Biotechnology & Biotechnological Equipment, 2014
Vol. 28, No. 3, 431–437, http://dx.doi.org/10.1080/13102818.2014.928499

Taylor & Francis
Taylor & Francis Group

ARTICLE; AGRICULTURE AND ENVIRONMENTAL BIOTECHNOLOGY

Tomato (*Solanum lycopersicum*) variety discrimination and hybridization analysis based on the 5S rRNA region

Yan-Lin Sun[a]*, Ho-Min Kang[b], Young-Sik Kim[c], Jun-Pill Baek[d], Shi-Lin Zheng[a], Jin-Jun Xiang[a] and Soon-Kwan Hong[e,f]

[a]*School of Life Sciences, Ludong University, Yantai, Shandong, P. R. China;* [b]*Department of Horticulture, Kangwon National University, Chuncheon, Korea;* [c]*Department of Plant & Food Science, Sangmyung University, Cheonan, Korea;* [d]*Agricultural and Life Science Research Institute, Kangwon National University, Chuncheon, Korea;* [e]*Department of Bio-Health Technology, Kangwon National University, Chuncheon, Korea;* [f]*Institute of Bioscience and Biotechnology, Kangwon National University, Chuncheon, Korea*

(*Received 10 November 2013; accepted 2 January 2014*)

The tomato (*Solanum lycopersicum*) is a major vegetable crop worldwide. To satisfy popular demand, more than 500 tomato varieties have been bred. However, a clear variety identification has not been found. Thorough understanding of the phylogenetic relationship and hybridization information of tomato varieties is very important for further variety breeding. Thus, in this study, we collected 26 tomato varieties and attempted to distinguish them based on the 5S rRNA region, which is widely used in the determination of phylogenetic relations. Sequence analysis of the 5S rRNA region suggested that a large number of nucleotide variations exist among tomato varieties. These variable nucleotide sites were also informative regarding hybridization. Chromas sequencing of Yellow Mountain View and Seuwiteuking varieties indicated three and one variable nucleotide sites in the non-transcribed spacer (NTS) of the 5S rRNA region showing hybridization, respectively. Based on a phylogenetic tree constructed using the 5S rRNA sequences, we observed that 16 tomato varieties were divided into three groups at 95% similarity. Rubiking and Sseommeoking, Lang Selection Procedure and Seuwiteuking, and Acorn Gold and Yellow Mountain View exhibited very high identity with their partners. This work will aid variety authentication and provides a basis for further tomato variety breeding.

Keywords: tomato variety; 5S rRNA region; variety discrimination; phylogenetic relationship; genetic diversity

Introduction

Ribosomal 5S RNA (5S rRNA) genes, which are present in multiple copies in the eukaryotic genome, are the most widely used gene family for the determination of phylogenetic relations among plant and animal species. In higher eukaryotes, 5S rRNA genes exist in tandem repeats; the number of repeats varies from less than 1000 to more than 75,000.[1,2] These genes contain the coding region and a non-transcribed spacer (NTS) region. The coding region of 5S rRNA genes is highly conserved and commonly 120 bp in length, whereas the NTS region is of different size in different species depending on the coding region and exhibits high variation. As reported in previous studies, NTS regions are very variable not only in nucleotide identity but also in sequence length in plants.[3,4] NTS lengths nearly closely vary among different loci in a genome, and high rates of nucleotide deletion in the NTS occur in some plant species.[3,5] Consequently, the 5S rRNA region, which exhibits coding region sequence conservation and high divergence within the NTS regions, has been considered a good model for studying the organization and evolution of multigene families in various plant species.[3,6,7]

Tomato (*Solanum lycopersicum*) is a major vegetable crop grown worldwide. Due to its good flavour and high nutrient value, the tomato is an economically important agricultural crop around the world.[8,9] With higher demand due to increased living standards, more tomato varieties have been bred, including Belle, Rally, Campari and Temptation. Based on their morphological and physiological characteristics, these new varieties have been determined to be distinct from all existing varieties 'in common knowledge' by at least one character by the relevant Variety Identification Department.[10] These tomato varieties are reported to meet established standards with respect to uniformity and stability of the characteristics used to demonstrate distinctness.[10] However, the number of registered tomato varieties remains too large to efficiently check the distinctness, uniformity and stability of their characteristics. To further discriminate tomato varieties and understand their phylogenetic relationships, we used molecular markers and DNA profiling techniques

