Journal of Plant Biology, December 2007, 50(6) : 687-691

SHORT COMMUNICATION

Identification of a Molecular Marker and Chromosome Mapping of the 5S rRNA Gene in *Allium sacculiferum*

Jun Hyung Seo[1], Byung Ha Lee[2], Bong Bo Seo[1], and Ho-Sung Yoon[1]*
[1]Department of Biology, Kyungpook National University, Daegu 702-701, Korea
[2]Agro-Biotechnology Education Center, Kyungpook National University, Daegu 702-701, Korea

The 5S rRNA gene in higher eukaryotes is organized into repeated units of tandem array that comprise a conserved 120-bp coding region and a non-transcribed spacer (NTS) of variable length with nucleotides. The allotetraploid genome of *Allium sacculiferum* consists of two unknown diploids (2n=32). Analyses have not been successful toward clarifying the origin of each genome due to their similar chromosome morphology and unmatched C-banding patterns. We PCR-amplified the coding and NTS regions of its 5S rRNA genes, cloned them into vectors, and determined their DNA sequences. Interestingly, the aligned sequences of the NTS clones could be divided into two distinctive groups based on the existence of a 3-bp CCT insertion/deletion at the beginning of the NTS region. This feature makes it an important genetic marker for distinguishing the origin of the *A. sacculiferum* chromosomes. Furthermore, by applying fluorescence *in situ* hybridization, we located the 5S rRNA gene loci on Chromosomes 5, 7, 8, 9, and 14; their distribution is unique to *A. sacculiferum*. These data support the idea that one set of this genome has originated from a CCT-containing close relative -- *A. deltoid-fistulosum* -- and that the NTS region may be used as a molecular marker for identifying parental lines for the allotetraploidity of *A. sacculiferum*.

Keywords: *Allium sacculiferum*, allotetraploid, 5S rRNA, molecular marker, sequence variation

In higher plants, nuclear-encoded ribosomal RNA genes are organized into two distinct multigene families that are clustered in two tandem arrays. One family is formed by moderate repeats of a transcription unit (45S rRNA) that includes three ribosomal RNA molecules located at a secondary constriction of particular chromosomes. The other family comprises 5S rRNA genes that are organized into tandem repetitive units located at one or more loci on at least one chromosome (Goldsborough et al., 1981; Samson and Wegnez, 1984; Lee et al., 1999; Baum and Bailey, 2001). Each repetitive unit includes a highly conserved 120-bp sequence coding the 5S rRNA plus a divergent non-transcribed region that is typically 100 to 900 bp long (Appels et al., 1980; Long and David, 1980; Flavell, 1986; Rogers et al., 1986; Rogers and Bendich, 1987; Specht et al., 1990; Lewin, 2004). 5S rRNA genes have been studied in several plant species (Mascia et al., 1981; Lee et al., 1999; Tronin et al., 1999; Do and Seo, 2000).

The phylogenetic relationship within *Allium* species has been examined using such molecular markers as RAPDs, RFLPs, and AFLPs (Klaas, 1998; Fritsch and Friesen, 2002), sequences from the nuclear ribosomal ITS region (Mes et al., 1999), and sequences from the 5S non-transcribed region (Do and Seo, 2000). Cronn et al. (1996) have reported that most nucleotide positions in the spacer region are expected to change because variants are selectively neutral and, consequently, those variant nucleotides can become fixed, causing definite interspecific differences to accumulate. By contrast, most mutations in the 5S gene are expected to be selectively neutral only when they occur in a subcritical position of repeats in the array. Therefore, fixed differences between species fail to accumulate in 5S genes despite their potential for undergoing mutations equal to those in the NTS sequences. The somatic chromosomes of *Allium sacculiferum* are known to be allotetraploid (2n=XXYY=32, plus zero to several B chromosomes with different morphology), comprising chromosomes from two unknown diploid species. Although genome analyses have been conducted to determine the composition of the original genomes for that species, no evidence has revealed the family relationship due to similar chromosome morphologies and unmatched C-banding patterns (Seo and Kim, 1989).

