Mol. Cells, Vol. 10, No. 1, pp. 18–24

Molecules and Cells

Chromosomal Localization and Sequence Variation of 5S rRNA Gene in Five *Capsicum* Species

Youn-Kyu Park, Kyung-Cheol Park, Cheol-Ho Park, and Nam-Soo Kim*
Division of Applied Plant Sciences, Kangwon National University, Chunchon 200-701, Korea.

(Received on August 27, 1999)

Chromosomal localization and sequence analysis of the 5S rRNA gene were carried out in five *Capsicum* species. Fluorescence *in situ* hybridization revealed that chromosomal location of the 5S rRNA gene was conserved in a single locus at a chromosome which was assigned to chromosome 1 by the synteny relationship with tomato. In sequence analysis, the repeating units of the 5S rRNA genes in the *Capsicum* species were variable in size from 278 bp to 300 bp. In sequence comparison of our results to the results with other *Solanaceae* plants as published by others, the coding region was highly conserved, but the spacer regions varied in size and sequence. T stretch regions, just after the end of the coding sequences, were more prominant in the *Capsicum* species than in two other plants. High G·C rich regions, which might have similar functions as that of the GC islands in the genes transcribed by RNA PolII, were observed after the T stretch region. Although we could not observe the TATA like sequences, an AT rich segment at −27 to −18 was detected in the 5S rRNA genes of the *Capsicum* species. Species relationship among the *Capsicum* species was also studied by the sequence comparison of the 5S rRNA genes. While *C. chinense*, *C. frutescens*, and *C. annuum* formed one lineage, *C. baccatum* was revealed to be an intermediate species between the former three species and *C. pubescens*.

Keywords: 5S rRNA Genes; *Capsicum*; FISH; Sequence Variation; Species Relationship.

Introduction

The genes for 5S ribosomal RNA (5S rRNA) are present in multiple copies in eukaryotic genome to constitute a multigene family. In higher eukaryotes, the 5S rRNA genes are organized in tandem repeat of a basic unit, 200–900 bp long, ranging from 1,000 to 50,000 copies (Goldsbrough *et al*., 1982; Long and Dawid, 1980). The gene is 120 bp long and is associated with a spacer with various sizes to form the repeating unit of the tandem arrays. Sequence conservation of coding region and high divergence in the spacer regions provided a good model for studying the organization and evolution of multigenes in different plant species (Cox *et al*., 1992; Gottlob-McHugh *et al*., 1990; Scoles *et al*., 1988).

Chromosomal localization of the 5S rRNA gene has been investigated in a number of plant species by *in situ* hybridization. Depending on the species, the 5S rRNA genes were located on a chromosome (Lapitan *et al*., 1991; Mascia *et al*., 1981) or scattered on several chromosomes in more than a single location (Ellis *et al*., 1991; Taketa *et al*., 1999). In addition, the chromosomal locations of the 5S rRNA genes were either at distinct sites from the major rDNA clusters, 45S rDNA or 18S-5.8S-26S rDNA, (Appels, 1980; Lapitan *et al*., 1991) or at adjacent sites to the major rDNA clusters (Brown and Carlson, 1997; Castilho and Heslop-Harrison, 1995) in one or more chromosomes.

In the genus of *Capsicum*, over 40 species have been known of which only five species are cultivated. Pepper is one of the most important fruit vegetable crops. Although many studies on the genetics such as inheritance of genes (Daskolov and Poulos, 1994; Lippert *et al*., 1965; Tanksley *et al*., 1988) and linkage studies (Kim *et al*., 1997; Lefebvre *et al*., 1995; Prince *et al*., 1993) have been reported, corresponding studies on the cytogenetics have not been produced extensively (Pickersgill, 1971; Tanksley

* To whom correspondence should be addressed.
Tel: 82-361-250-6416; Fax: 82-361-256-9942
E-mail: kimnamsu@cc.kangwon.ac.kr

Abbreviations: BSA, bovine serum albumin; EDTA, ethylenediamine tetra acetic acid; FISH, fluorescent *in situ* hybridization;. FITC, fluorescein isothiocyanate; PCR, polymerase chain reaction.

et al., 1988). Localization and sequence characterization of the rRNA gene families have not been attempted in pepper, in contrast with the vast information available on these genes in tomato (Lapitan *et al*., 1991; Tanksley *et al*., 1988), which is one of the closest plant species to pepper. Therefore, we have undertaken an experiment to localize the 5S rRNA gene families on the chromosome by *in situ* hybridization and to associate the species relationships of the five cultivated *Capsicum* species by comparing 5S rDNA sequences.

Materials and Methods

Plant materials and genomic DNA extraction Five *Capsicum* species, *C. annuum*, *C. chinense*, *C. frutescens*, *C. baccatum*, and *C. pubescens*, were kindly provided by Professor Byung-Soo Kim, Kyungbook National University, Taegu, Korea. The preparation of the plant genomic DNA was carried out with a slight modification of the method of Edwards *et al*. (1991). Briefly, several pieces of leaf discs from young plants were crushed in a 1.5 ml microtube with liquid nitrogen. After adding 750 μl of extraction buffer (100 mM Tris pH 8.0, 500 mM NaCl, 50 mM EDTA, 1.25% SDS) to each tube, they were mixed by inverting the tubes several times. An equal volume of chloroform was then added, mixed, and incubated for 20 min at 60°C. The tube was spun at 13K for 15 min and the supernatant was transferred to a new tube to mix with an equal volume of ice-cold isopropyl alcohol to precipitate the genomic DNA by centrifugation for 15 min at 13K. The genomic DNA was dissolved with 200 μl of TE buffer containing RNase (20 μg/ml).