*Corresponding author. Email: ylsun04@mails.gucas.ac.cn

that have been widely used for the systematic identification of plants, fungi and even animals.[11,12] Molecular methods provide results that are independent of environmental influences and high levels of polymorphism. Therefore, in this study, we aimed to analyse the phylogenetic relationship of 26 tomato varieties based on sequence variation analysis of 5S rRNA genes. The observed sequence variation can elucidate phylogenetic and hybridization information.

Materials and methods

Plant materials

Twenty-six tomato varieties registered at the Korea Seed & Variety Service were investigated in the present study. All tomato varieties investigated in this study were provided by Kangwon National University, Chuncheon, Korea. The morphological characteristics and other traits of the tomatoes including fruit size, colour and viral resistance are shown in Table 1. Fresh mature leaves from these tomato varieties were sampled and immediately stored in liquid nitrogen until DNA extraction.

Isolation of DNA, polymerase chain reaction (PCR) amplification and sequencing

DNA extractions were performed using the modified cetyltrimethylammonium bromide (CTAB) method.[13] The 5S rRNA gene was amplified using the 5SF (5′-CGGTGCATTAATGCTGGTAT-3′) and 5SR (5′-CCATCAGAACTCCGCAGTTA-3′) primer set [14] in a 20 μL polymerase chain reaction (PCR) reaction. PCR was performed using a Gene Amp 9700 PCR system (Applied Biosystems Incorporated, Warrington, Cheshire, UK) with the following reaction components: 1 μL of template DNA (approximately 1−100 ng), 10 × Ex Taq Buffer (TaKaRa Bio Inc., Japan), 200 μmol L^{-1} of each desoxynucleoside triphosphate (dNTP), 0.1 μmol L^{-1} of each primer, and 0.1 μL of TaKaRa Ex Taq (5 units μL^{-1}, TaKaRa Bio Incorporated, Japan). The PCR protocol included an initial denaturation step of 94 °C for 1 min, followed by 35 cycles of denaturation at 95 °C for 1 min, annealing at 53 °C for 1 min and extension at 72 °C for 1.5 min; the reaction concluded with a final extension step at 72 °C for 5 min. The amplification products were verified by electrophoresis in a 1.0% agarose gel and purified

Table 1. Description of 26 tomato varieties including fruit shape, colour, viral resistance and NCBI GenBank accession numbers of the 5S rRNA genes.

No.	Tomato variety	Fruit shape	Fruit colour	Viral resistance	NCBI GenBank Acc. No.
(1)	Black Kiss 20	Round	Black	–	–
(2)	Black Kiss 20	Round	Black	–	–
(3)	Mini Chal	Chinese date-shaped	Red	–	–
(4)	Vitamini	Chinese date-shaped	Red	–	KF156909
(5)	Rikopin 9	Chinese date-shaped	Red	–	–
(6)	Rikopin 9	Chinese date-shaped	Red	–	KF156910
(7)	Sseommeoking	Chinese date-shaped	Red	–	KF156911
(8)	Yellow Mountain View	Chinese date-shaped	Yellow	–	KF603895
(9)	Acorn Gold	Chinese date-shaped	Yellow	–	KF603896
(10)	Gold Sugar	Chinese date-shaped	Yellow	–	KF156912
(11)	Saenggeurinbichwibol	Chinese date-shaped	Grass green	–	KF156913
(12)	Abstract Saenggeurin	Round	Grass green	–	KF156914
(13)	Seuwiteuking	Round	Red	–	KF156915
(14)	Cutie	Round	Red	–	KF156916
(15)	Rubiking	Round	Red	TY Resistance	KF156917
(16)	Ten Ten	Round	Red	–	–
(17)	Yo-Yo Captain	Round	Red	–	–
(18)	Unicorn	Round	Red	–	KF156918
(19)	Hoyong	–	Peachblow	–	KF156919
(20)	Lang selection procedure	–	Peachblow	–	KF156920
(21)	Rafito	–	Peachblow	–	KF156921
(22)	Nice Def	Round or oval	Deep red	TY Resistance	–
(23)	Max Thailang	Heart-shaped	Deep red	TY Resistance	KF156922
(24)	Madison	Quail egg-shaped	Deep red	–	–
(25)	Campari	Golf ball-shaped	Deep red	–	–
(26)	Temptation	Flat round to round	Deep red	–	–

