University of Nebraska - Lincoln
DigitalCommons@University of Nebraska - Lincoln

West Central Research and Extension Center, North Platte

Agricultural Research Division of IANR

1-1-2013

Multiple Divergent ITS1 Copies Were Identified in Single Tomato Genome Using DGGE Analysis

Bo Liu
West Central Research and Extension Center, University of Nebraska–Lincoln, North Platte, bliu5@unl.edu

Frank Louws
North Carolina State University, frank_louws@ncsu.edu

Follow this and additional works at: http://digitalcommons.unl.edu/westcentresext

Liu, Bo and Louws, Frank, "Multiple Divergent ITS1 Copies Were Identified in Single Tomato Genome Using DGGE Analysis" (2013). *West Central Research and Extension Center, North Platte.* Paper 76.
http://digitalcommons.unl.edu/westcentresext/76

This Article is brought to you for free and open access by the Agricultural Research Division of IANR at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in West Central Research and Extension Center, North Platte by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

Published in *Plant Molecular Biology Reporter* 31 (2013), pp. 272–279; doi: 10.1007/s11105-012-0500-0
Copyright © 2012 Springer-Verlag. Used by permission.
Published online August 4, 2012.

Multiple Divergent ITS1 Copies Were Identified in Single Tomato Genome Using DGGE Analysis

Bo Liu[1] and Frank Louws[2]

1. West Central Research and Extension Center, University of Nebraska–Lincoln, North Platte, NE 69101-7751, USA
2. Department of Plant Pathology, North Carolina State University, Raleigh, NC 27695-7616, USA

Corresponding author — Bo Liu, email bliu5@unl.edu

Abstract
The intra-genomic variation in the internal transcribed spacer (ITS) region has led to misleading conclusions in the evolutionary analysis of plants; understanding this variation is critical for correct evolutionary analysis based on ITS sequences. To reveal the ITS variation in tomato, entire copies of ITS1 sequences within tomato species were separated using denaturing gradient gel electrophoresis (DGGE) and DNA sequence analysis. ITS1 copies varied significantly in sequence composition, but not in sequence length within the same tomato cultivar. DNA sequence similarity of the ITS1 copies was 77–100%. Moreover, AT and GC contents in ITS1 copies from each tomato cultivar were significantly different, ranging from 50.4 to 64.3% for GC and from 35.7 to 49.6% for AT. However, the length variation of ITS1 was insignificant, ranging from 279 to 282 bp. Multiple copies of divergent ITS1 present in the tomato genome indicate that some copies may be paralogues. In conclusion, DGGE technique is a reliable and novel approach to reveal the entire ITS copy variation and the possible evolutionary relationship of tomato.

Keywords: divergent ITS copy, DGGE, GC content, *Solanum*, tomato

Introduction

The most widely used DNA marker in plant systematics is the ITS1-5.8S-ITS2 region of the nuclear ribosomal RNA (rRNA) transcription unit (Baldwin et al. 1995; Feliner and Rosselló 2007; Marshall et al. 2001). By 1998, early surveys of the internal transcribed spacer (ITS) region in angiosperm species have shown that 2,900 sequences were deposited in GenBank (Hershkovitz et al. 1999). One third of the published reports for phylogenetic analysis in the last 5 years were based exclusively on ITS regions (Álvarez and Wendel 2003). A recent study showed that publicly available ITS sequences have tripled since 2003; the number of Embryophyta sequences has gone from 23,937 in 2003 to 74,866 in 2007 (Calonje et al. 2009).

Using the ITS region to infer phylogenetic relationship at the species or the generic level has several advantages: biparental inheritance, easy polymerase chain reaction (PCR) amplification with several universal primers, moderate size for easy sequencing, multiple sets with thousands of copies or paralogues, and sufficient variation for phylogenetic comparison (Álvarez and Wendel 2003; Mayol and Rosselló 2001, Poczai and Hyvönen 2010). However, the existence of highly divergent rRNA types within a single genome can cause homogenization to proceed extremely slowly, particularly when the rRNA types are located on different chromosomes. Under such situations, silenced rRNA loci can evolve independently as pseudogenes (Baldwin et al. 1995). These pseudogenes and recombination between paralogues will result in great variation in the ITS region; therefore, using the ITS region to refer evolutionary analysis will lead to erroneous conclusions (Buckler et al. 1997; Sanderson and Doyle 1992).

Intraspecies ITS variation has been discovered in several plant species. This intraspecies variation has caused homoplasy and low bootstrap value support (Alvarez and Wendel 2003; Bohs 2007; Feliner and Rosselló 2007). Razafimandimbison et al. (2004) have revealed high levels of ITS polymorphism in three tropical tree species [*Adinauclea fagifolia* (Teijsm & Binn. Ex havil. Ridsdale), *Haldina cordifolia* (Roxb.) Ridsdale, and *Mitragyna rubrostipulata* (K. Schum.) Havil.] of the coffee (Rubiaceae) family (30, 40, and 14%, re-

spectively) and attributed the polymorphism to the existence of highly diverse putative pseudogenes. Mayol and Rosselló (2001) compared ITS sequences of the genus *Quercus* from two different laboratories, who reported conflicting hypothesis about phylogeny. One laboratory showed that high levels of ITS variation was caused by paralogues with higher rates of substitution and lowered secondary structure stability. Intraspecific variation levels were over 42% in ITS2 for *Quercus suber* and up to 27% in ITS1 for *Quercus rubra.* The contrasting results were due to analyzing differing subsets of the ITS1 region. Hughes et al. (2002) studied a total of 87 sequences from 65 accessions of Fabaceae and have identified 26 of them as pseudogenes. Buckler and Holtsford (1996) characterized 78 ITS cloned sequences in maize (*Zea mays*, including 13 clones that were amplified without DMSO. They found four sequences to be pseudogenes with a low GC content (62–65%) compared with the normal gene GC content of 70–73%. Moreover, Rausche et al. (2002) investigated intra-genomic ITS diversity in the allopolyploid complex of *Glycine tomentella* using repeat-specific PCR primers in a mismatch amplification of low-copy ITS repeats. They also found that significantly variable ITS copies existed in *G. tomentella.*