As in many tandem-repeating multigene families, numerous copies of ribosomal DNA families can undergo concerted evolution because of homogenizing forces. This can result in sequences of all gene copies from such a family being virtually identical within a species, despite the presence of normal levels of divergence between orthologous genes in different species (Brown et al., 1972; Dover, 1982; Wendel et al., 1995; Liu et al., 2003; Yoon, 2003). Nevertheless, in several plant groups, sequence heterogeneity among 5S rDNA repeats within genomes has been reported, indicating that homogenizing forces have not been strong enough to overcome the processes that generate such variations (Kellogg and Appels, 1995; Cronn et al., 1996; Campbell et al., 1997).

In a phylogenetic study, polymorphisms of informative factors, e.g., DNA sequence variations, play an important role when analyzing the molecular relationship within or among genera. Two typical types of polymorphism can be described: 1) an autapomorphic site with only a single, unique nucleotide aligned to other nucleotides; or 2) a synapomorphic site, which is usually preferred as an informative factor in phylogenetic evaluations where it is considered the main molecular factor when classifying the relationships between species, providing sequence variations in nucleotides for at least two accessions from the aligned sequences (Maughan et al., 2006).

To overcome the difficulties of identifying the chromo-

*Corresponding author; fax +82-53-953-3066
e-mail hyoon@knu.ac.k

somes with C-bands in satellite and terminal regions, fluorescence *in situ* hybridization (FISH) has been widely performed to identify the exact localization of species-specific repetitive, ribosomal genes and unique sequences (Mukai et al., 1990; Maluzynska and Heslop-Harrison 1991; Leitch and Heslop-Harrison, 1993; Do et al., 1999). Multigene families, e.g., 5S rDNA, 18S-5.8S-26S rDNA, and highly repetitive DNA sequences, have now been mapped to indicate more significant molecular factors with the corresponding chromosomal relationships between species (Lapitan et al., 1989; Mukai et al., 1990; Leitch and Heslop-Harrison, 1993; Do et al., 1999).

In the current study, we hypothesized that polymorphic parental types might be maintained in the hybrid genome of the allotetraploid for *A. sacculiferum*, and that it may be possible to identify different genomes by analyzing multiple numbers of 5S rDNA units and NTS sequences in this species. Our aim was to characterize the structural features of the 5S rDNA repeat unit via sequencing analysis and FISH.

MATERIALS AND METHODS

Plant Materials, Chromosome Preparation, and DNA Extraction

We collected 76 plants of *A. sacculiferum* from Uedo, a South Sea island of Korea, and grew them in a greenhouse. From these, fresh root tips were excised and transferred to ice water for 21 h at 0°C to allow for the metaphase to accumulate before being fixed in ethanol-acetic acid (3:1). Cytological preparations for chromosome analyses were obtained using the methods described by Mukai et al. (1993). Genomic DNA was extracted from fresh young leaves according to the protocol of Rogers and Bendich (1988).

Amplification and Cloning of 5S rRNA Genes

PCR was used to amplify the 5S rRNA genes, including the coding and NTS regions. We designed primer sequences of the forward primer (corresponding to nucleotides 30 to 47: 5′-GGATCCCATCAGAACTCC-3′) and the reverse primer (nucleotides 11 to 29: 5′-GGTGCTTTAGTGCTGGTAT-3′) based on previous reports of the conserved coding region for the 5S rRNA gene repeat units. The reaction mixture (25 μL) contained 10 ng of template (i.e., total genomic DNA); 5 pM of each primer; 200 mM each of dATP, dGTP, dCTP, and dTTP; and 2.5 U of Ex Taq polymerase in a 1X Ex Taq buffer with 2 mM $MgCl_2$. PCR was performed in a thermal cycler (Takara, Japan). Amplification involved a preliminary 2-min denaturation at 94°C; then 1 min at 94°C (denaturation), 1 min at 55°C (annealing), and 1 min at 72°C (extension); and a final extension for 7 min at 72°C, followed by slow cooling to 4°C. PCR products were visualized by electrophoresis in 1% agarose gels, and stained with ethidium bromide. Integral multiples of the smallest size were found among the DNA fragments within a set of discrete sizes. A prominent band corresponding to the full length of a single unit from the 5S rRNA genes was isolated with a Gel extraction kit (Qiagen, USA), ligated into a pMD 18-T vector (Takara, Japan), and introduced into DH5a *E. coli* competent cells by transformation (Hanahan, 1983). Recombinant plasmids were obtained using a Miniprep kit (Qiagen, USA).