Notes: '–': undetected or undetermined. 'TY Resistance': Tomato yellow leaf curl virus resistance.

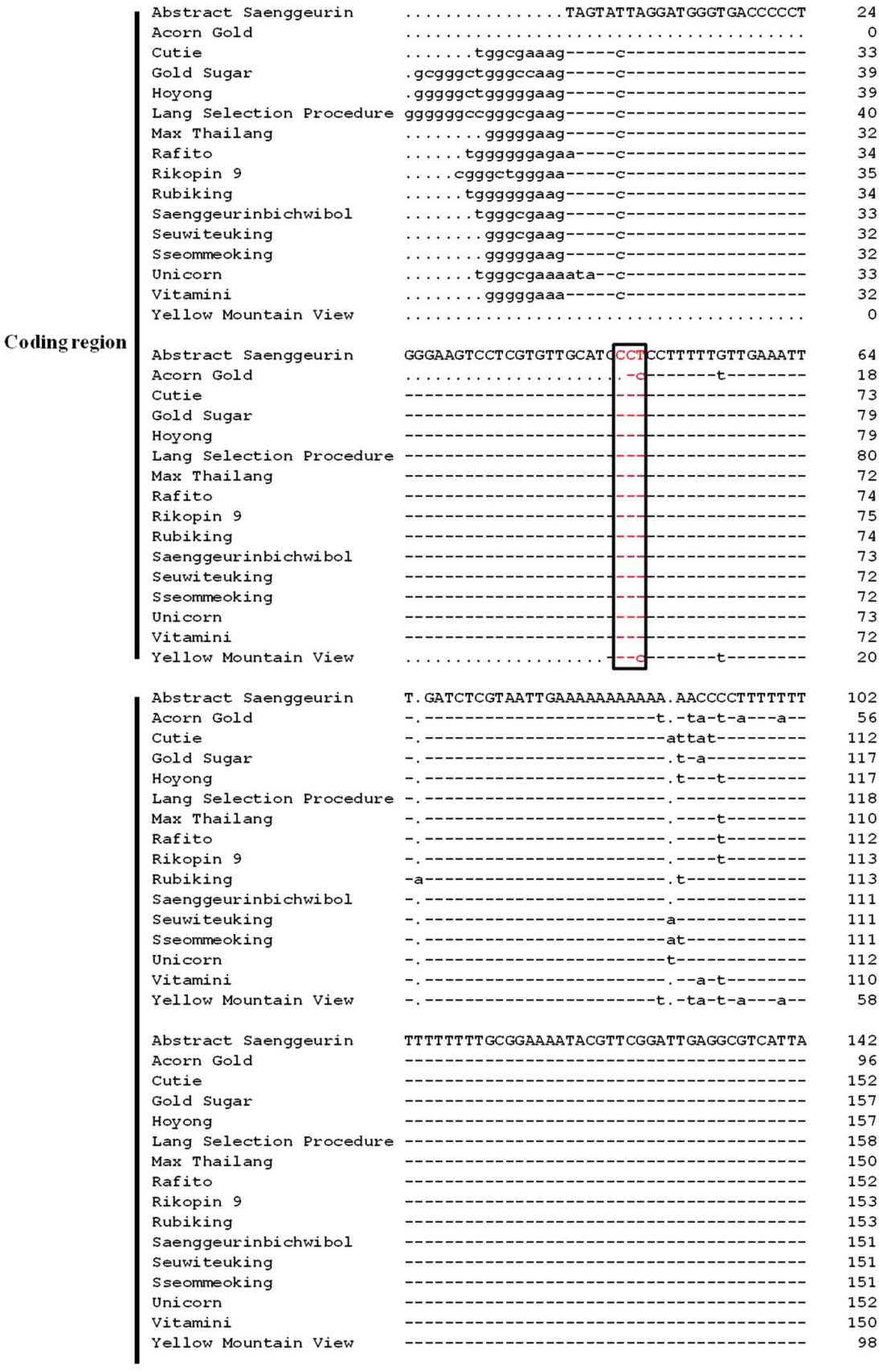

Figure 1. DNA alignment of the 5S rRNA gene sequences of 16 tomato varieties including the partial coding region and the non-transcribed spacer (NTS) region. The nucleotide arrays marked with red frames are the end sites of coding regions. The NTS region begins directly following the end site. Dots mark the deletion of a nucleotide, and short lines indicate that a nucleotide is identical to that of the reference sequence.

Non-Transcribed Spacer (NTS) region

```
Abstract Saenggeurin     GGATATGGGATGGTGGCGTCGGGGATGGGCGTGACGGGCG   182
Acorn Gold               ----------------------------------------   136
Cutie                    ----------------------------------------   192
Gold Sugar               ----------------------------------------   197
Hoyong                   ----------------------------------------   197
Lang Selection Procedure ----------------------------------------   198
Max Thailang             ----------------------------------------   190
Rafito                   ----------------------------------------   192
Rikopin 9                ----------------------------------------   193
Rubiking                 ----------------------------------------   193
Saenggeurinbichwibol     ----------------------------------------   191
Seuwiteuking             ----------------------------------------   191
Sseommeoking             ----------------------------------------   191
Unicorn                  ----------------------------------------   192
Vitamini                 ----------------------------------------   190
Yellow Mountain View     ----------------------------------------   138

Abstract Saenggeurin     TCGTCGTGCGTCGGTGCGTGGAGGGTTTTAAAGCGGGGGG   222
Acorn Gold               ----------------------------------------   176
Cutie                    ----------------------------------------   232
Gold Sugar               ----------------------------------------   237
Hoyong                   ----------------------------------------   237
Lang Selection Procedure ----------------------------------------   238
Max Thailang             ----------------------------------------   230
Rafito                   ----------------------------------------   232
Rikopin 9                ----------------------------------------   233
Rubiking                 ----------------------------------------   233
Saenggeurinbichwibol     ----------------------------------------   231
Seuwiteuking             ----------------------------------------   231