The abundance of ITS sequence variability within plant species suggests that some plant classes have a long history of duplication, incomplete homogenization, and pseudogenization in their genomes (Bortiri et al. 2001; Cronn et al. 2002). ITS sequence variations have been also found in other eukaryotes including beetles (Vogler and DeSalle 1994), yellow monkey flowers *Mimulus* (Ritland and Straus 1993), coral *Acropora* (Marquez et al. 2003; Odorico and Miller 1997), the fungus *Fusarium* (O'Donnell and Cigelnik 1997), and *Trichaptum abietihum* (Ko and Jung 2002), *Scutellospora* (Hijri et al. 1999), *Gigaspora* (Zeze et al. 1997), sponges (Wörheide et al. 2004), and Ascochyta (Fatehi and Bridge 1998).

The existence of diverse ITS copies suggests that multiple clones from a single individual need to be sequenced (Buckler and Holtsford 1996) in order to reveal this variation. PCR conditions must be varied or specifically designed primers must be utilized to obtain the divergent copies (Jason et al. 2002; Rausche et al. 2002). However, it is unknown how many clones have to be cloned to detect the ITS variation. Furthermore, lack of sequence availability in GenBank and other public databases would preclude primer design to target different ITS copies. Likewise, fluorescent in situ hybridization or genomic in situ hybridization data may not exist to reveal array number and chromosomal distribution for rRNA. Therefore, exploring new approaches to quickly locate and analyze divergent ITS copies is crucial for both accurate plant evolutionary analysis and plant progenitor or lineage assignment.

Denaturing gradient gel electrophoresis (DGGE) might be a potentially sensitive technique for separating variable ITS copies in the plant genome. DGGE has been successfully used to study point mutations directly from human cell (Michikawa et al. 1999); detect single-base substitutions, small deletions and insertions from other genes (Tuddenham et al. 1994; Sheffield et al. 1989); and characterize microbial communities in normal and extreme environments (Liu et al. 2008; Muyzer et al. 1993). Denatured DNA that differs in nucleotide sequence is separated by acrylamide gel electrophoresis in the presence of a linear gradient of denaturants (Muyzer et al. 1993). Purified DNA fragments amplified from individual DGGE bands can be sequenced and identified by Blast search (Altschul et al. 1997).

The objective of this study was to explore the multiple divergent ITS1 regions present in tomato (*Solannum lycopersicum* L.) using DGGE analysis to determine how many different ITS1 copies exist in each tomato line and hybrid. Since tomato has long been a classic model species for plant genetics and evolutionary genomics (Bohs 2007; Rivard and Louws 2008), we decided to choose one commercial cultivar and three root stocks for this study. We believe that the information obtained will be highly valuable for tomato breeding and cultivar comparison.

Materials and Methods

Plant Material

One cultivar and three root stocks of tomato (*S. lycopersicum*) were chosen for the characterization of ITS1 copies and phylogenetic analysis (Table 1).

DNA Extraction and Primer Design

Total genomic DNA was extracted from dried leaves of 2-week-old seedlings using DNeasy Plant Mini Kit following the manufacturer's instructions (Qiagen, Valencia, CA, USA). PCR primers were designed based on representative sequences of ITS deposited in GenBank. The following ITS1 sequences representing various *S. lycopersicum* (AB373816, FJ998172, FJ998171, AB373815, AB373814, EU760390, DQ314157, AY552528, AF244747, AJ300200, FJ998171, AB373815, AB373814, DQ001746, AJ300201, and EU760392) were retrieved from GenBank (National Center for Biotechnology information; http://www.ncbi.nlm.nik.gov), aligned using ClustalX (default setting; Thompson et al. 1997), and *S. lycopersicum* genus-specific primers [forward primer ToF: GGAAGGAGAAGTCGTAACAAGG (22 bp, which corresponded with GenBank sequence of DQ314157 from bases 14 to 35); reverse primer ToR: GTTCTTCATCGATGCGAGAG (20 bp, which corresponded with GenBank sequence of

Table 1. Information on the tomato cultivar and root stocks used in this study

Plant material	Parental species	Availability	Seed company
German Johnson	*S. lycopersicum* × *Solanum sp.*	Commercial	Reimer Seeds
Big Power	*S. lycopersicum* × *Solanum sp.*	Commercial	Rij Zwaan
Beaufort	*S. lycopersicum* × *S. habrochaites*	Commercial	De Ruiter Seeds
Maxifort	*S. lycopersicum* × *S. habrochaites*	Commercial	De Ruiter Seeds

DQ314157 from bases 319 to 338)] were designed based on the homologous regions specific to ITS1 sequences in *S. lycopersicum*. A fragment of 280 bp was amplified.

PCR Amplification

Two microliters of plant DNA (20 ng/µl) was used for PCR amplification. Each 50-µl reaction mixture contained 5 µl of 10× PCR buffer (Invitrogen, Carlsbad, CA, USA), 2.5 µl of deoxynucleoside–triphosphate mix (2.5 mM each), 2 µl bovine serum albumin (in milligrams per milliliter), 2 µl of both forward and reverse primers (2.5 uM), and 0.2 µl Taq polymerase (5 U/ml, Invitrogen). A 40-base GC clamp was attached to reverse primer ToR for DGGE analysis (Muyzer et al. 1993). PCR parameters were 94 °C for 2 min, 94 °C for 1 min, 55 °C for 1 min, and 72 °C for 2 min, for a total of 30 cycles, with a final extension at 72 °C for 10 min.

