynoecious cucumber) had to be isolated and analyzed. The genomic sequences as well as the 5′- and 3′-flanking regions of Cs-ACS1 and Cs-ACS1G were isolated by PCR-based genomic methods, and the gene isolation strategy is depicted in Fig. 2A.

Sequence comparison between the 5′-flanking region of Cs-ACS1 and Cs-ACS1G revealed that the two sequences are identical up to 607 bp upstream of the ATG and thereafter the identity between the two genes ends (Fig. 2B). The point of divergence between the two genes (at position -410) was termed by us the ‘identity border’ (Fig. 2B). Transcription of Cs-ACS1 is initiated 197 nucleotides upstream of the ATG start codon (Fig. 2B, according to Shiomi et al. 1998, Mathooko et al. 1999). Thus the identical region spans the 5′-untranslated region (UTR) and 410 bp upstream of the transcription initiation site, confirming results published by Mibus and Tatlioglu (2004) (Fig. 2B and Supplementary Fig. S1). The region between the transcription initiation site and the ‘identity border’ (from position -1 to -410) is referred to as the proximal promoter. The sequence region that is divergent between the Cs-ACS1 and Cs-ACS1G genes (upstream of the ‘identity border’) is referred to as the 5′-distal Cs-ACS1 or 5′-distal Cs-ACS1G, respectively.

To study the extent of similarity/divergence between the two genes, we isolated the region corresponding to the Cs-ACS1 and Cs-ACS1G transcript (2,058 bp). From the DNA fragments that were amplified, a single contig was constructed for each gene: 7.659 kb corresponding to the non-sex-specific Cs-ACS1 (isolated from monoecious and gynoecious plants) and 7.713 kb corresponding to the gynoecious-specific Cs-ACS1G (isolated from gynoecious plants) (Fig. 2B). The identity detected between the two genes at the proximal promoter and the 5′ UTR continued throughout the CDS including the intron sequences, and extended into the 3′ UTR and 1,153 bp further downstream (Fig. 2B). This implies that the gynoecious genotype ( FF ) arose by a duplication event in monoecious cucumber. The duplication involved a genomic region of at least 3.6 kb and possibly a much larger fragment, as can be observed in the DNA blot analyses (Figs. 1, 2C and D). When using a probe derived from the 3′ part of the gene, the digestion of the genomic DNA with EcoRI yields two high molecular DNA fragments (~15 kb) in the gynoecious genome and a single band in the monoecious genome (Fig. 2C). Within the sequences we isolated there is no EcoRI restriction site. Thus, the two gene copies might reside on two separate chromosomes or on the same chromosome with an EcoRI restriction site between them. Digestion of the genomic DNA with SacI and PvuII yields a single DNA fragment (~20 and 40 kb), with a stronger signal in the gynoecious line suggesting the presence of a second fragment of similar size (Fig. 2C). Thus, we cannot conclude whether Cs-ACS1 and Cs-ACS1G reside on the same DNA fragment, but if
The cucumber *Female*-specific *Cs-ACS1G* and gynoecy