Cloning and sequencing of the 5S rRNA genes Two oligonucleotide primers, P1 5′-GATCCCATCAGAACTCC-3′ (17mer) and P2 5′-GGTGCTTTAGTGCTGGTAT-3′ (19mer), were used to amplify the 5S rRNA genes in the *Capsicum* species. The sequences of the primers were derived from the highly conserved regions in coding regions (Cox *et al*., 1992; Venkateswarlu, 1991). The PCR reaction mixture comprised of 1 ng template DNA, 0.5 pmol of each primer, 200 μM each dNTP, 2.5 μl of 10× reaction buffer, 1 unit of *Taq* DNA polymerase (Dynazyme, Finland). Thirty amplification cycles were performed, each involving denaturation at 94°C for 1 min, annealing at 57°C for 1 min, and extension at 72°C for 2 min. The amplified 5S rDNA fragments were cloned into the pGEM-T easy vector and sequenced with an LKB automatic DNA sequencer using Cy5™ AutoCycle™ Sequencing Kit (Pharmacia LKB Biotech). The sequence alignment and comparison were performed with the DNASTAR-PC program.

***In situ* hybridization** Chromosome preparation was made with actively growing root-tips using the squashing method with a drop of 45% acetic acid. After the cover slides were removed using the quick freezing method, the slides were dehydrated with a series of EtOH (70%/80%/95%) for 5 min each, and air-dried. The slides were treated with RNase (100 μg/ml in 2× SSC) for 1 h at 37°C, washed with 2× SSC for 5 min at room temperature, and dehydrated again with the EtOH series. Chromosomal DNAs were denatured by dipping the slides in denaturing solution (70% formamide in 2× SSC) for 2.5 min at 70°C and dehydrated again as above. The probe was made with Biotin-14-dATP by nick translation. Probe mixture (25 ng labeled 5S rDNA, 250 ng yeast tRNA, 250 ng herring sperm DNA, 2× SSC, 50% formamide, 100 mM phosphate, 1× Denhardt's solution, pH 6.5 and 10% dextran sulfate) was also denatured at 70°C for 5 min and cooled in ice, and 15 μl of the denatured probe mixture was applied to each slide, covered with cover slide, and sealed with rubber cement. The chromosomal DNAs and probe DNA were denatured again at 70°C for 3 min, cooled to room temperature, and incubated in a humid chamber overnight at 37°C for hybridization. After the cover slides were removed, the slides were washed in 50% formamide in 2× SC 5 min/3 times and 0.5× SSC 5 min/3 times at 37°C. After treating with blocking agent (3% BSA in 4× SSC, 200 μl/slide) for 20 min at 37°C, a drop of FITC-avidin (Borehringer Mannheim, 5 μg/ml in 1% BSA, 4× SSC, 0.1% Tween-20) was applied to each slide to incubate for 30 min at 37°C. Then excess fluorescent-conjugate was removed with washing solution (4× SSC, 0.1% Tween-20) for 3 min/3 times at 42°C. A drop of biotinylated anti-avidin (Vector Laboratories, 5 μg/ml in 1% BSA, 4× SSC, 0.1% Tween-20) was applied to each slide and incubated for 30 min at 37°C. This solution was then washed with the same washing solution as above. The FITC-avidin was treated and washed as for signal amplification above. The slides were rinsed briefly in distilled water and mounted with 20 μl of antifading agent (Vector Shield) containing propium iodine (0.5 μg/ml). The signals were visualized and captured with an Axiophot fluorescent microscope (Zeiss).

Results and Discussion

Amplifying and cloning the 5S rRNA genes PCR amplification of the 5S rDNA in the five *Capsicum* species are shown in Fig. 1. Three prominent bands with approximate sizes of 300, 600, 900 bp, which were revealed to be monomer, dimer, trimer of the 300 bp repeating unit in subsequent sequencing analysis, appeared in all five *Capsicum* species. The PCR products were directly cloned into the pGEM-T easy vector and the

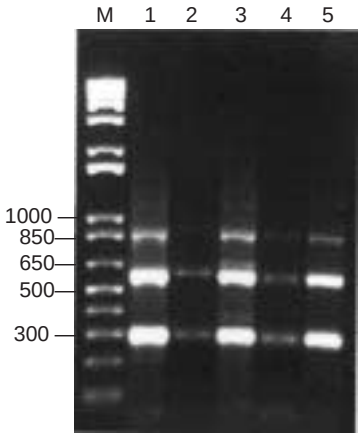

Fig. 1. PCR amplified fragments of the 5S rRNA genes in the *Capsicum* species. M, size marker; 1, *C. annuum*; 2, *C. chinense*; 3, *C. frutescens*; 4, *C. baccatum*; 5, *C. pubescens.*

clones with an approximate insert size of about 300 bp were selected for further characterization by *in situ* hybridization and sequencing analysis.