Sseommeoking             ----------------------------------------   231
Unicorn                  ----------------------------------------   232
Vitamini                 ----------------------------------------   230
Yellow Mountain View     ----------------------------------------   178

Abstract Saenggeurin     CGGGCTAGGGCGTTGGGAGGAAGGTTGTGTT.AATAGATT   261
Acorn Gold               ------------------------------t--------   216
Cutie                    ------------------------------.--------   271
Gold Sugar               ------------------------g-tgtg-t--------   277
Hoyong                   ------------------------------.--------   276
Lang Selection Procedure ------------------------------t--------   278
Max Thailang             ------------------------------.--------   269
Rafito                   ------------------------------.--------   271
Rikopin 9                ------------------------------.t-a-a---   272
Rubiking                 ------------------------------t--------   273
Saenggeurinbichwibol     ------------------------------t--------   271
Seuwiteuking             ------------------------------t--------   271
Sseommeoking             ------------------------------t--------   271
Unicorn                  ------------------------------t--------   272
Vitamini                 ------------------------------t--------   270
Yellow Mountain View     ------------------------------t--------   218

Abstract Saenggeurin     TA.GAGTGCAATAAAT                          276
Acorn Gold               -ta---g----g....                          228
Cutie                    --.--------.....                          281
Gold Sugar               -..agag-----g---                          291
Hoyong                   --.---------t---                          291
Lang Selection Procedure -..agag-----....                          288
Max Thailang             --.------------.                          283
Rafito                   -taag-g-------t.                          286
Rikopin 9                --.-----------..                          285
Rubiking                 --.-------------                          288
Saenggeurinbichwibol     a..-----------t.                          284
Seuwiteuking             -..agag-----....                          281
Sseommeoking             --.--------.....                          281
Unicorn                  -t.agag-----g--.                          286
Vitamini                 --.---g-------..                          283
Yellow Mountain View     -ta---gt---g-t-.                          233
```