Fig. 2. Analysis of the full-length genomic sequence of *Cs-ACS1* and *Cs-ACS1G* and their 5' - and 3'-flanking regions. (A) Schematic representation of the strategy employed to isolate genomic sequences of *Cs-ACS1* and *Cs-ACS1G* (see details in Materials and Methods). An arrow marks the location and direction of the primer; lines between arrows, or extended from it, indicate the length of the fragment obtained by either inverse-PCR (IPCR) or direct PCR (PCR) as indicated above the line. A dashed line separates the strategy employed for *Cs-ACS1* and that for *Cs-ACS1G*. (B) Schematic representation the full-length genomic sequence of gynoecious sex-specific *Cs-ACS1G* (7,713 bp) and of the non-sex specific *Cs-ACS1* (7,659 bp). The +1 indicate the start of transcription and 'Identity border' at –410 marks the end of sequence identity; the exons are filled blocks, the introns are lines between exons; the 5' and 3' UTR are open blocks. The proximal promoter region is represented by an open box; the 5'-distal *Cs-ACS1* is represented by a white box and the 5'-distal *Cs-ACS1G* is represented by a checkered box. Letters above indicate restriction enzyme sites as follows: A, AccI; B, BstNI; H, HindIII; M, MspI; P, PvuII; S, SacI; X, XbaI. The size (kb) of DNA that corresponds to the DNA fragments identified by the respective probes in Fig. 1, and (C) and (D) is marked on the arrows between the appropriate restriction sites. P-CD, P-5'CD, P-3', P-XbaI and P-AccI represent the location of the DNA probes used in Fig. 1, and (C) and (D). (C) DNA blot hybridization analysis of near-isogenic monoecious (M; ffMM) and gynoecious (G; FFMM) cucumber. Genomic DNA (10 μg per lane) was digested with the restriction enzymes indicated in the figure (no EcoRI restriction site is present in the sequenced DNA fragments) and separated on a 0.7% agarose gel. The blot was probed with a probe spanning the third exon and the 3'-flanking region of *Cs-ACS1/G* (P-3'). (D) DNA blot hybridization analysis of near-isogenic monoecious (M; ffMM) and gynoecious (G; FFMM) cucumber. Genomic DNA (10 μg per lane) was digested with the restriction enzymes indicated in the figure, and separated on a 0.8% agarose gel. The blot was probed with a *Cs-ACS1G*-specific probe (P-XbaI) or a *Cs-ACS1*-specific probe (P-AccI). The location of the probes on the respective genes is marked in (B).
they do these results indicate that the distance between them is at least 20 kb.

Based on the gene sequences, a probe derived from the CDS region, that is identical in both genes (such as 5'-CDS, Fig. 2), will detect two distinct MspI fragments in gynoecious genomic DNA corresponding to the Cs-ACS1 and Cs-ACS1G genes as depicted in Fig. 1. The MspI Cs-ACS1G fragment (3,590 bp) spans the region from the start of the third exon to the MspI restriction site in the 5'-distal Cs-ACS1G, and the MspI Cs-ACS1 fragment (2,516 bp) contains the region from the start of the third exon to the MspI restriction site in the 5'-distal Cs-ACS1 (Figs. 1, 2B). The sex genotype specificity of each sequence was examined by DNA blot hybridization analysis (Fig. 2D). As expected, the 5'-distal Cs-ACS1-specific AccI probe identified the 2.5 kb MspI fragment in both genotypes, and the 5'-distal Cs-ACS1G-specific XbaI probe identified the expected 3.5 kb gynoecious-specific MspI DNA fragment (Fig. 2B). However, the gynoecious-specific XbaI probe recognized a DNA fragment of similar size in both monoecious (ff) and gynoecious (FF) cucumber (Fig. 2D). Additional probes from the 5'-distal Cs-ACS1G region confirmed this finding (data not shown). We interpret this as an indication that the monoecious genome contains a region that is highly homologous if not identical to the Cs-ACS1G distal promoter (covered by the XbaI probe, Fig. 2). This genomic fragment was termed the GynoeciousMonoeccious (GM) fragment.

Cloning the GM fragment from the monoeccious genome

To investigate further the Cs-ACS1 duplication event and the origin of Cs-ACS1G, we isolated the GM DNA sequences from monoecious plants (term M-GM, Fig. 3). The M-GM consists of two zones: a ‘GM zone’ that shows no similarity to either of the Cs-ACS1 or Cs-ACS1G sequences we have isolated (980 bp) and a ‘Cs-ACS1G zone’ that is homologous to the 5'-distal Cs-ACS1G (from position −410 to −3,057) (Fig. 3). As depicted in Fig. 3, the ‘identity border’ that was identified above as the divergence point between Cs-ACS1 and Cs-ACS1G is also the border separating the two zones of the M-GM fragment (Figs. 2B, 3). A GM fragment that has the same sequence characteristics as M-GM was amplified from the gynoecious genome and termed G-GM. Sequence alignment showed that M-GM and G-GM are identical along the ‘GM zone’.