Fluorescent *in situ* hybridization Fluorescent *in situ* hybridization was carried out to localize the 5S rRNA gene locus (Fig. 2). In all five species, there was only one set of bright hybridization signals in the short arm of a set of chromosomes. The single locus in each species was expected based on the fact that only a single repeating unit of 300 bp appeared in PCR amplification. The chromosome carrying the 5S rRNA gene was assigned to chromosome 1 by synteny relationship with tomato, since the locus information was not available in both the cytological map and the linkage map in pepper. In tomato, Lapitan *et al.* (1991) showed the location of the 5S rRNA locus in the middle of the short arm of chromosome number 1 and placed this locus in linkage group 1. Conservation on a single locus in a chromosome across the species in *Solanaceae* is unusual compared to the *Gramineae* species. High variation in chromosomal locations, number of loci, and number of chromosomes carrying the 5S rRNA genes were reported in barley (Taketa *et al*., 1999), *Phaseolus* (Danna *et al*., 1996), *Aegilops* (Castilho and Heslop-Harrison, 1995), and wheat (Dvořák *et al*., 1989; Kellogg and Appels, 1995; Lagudah *et al*., 1989). By *in situ* hybridization, major rRNA genes (45S rRNA) were localized on chromosome 5 in *C. annuum* and chromosomes 5 and 6 in *C. chinense* (Tanksley *et al*., 1988). However, our study revealed that the 5S rRNA genes were not on the chromosomes with secondary constrictions which are the sites of the major rRNA gene families expressed. Recently, Park and Kim (1999) have shown that the 5S rRNA genes were not on the same chromosomes as the 45S rRNA genes in *C. annuum*, *C. chinense*, and *C. frutescens*. However the 5S rRNA genes were proximal locations to the 45S rRNA gene loci in the short arm of chromosome 1 in *C. baccatum* and *C. pubescens* in their report. Physical linkages of the 5S rRNA locus with the major rDNA loci were known in many other plant species (Badaeva *et al*., 1996; Brown and Carlson, 1997; Castilho and Heslop-Harrison, 1995; Taketa *et al*., 1999).

5S rRNA gene sequence analysis Sequences revealed the precise size of the 5S rDNA, 294 bp in *C. annuum*, 300 bp in *C. baccatum*, 298 bp in *C. chinense*, 296 bp in *C. frutescens*, and 278 bp in *C. pubescens*, respectively (Fig. 3). The pepper seemed to have the shortest 5S rDNA repeating units among the *Solanaceae* 5S rRNA gene families which were to have 324–329 bp in potato (Zanke *et al*., 1995), 345 bp in tomato (Venkateswarlu *et al*., 1991), and 451 bp in petunia (Frasch *et al*., 1989) (Table 1). GC content in the coding region was higher than the spacer region (Table 1). Sequence homology in the 5S rRNA genes among the five *Capsicum* species varied from 69.4% between *C. frutescens* and *C. pubescens* to 89.2%