DGGE Analysis

DGGE was performed with a Decode universal mutation detection system (Bio-Rad Laboratories, Hercules, CA, USA). Seven microliters of the PCR product was loaded onto an 8% acrylamide gel (acrylamide/bis solution, 37.5:1; Bio-Rad) with a linear chemical gradient ranging from 20% to 70% denaturant [7 M urea and 40% (v/v) formamide]. Gels were run for 7 h at 104 V. The acrylamide gels were kept at 60 °C in 1× TAE buffer. The gel was stained with SYBR green I nucleic acid gel stain (1:10,000 dilution; Molecular Probes, Eugene, OR, USA) and photographed on a UV transilluminator (Liu et al. 2008).

Excision and Re-amplification of the DGGE Bands and DNA Sequence Analysis

The middle portions of the major DGGE bands were carefully excised on a UV transilluminator table and transferred to 50 µl distilled water. The DNA was allowed to diffuse for 48 h at 4 °C, and 0.5 µl was used as the template for PCR amplification with GC-clamped primers under the conditions noted previously. The re-amplified PCR products were rerun on DGGE gel to affirm the band position. The confirmed bands were additionally amplified using non-GC-clamped primers following the same PCR conditions. The final PCR product was purified using a Quick PCR purification kit (Qiagen) and directly sequenced using the ToF and ToR primers at the Genome Research Laboratory at the North Carolina State University, Raleigh, NC, USA. Sequencing reactions were carried out with ABI PRISM Dye Terminator Cycle Sequence Ready Reaction Kit (Applied Biosystems Inc., Foster City, CA, USA).

Table 2. Similarities of tomato ITS1 sequences match to closest relatives in GenBank

Bands	Length (bp)	Accession no.[a]	Similarity (%)	Accession no.[b]	Species
Band1	279	GU815245	86	AB373816	*S. lycopersicum* clone ITS1
Band2	278	GU815246	87	EU760390	*S. lycopersicum* clone ITS1
Band3	279	GU815247	89	EU760390	*S. lycopersicum* clone ITS1
Band4	281	GU815248	87	AB373816	*S. lycopersicum* clone ITS1
Band5	278	GU815249	86	EU760390	*S. lycopersicum* clone ITS1
Band6	270	GU815250	90	AC215351	*S. lycopersicum* clone ITS1
Band7	279	GU815251	92	EU760390	*S. lycopersicum* clone ITS1
Band8	279	GU815252	77	EU760390	*S. lycopersicum* clone ITS1
Band9	279	GU815253	84	AB373816	*S. lycopersicum* clone ITS1
Band10	279	GU815254	91	AB373816	*S. lycopersicum* clone ITS1
Band11	281	GU815255	91	AB373816	*S. lycopersicum* clone ITS1
Band12	282	GU815256	91	AB373816	*S. lycopersicum* clone ITS1
Band13	280	GU815257	100	AB373816	*S. lycopersicum* clone ITS1

a. GenBank accession numbers of the ITS1 sequences of tomato cultivars deposited from this research
b. GenBank accession numbers of the ITS1 sequences of tomato cultivars in GenBank matched to the sequences from this research

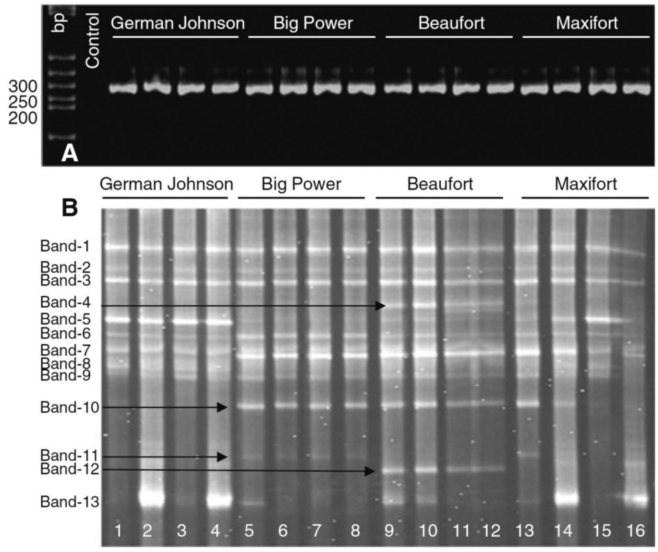

Figure 1. a) An approximately 280-bp fragment was amplified using the primers ToF and ToR designed based on the ITS1 region of the representative sequences of tomato cultivars (*S. lycopersicum*) from GenBank. b) Amplified DNA shown in (a) were run on DGGE gel and different ITS1 copies (DGGE bands) of the tomato cultivars were separated, which showed a significant variation of the ITS1 copies in the same species. Different cultivars not only share the identical sequence of the ITS1 copies but also possess different sequences of the ITS1 copies.

BLAST Search for Species Identification

Representative sequences derived from the excised bands of DGGE gels were compared with sequences deposited in GenBank using BLASTn (default setting) algorithm (Altschul et al. 1997) to search for close evolutionary relatives. The sequence datasets were aligned using the program ClustalX (Thompson et al. 1997). All representative sequences have been deposited to GenBank with accession numbers from GU815245 to GU815257 (Table 2 and Figure 2).

Results

DGGE Analysis

Approximately 280-bp fragments were amplified using tomato ITS1 specific primers (ToF and ToR) from one commercial cultivar and three root stocks of tomato (Figure 1a). PCR products from each cultivar were further separated into multiple bands using DGGE. The different bands on the DGGE gel represent different ITS1 copies in the tomato genome. Commercial cultivar German Johnson possesses eight bands, indicating that this cultivar has eight divergent ITS1 copies; likewise, Big Power root stock has ten bands, indicating ten divergent ITS1 copies; Beaufort root stock has 12 bands, indicating 12 divergent ITS1 copies; and Maxifort root stock has ten, bands indicating ten divergent ITS1 copies. Moreover, different cultivars not only share the same band patterns but also possess their own unique bands (Table 1). For example, German Johnson and the three root stocks share bands 1, 2, 3, 6, 7, and 8. Beaufort possesses the unique bands 4 and 12. Big Power and Maxifort possess band 11. Big Power, Beaufort, and Maxifort share band 10 (Figure 1b and Table 3).