The 5'-distal Cs-ACS1G is identical to the ‘ACS-ACS1G zone’ of G-GM and 96% identical to the corresponding zone of the M-GM. These results are in accordance with the DNA blot hybridization analysis showing that the monoecious genome contains sequences similar to that of the 5'-distal Cs-ACS1G region (P-XbaI, Fig. 2D). Thus, a probe derived from the 5'-distal Cs-ACS1G (P-XbaI, Figs 2, 3) will

---

![Fig. 3](https://academic.oup.com/pcp/article-abstract/47/9/1217/232950) Schematic structure of the GM fragment that is homologous to the 5'-distal Cs-ACS1G and to the branched-chain amino acid transaminase (BCAT) genes. M-GM was isolated from monoecious cucumber (ff) and G-GM from gynoecious cucumber (FF) by the DNA walking procedure and direct PCR. Arrows below the DNA scheme mark the primer location. For details concerning the GM isolation, see Materials and Methods. The gray filled boxes represent sequences highly identical to the conceptual translation of exon 10 (Ex10), exon 9 (Ex9) and exon 8 (Ex8) of genes encoding BCAT protein. For restriction enzymes and the Cs-ACS1G structure details, see the legend of Fig. 2B.
detect two MspI fragments in the gynoecious genome, the Cs-ACS1G (3,590 bp) and the G-GM fragment (2,467 bp), while in the monoecious genome a single DNA fragment is detected corresponding to the M-GM fragment (2,467 bp).

Analysis of the translated GM against the translated database (tblastx, Altschul et al. 1997) revealed that the GM fragments contain sequences similar to exons 8–10 of genes encoding BCAT genes (Figs 3, 4 and Supplementary Fig. S2). At the protein level, exons, 9 and 10 of an Arabidopsis BCAT gene (AY087619.1) are 89, 87 and 69% identical to the corresponding sequence in GMs, respectively. Thus, the GM fragments span the last three exons and possibly represent a partial genomic sequence of a cucumber BCAT. The ‘identity border’ is located in the intron separating exons 8 and 9 of BCAT, thus the 5′-distal Cs-ACS1G contains sequences that correspond to exons 9 and 10 of BCAT (at positions −1,276 and −1,791 bp; Figs. 3, 4 and Supplementary Fig. S2). Collectively the data indicate that an event of gene duplication and recombination occurred between Cs-ACSI and a BCAT gene. The recombination site can be putatively assigned to the ‘identity border’ identified in the 5′-flanking region of Cs-ACSI and Cs-ACS1G, and in the GMs.

Putative cis-acting regulatory elements identified in the Cs-ACSI and Cs-ACS1G promoters

To avoid the region of Cs-ACS1G that is similar to BCAT, 1.0 kb upstream of the transcription initiation sites of Cs-ACSI and Cs-ACS1G were analyzed for the presence of putative cis-acting regulatory elements using the PLACE database (Higo et al. 1999). This region contains the proximal promoter region that is identical in both genes (from position −1 to −410) and the first 590 bp from the 5′-distal region that is unique to each of the two genes (from position −411 to −1,000). Various potential cis-regulatory elements were identified and the most notable ones are listed in Table 1. In the proximal promoter region identical in both genes, a TATA box signal (TATAAT) is located at position −46 (Supplementary Fig. S1). Two putative auxin-responsive factor (ARF)-binding sites (AuxRE), and an ABA-responsive element (ABRE) reside in the proximal

![Fig. 4 Amino acid sequence alignment of GM (derived amino acids of the BCAT exons 8–10 region) with the full-length rice BCAT (NM_197648) and Arabidopsis BCAT (AY087619). Lines above the sequences mark the exon region. Black and gray boxes indicate identities and similarities among the different proteins, respectively.](https://academic.oup.com/pcp/article-abstract/47/9/1217/2329590)
promoter region, which is identical in both promoters (Table 1).