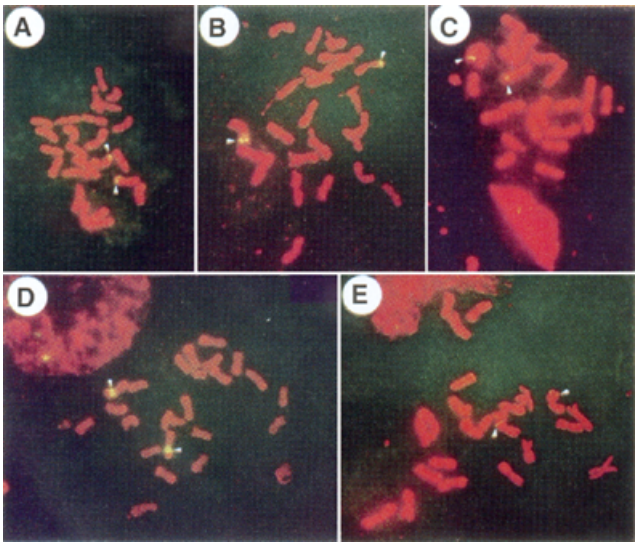

Fig. 2. Chromosomal locations of the 5S rRNA genes in *Capsicum* species. The arrow heads indicate the hybridization sites of the 5S rDNA probe. **A.** *C. annuum*; **B.** *C. chinense*; **C.** *C. frutescens*; **D.** *C. baccatum*; **E.** *C. pubescens*.

```
Consensus     GGATGCGATC  ATACCAGCAC  TAAAGCACCG  GATCCCATCA  GAACTCCGAA
C.annuum      -T--------  ----------  ----------  ----------  ----------
C.baccatum    ----------  ----------  ----------  ----------  ----------
C.chinense    ----------  ----------  ----------  ----------  ----------
C.frutescens  ----------  ----------  ----------  ----------  ----------
C.pubescens   ----------  ----------  ----------  ----------  ----------

Consensus     GTTAAGCGTG  CTTGGGCGAG  AGTAGTACTA  GGATGGGTGA  CCCCCTGGGA
C.annuum      -----C----  ----------  ----------  ------C---  -------C--
C.baccatum    ----------  ----------  ----------  ----------  ----------
C.chinense    ----------  ----------  ----------  ----------  ----------
C.frutescens  ----------  ----------  ----------  ----------  ----*-----
C.pubescens   ----------  ----------  ----------  ----------  ----------

Consensus     AGTCCTCGTG  TTGCATCCCT  T*********  TCTTTTTTTT  TTGTCCTTTA
C.annuum      ----------  -------G--  ----------  CT--------  -***------
C.baccatum    ----------  ----------  -CCTCTTTTT  ----------  --TC-----G
C.chinense    ----------  ----------  ----------  ----------  ----------
C.frutescens  ----------  ----------  ----------  ----------  --*-------
C.pubescens   ----------  ----------  -C--------  ********---  -----****G

Consensus     AAATTCGGTT  TAATTTTGCC  GGTTC**GAT  CGGCGAATCT  TTTTG*TTTT
C.annuum      ----------  ----------  ----------  ----------  ----------
C.baccatum    ----C--A--  -------T--  ----------  ----------  ----------
C.chinense    ----------  --------T-  ----------  ---A------  ----------
C.frutescens  --T-------  --------TT  A---------  ----------  ----------
C.pubescens   --------G-  -G--------  --C--CT--C  ---T----T-  -C-C-GC-C-

Consensus     TCCTCTTGGC  GGAAACGAGA  AAGGGG*AAA  GGCGCAGGGG  CGTGCGTGCG
C.annuum      ----------  --------C-  ------GC--  ----------  --------A-
C.baccatum    --G-------  ----------  --T----***  ----***A-T  --C-------
C.chinense    ----------  -----T--A-  ------C---  ---A------  --C------A
C.frutescens  --------A-  -A--------  ------A-G-  -------T--  ---T------
C.pubescens   ----------  -----G-C--  G-C----***  ----***---  --A-***---

Consensus     GGGTGTGACG  GGCGGCGCAG  GAGAATGAGG  CTTAATAGAA  TTTGGAGTGC
C.annuum      ----------  -------G--  ----*---A-  -AA-------  ----------
C.baccatum    ----------  ----------  -----G----  ---------T  ----------
C.chinense    ----------  ----------  --------T-  ------T---  ----------
C.frutescens  ----------  A---------  ----------  -------A--  ----------
C.pubescens   A--C-***--  -----G----  ----GG----  ----------  --GA---C--

Consensus     TGGGAATGAC
C.annuum      ----------   (294)
C.baccatum    ----------   (300)
C.chinense    ---------T   (298)
C.frutescens  ------A---   (296)
C.pubescens   ----------   (278)
```

Fig. 3. Sequences of the 5S rDNA of the five different *Capsicum* species. Deleted bases are designated by asterisks. These sequences have been deposited in GenBank as Accession Nos. AF217950, AF217951, AF217952, AF217953, and AF217954.

Table 1. Length, GC-content of the 5S rDNA of the five different *Capsicum* species, potato, and tomato.

Species	Length (bp)			GC-content (%)		
	5S rDNA repeat	Coding region	Spacer	5S rDNA repeat	Coding region	Spacer
C. annuum	294	120	174	51.4	55.0	48.9
C. baccatum	300	120	180	50.1	55.0	47.8
C. chinense	298	120	178	49.3	55.0	45.5
C. frutescens	296	119	177	48.3	55.5	43.5
C. pubescens	278	120	158	57.6	55.0	59.5
Potato[a]	324	119	205	50.0	56.3	46.3
Tomato[b]	355	120	235	49.3	55.8	46.0

[a] According to Zanke *et al*. (1995)
[b] According to Venkateswarlu *et al*. (1991)

between *C. chinense* and *C. frutescens*. If the sequences of the *Capsicum* species were compared to the 5S rRNA genes in other *Solanaceae* species such as potato and tomato, *Capsicum* species showed higher homology with tomato than potato. While the coding region was highly conserved (by as much as 91.6% to 100%), spacer sequences were not highly conserved (Table 2).

The coding sequence started with GGA and ended with CCT to make 120 bp in *C. baccatum*, *C. chinense*, *C. frutescens*, and *C. pubescens*. However, *C. annuum* started with GGT and ended with GCT. The length of the coding sequence of *C. frutescens* was 119 bp since a cytosine at the position of 95 from the transcription starting site was deleted. Variable sequences in the start sites have been reported in other plant species such as GGA or TGA in *Gramineae* species (Baum and Johnson, 1996), AGG in legume species (Ellis *et al*., 1988; Gottlob-McHugh *et al*., 1990), GGA in *Solanaceae* (Zanke *et al*., 1995), AGA in petunia (Frasch *et al*., 1989), and GGG in *Brassica* species (Kim *et al*., 1998). The variable sequences in the terminal region of the coding region were also reported as TCT in soybean (Gottlob-McHugh *et al*., 1990) and pea (Ellis *et al*., 1988), GTG or TCC in barley (Baum and Johnson, 1996). Among the five *Capsicum* species, *C. annuum* showed the most variable sequences in the coding region. While five bases were changed in *C. annuum*, perfect homology of the 120 bases in coding region was found between *C. baccatum*, *C. chinense*, and *C. pubescens* (Table 2, Fig. 3). Except for one base missing at the position of 95 from the start site, *C. frutescens* also showed perfect homology with the latter three species. All five base changes in *C. annuum* were transversion, one A→T, three G→C, one C→G. Kellogg and Appels (1995) also noted higher transversion mutations in 5S rRNA genes in *Triticeae*. Ellis *et al*. (1988) showed the sequence variation of the 5S rRNA genes in 19 different plant species. It is interesting to note that the base change sites were not random in the coding region. Among the 120 bp of the 5S rDNA nucleotides, the nucleotides in 36 sites were found to have changed in our study. The sites of the base changes in the *C. annuum* were also in five sites among 36 sites, which might indicate that there would be some selection pressure for other regions. Instead of the protein coding function, the rRNA nucleotides make the nucleolus by the formation of secondary structure and association of small nuclear proteins. Therefore, the conserved nucleotides in the 5S rRNA coding region might be involved in forming the nucleolus by secondary structure or associating with proteins. The 5S rRNA genes are transcribed by RNA PolIII. This polymerase does not bind to the DNA directly, but rather depends on the interaction with tanscription factor IIIA (TFIIIA) (Hayes and Tullius, 1992). The TFIIIA binds to the newly synthesized RNA molecule (Clemens *et al*., 1993). Therefore, this 5S rRNA may not be transcribed efficiently to be screened out, if the nuceotides associating with the TFIIIA are mutated (Kellogg and Appels, 1995).