Sequencing Analyses and Alignments

To identify the nucleotide sequences for both the coding and the NTS regions in the 5S rRNA gene unit, DNA sequences of the inserts that originated from 40 randomly selected colonies were determined by the chain termination method through a commercial sequencing service (Genotech, Korea). These sequences were then aligned by the Clustal W multiple alignment program. The Bio-edit program was used for visual-editing to aid in dissecting homologies and specific variations (Thompson et al., 1994; Reddy et al., 2003).

Fluorescence *in Situ* Hybridization

Fluorescence *in situ* hybridization (FISH) was used to ascertain the physical location of the 5S rRNA gene unit. Chromosome samples were prepared by the 1% acetic-carmine method of Mukai et al. (1990). For probe-labeling, sequenced plasmid DNA with the insert of the 5S rRNA gene was amplified with digoxigenin-11-dUTP (Roche Diagnostics, Germany). The reaction mixture and PCR conditions were the same as those described for the experiments with genomic DNA, except for the following substitutions: 140 μM dTTP and 60 mM digoxigenin-11-dUTP instead of 200 μM dTTP. Chromosomal DNA on slides was denatured for 2 min at 67°C in 70% formamide-20XSSC (1XSSC is 0.15 M NaCl with 0.015 M sodium citrate; pH 7.0), then dehydrated immediately in an ethanol series (70%, 95%, and 100%) for 5 min each at –20°C. Afterward, the probe mixture (PCR products with digoxigenin-11-dUTP, formamide, 50% dextran sulfate, 20XSSC, and 500 mg mL^{-1} of salmon sperm DNA) was applied to the slides before hybridizing them at 37°C for 12 h in a humidified chamber. These hybridized slides were washed in a series of 2XSSC, 50% formamide, 2XSSC, 1XSSC and 4XSSC, then incubated at 37°C for 1 h, followed by staining with anti-digoxigenin-rhodamine (Roche Diagnostics, Germany). The slides were again washed with 2XSSC, 4XSSC/0.1% triton-100, and 4XSSC, before overlaying cover slips with counterstaining by 4,6-diamino-2-phenylindole (DAPI) and the Vectashield mounting medium (Vector, USA). Signal detection for each slide was examined with an epifluorescence microscope (Zeiss Axiophot, Germany).

RESULTS AND DISCUSSION

The PCR products of the 5S rRNA gene were cloned into the pMD 18-T vector. Clones with an approximate insert size of 360 bp were selected for further characterizations. A total of 10 clones with inserts were isolated (pAsac in Fig. 1) and the DNA sequence analysis revealed two distinctive groups, ASAC1 and ASAC2 (Fig. 2). All the clones examined were unique and varied in size. The coding regions of the 5S rDNA showed high homology with previously identified 5S rRNA genes from *Pinus* (Liu et al., 2003), *Capsicum* (Park