Expression of Cs-ACS1/G during flower development

Previously we have shown that Cs-ACS1/G gene expression is induced following auxin treatment; however, the gene’s transcript level was not detectable in shoot apices or leaves of monoecious or gynoecious plants (Trebitsh et al. 1997). To analyze the expression of Cs-ACS1/G, semi-quantitative reverse transcription–PCR (RT–PCR) was performed using the expression of C. sativus ubiquitin extension protein as a reference standard gene. Tissue was collected from near-isogenic cucumber lines that differ in the F locus. The relative abundance of Cs-ACS1/G mRNA was compared between monoecious (ff) and gynoecious (FF) shoot apices that contain young bisexual floral buds and floral buds at early stages of sex differentiation. mRNA abundance was also compared between young male and female floral buds and male and female flowers at anthesis (Fig. 5). Cs-ACS1 mRNA was detected in monoecious shoot apices and increased during male flower development up to flowers at anthesis where the highest transcript level was detected. Monoecious plants possess solely the Cs-ACS1 gene, thus we can conclude that the promoter of this gene is active. Gynoecious plants possess both Cs-ACS1 and Cs-ACS1G; however, if the promoter of Cs-ACS1G is inactive, the transcript abundance, in tissue collected from monoecious and gynoecious plants, might be expected to be similar. As depicted in Fig. 5, a considerably higher transcript level is observed in gynoecious shoot apices than in monoecious ones. Similarly, higher gene expression was observed in young female flower buds and flowers at anthesis than in the male ones (Fig. 5). We take this as an indication that the Cs-ACS1G promoter is indeed active and that the gene has a role in gynoecium development.

Discussion

In eukaryotes, gene and genome duplication events are common phenomena and a source of new genes given that genes are often retained after duplication due to sub- and neofunctionalism (Moore and Purugganan 2005). Members of the ACS gene family, as well as duplicate genes, can be differentially regulated, and each family member can be responsive to multiple factors (Huang et al. 1991, Rottmann et al. 1991, Destefano-Beltran et al. 1995, Liang et al. 1995, Bekman et al. 2000, Tsuchisaka and Theologis 2004). In addition, post-translational modifications at the N- or C-termini of the protein can modulate the activity of ACS enzymes (Li et al. 1996, Chae et al. 2003). To gain insight into the origin of the F allele and the mechanism by which

<table>
<thead>
<tr>
<th>Response</th>
<th>cis-element</th>
<th>Sequence(^a)</th>
<th>Proximal</th>
<th>Distal</th>
<th>Distal</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>promoter</td>
<td>Cs-ACS1</td>
<td>Cs-ACS1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>position</td>
<td>promoter</td>
<td>promoter</td>
</tr>
<tr>
<td>Auxin response</td>
<td>AuxRe [TGTCTC]</td>
<td>TGTCTC</td>
<td>−158</td>
<td>−288(^b)</td>
<td></td>
</tr>
<tr>
<td>ABA response</td>
<td>ABRE [ACGT(C/G)(C/G)]</td>
<td>ACGTGGCC</td>
<td>−67</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[(C/T)ACGTGGCC]</td>
<td>TACGTGGC</td>
<td>−68</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[ACGTG]</td>
<td>ACGTG</td>
<td>−70</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transcriptional</td>
<td>ARRI [(A/C/G/T) GATT]</td>
<td>GGATT [A/T/C]GATT</td>
<td>−59, −116(^b), −128(^b), −192(^b), −262(^b), −335(^b), −377(^b)</td>
<td>+11 positions</td>
<td>+21 positions</td>
</tr>
<tr>
<td>regulators</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CarG [C(8*A/T)G]</td>
<td>CAATTATTTG,</td>
<td>−668,</td>
<td>−837</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>CATTTTATTTG</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>DOF [AAAG]</td>
<td>AAAG</td>
<td>−93, −284(^b)</td>
<td>+36 positions</td>
<td>+20 positions</td>
</tr>
<tr>
<td></td>
<td>MYB [(A/T)AACCA]</td>
<td>TAACCA,</td>
<td>−121, −309, −392</td>
<td>+7 positions</td>
<td>+10 positions</td>
</tr>
<tr>
<td></td>
<td>or [TAACAAA]</td>
<td>AAACCA,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>TAACAAA</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^a\)Position indicated from the 5' to the 3' end.
\(^b\)Sequence on the complementary strand.
The cucumber Female-specific Cs-ACS1G and gynoecy

The higher transcription level observed in female flowers was detected in monoecious plants lacking the Cs-ACS1 gene, indicating that the promoter of this gene is active. The promoters of both genes, could indicate that the promoters of both genes are active (Fig. 5). The additional gynoecious sex-specific Cs-ACS1G might lead to enhanced ethylene production and feminization, due to a double gene dosage effect. Alternatively, regulatory elements within the promoter, as well as long distance- and trans-acting regulatory elements, might be involved in gene-specific regulation resulting in a differential temporal and spatial expression pattern between the two genes.