Sequence Identification Based on Blast Search

Blast search showed that all band sequences (ITS1 copies) from the DGGE gel were from tomato. Only the sequence of band 13 revealed 100% identity with sequence

Table 3. Distribution of ITS1 copies and similarity of the ITS1 copies in each cultivar

Bands	Length (bp)	Similarity (%)[a]	Difference (%)[b]	German Johnson	Big Power	Beaufort	Maxifort
Band 1	279	86	14	+	+	+	+
Band 2	278	87	13	+	+	+	+
Band 3	279	89	11	+	+	+	+
Band 4	281	87	13	−	−	+	−
Band 5	278	89	11	−	−	+	−
Band 6	270	89	11	+	+	+	+
Band 7	279	92	8	+	+	+	+
Band 8	279	77	23	+	+	+	+
Band 9	279	84	16	+	+	−	+
Band 10	279	91	9	−	+	+	+
Band 11	281	84	16	−	+	+	+
Band 12	282	91	9	−	−	+	−
Band 13	280	100	0	+	+	+	+
Similarity (%)[c]				77–100	77–100	77–100	77–100

a. Similarity of sequences of the ITS1 region (represented by each DGGE band) compared with the sequences of band 13
b. Difference of sequences of the ITS1 region (represented by each DGGE band) compared with the sequences of band 13
c. Similarity range among the sequences of the ITS1 copies in each cultivar