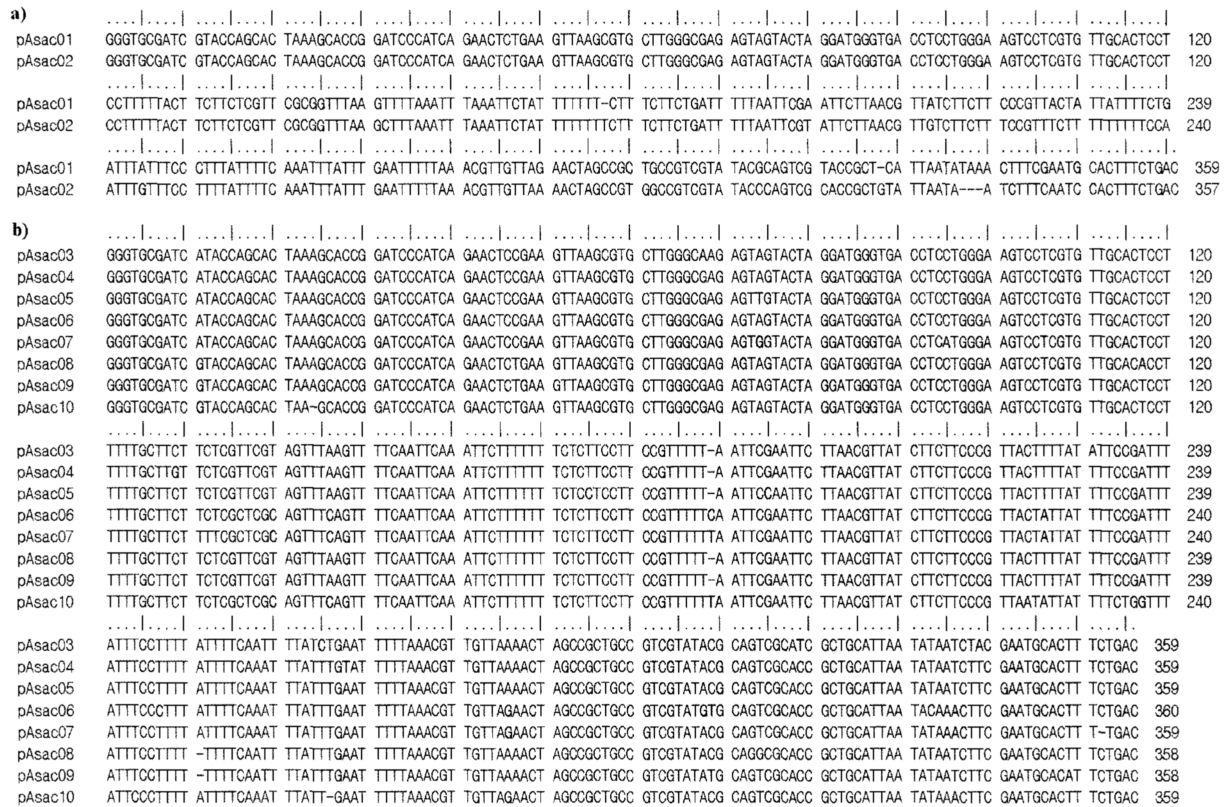

Figure 1. Clones containing 5S rDNA from *A. sacculiferum* were separated into 2 distinctive groups: **a**) aligned sequences of 5S rRNA gene containing 3 bp of CCT, with polymorphic sites indicated by gray background; and **b**) aligned sequences of 5S rRNA gene without 3 bp of CCT, with polymorphic sites indicated by gray background.

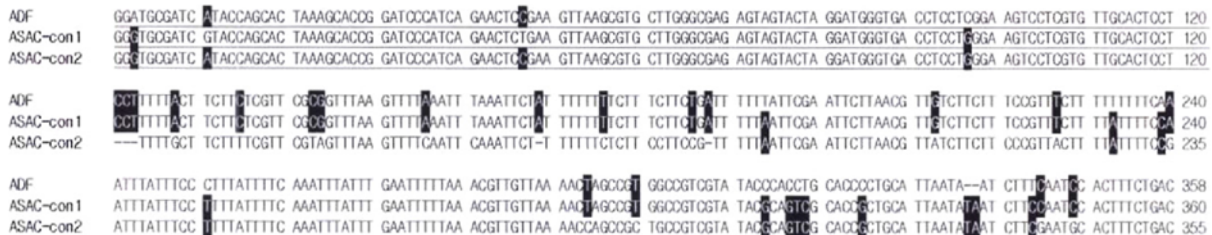

Figure 2. Comparison of 2 consensus sequences for 5S rRNA gene from *A. sacculiferum* with 5S rDNA repeat units from *A. deltoid-fistulosum*. ASAC-con1 stands for consensus sequence from group of clones containing 3 bp of CCT; clones designated for ASAC-con2 do not have such sequences. The first 120 bp (underlined) indicate coding region from 5S rRNA gene. Deleted bases are represented by dash lines, and sequence differences among 3 groups are shown by black rectangles.

et al., 2000), and the Poaceae family (Baum and Johnson, 1999). We examined 120 bp of the coding regions and found that they were highly conserved, starting with 5'-GGGTGCGA and ending with GCACTCCT-3' (positions 1 to 120 in Fig. 1). Four polymorphic sites were located in ASAC2 whereas no variations were found in ASAC1 (Table 1). DNA sequences for all 5S rRNA coding regions were highly homologous.