Therefore, we compared the 5′-flanking region of Cs-ACS1 and Cs-ACS1G. In agreement with published data (Mibus and Tatlioglu 2004), we show that the identity between the twin genes extends beyond the 5′ UTR for another 410 bp upstream of the transcription initiation site (Fig. 2, Supplementary Fig. S1). The region surrounding the TATA box signal (−4 to −71) located at position −46 shows high sequence identity (91%) to the corresponding region in the auxin-induced Cm-ACS gene isolated from Cucurbita maxima (Nakagawa et al. 1991). In addition to the putative AuxRE element reported previously, we have identified a second site located at position −288 (Table 1, Mibus and Tatlioglu 2004). The AuxRE element is found more frequently in the 5′-flanking region of auxin- and brassinosteroid-regulated genes in Arabidopsis (Hagen and Guilfoyle 2002). Both hormones promote femaleness in cucumber, and it was suggested that their feminizing effect is due, at least partly, to their induction of ethylene production (Takahashi and Jaffe 1984, Trebitsh et al. 1987, Papadopoulou and Grumet 2005). Expression of the Cs-ACS1/G gene and ACS enzyme activity are both promoted by auxin (Trebitsh et al. 1987, Trebitsh et al. 1997). Furthermore, inhibitors of ethylene synthesis or perception can negate the F locus action in F− plants, and auxin-induced femaleness in monoecious ff plants (Takahashi and Jaffe 1984, Trebitsh et al. 1997). Thus, the hypothesis that the feminizing effect of auxin is via ethylene is further substantiated by the presence of putative ARF-binding sites in the proximal promoter region of Cs-ACS1/G.

Several putative binding sites for translational activators or enhancers, such as ARR1, DOF and MYB, were also detected in the proximal promoter, the 5′-distal Cs-ACS1G and the 5′-distal Cs-ACS1 regions (Table 1). These factors have been shown to regulate the transcription of several plant genes in response to a wide range of cues including hormonal signals (Jin and Martin 1999, Oka et al. 2002, Yanagisawa 2004). The existence of multiple hormonal and developmental putative cis-acting regulatory elements is in agreement with physiological studies in which the activity of the F allele, and thereby gynoecy, is amenable to various signaling cascades (developmental, hormonal and environmental). Furthermore, as manifested by the response of monoecious (ff) plants, the f allele is
responsive to female inductive factors, thus mimicking the dominant $F$ allele (reviewed in Frankel and Galun 1977, Malepszy and Niemirowicz-Szczytt 1991, Perl-Treves 1999). Thus, the physiological and molecular data concerning the $F/f$ alleles and the $Cs-ACS1G/Cs-ACS1$ genes, in particular the identity of the proximal promoter, favor the hypothesis that ethylene production, and consequently femaleness, is induced by a double gene dose response. On the other hand, we cannot rule out the possibility that following gene duplication and the acquisition of a unique 5'-distal region, the $Cs-ACS1G$ gene adopted a temporal and spatial expression pattern unlike that of $Cs-ACS1$. In this case, $Cs-ACS1G$ might have assumed a different and/or overlapping developmental role to that of $Cs-ACS1$, such as regulation of femaleness. A differential transcriptional regulation of $Cs-ASC1$ and $Cs-ACS1G$ is expected to originate from the gene-specific 5'-distal flanking region (Fig. 2). Most intriguing is the putative CArG domain found in the 5'-distal region of the gynoecious sex-specific $Cs-ACS1G$ but not in that of $Cs-ACS1$ nor in the proximal promoter region. The CArG motif binds MADS domain transcription factors such as those involved in floral organ development or hormone biosynthesis (Becker and Theissen 2003, Tang and Perry 2003, Wang et al. 2004). The $Cs-ACS1G$ gene could be a direct target of MADS box proteins and be regulated by homeotic genes that determine floral organ identity. This suggestion is supported by the finding that $CUM1$, the cucumber homolog of $AGAMOUS$ (a class C gene), is down-regulated in anther primordia and coincides with the appearance of the carpel primordia that is followed by the arrest of stamen development in female flowers of cucumber (Kater et al. 1998, Bai et al. 2004).