Figure 2. Sequence alignment of the different copies of ITS1 sequences of tomato cultivars (*S. lycopersicum*) from the excised DGGE bands, which showed significant variations in the sequence of ITS1 copies within the same species. The GenBank accession number of each band on the DGGE gel was listed on the left. **a:** Accession numbers of the DGGE bands deposited in GenBank. The underlined sequence is the sequence of band 13, which is 100% similar to AB373816 from GenBank

```
TTTCCATAGGTTAACCTGCTGAAGGATCATTGTCAAAACTTGTACAGCAGAATGACCTGC Band7  GU815251[a]
TTTCCATAGGTTAACCTGCTGAAGGATCATTGTCAAAACTTGTACAGCAGAATGACCTGC Band3  GU815247
TTTCCATAGGTGATCCTGGGGAAGGATCATTGTCAAAACTTGTACAGCAGAATGACCTGC Band2  GU815246
TTTCCATAGGTGAACCTGGGGAAGGATCATTGTCGAAACCTGTACAGTAGAATGACCTAC Band6  GU815250
TTTCCATAGGTCATCCTACGGAAGGATCATTGTCAAAACCTGTACTGTAGAATGACCTAC Band1  GU815245
TTTCCGTAGGTGAACCTGCGGAAGGATCATTGTTAAAACCTGTTCAGCAGAACGACCTGC Band8  GU815252
TTTCCATAGGTTAACC-GGGGAAGGATCATTCTAGAAACCTGTACAGCAGAACGCCCTGC Band9  GU815253
TTTCCGTAGGTGAACCTGCGGAAGGATCATTGTCGAAACCTGCACAGCAGAACGACCCGC Band13 GU815257
TTTCCGTAGGTGAACATGCGGAACGATCATTATCGAAACCTGCAAAGAAGAACGACCTAT Band5  GU815249
TTTCCGTAGGTGAACCTGCGGAAGGATCATTGTCGAAACCAGCACTGCATAACGACCCGC Band12 GU815256
TTTCCGTAGGTGAACCTGCGGAAGGATCATTGTCGAAACCAGCACTGCACAATGACCTGC Band11 GU815255
TTTCCGTAGGTGAACCTGCGAAAGGATTATTCTCGAAACCTGCACAGCAGAACGACCCGC Band10 GU815254
TTTCCGTAGGTGAACCTACGGAAGGATCATTATCGAAACCAGCAAAGCAGAACAACCTGC Band4  GU815248
***** ***** * *      ** *** *** *  ****  *    * * **   **
GAACTCGTTTTTAACACCTGGGGGCGGTGCTCGCTCGTTGTGCGCCTCCCTCCTGTCGCC Band7  GU815251
GAACTCGTTTTTAACACATGGGGGCGGTGCTCGCTCGTTGTGCGCCTCCCTCTTGTCGCC Band3  GU815247
GAGCTCATTTTTA-CACCTGGGGGCGGTGCTTGCTCGTTGTGCGCCTCCCTCCTGTCGCC Band2  GU815246
GAGCTCATTTTAAACACCTGGGGGCGGTGCTTGCTAGTTATGCGCCTCCCTCCTGTCACC Band6  GU815250
GAGCTCATTTTAAACACCTGGGGGCGGCGCTTGCTAGTTTTGCGCCTCCCTCCTGTAGCC Band1  GU815245
GAACTCGTTTTAAACACCTGGGGGCGGCGCTCGCTCGTTGTGCGCCTCCCTCCTGTCGCC Band8  GU815252
AAACTCGTTTTAAACACCTGGGG-TGGGGGTCGCTCGTTGTGCGCCTCCCTCCTGTCGCC Band9  GU815253
GAACTCGTTTTAAACACC-GGGGGCGGCGCTCGCTCGTCGCGCGCCTCCC-CCCGTCGCC Band13 GU815257
GAACACGTTTTAAACACTTGGGGGAGGTGC-----------GCGCCTCCCTCCTGTCGCC Band5  GU815249
GAACTTGTTTTAAACAC-TTGGAGTGGCGCGTGCTCGTCGTGCGCCTCCCTCCCGTCTCC Band12 GU815256
GAACTTGTTTTAAACAC-TTGGGGTGGCGCGTGCTCGTCGTGCGCCTCCCTCCCGTCTCC Band11 GU815255
GAACTCATTTTATACAC-TGGGGGCGGGGCTTGCTCATCGCGCTCCTCCCTCCCGTCGCC Band10 GU815254
TAACTCATTTTAAACAC-TGGGGGGAGCGCTTGCTCGTCGCACGCCTCCCTCCCAATGCC Band4  GU815248
* *   ****   ***   **    * *              * ****** *      **
CGACGCGCGCAAGGTCTTCGGGCGACCACCGAACCCCGGCGCGGAAAGC-CCCAAGGAAT Band7  GU815251
CGACACACGCAATGTCTTCGGGAGACCACCGAACATCGGCGCGGAAAGC-CCCAAGGAAT Band3  GU815247
CGACGCGCCCAATGTCTTCGGGCGACCACCGAACCCCGGCGCGGAAAGC-CCCAAGGAAT Band2  GU815246
CGACGCGCGCAAGCTCTTCGGGCGACCACCGAACCCCGGCGCGGAAAGC-GCCAAGGAAT Band6  GU815250
CGACGCGCCCAAGCTCTTCGGGCGACAACCGAACCCCGGCACGGAAAGC-GCCAAGGAAT Band1  GU815245
CGACGCGCGCAAGGTCTTCGGGCGACCACCGAACCCCGGCGCGGAAAGC-GCCAAGGAAT Band8  GU815252
CGACGCGCGCAAGCTCTTCGGGGGACCACCGAACCCCTGCGCGGAAAGC-GCCAAGGAAT Band9  GU815253
CGAGGCGCGCAAGCTCTTCGGGCGACCAACGAACCCCGGCGCGGAAAGC-GCCAAGGAAT Band13 GU815257
CAACACGCGCAAAATCTTCGTGCGACCAACAAACCCGGGCGCGGAAAAC-ACCAAGGAAT Band5  GU815249
CGAGGCGTGAAAGCTCTTCGATCGACCAACGAATCCAAGCGCAGAAAGCCGCCAAGGAAT Band12 GU815256
CGAGGGGTGAAAGCTCTTCGATCGACCAACGAATCCAAGCGCAGAAAGCCGCCAAGGAAT Band11 GU815255
CGACGCGCGCAAGCTCTTCGGTCGACCAACGAACCCTGATGCGGAAAGC-GCCAAGGAAT Band10 GU815254
CGAGGAGAGCAAGCTCTTCGGTCTACCAACGAACCCCACCTCGGAAAAC-GCCAAAAAAT Band4  GU815248
* *       **  ******    ** * * **        * **** *  ****  ***
ACTACAATCGACAGCCCTCCCC-TCGCTCCCTGTTCACGGATCATG--GGGGGATGCGCG Band7  GU815251
ACTACAATCGATAGCCCTATCC-TCGTGCCTTGTTCACGGATCATG--GGGGGATGCGCG Band3  GU815247
ACTACAATCGACAGCCCTCCCC-TCGCGCCCTGTTCACGGATCATG--GGGGGATGCGCG Band2  GU815246
ACTACAATCGACAGCCCTCCCT-TCCCTCCCTGTTCACGGATCATG--GGGGGATGCGCG Band6  GU815250
ACTACAATAGACAGCCCTCCCC-TCGCTCCCTGTTCACGGATCATG--GGGGGATGTGAG Band1  GU815245
ACTACAATCGACAGCCCTCCCC-TCGCGCCCTGTTCACGGATCATG--GGGGGATGCCCG Band8  GU815255
ACTACAATCCACAGCCCTCCCC-TTCCGCCCTGTTCATGGATCATG--GGGGGATGCGCG Band9  GU815253
ACTACAATCGACAGCCCTCCCCCTCGCGCCCCGTTCGCGGATCGTGCGGGGGGAAGCGCG Band13 GU815257
ACTAAAATTGATAGCCCTACCC-TCGCGCCCCATTTGCGGATCATGCAGAGGGACACACG Band5  GU815249
ACTACAATCGACAGCCCTCCATTTCGCTCCCCATTCGCGGATTGGGTGGGGAGAAGCGTG Band12 GU815256
ACTACAATCGACAGCCCTCCCTTTCGCTCCCCATTCGGGGATTGGGGGGGGAGAAGCGTG Band11 GU815255
ACTACGATCGATAGCCCTCCCCCTCGCGCCACATTCGCGGATTGGGCGGGGGGAAGCACA Band10 GU815254
ACTACAATCAACATCCCTCCCCCTCGCGGCCTATTCCCGGATCCTGCGGGTGGAAGCGCA Band4  GU815248
****  **  * * ****     *     *   **   ****   *  *   **
CTGCTCTGATAACACAAACAACTCTCGACAACGAATATCTCGG Band7  GU815251
TTGCTCTGATAACACAAACAACTCTCGACAACGAATATCTTAG Band3  GU815247
CTGCTCTGATAACACAAACAACTCTCGACAACGAATATCTCGG Band2  GU815246
CTGCTCTGATAACACAAACAACTCTTGACAAGGAATATCTCGG Band6  GU815250
CTGCTCTGATAACACAAATGACTCTCGACAACGAATATCTCGG Band1  GU815245
TTGCTCTGATAACCCAAACAACTCTCGACAACGAATATCTCGG Band8  GU815255
CTGCTCTGATAACACAAACGACTCTCTACAACAAATATTTCCT Band9  GU815253
CTGCTCTGTTAACACAAACGACTCTCGGCAACGGATATCTCGG Band13 GU815257
TTGCTCTCTAAACACAAATGACTCTTGGCAACAGATATTTTGG Band5  GU815249
TTGCTCTATAAACAGAAATGATTCTCGGCAATGGATATCTCGT Band12 GU815256
TTGCTCTCTAAACAGAAATGATTCTCGGCAATGGATATCTCGT Band11 GU815255
CTACTCTGTAAACACAAATGACTCTGGGAAACAGATATCTCAG Band10 GU815254
CTGCTCTGTAAACACAAATGACTCTCGGCAACGGATATCCTGG Band4  GU815248
 * ****   ***  ***  * ***    **   ****
```

AB373816 in GenBank. Other DGGE band sequences (ITS1 copies) in each cultivar had only partial similarity to sequences in GenBank (91 to 77%; Table 2).