As reported previously by others, the spacers were highly variable in sequence and length among the *Capsicum* species in contrast to the high conservation of the coding regions (Fig. 4). The shortest one was *C. pubescens* with 158 bp and the longest one was *C. baccatum* with 180 bp (Table 1). The length of the spacers in *Capsicum* species seemed to be shorter than other 5S rRNA genes in *Solanaceae* species (Zanke *et al*., 1995). *C. baccatum* and *C. pubescens* showed highly variable base insertion/deletion mutations in the spacer. While nine extra bases from the site of 122 to 130 were present, four bases between 227 and 230 and three bases between 235 and 237 were deleted in *C. baccatum*. In *C. pubescens*, although one extra base at the site of 122 and two bases at sites of 176 and 177 were present, deletions of 28 bases in multiple sites made the spacer of 5S rRNA gene 158 bp long. High rates of deletion mutations in the spacer regions were also reported in other plant species (Baum and Johnson, 1995; Cox *et al*., 1992). In the presence of several 5S rDNA loci in a genome in wheat and barley, the length differences in the spacer region differentiate into the several loci (Cox *et al*., 1992; Dvořák *et al*., 1989).

Table 2. Homology of the 5S rDNA sequences of the five different *Capsicum* species, potato, and tomato.

Homology (%)		*C. annuum*	*C. baccatum*	*C. chinense*	*C. frutescens*	*C. pubescens*	Potato
5S rDNA	*C. baccatum*	83.3		–	–	–	–
repeated	*C. chinense*	88.8	85.9		–	–	–
	C. frutescens	85.7	82.4	89.2		–	–
	C. pubescens	70.9	75.2	73.0	69.4		–
	Potato[a]	56.8	57.0	56.0	55.7	61.5	
	Tomato[b]	59.2	62.3	60.4	58.4	64.4	70.1
Coding	*C. baccatum*	95.8		–	–	–	–
region	*C. chinense*	95.8	100		–	–	–
	C. frutescens	91.6	95.8	95.8		–	–
	C. pubescens	95.8	100	100	95.8		–
	Potato[a]	90.0	94.2	94.2	99.2	94.2	
	Tomato[b]	94.2	98.3	98.3	94.1	98.3	94.2
Spacer	*C. baccatum*	76.4		–	–	–	–
	C. chinense	86.2	79.2		–	–	–
	C. frutescens	84.5	75.1	85.3		–	–
	C. pubescens	52.5	58.2	54.4	51.9		–
	Potato[a]	35.6	35.0	36.5	34.5	36.7	
	Tomato[b]	31.0	38.9	37.1	36.7	39.9	56.1

[a] According to Zanke *et al*. (1995)
[b] According to Venkateswarlu *et al*. (1991)

Like other terminators, the most conspicuous feature was the A·T rich region with T residues predominating in the spacer region just after the G·C rich 3′-end of the coding region (Fig. 3). The A·T rich segments were also known in 5S rRNA genes in other species (Baum and Johnson, 1996; Ellis *et al*., 1988; Venkateswarlu *et al*., 1991) as well as the genes transcribed by RNA PolII (Platt, 1986; Rosenberg and Court, 1979). In comparison with the spacer sequences from other plants in *Solanaceae* as published in elsewhere, potato and tomato were similar to *Capsicum* in having the feature of T stretch residues just after the coding sequences, with high variation in length of the T residues and sequences (Fig. 4). In *Capsicum* species, CGG or CGGCG sequences were found in the middle of the T residues, which were not prominent in potato and tomato. This A·T rich segment region was the most variable in the whole 5S rRNA repeating unit.