In contrast, the sequences of the NTS region were variable (see also Park et al., 2000; Liu et al., 2003). Based on the two distinctive groups, 22 polymorphic sites were found in the NTS of ASAC1, including 1 deletion; these were all autapomorphic. In contrast, ASAC2 showed 24 sites, including 4 deletions and 1 synapomorphie (Table 1). Despite these numbers of polymorphic sites, two reasonable consensus sequences could be constructed as ASAC-con1 and ASAC-con2 by reading the major group of sequences within each polymorphic site. A significant informative factor of 3 bp of insertion/deletion (CCT) was found at the beginning of the spacer; 80% of the clones showed the deletion in ASAC-con2, but none appeared in ASAC-con1 (Fig. 2). The aligned consensus sequences in the two groups revealed variations at 20 sites, whereas the coding region was highly conserved except for the base substitution of guanine with alanine at Position 9. The polymorphism of 5S rRNA coding and the variations in NTS regions within many plant and animal species have been widely used as genetic markers to distinguish closely related species, subspecies, lines, and hybrids, as well as to analyze evolutionary relationships (Crisp et al., 1999). Therefore, when one considers the composition of the original genome for this species, we might presume that each group came from different origins.

Table 1. Sequence analysis and autapomorphic, synapomorphic, and deletion nucleotide polymorphism in the coding and NTS regions from *Allium sacculiferum*.

Taxon	Code	Clones isolated	Unique clone	Size of clones isolated	Nucleotide Polymorphism					
					Autapomorphie		Synapomorphie		Deletion	
					Coding	NTS	Coding	NTS	Coding	NTS
Allium sacculiferum	ASAC1	2	2	357 - 359	0	21	0	0	0	1
	ASAC2	8	8	358 - 360	4	18	0	1	1	4

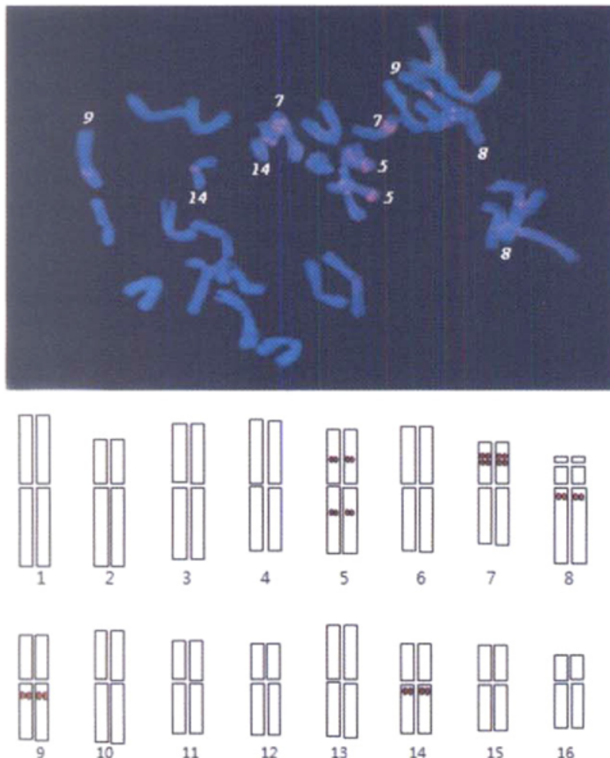

Figure 3. Chromosomal localization of 5S rRNA gene loci in *A. sacculiferum*. Digoxigenin-labelled 5S rRNA genes were detected by anti-digoxigenin-rhodamine conjugated (Spectrum-Red) on DAPI (blue)-counterstained metaphase chromosomes. Idiogram represents location of 5S rRNA gene loci in species (circles) on each chromosome.

Moreover, the deletion of 3 bp in the NTS region can be very useful when distinguishing between related species and, thus, identifying the origin of *A. sacculiferum*.

Fluorescent *in situ* hybridization was carried out to localize the 5S rRNA gene locus in *A. sacculiferum*. Figure 3 indicates the patterns of the 5S rRNA gene on the metaphase, as well as the idiogram for chromosomes carrying the 5S rRNA gene loci in that species. Signals were detected at Chromosomes 5, 7, 8, 9, and 14, and distribution of the 5S rRNA loci was unique to *A. sacculiferum*. Due to the highly conserved region of 5S rRNA coding and the similar sequence pattern of NTS, we were unable to determine the two isolated distinctive groups by FISH.