Similarity of the ITS1 Copies in Each Cultivar

A total of 13 major bands (ITS copies) were separated using DGGE for all tested tomato plants (Figure 1b). The bands were excised and re-amplified using primers without a GC clamp, followed by DNA sequence analysis. The results showed no significant variation in the length of ITS1 copies among each cultivar and root stock, ranging from 278 to 282 bp. The exception is band 6 with only 270 bp due to an 11-bp deletion in the middle portion of the ITS1 copy (Table 2 and Figure 2).

There was a significant sequence difference among ITS1 copies in each cultivar and root stock. The sequence of band 13, which has the highest GC content (Table 4), was dissimilar to other ITS1 band sequences by 8–23%. Band 7 was most similar to band 13, with a

variation of only 8%, whereas band 8 was the most different with dissimilarity at 23% (Table 2). Furthermore, the variability of ITS1 sequence among cultivar and root stocks of tomato was between 8 and 23%.

Comparison of AT and GC Contents Among ITS1 Copies

The GC content of band 13 was the highest (64.3%) and the AT content was the lowest (35.7%. On the other hand, the lowest GC content (50.4%) and concomitant highest rate of AT (49.6% was in band 6. The rest of the bands' GC composition (51.6–59.1%) was lower than band 13 and the AT composition (42.4–48.4%) higher than band 13 (Table 4).

Discussion

An ITS1 region with 9–13 divergent ITS1 subtypes were found in tomato cultivar and root stocks. The same ITS1 copy sequence is often shared by different cultivars, but each cultivar may also possess unique ITS1 subtypes. Some ITS1 sequences have significantly higher AT contents. These may be nonfunctional pseudogenes which have been reported to be AT-rich relative to the GC (Li and Graur 1991). The ITS region has high rates of substitutions and deletions, which reflect the reduced thermodynamic stability of the RNA structure (Sang et al. 1995) and may give rise to these pseudogenes. Pseudogenes discovered from *Q. rubra*, *Quercus acutissima*, and *Q. suber* have a remarkably lower GC content and an average ITS sequence dissimilarity of 17.29% (Manos et al. 1999). Moreover, most of the ITS sequences of *A. fagifolia*, *H. cordifolia*, and *M. rubrostipulata* were found to be putative pseudogenes. Five sequences of *A. fagifolia* are assumed functional alleles because they possess a higher GC content of 66.1–67.48% (Razafimandimbison et al. 2004).

The ITS loci may be widespread, but under-detected within angiosperms. Mayol and Rossello (2001) noticed that a phylogenetic analysis of *Quercus* based on ITS sequences resulted in conflicting hypothesis due to the presence of two highly divergent sequences within several *Quercus*. The sequence divergence of the ITS region in *Quercus* was far greater than the usual <5% reported in the literature. Moreover, pseudogenes were found in two oak species, *Quercus petrea* and *Quercus robur*, which hybridize commonly and share three divergent rRNA types, two of which were demonstrated to be pseudogenes (Mui et al. 2001).

Significantly divergent ITS copies occur in plant genomes, perhaps arising from incomplete concerted evolution. Under incomplete concerted evolution, some copies of the tandem arrays become non-functional pseudogenes and further intensify rRNA pattern diversity

Table 4. Statistics of G, C, A, and T in each ITS1 copy in tomato cultivars

Bands	Length (bp)	A (bp)	Percentage	T (bp)	Percentage	C (bp)	Percentage	G (bp)	Percentage	A and T (bp)	Percentage	C and G (bp)	Percentage
Band 1	279	68	0.244	60	0.215	85	0.305	66	0.237	128	0.459	151	0.541
Band 2	278	60	0.216	58	0.209	89	0.320	71	0.255	118	0.424	160	0.576
Band 3	279	69	0.247	66	0.237	78	0.280	66	0.237	135	0.484	144	0.516
Band 4	281	77	0.274	53	0.189	91	0.324	60	0.214	130	0.463	151	0.537
Band 5	278	63	0.227	59	0.212	91	0.327	65	0.234	122	0.439	156	0.561
Band 6	270	80	0.296	54	0.200	76	0.281	60	0.222	134	0.496	136	0.504
Band 7	279	66	0.237	58	0.208	85	0.305	70	0.251	124	0.444	155	0.556
Band 8	279	62	0.222	58	0.208	90	0.323	69	0.247	120	0.430	159	0.570
Band 9	279	60	0.215	54	0.194	92	0.330	73	0.262	114	0.409	165	0.591
Band 10	279	60	0.215	54	0.194	94	0.337	71	0.254	114	0.409	165	0.591
Band 11	281	68	0.242	53	0.189	89	0.317	71	0.253	121	0.431	160	0.569
Band 12	282	68	0.241	63	0.223	80	0.284	71	0.252	131	0.465	151	0.535
Band 13	280	57	0.204	43	0.154	99	0.354	81	0.289	100	0.357	180	0.643

(Koch et al. 2003; Li and Graur 1991; Wendel et al. 1995). Some rRNA types brought together by hybridization may be silenced by nuclear dominance and subsequently evolve as pseudogenes. Alternatively, hybridization may cause chromosomal rearrangement and relocation of rRNA copies to different chromosomal positions and, thus, reduce the homogenizing effect of concerted evolution (Mui et al. 2001). Concerted evolution is probably responsible for the complex patterns following the merging of ITS repeats within a single genome (Álvarez and Wendel 2003).