Following the A·T rich segment, highly conserved G·C rich segments appeared (Table 1, Fig. 3). Since the genes transcribed by RNA PolII also have GC box in the upstream region of the TATA box, the G·C rich segments may have similar functions for RNA PolIII to transcribe 5S rRNA genes. The potato and tomato showed extra sequences in the regions of between 95 and 105, between 162 to 169, and between 184 and 195 from the spacer start site (Fig. 4). Although we could not find TATA-like sequences in the upstream region from the transcription start site in *Capsicum*, the eleven A·T rich bases from 210 to 220 may have a similar function as TATA. In soybean

```
Consensus   TC********  TCTTTTTT*T  TTTGTCCTT*  TAAAATTCGG  TTTAAT**TT
C.annuum    -*--------  CT--------  --***-----  ----------  ----------
C.baccatum  --CTCTTTTT  ----------  ---TC-----  -G----C--A  ----------
C.chinense  -*--------  ----------  ----------  ----------  ----------
C.frutescens -*-------  ----------  ---*------  ----T-----  ----------
C.pubescens ----------  *******---  ------***-  *G--------  G--G------
Potato      C-TTTTTGTC  GA**--CGG-  *-C--G*-AA  --G--AAAAA  -A-T--TA--
Tomato      C-TTTTTGTT  GAAA---GA-  C-C--AA--G  A----AAAAA  -A--C-CA—

Consensus   TGTCGGTTC*  *GATCGGCGA  ATCTTTTTG   TTT****TTC  CTCT******
C.annuum    --C-------  ----------  ----------  ----------  ----------
C.baccatum  -TC-------  ----------  ----------  ----------  G---------
C.chinense  ----------  -------A--  ----------  ----------  ----------
C.frutescens ---TA-----  ----------  ----------  ----------  ----------
C.pubescens --C---C--C  T--C---T--  --T--C-C-G  C-C-------  ----------
Potato      -A-TTT--T-  -****A-AAC  -CTCG--C--  ***----***  *A--AACATA
Tomato      -A-TTT--TT  TTTG---AA-  --ACG--C-G  A--GAGGCGT  -AT-AGGATA

Consensus   *****TGGCG  GAAACGAGAA  AGGGGGTGA*  *GGCG****C  ATGGGC**GT
C.annuum    ----------  -------C--  ------CA--  ----------  -G--------
C.baccatum  ----------  ----------  -T---****-  ---------*  **A-T----C
C.chinense  ----------  ---T---A--  -----CAA--  ----A-----  -G-------C
C.frutescens -------A--  A---------  -----AA---  ----------  T---------
C.pubescens ----------  ----G-C--G  -C----***-  ---------*  **-------A
Potato      TGGG-*****  *CGTT-GAG-  T--*C-----  G----GCGT-  ---C-TCG--
Tomato      TGGGA---T-  -CGT--G-G-  T---C----C  G----TCGT-  GT—C-TCG--

Consensus   GCGTGCGGGG  T********G  TGACGGGCGG  CGC*******  *****AGGAG
C.annuum    -----A----  ----------  ----------  --G-------  ----------
C.baccatum  ----------  ----------  ----------  ----------  ----------
C.chinense  ------A---  ----------  ----------  ----------  ----------
C.frutescens ----------  ----------  -----A----  ----------  ----------
C.pubescens -***---A--  C---------  ***-------  G---------  ----------
Potato      ---C-T--A-  GCTAGGTCG-  --GG----A-  G*-TAGGGCG  TTGGG----A
Tomato      -----GA---  -TTTAA--A-  C-GG------  G—TAGGGCG   TTGGG----A

Consensus   AGTGATGCTT  AATAGAATTT  GGAGTGCTGG  GAATGAC
C.annuum    -*---A--AA  ----------  ----------  -------  (174)
C.baccatum  -AG--G----  ------T---  ----------  -------  (180)
C.chinense  -A--------  ---T------  ----------  ------T  (178)
C.frutescens -A---G----  ----A-----  ----------  ---A---  (177)
C.pubescens --G--G----  ---------G  A---C-----  -------  (158)
Potato      G-A-G--T--  ----------  AA------AT  ------T  (205)
Tomato      G--TG--T--  ------T---  A------AAT  -------  (235)
```

Fig. 4. Sequence comparison in the 5S rDNA spacers in the five *Capsicum* species with potato and tomato. The sequences of the potato and tomato were derived from Zanke *et al*. (1995) and Venkateswarlu *et al*. (1991), respectively. The numbers at the end denote the length of the spacer in each species. Deleted bases are designated by asterisks.

5S rRNA gene, Gottlob-McHugh *et al.* (1990) argued that the AATATA in the spacer region was the TATA-like sequence since 5S rRNA genes from other legume species also had the same sequences (Ellis *et al*., 1988; Hemleben and Wert 1988; Rafalski *et al*., 1982). Since several conserved sequences above and below the transcription start site were known to control the transcription of the 5S rRNA genes (Pieler *et al*., 1987), this might be a loose requirement for the TATA sequence for transcription.

Species relationship in *Capsicum* species Morphological characters have been used for species identification and species relationship in the *Capsicum* species (Andrews, 1984). In studies of several isozymes, McCleod *et al*. (1979) could not differentiate between *C. annuum*, *C. frutescens*, and *C. chinense*. In a breeding experiment by intercrossing between *Capsicum* plants, *C. annuum*, *C. frutescence*, and *C. chinense* were proposed as being a single unit genetically linked from a wild ancestral gene pool (Eshbaugh, quoted by Andrews, 1984). *C. baccatum* and *C. pubescens*, however, formed separate units in the lineage, respectively.

By comparing the 5S rRNA gene sequences, Eshbaugh's proposal was substantiated since *C. chinense*, *C. frutescens*, and *C. annuum* formed a separate lineage from *C. baccatum* and *C. pubescens* (Fig. 5). Although sequences in the coding regions of *C. pubescens*, *C. baccatum*, and *C. chinense* showed perfect homology, *C. chinense* and *C. frutescens* were the closest two species with 89.2% base sharing of total 5S rDNA repeated (Table 2). The close phylogenetic relationship between these two species may be backed up by the geographical overlapping of their origins in South America (Andrews, 1984). The primitive form of *C. frutescens* was known to be the ancestor of *C. chinense* (http://easyweb.easynet.co.uk/~gcaselton/chile/variety.html). The next closest was *C. annuum* with 88.8% base sharing with *C. chinense* and 85.7% sharing with *C. frutescens*. *C. baccatum* might be the intermediate species between the three clustered species and *C. pubescens*. *C. pubescens* showed the most divergence in 5S rRNA gene sequences among the five *Capsicum* species. *C. pubescens* also showed distinct morphological features. While the other four *Capsicum* species have white flowers and glabrous leaves and stems, *C. pubescens* has purple flowers and pubescent leaves and stems.