Figure 1. Continued

before DNA sequence analysis, using a QIAquick PCR Purification Kit (QIAGEN, Korea) or Gel Purification Kit (QIAGEN, Korea) according to the manufacturer's instructions. Purified PCR products were then sequenced at MACROGENE Advancing through Genomics (Korea, http://dna.macrogen.com/kor/).

Sequence editing and alignment

Sequencing results were edited and assembled using the software DNAMAN version 6.0 (Lynnon Biosoft Corporation, USA, www.lynnon.com). Analogues of the identified sequences and nucleotide sequence comparisons were detected using Basic Local Alignment Search Tool (BLAST) network services against several databases (http://www.ncbi.nlm.nih.gov/). The phylogenetic relationships were analysed based on the multiple sequence alignment of the ITS1-5.8S-ITS2 region using the software DNAMAN version 6.0. Assembled sequences were deposited in the National Center for Biotechnology Information (NCBI, http://www.ncbi.nlm.nih.gov/). The NCBI GenBank accession numbers of the tomato varieties investigated in this study are shown in Table 1.

Results and discussion

PCR amplification of the 5S rRNA gene

The 5SF and 5SR primers used for PCR amplification of the 5S rRNA gene are located in the middle of the highly conserved coding region and downstream of the NTS region, respectively. The sequences of the PCR products were aligned using DNAMAN version 6.0 (Figure 1). Among the samples of the 26 investigated tomato varieties, only 16 DNA sequences were amplified successfully. The remaining sequences were often contaminated with fungal and bacterial DNA, strongly affecting the PCR amplification of common genes, i.e. the 5S rRNA gene. Based on our sequencing results, the 5S rRNA genes of each tomato variety contained a partial coding region (approximately 38−54 bp) and the NTS region (approximately 229−238 bp, Figure 1).

Sequence analysis of the 5S rRNA gene

As reported previously, the 5S rRNA gene encodes a highly conserved region approximately 120 bp in length, [15] which is commonly used for more accurate sequence alignment. Non-incident shifting of nucleotides in the highly conserved coding region can be used as a reference in the sequence analysis programme. Within the 5S rRNA gene, the coding region proved to be a valuable target for the study of phylogenetic relationships due to its highly conserved sequence. Moreover, the NTS region was informative for the study of phylogenetic relationships at the interspecific and intergeneric levels due to the faster rate of divergence in comparison to the highly conserved coding region. Differences in spacer regions are generally considered to result from duplication or deletion events and are mostly accumulated in the middle spacer region. [2,7]

In Figure 1, the red frame denotes the end site of the coding region. Excluding Abstract Saenggeurin and Yellow Mountain View, all other tomato varieties ended with CCT in the coding region of the 5S rRNA gene. All NTS regions of the tomato varieties investigated in this study started with CCT, and this nucleotide site was aligned with the start codon of the NTS of the 5S rRNA gene. A phylogenetic tree was constructed based on the sequence variation of the NTS region (Figure 2). Some tomato varieties exhibited very high identity, such as Rubiking and Sseommeoking, Lang Selection Procedure, and Seuwiteuking, Acorn Gold and Yellow Mountain View (Figure 2).

Hybridization analysis based on the 5S rRNA gene

Multiple copies of 5S rRNA gene families, as with other multigene families, can undergo concerted evolution due to homogenizing forces that lead to a high level of identity of all gene copies within the species and intraspecies.