Pseudogenes and paralogues can be used to track parental relationships and hybrid speciation (Sang et al. 1995). In this study, the ITS1 copies in root stocks Beaufort, Big Power, and Maxifort are more similar than ITS1 in the commercial variety German Johnson. These root stocks, therefore, might be more closely related and possibly originated from a common ancestor as compared to German Johnson. Early research showed that the ITS region was successfully used to evaluate parental relationships and hybrid speciation in the *G. tomentella* complex (Rausche et al. 2002). Similarly, the lineage relationship of natural populations of *Arabis divaricarpa* was elucidated by detecting multiple intra-individual ITS copies in several *A. divaricarpa* accessions in the parental species, *Arabis holboellii* and *Arabis drummondii* (Koch et al. 2003).

Our study is the first to characterize the entire ITS1 set between tomato cultivars using DGGE analysis. If heterogenic sequences exist in the ITS region of the plant genome, these variable copies might give misleading results when used in phylogenetic analysis. Therefore, the evolutionary relationship based on ITS region with multiple divergent copies in early work, especially in eukaryotes, should be reevaluated.

Public nucleotide databases contain thousands of plant ITS sequences (Gemeinholzer et al. 2006), providing an invaluable source of raw DNA sequence data for taxonomic identification. However, the percentage of ITS pseudogenes is unknown. Furthermore, it is difficult to differentiate sequences of functional vs. non-functional genes from ITS sequences submitted to Genbank. Nevertheless, we believe that this will not affect primer design for detecting multiple divergent ITS sequences. Phylogeneticists and taxonomists can simply download representative sequences of ITS regions from GenBank, align the sequences, and design universal primers for DGGE analysis.

This research demonstrates that divergent ITS1 copies exist in tomato and are not due to preferential amplification of the ITS1 region. The same DGGE band patterns were repeatedly amplified using DNA extracted from leaves, roots, and seeds of tomato, respectively (data not shown). The ITS2 region might also have multiple divergent copies in the tomato genome, and further DGGE analysis is needed to clarify this.

Acknowledgments — This research was supported in part by grants from the USDA Sustainable Agriculture Research and Education (SARE) Program and in part by the North Carolina State University Agricultural Research Service. We would like to express our appreciation to Jinping Sun and Cary Rivard for their technical assistance and Stella Chang for editing the early version of the manuscript. We also would like to thank Dr. Bruce Baldwin for his suggestion through e-mail communication.

References

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. Nucleic Acids Res 25:3389–3402

Álvarez I, Wendel JF (2003) Ribosomal ITS sequences and plant phylogenetic inference. Mol Phyl Evol 29:417–434

Baldwin BG, Sanderson MJ, Porter JM, Wojciechowski MF, Campbell CS, Donoghue MJ (1995) The ITS region of nuclear ribosomal DNA: A valuable source of evidence on angiosperm phylogeny. Annual Mo Bot Garden 82:247–277

Bohs L (2007) Phylogeny of the Cyphomandra clade of the genus *Solanum* (Solanaceae) based on ITS sequence data. Taxon 56:1012–1026

Bortiri ES, Oh H, Jiang J, Baggett S, Granger A, Weeks C, Buckingham M, Potter D, Parfitt DE (2001) Phylogeny and systematics of *Prunus* (Rosaceae) as determined by sequence analysis of ITS and the chloroplast trnL-trnF spacer DNA. Syst Bot 26:797–807

Buckler ES, Holtsford TP (1996) *Zea* ribosomal repeat evolution and substitution patterns. Mol Biol Evol 13:623–632

Buckler ES, Timothy AI, Holtsford P (1997) The evolution of ribosomal DNA: Divergent paralogues and phylogenetic implications. Genetics 145:821–832

Calonje M, Martın-Bravo S, Dobes C, Gong W, Jordon-Thaden I, Kiefer C, Kiefer M, Paule J, Schmickl R, Koch MA (2009) Non-coding nuclear DNA markers in phylogenetic reconstruction. Plant Systemat Evol 282:257–280

Cronn RC, Small RL, Haselkorn T, Wendel JF (2002) Rapid diversification of the cotton genus (*Gossypium*: Malvaceae) revealed by analysis of sixteen nuclear and chloroplast genes. Am J Bot 89:707–725

Fatehi J, Bridge P (1998) Detection of multiple rRNA-ITS regions in isolates of *Ascochyta*. Mycol Res 102:762–766

Feliner GN, Rosselló JA (2007) Better the devil you know? Guidelines for insightful utilization of nrDNA ITS in species-level evolutionary studies in plants. Mol Phyl Evol 44:911–919

Gemeinholzer B, Oberprieler C, Bachmann K (2006) Using GenBank data for plant identification: Possibilities and limitations using the ITS1 of Asteraceae species belonging to the tribes Lactuceae and Anthemideae. Taxon 55:173–187

Hershkovitz MA, Zimmer EA, Hahn WJ (1999) Ribosomal DNA sequences and angiosperm systematics. In: Hollingsworth PM, Bateman RW, Gornall RJ, eds., Molecular systematics and plant evolution. Taylor & Francis, London, pp 268–453

Hijri M, Hosny M, van Tuinen D, Dulieu H (1999) Intraspecific ITS polymorphism in *Scutellospora castanea* (Glomales, Zygomycota) is structured within multinucleate spores. Fung Gen Biol 26:141–151

Hughes CE, Bailey CD, Harris SA (2002) Divergent and reticulate species relationships in Leucaena (Fabaceae) inferred from multiple data sources: Insights into polyploid origins and nrDNA polymorphism. Am J Bot 89:1057–1073

Jason TR, Doyle JJ, Brown HD (2002) Internal transcribed spacer repeat-specific primers and the analysis of hybridization in the analysis of hybridization in the *Glycine tomentella* (Leguminosae) polyploidy complex. Mol Ecol 11:2691–2702