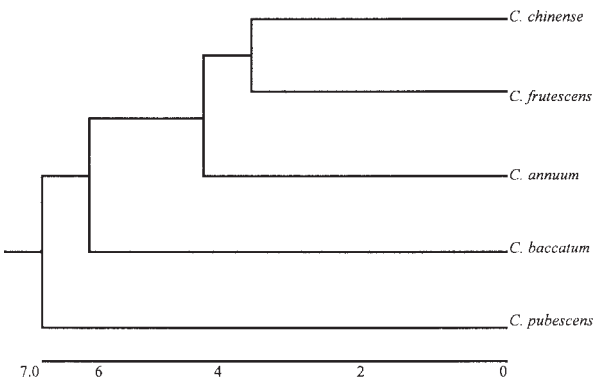

Fig. 5. A model depicting the evolution of *Capsicum* species based on the 5S rRNA gene sequences.

Acknowledgments This work was carried out with a grant from the R&D Promotion Center for Agriculture, Forestry, and Fishery in Korea.

References

Andrews, J. (1984) *Peppers, The Domesticated Capsicums*, University of Texas Press, Austin, USA.

Appels, R., Gerlach, W. L., Dennis, E. S., Swift, H., and Peacock, W. J. (1980) Molecular and chromosomal organization of DNA sequences coding for ribosomal RNAs in cereals. *Chromosoma* **78**, 293–311.

Badaeva, E. D., Friebe, B., and Gill, B. S. (1996) Genome differentiation in *Aegilops*. 2. Physical mapping of 5S and 18S–26S ribosomal RNA gene families in diploid species. *Genome* **39**, 1150–1158.

Baum, B. R. and Johnson, D. A. (1996) The 5S rRNA gene units in ancestral two-rowed barley (*Hordeum spontaneum* C. Koch) and bulbous barley (*H. bulbosum* L.): sequence analysis and phylogenetic relationships with 5S rDNA units of cultivated barley (*H. vulgare* L.). *Genome* **39**, 140–149.

Brown, G. R. and Carlson, J. E. (1997) Molecular cytogenetics of the genes encoding 18S-5.8S-26S rRNA and 5S rRNA in two species of spruce (*Picea*). *Theor. Appl. Genet.* **95**, 1–9.

Castilho, A. and Heslop-Harrison, J. S. (1995) Physical mapping of 5S and 18S–25S rDNA and repetitive sequences in *Aegilops umbellulata*. *Genome* **38**, 91–96.

Clemens, K. R., Wolf, R. V., McBryant, S. J., Zhang, P., Liao, X., Wright, P. E., and Gottesfeld, J. M. (1993) Molecular basis for specific recognition of both RNA and DNA for a zinc finger proteins. *Science* **260**, 530–533.

Cox, A. V., Bennet, M. D., and Dyer, T. A. (1992) Use of the polymerase chain reaction to detect spacer size heterogeneity in plant 5S-rRNA gene clusters and to locate such clusters in wheat (*Triticum aestivum* L.). *Theor. Appl. Genet.* **83**, 684–690.

Danna, K. J., Workman, R., Coryell, V., and Keim, P. (1996) 5S rRNA genes in tribe Phaseoleae: array size, number, and dynamics. *Genome* **39**, 445–455.

Daskalov, S. and Poulos, J. M. (1994) Updated *Capsicum* gene list. *Capsicum and Eggplant Newsl.* **13**, 15–26.

Dvořák, J., Zhang, H.-B., Kota, R. S., and Lassner, M. (1989) Organization and evolution of the 5S ribosomal RNA gene family in wheat and related species. *Genome* **32**, 1003–1009.

Edwards, K., Johnstone, C., and Thomson, C. (1991) A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. *Nucleic Acids Res.* **19**, 1349.

Ellis, T. H. N., Lee, D., Thomas, C. M., Simpson, P. R., Cleary, W. G., Newman, M. A., and Burcham, K. W. G. (1988) 5S rRNA Genes in *Pisum*: sequence, long range and chromosomal organization. *Mol. Gen. Genet.* **214**, 333–342.

Frasch, M., Wenzel, W., and Hess, D. (1989) The nucleotide sequences of nuclear 5S rRNA genes and spacer regions of *Petunia hybrida*. *Nucleic Acids Res.* **17**, 2857–2857.

Goldsbrough, R. B., Ellis, T. H. N., and Lomonossoff, G. D.

(1982) Sequence variation and methylation of the flax 5S RNA genes. *Nucleic Acids Res.* **10**, 450–4514.

Gottlob-McHugh, S. G., Levesque, M., MacKenzie, K., Olson, M., Yarosh, O., and Johnson, D. A. (1990) Organization of the 5S rRNA genes in the soybean *Glycine max* (L.) Merrill and conservation of the 5S rDNA repeat structure in higher plant. *Genome* **33**, 486–494.

Hayes, J. J. and Tullius, T. D. (1992) Structure of the TFIIIA-5S DNA complex. *J. Mol. Biol.* **227**, 407–417.