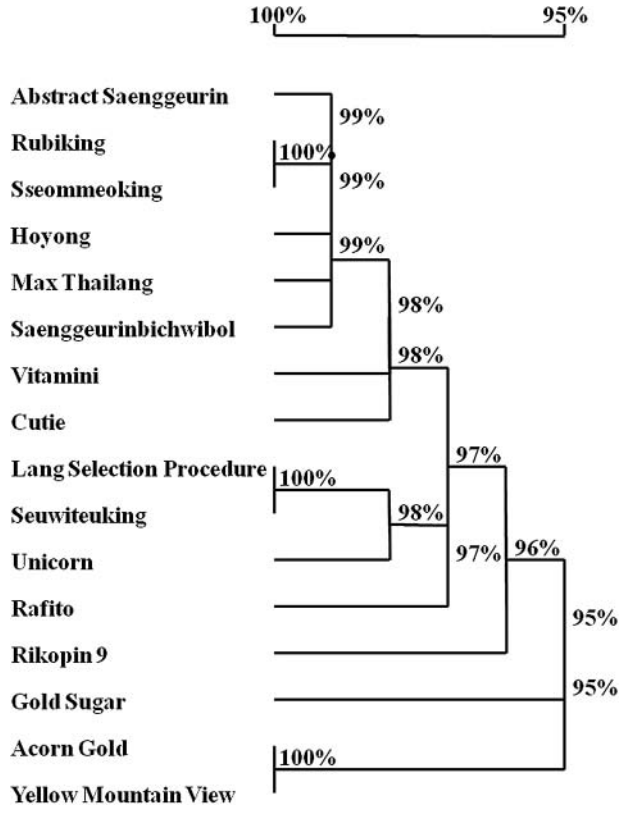

Figure 2. Phylogenetic tree constructed based on the 5S rRNA gene sequences of 16 tomato varieties.