Ko KS, Jung HS (2002) Three nonorthologous ITS1 types are present in a polypore fungus *Trichaptum abietinum*. Mol Phyl Evol 23:112–122

Koch MA, Dobes C, Mitchell-Olds T (2003) Multiple hybrid formation in natural populations: Concerted evolution of the internal transcribed spacer of nuclear ribosomal DNA (ITS) in North American *Arabis divaricarpa* (Brassicaceae). Mol Biol Evol 20:338–350

Li WH, Graur D (1991) Fundamentals of molecular evolution. Sinauer Associates, Sunderland, pp 1–284

Liu B, Glenn D, Buckley K (2008) *Trichoderma* communities in soils from organic, sustainable and conventional farms, and their relation with southern blight of tomato. Soil Biol Biochem 40:1124–1136

Manos P, Doyle JJ, Nixon KC (1999) Phylogeny, biogeography, and processes of molecular differentiation in *Quercus* subgenus *Quercus* (Fagaceae). Mol Phyl Evol 12:333–349

Márquez LM, Miller DJ, MacKenzie JB, van Open MAH (2003) Pseudogenes contribute to the extreme diversity of nuclear ribosomal DNA in the hard coral *Acropora*. Mol Biol Evol 20:1077–1086

Marshall JA, Knapp S, Davey MR, Power JB, Cocking EC, Bennett MD, Cox AV (2001) Molecular systematics of *Solanum* section Lycopersicum (*Lycopersicon*) using the nuclear ITS rDNA region. Theor Appl Genet 103:1216–1222

Mayol M, Rosselló JA (2001) Why nuclear ribosomal DNA spacers (ITS) tell different stories in *Quercus*. Mol Phyl Evol 19:167–176

Michikawa Y, Mazzucchelli F, Bresolin N, Scarlato G, Attardi G (1999) Aging-dependent large accumulation of point mutations in the human mtDNA control region for replication. Science 286:774–779

Muyzer G, De Waal EC, Uitterlinden AG (1993) Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction amplified genes for 16S rRNA. Appl Environ Microb 59:695–700

Mui G, Fleming CC, Schlotterer C (2001) Three divergent rDNA clusters predate the species divergence in *Quercus petraea* (Matt.) Liebl. and *Quercus robur* L. Mol Biol Evol 18:112–119

O'Donnell K, Cigelnik E (1997) Two divergent intragenomic rDNA ITS2 types within a monophyletic lineage of the fungus *Fusarium* are nonorthologous. Mol Biol Evol 7:103–116

Odorico DM, Miller DJ (1997) Variation in the ribosomal internal transcribed spacers and 5.8S rDNA among five species *Acropora* (Cnidaria; Scleractinia): patterns of variation consistent within reticulate evolution. Mol Biol Evol 14:465–473

Poczai P, Hyvönen J (2010) Nuclear ribosomal spacer regions in plant phylogenetics: Problems and prospects. Mol Biol Rep 37:1897–1912

Razafimandimbison SG, Kellogg EA, Bremer B (2004) Recent origin and phylogenetic utility of divergent ITS putative pseudogenes: A case study from Naucleeae (Rubiaceae). Syst Biol 53:177–192

Rausche JT, Doyle JJ, Brown HD (2002) Internal transcribed spacer repeat-specific primers and the analysis of hybridization in the *Glycine tomentella* (Leguminosae) polyploid complex. Mol Ecol 11:2691–2702

Ritland CE, Straus NA (1993) High evolutionary divergence of the 5.8S ribosomal DNA in *Mimulus glaucescens* (Scrophulariaceae). Plant Mol Biol 22:691–696

Rivard C, Louws FJ (2008) Grafting to manage soilborne diseases in heirloom tomato production. Hortic Sci 43:2104–2111

Sanderson MJ, Doyle JJ (1992) Reconstruction of organismal and gene phylogenies from data on multigene families: Concerted evolution, homoplasy, and confidence. Syst Biol 41:4–17

Sang T, Crawford DJ, Stuess TF (1995) Documentation of reticulate evolution in peonies (*Paeonia*) using internal transcribed spacer sequences of nuclear ribosomal DNA: Implications for biogeography and concerted evolution. Proc Natl Acad Sci U S A 92:6813–6817

Sheffield VC, Cox DR, Lerman LS, Myers RM (1989) Attachment of a 40-base-pair G+C-rich sequence (GC-clamp) to genomic DNA fragments by the polymerase chain reaction results in improved detection of single-base changes. Proc Natl Acad Sci U S A 86:232–236

Thompson JD, Higgins DG, Gibson TJ (1997) ClustalX 1.5, software for Power PC. European Molecular Biology Organization, Heidelberg

Tuddenham EGD, Schwaab R, Seehafer J, Millar DS, Gitschier J, Higuchi M, Bidichandani S, Connor JM, Hoyer LW, Yoshioka A, Peake IR, Oiek K, Kazazian HH, Lavergne JM, Giannelli F, Antonarakis SE, Cooper DN (1994) Haemophilia A: Database of nucleotide substitutions, deletions, insertions and rearrangements of the factor VIII gene, second edition. Nucleic Acids Res 22:4851–4868

Vogler AP, DeSalle R (1994) Evolution and phylogenetic information content of the ITS-1 region in the tiger beetle *Cicindela dorsalis*. Mol Biol Evol 11:393–405

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

Wörheide G, Nichols SA, Goldberg J (2004) Intragenomic variation of the rDNA internal transcribed spacers in sponges (phylum Porifera): Implications for phylogenetic studies. Mol Phyl Evol 33:816–830

Zeze A, Sulistyowati E, Ophel-Keller K, Barker S, Smith S (1997) Intersporal genetic variation of *Gigaspora margarita,* a vesicular arbuscular mycorrhizal fungus, revealed by M13 minisatellite-primed PCR. Appl Environ Microb 63:676–678