Our data from the sequencing analysis present evidence that one set of the *A. sacculiferum* genome has originated from a CCT-containing close relative species, *A. deltoid-fistulosum*. Therefore, this demonstrates that the 5S rRNA gene is useful for identifying parental lines for allotetraploids. Although the significant patterning of 5S rRNA in our two groups of *A. sacculiferum* could not be localized via FISH because of the highly conserved sequences in both regions for coding and NTS, the information obtained from each consensus sequence will be very beneficial when analyzing more significant relationships during the molecular phylogenetic study of *Allium* species. Further research should focus on finding an origin and sequencing for each genome. FISH analyses of the 5S rRNA in this species also will enable us to confirm the relationship of other closely related species in this genus.

ACKNOWLEDGEMENTS

This work was supported by a grant (20070301-034-021-007-01-00) from the BioGreen 21 Program, Rural Development Administration, Republic of Korea.

Received August 21, 2007; accepted November 7, 2007.

LITERATURE CITED

Appels R, Gerlach WL, Dennis ES, Swift H, Peacock WJ (1980) Molecular and chromosomal organization of DNA sequences coding for the ribisomal RNAs in cereals. Chromosoma 78: 293-311

Baum BR, Bailey LG (2001) The origin of A genome donor of wheats (*Triticum*: Poaveae) - A prospective based on the sequence variation of the 5S DNA gene units. Genet Resource Crop Evol 51: 183-196

Baum BR, Johnson DA (1999) The 5S rRNA gene in wall barley (*Hordeum murinum* L. sensu lato): Sequence variation among repeat units and relationship to the Y haplome in the genus *Hordeum* (Porceae: Triticeae). Genome 42: 852-866

Brown DD, Wensink PC, Jordan E (1972) A comparison of the ribosomal DNAs of *Xenopus laevis* and *Xenopus mulleri*: The evolution of tandem genes. J Mol Biol 63: 57-72

Campbell CS, Wojciechowski MF, Baldwin BG, Alice A, Donoghue MJ (1997) Persistent nuclear ribosomal DNA sequences polymorphism in the *Amelanchier* agamic complex (Rosaceae). Mol Biol Evol 14: 81-90

Crisp MD, Gilmore, SR, Weston PH (1999) Phylogenetic relationships of two anomalous species of *Pultenaea* (Fabaceae: Mirbelieae), and description of a new genus. Taxon 48: 701-704

Cronn RC, Zhao X, Paterson AH, Wendel JF (1996) Polymorphism and concerted evolution in a tandemly repeated gene family: 5S ribosomal DNA in diploid and allopolyploid cottons. J Mol Evol 42: 685-705

Do GS, Seo BB (2000) Phylogenetic relationships among *Allium* subg. *Rhizirideum* species based on the molecular variation of 5S rRNA genes. Kor J Biol Sci 4: 77-85

Do GS, Seo BB, Ko JM, Lee SH, Park JH, Kim IS, Song SD (1999) Analysis of somaclonal variation through tissue culture and chromosomal localization of rDNA sites by fluorescent in situ hybridization in wild *Allium tuberosum* and a regenerated variant. Plant Cell Tiss Org Cult 57: 113-119

Dover GA (1982) Molecular Drive: A cohesive mode of species evolution. Nature 299: 111-117

Flavell RB (1986) Repetitive DNA and chromosomes. Phil Trans Royal Soc London B Biol Sci 312: 227-242

Fritsch RM, Friesen N (2002) Evolution, domestication and taxonomy. *In* HD Ravinovitch, L Currah, eds, *Allium* Crop Science: Recent Advances. CAB international, Wallingford, England, pp 5-27

Goldsborough PB, Ellis TH, Cullis CA (1981) Organization of the 5S RNA genes in flax. Nucl Acids Res 9: 5895-5904

Hanahan D (1983) Studies on transformation of *Escherichia coli* with plasmids. J Mol Biol 166: 557-580

Kellogg EA, Appels R (1995) Intraspecific and interspecific variation in 5S RNA genes are decoupled in diploid wheat relatives. Genetics 140: 325-343

Klaas M (1998) Applications and impact of molecular markers on evolutionary and diversity studies in the genus *Allium*. Plant Bleed 117: 297-308

Lapitan NLV, Ganal MW, Tanksley SD (1989) Somatic chromosome karyotypes of tomato based on in situ hybridization of the TAGI satellite repeat. Genome 32: 992-998

Lee SH, Do GS, Seo BB (1999) Chromosomal localization of 5S rRNA gene loci and the implications for relationships within the *Allium* complex. Chrom Res 7: 89-93