In view of the study presented herein (Figs 3, 4 and Supplementary Fig. S2), we postulate that in monococious cucumber an event of gene duplication and an inter- or intrachromosomal recombination and rearrangement, between a $BCAT$ gene and $Cs-ACS1$, gave rise to $Cs-ACS1G$, the $F$ locus, and thereby to gynoecy in cucumber. Following this event, a $BCAT$ gene is still functional in gynoecious cucumber and the transcript level and size do not differ from those of the monococious $BCAT$ (data not shown). The border of the recombination between $ACS1$ and the $BCAT$ can be assigned to the ‘identity border’ that marks the point of divergence between the $Cs-ACS1$ and $Cs-ACS1G$ sequences. Furthermore, based on the identity between the two genes, we suggest that the gene duplication occurred relatively recently.

Duplication of $ACS$ genes has been reported in several plant species such as $Arabidopsis$, tomato, potato, $Lupinus$ and $Cucurbita pepo$ (Huang et al. 1991, Rottmann et al. 1991, Destefano-Beltran et al. 1995, Liang et al. 1995, Bekman et al. 2000); however, no complete identity between the ‘twin genes’ was observed. The $Arabidopsis$ $At-ACS1$ and $At-ACS3$ show the highest degree of identity (99.6%); however, $At-ACS3$ is an inactive truncated version of $At-ACS1$ that lacks the fourth exon of $At-ACS1$ (Liang et al. 1995). The above-mentioned $ACS$ twin genes reside on a genomic fragment of 8–9 kb separated by 2–4 kb. Interestingly, $At-BCAT3$ and $At-ACS9$ reside on chromosome 3 of $Arabidopsis$ with a distance of 9 kb between them. Among the $Arabidopsis$ $ACS$ gene family members, $ACS9$ shows the highest homology to $Cs-ACS1/G$ and is induced by auxin, in a developmentally dependent manner (Tsuchisaka and Theologis 2004). Moreover, $ACS9$ has the most limited range of expression, concerning plant organs and response to various inducers of ethylene production, but is highly expressed in stigma, thus implicating ethylene in gynoecium development of hermaphrodite flowers, as also shown in tobacco (De Martinis and Mariani 1999, Tsuchisaka and Theologis 2004). Presently we cannot establish whether $Cs-ACS1$ and $Cs-ACS1G$ reside on the same chromosome, but chromosomal location could be determined in the future either by chromosome fluorescence in situ hybridization (FISH) or using a new polymorphism that might be found between the gene copies (based on the newly isolated sequences). If indeed the $BCAT$ and $Cs-ACS1$ genes of cucumber reside on the same chromosome, it might be another example of synteny and conservation of gene order between different taxa (Salse et al. 2004).

At present, we cannot rule out the possibility that a gene other than $Cs-ACS1G$ is located near $Cs-ACS1G$ at the same chromosomal position as the $F$ gene. However, taking into consideration that ethylene mimics the activity of the $F$ locus and that $Cs-ACS1G$ co-segregates with the $F$ locus, our results strongly suggest that the $Cs-ACS1G$ gene is indeed the $F$ gene. In that case, and if indeed the two genes reside on the same chromosome, we hypothesize that following the duplication of $Cs-ACS1$, the position on the chromosome occupied by $Cs-ACS1$ (in monococious $ff$) is now occupied by both $Cs-ACS1$ and $Cs-ACS1G$ in the gynoecious ($FF$) genome. Thus the term ‘allele’ in the classical sense might not apply to the $F$ locus and it should be better described as consisting of two idiomorphs $F$ and $f$. The term idiomorph was coined to describe the mating type locus ($MAT$) in $Neurospora crassa$ that can be occupied by two alternative (non-similar) sequences according to the mating type (Metzenberg and Glass 1990, Turgeon 1998).