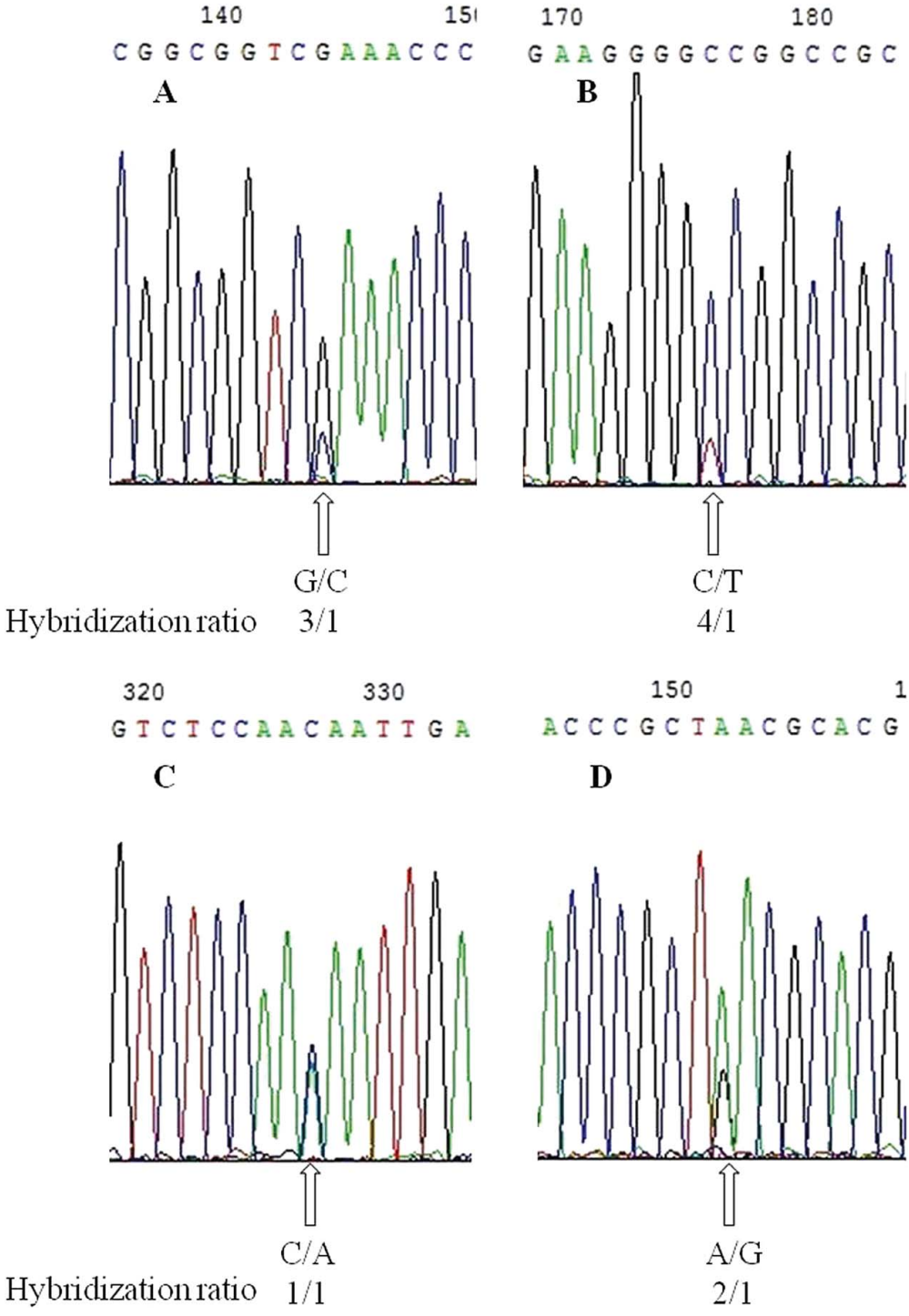

Figure 3. Hybridization analysis of tomato varieties based on the Chromas sequencing result. Yellow Mountain View (No. 8) and Seuwiteuking (No. 13) were used as examples. Hybridized nucleotide sites are marked with arrows under the Chromas sequencing result, and the nucleotides present and the hybridization ratio are shown below. Hybridized nucleotide sites in Yellow Mountain View sequences (A, B and C); a hybridized nucleotide site in a Seuwiteuking sequence (D).

Sequence differences result from normal levels of divergence between orthologous genes in different species. These differences are associated with the number of repeats in an array, the identity of natural selection, and the effective population size,[16−18] and therefore represent hybridization information among varieties.[19,20] Our results also strongly support this model. For example, when we sampled Yellow Mountain View and

Seuwiteuking tomato varieties, we found three hybridized variable nucleotide sites in the 5S rRNA gene in Yellow Mountain View but only one site in Seuwiteuking (Figure 3). The nucleotide variations were all caused by the use of different father and mother plants for variety hybridization. The occurrence of appearance of variable nucleotide sites depended on the degree of the hybridization present and the varieties used for hybridization. The degree of hybridization was apparent from the number of variable nucleotide sites and the hybridization ratio (Figure 3). Sequence variation of the NTS region of the 5S rRNA gene and the phylogenetic relationships among the tomato varieties clearly showed their homogenization, which provides a basis for tomato hybridization and breeding.

Conclusions

In this study, we attempted to analyse the phylogenetic relationships among 26 tomato varieties based on their 5S rRNA gene sequences, although 5S rRNA gene sequences of only 16 tomato varieties were amplified and used to analyse the phylogenetic relationships. The observed sequence variation of the 5S rRNA genes also represents information regarding the hybridization of the various tomato varieties studied. This work helps to further our understanding of the phylogenetic relationships among tomato varieties and provides a basis for tomato variety breeding.

Funding

This work was supported by Export Promotion Technology Development Program, Ministry of Agriculture, Food and Rural Affairs, Republic of Korea, New Talent Introduction Project, Ludong University [grant number LY2012008] and Outstanding Young Scientist in Shandong Province Award Fund Project [grant number BS2013SW021].