Leitch IJ, Heslop-Harrison JS (1993) Physical mapping of four sites of 5S ribosomal DNA sequences and one site of the x-amylase 2 gene in barley (*Hordeum vulgare*). Genome 36: 517-523

Lewin B (2004) Genes III. Prentice Hall, Pearson Education, Inc., Upper Saddle River, NJ, USA

Liu ZL, Zhang D, Wang XQ, Ma XF, Wang XR (2003) Intragenomic and interspecific 5S rDNA sequence variation in five Asian pines. Amer J Bot 90: 17-24

Long EO, David IB (1980) Repetitive genes in eukaryotes. Annu Rev Biochem 49: 727-764

Maluszynska J, Heslop-Harrison JS (1991) Localization of tandemly-repeated DNA sequences in *Arabidopsis thaliana*. Plant J 1: 159-166

Maughan PJ, Kolano BA, Maluszynska J, Coles ND, Bonifacio A, Rojas J, Coleman CE, Stevens MR, Fairbanks DJ, Parkinson SE, Jellen EH (2006) Molecular and cytological characterization of ribosomal RNA genes in *Chenopodium quinoa* and *Chenopodium berlandieri*. Genome 49: 825-839

Mes THM, Fritsch RM, Pollner S, Bachmann SK (1999) Evolution of the chloroplast genome and polymorphic ITS regions in *Allium* subg. Melanocrommyum. Genome 42: 237-247

Mascia PN, Rubenstein I, Philips RL, Wang AS, Xiang LZ (1981) Localization of the 5S rRNA genes and evidence for diversity in the 5S rRNA region of maize. Gene 18: 7-20

Mukai Y, Endo TR, Grill BS (1990) Physical mapping of the18S-26S rRNA multigene family in common wheat. J Hered 81: 290-295

Mukai Y, Nagahara Y, Yamamoto M (1993) Simultaneous discrimination of the three genomes in hexaploid wheat by multicolor fluorescence in situ hybridization using total genomic and highly repeated DNA probes. Genome 35: 489-495

Park YK, Park HC, Park CH, Kim NS (2000) Chromosomal localization and sequence variation of 5S rRNA gene in five *Capsicum* species. Mol Cells 10: 18-24

Reddy AD, Suh SB, Ghaffari R, Singh NJ, Kim DJ, Han JH, Kim KS (2003) Bioinformatics analysis of SARS proteins and molecular dynamics simulated structure of an alpha-helix motif. Bull Kor Chem Soce 24: 899-900

Rogers SO, Bendich AJ (1987) Ribosomal gene in plants, variability in copy number and in the intergenic spacer. Plant Mol eBiol 9: 509-520

Rogers SO, Bendich AJ (1988) Extraction of DNA from plant tissue, *In* SB Gelvin, RA Schilperoort, eds, Plant Molecular Manual A6. Kluwer Academic Publishers, Dordrecht, pp 1-10

Rogers SO, Honda S, Bendich AJ (1986) Variation in the ribosomal RNA genes among individuals of *Vicia fava*. Plant Mol Biol 6: 339-345

Samson ML, Wegnez M (1984) The 5S ribosomal genes in *Drosophila melanogaster* species subgroup. Nucleotide sequence of a 5S unit from *Drosophila simulance* and *Drosophila teissieri*. Nucl Acids Res 12: 1003-1014

Seo BB, Kim HH (1989) Giemsa C-banded karyotypes in two diploid and two tetraploid *Allium* species. Kor J Bot 32: 181-188

Specht T, Wolters J, Erdmann VA (1990) Compilation of 5S rRNA and 5S rRNA gene sequences. Nucl Acids Res 18: 2215-2230

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: Improving the sensitivity of progressive multiple alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucl Acids Res 16: 486-499

Tronin J, Grandemange C, Favre J (1999) Two highly divergent 5S rDNA unit size classes occur in composite tandem array in European larch (*Larix decidua* Mill.) and Japanese larch (*Larix kaempferi* (Lamb) Carr.). Genome 42: 837-848

Wendel JF, Schnabel A, Seelanan T (1995) Bidirectional interlocus concerted evolution following allopolyploid speciation in cotton (*Gossypium*). Proc Natl Acad Sci USA 92: 280-284

Yoon H-S (2003) A floral meristem identity gene influences physiological and ecological aspect of floral organogenesis. J Plant Biol 46(4): 271-276