Here we provide evidence that the female-specific $Cs-ACS1G$ originated from a gene duplication and an inter- or intrachromosomal recombination and rearrangement between a $BCAT$ gene and $Cs-ACS1$. Based on the identity between the non-sex-specific $Cs-ACS1$ and the sex-specific $Cs-ACS1G$ in the region downstream of the
DNA blot hybridization analysis

Genomic DNA was extracted from young leaves as described by Chetalat et al. (1995). DNA blot hybridization analysis was performed as previously described (Trebitsh et al. 1997).

Materials and Methods

Plant materials and growth conditions

Monococious (C. sativus L. var sativus cv Beit Alpha; f3M) and gynoecious (C. sativus L. var sativus cv Beit Alpha; FFMM) seeds were obtained from HAZERA GENETICS Ltd. (Berurim M.P. Shikmim, Israel). These are near-isogenic cucumber lines, i.e. the genomes are nearly identical except for the Female (F) gene. Under our experimental conditions, the monococious line produces male flowers while the gynoecious line produces female flowers only.

Seeds were germinated and grown in trays containing a mixture of peat and vermiculite (1:1) in a growth chamber under a 16:8 h photoperiod and 27 °C. Young leaves, plant growth apices or floral buds were harvested, frozen in liquid nitrogen and stored at −70 °C for genomic DNA extraction or RNA extraction.

DNA blot hybridization analysis

Genomic DNA was extracted from young leaves as described by Chetalat et al. (1995). DNA blot hybridization analysis was performed as previously described (Trebitsh et al. 1997).

Isolation of Cs-ACS1 and Cs-ACS1G and their 3'- and 5'-flanking sequences

The 5'-flanking region of the previously isolated Cs-ACS1 (U59813, Trebitsh et al. 1997) was amplified by inverse PCR (IPCR) on genomic DNA extracted from gynoecious and monococious cucumber according to Benkel and Fong (1996). First we cloned and sequenced the 2.5 kb MspI genomic fragment of Cs-ACS1 and the approximately 3.5 kb MspI genomic fragment of Cs-ACS1G. Genomic DNA was digested with MspI, self-ligated and amplified by PCR (Expand Long Template PCR System, Roche Diagnostics, Mannheim, Germany), using two back to back primers P1S and P1A. The primers were derived from the partial genomic sequence we determined previously (Fig. 1 and U59813, Trebitsh et al. 1997). For further upstream sequences, IPCR amplifications were performed with primers derived from the newly obtained gene-specific sequences: on gynoecious DNA digested with BstNI, and primers derived from the Cs-ACS1G sequences (P2S and P2A); and on monococious and gynoecious DNA digested with HindIII, and primers derived from the Cs-ACS1 sequences (PSS and P5A). The gene isolation strategy is depicted in Fig. 2A.

A complete sequence of Cs-ACS1 mRNA was isolated previously and named by the authors Cs-ACS1 (AB006805, Shiomi et al. 1998, Mathooko et al. 1999). To avoid confusion, we refer to this sequence as Cs-ACS1 mRNA. The full-length genomic sequences corresponding to the CDS were amplified by direct PCR using a primer from the 3’ UTR of the Cs-ACS1 mRNA (P3A) conjugated to a forward primer derived from the Cs-ACS1G-specific region (P3S) or a primer derived from the Cs-ACS1G-specific region (P2S). A total of 1,153 bp downstream of the Cs-ACS1/G 3’ UTR were obtained from gynoecious and monococious genomic DNA digested with AccI. The IPCR was performed with two back to back primers derived from the third exon of Cs-ACS1 (P4S and P4A). DNA fragments obtained by direct PCR or IPCR were cloned into the pGEMT-easy vector (Promega, Madison, WI, USA) and sequenced. Primer locations are indicated in Fig. 2; for primer sequences, see Supplementary Fig. S3.

Isolation of G-GM and M-GM

To isolate the GM fragment, we first performed a nested DNA walking reaction. The DNA walking reaction was performed on genomic DNA extracted from monococious cucumber, using the DNA Walking SpeedUp kit (Seegene, Seoul, Korea) and the Expand Long Template PCR System polymerase (Roche Diagnostics, Mannheim, Germany), according to the manufacturer’s instructions. Two nested primers in the forward direction, derived from the 5’-distal Cs-ACS1G (P2S and P7S), were used in conjugation with the primers provided in the kit. To isolate the 5’ sequences of the GM fragment, a direct PCR was performed on both gynoecious and monococious genomic DNA, using a primer in the sense direction derived from the 5’-distal Cs-ACS1 gene-specific region (P6F) and a primer in the antisense direction derived from the GM-specific region (P6R) obtained by the DNA walking reaction.