References

[1] Campell BR, Song Y, Posch TE, Cullis CA, Town CD. Sequence and organization of 5S ribosomal RNA-encoding genes of Arabidopsis thaliana. Gene. 1992;112:225–228.
[2] Sastri DC, Hilu K, Appels R, Lagudah ES, Playford J, Baum BR. An overview of evolution in plant 5S DNA. Plant Syst Evol. 1992;183:169–181.
[3] Cox AV, Bennet MC, Dyer TA. Use of the polymerase chain reaction to detect spacer size heterogeneity wheat (*Triticum aestivum* L.). Theor Appl Genet. 1992;83: 684–690.
[4] Dvořák J, Zhang HB, Kota RS, Lassner M. Organization and evolution of the 5S ribosomal RNA gene family in wheat and related species. Genome. 1989;32:1003–1009.
[5] Baum BR, Johnson DA. 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. 1996;39:140–149.
[6] Gottlob McHugh SG, Levesque M, MacKenzie K, Olson M, Yarosh O, Johnson DA. 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. 1990;33:486–494.
[7] Scoles GJ, Gill BS, Xin ZY, Clarke BC, McIntyre CL, Chapman C, Appels R. Frequent duplication and deletion events in the 5S RNA genes and the associated spacer regions of the Triticeae. Plant Syst Evol. 1988;160:105–122.
[8] Robertson LD, Labate JA. Genetic resources of tomato (*Lycopersicon esculentum* var. *esculentum*) and wild relatives. In: Razdan MK, Mattoo AK, editors. Genetic improvement of *Solanaceous* crops. vol. 2: Tomato. (New Hampshire): Science Publish Enfield; 2007. p. 25–75.
[9] Schauer N, Zamir D, Fernie AR. Metabolic profiling of leaves and fruit of wild species tomato: a survey of the *Solanum lycopersicoides* complex. J Exp Bot. 2005; 56:297–307.
[10] Bredemeijer GMM, Cooke RJ, Ganal MW, Peeters R, Isaac P, Noordijk Y, Rendell S, Jackson J, Röder MS, Wendehake K, Dijcks M, Amelaine M, Wickaert V, Bertrand L, Vosman B. Construction and testing of a microsatellite database containing more than 500 tomato varieties. Theor Appl Genet. 2002;105:1019–1026.
[11] Heubl G. New aspects of DNA-based authentication of Chinese medicinal plants by molecular biological techniques. Planta Medica. 2010;76:1963–1974.
[12] Li DZ, Liu JQ, Chen ZD, Wang H, Ge XJ, Zhou SL, Gao LM, Fu CX, Chen SL. Plant DNA barcoding in China. J Syst Evol. 2011;49:165–168.
[13] Doyle JJ, Doyle JL. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull. 1987;19:11–15.
[14] Hizume M. Chromosomal localization of 5S rRNA genes in *Vicia faba* and *Crepis capillaries*. Cytologia. 1993;58:417–421.
[15] Seo JH, Lee SY, Seo BB. Genome analysis using sequence variation and localization of tandem repeats of 5S rRNA gene in *Allium wakegi*. Korean J Breed Sci. 2007; 39:70–76.
[16] Basten CJ, Ohta T. Simulation study of a multigene family, with special reference to the evolution of compensatory advantageous mutations. Genetics. 1992;132:247–252.
[17] Schlötterer C, Tautz D. Chromosomal homogeneity of Drosophila ribosomal DNA arrays suggests intrachromosomal exchanges drive concerted evolution. Curr Biol. 1994;4:777–783.
[18] Smith GP. Evolution of repeated DNA sequences by unequal crossover. Science. 1976;191:528–535.
[19] Rauscher JT, Doyle JJ, Brown AH. Internal transcribed spacer repeat-specific primers and the analysis of hybridization in the *Glycine tomentella* (Leguminosae) polyploidy complex. Mol Ecol. 2002;11:2691–2702.
[20] Sarge KD, Maxwell ES. Intermolecular hybridization of 5S rRNA with 18S rRNA: identification of a 5′-terminally-located nucleotide sequence in mouse 5S rRNA which base-pairs with two specific complementary sequences in 18S rRNA. Biochim Biophys Acta. 1991;1088:57–70.