Hemleben, V. and Werts, D. (1988) Sequence organization and putative regulatory elements in the 5S rRNA genes of two higher plants (*Vigna radiata* and *Matthiola incana*). *Gene* **62**, 165–169.

Kellogg, E. A. and Appels, R. (1995) Intraspecific and interspecific variation in 5S RNA genes are decoupled in wheat relatives. *Genetics* **140**, 325–343.

Kim, B.-D., Kang, B.-C., Nam, S.-H., Kim, B.-S., Kim, N.-S., Lee, M.-W., and Ha, K.-S. (1997) Construction of a molecular linkage map and development a molecular breeding technique. *J. Plant Biol.* **40**, 156–163.

Kim, S.-Y., Lim, Y.-P., and Bang, J.-W. (1998) Cytogenetic analysis of *Brassica campestris* var. *pekinensis* using C-banding and FISH. *Kor. J. Genet.* **20**, 285–294.

Lagudah, E. S., Clarke, B. C., and Appels, R. (1989) Phylogenetic relationships of *Triticum tauschii*, the D-genome donor to hexaploid wheat. 4. Variation and chromosomal location of 5S DNA. *Genome* **32**, 1017–1025.

Lapitan, N. L., Ganal, M. W., and Tanksley, S. D. (1991) Organization of the 5S ribosomal RNA genes in the genome of tomato. *Genome* **34**, 509–514.

Lefebvre, V., Palloix, A., Caranta, C., and Pochard, E. (1995) Construction of an intraspecific integrated linkage map of pepper using molecular markers and doubled-haploid progenies. *Genome* **38**, 112–121.

Lippert, L. F., Bergh, B. O., and Smith, P. G. (1965) Gene list for pepper. *J. Hered.* **56**, 30–34.

Long, E. O. and Dawid, I. B. (1980) Repeated gene in eukaryotes. *Annu. Rev. Biochem.* **43**, 727–764.

Mascia, P. N., Rubenstein, I., Phillips, R. L., Wang, A. S., and Xiang, L. Z. (1981) Localization of the 5S rRNA genes and evidence for diversity in the 5S rDNA region in maize. *Gene* **15**, 7–20.

McCleod, M., Eshbaugh, W. H., and Guttman, S. I. (1979) A preliminary biochemical systematic study of the genus *Capsicum-Solanaceae*; in *The Biology and Taxonomy of the Solanaceae*, Hawkes, J. G., and Lester, R. N., and Skelding, A. D. (eds.), pp. 102–120, Academic Press, New York.

Park, Y.-K. and Kim, N.-S. (1999) Species relationship in *Capsicum* species by molecular karyotyping and 5S rRNA gene sequence. The 11th Annual Meeting of the Korean Society for Molecular Biology, Seoul National University, Seoul, Korea, H-81.

Pickersgill, B. (1971) Relationships between weedy and cultivated forms in some species of chilli peppers (genus *Capsicum*). *Evolution* **25**, 683–691.

Pieler, T., Hamm, J., and Roeder, R. G. (1987) The 5S gene internal control region is composed of three distinct sequence elements, organized as two functional domains with variable spacing. *Cell* **48**, 91–100.

Platt, T. (1986) Transcription termination and regulation of gene expression. *Annu. Rev. Biochem.* **55**, 339–372.

Prince, J. P., Pochard, E., and Tanksley, S. D. (1993) Construction of a molecular linkage map of a pepper and a comparison of synteny with tomato. *Genome* **36**, 404–417.

Rafalski, J. A., Wiewiorowski, M., and Soll, D. (1982) Organization and nulceotide sequence of nuclear 5S rRNA genes in yellow lupin (*Lupinus leteus*). *Nucleic Acids Res.* **10**, 7635–7642.

Rosenberg, M. and Court, D. (1979) Regulatory sequences involved in the promotion and temination of RNA transcription. *Annu. Rev. Genet.* **13**, 319–354.

Scoles, G. J., Gill, B. S., Xin, Z.-Y., Clarke, B. C., McIntyre C. L., Chapman, C., and Appels, R. (1988) Frequent duplication and deletion events in the 5S RNA genes and the associated spacer regions of the *Triticeae*. *Plant Syst. Evol.* **160**, 105–122.

Taketa, S., Harrison, G. E., Heslop-Harrison, J. S. (1999) Comaparative physical mapping of the 5S and 18S-25S rDNA in nine *Hordeum* species and cytotypes. *Theor. Appl. Genet.* **98**, 1–9.

Tanksley, S. D. (1984) Linkage relationships and chromosomal locations of enzyme-coding genes in pepper, *Capsicum annuum*. *Chromosoma* **89**, 352–360.

Tanksley, S. D., Bernatzky, R., Lapitan, N. L., and Prince, J. P. (1988) Conservation of gene repertoir but not gene order in pepper and tomato. *Proc. Natl. Acad. Sci. USA* **85**, 6419–6423.

Venkateswarlu, L., Lee, S. W., and Nazar, R. N. (1991) Conserved upstream sequence elements in plant 5S ribosomal RNA-encoding genes. *Gene* **105**, 249–253.

Zanke, C., Borisjuk, N., Rouss, B., Schilde-Rentschler, L., Ninnemann, H., Hemleben, V. (1995) A species specific oligonulceotide of the 5S rDNA spacer and species-specific elements identify symmetric somatic hybrids between *Solanum tuberosum* and *S. pinnatisectum*. *Theor. Appl. Genet.* **90**, 720–726.