DNA sequence analysis

Three independent positive clones for each PCR-amplified fragment were sequenced in both directions to ensure sequence authenticity, and for each gene a single contig was constructed using the Sequencher program V. 4.5 (Gene Codes Co., Ann Arbor, MI, USA). Sequence analysis was carried out using BLAST searches (http://www.ncbi.nlm.nih.gov/BLAST/) (Altschul et al. 1997). Nucleotide and amino acid sequences were aligned using the MultAlin program (http://prodes.toulouse.inra.fr/multalin/multalin.html) and ClustalW at the EBI web server (http://www.ebi.ac.uk/cgi-bin/clustalw) (Corpet 1988, Thompson et al. 1994). Identification of putative cis-regulatory elements was performed using the PLACE database (http://www.dna.affrc.go.jp/PLACE/) (Higo et al. 1999).

Expression analysis of Cs-ACS1/G

Total RNA was extracted from different plant tissues by the EZ-RNA extraction kit according to the manufacturer’s instructions (Biological Industries, Beit Haemek, Israel). cDNA was synthesized from 5 µg of total RNA and M-MLV Reverse Transcriptase RNase H minus (Fermentas, Vilnius, Lithuania) and apoly(T) primer in a volume of 20 µL PCR with Taq DNA polymerase (Fisher Biotech Australia, Wembley, Australia) and 1 µl of the cDNA reaction was performed to amplify gene-specific products of Cs-ACS1/G and the cucumber ubiquitin gene as a reference standard gene. Cs-ACS1/G gene-specific primers were derived from the 5’ UTR (P5S) and the third exon (P8A) region (the primer location is indicated in Fig. 2; for primer sequences, see Supplementary Fig. S3). Ubiquitin gene-specific primers (P8S and P9A) were based on the published sequence of C. sativus ubiquitin extension protein (accession No, AY372537, Kim et al. 2004). The Cs-ACS1/G gene-specific primers span the two introns so that genomic DNA contamination would be detected. In all the experiments conducted, no genomic contamination was observed. These primers specifically amplified the target gene sequences (data not shown).

A 5 µl aliquot of the PCR products was separated on a 1.5% agarose gel and blotted onto Nytran Plus membrane (Schleicher & Schull, Dassel, Germany). The ubiquitin PCR product is approximately 500 nucleotides smaller than the Cs-ACS1/G PCR product. Probes corresponding to Cs-ACS1G and ubiquitin cDNAs were radiolabeled with [32P]dCTP by random priming (Random Primer DNA Labeling Mix, Biological Industries, Beit Haemek, Israel).
DNA blot hybridization and detection was performed as previously described (Trebstch et al. 1997). Following exposure to a phosphomager screen (FujiFilm BAS-1800, Tokyo, Japan), the band intensity was measured by an imaging densitometer (Image Gauge V3.0, FujiFilm, Tokyo, Japan). The ratio between the signal intensity of Cs-ACS1/G and that of ubiquitin represents the transcript level of Cs-ACS1/G and is expressed in arbitrary units. The relative expression level between tissue samples is calculated relative to the expression level in monoecious plant apices, whose transcript level was considered as 100%. Each experiment was performed on three independent RNA preparations.

Supplementary material
Supplementary material mentioned in the article is available to online subscribers at the journal website www.pcp.oupjournals.org.

Acknowledgments
This work was supported by the Chief Scientist of the Israeli Ministry of Agriculture and Rural Development (Project 854-0488-04). We thank Professor Dina Raveh for her critical reading of the manuscript.

References


Oka, A., Sakai, H. and Iwakoshi, S. (2002) His-Asp phosphorylation signal transduction in higher plants: receptors and response regulators for...


Aminocyclopropane-1-carboxylate synthase in tomato is encoded by a multigene family whose transcription is induced during fruit and floral senescence. *J. Mol. Biol.* 222: 937–961.


(Received May 19, 2006; Accepted July 21, 2006)