Asia-Pacific Journal of Science and Technology: Volume: 28. Issue: 03. Article ID.: APST-28-03-16. Research Article

APST

Asia-Pacific Journal of Science and Technology

https://www.tci-thaijo.org/index.php/APST/index

Published by the Research and Graduate Studies,
Khon Kaen University, Thailand

A mutation in acceptor splice site of GA3ox homolog *Cla015407* gene confers a dwarf phenotype in watermelon (*Citrullus lanatus* L.)

Namfon Chomkaeo[1], Kanokwan Janphen[1], Marilyn Hinlo[2], Nate Petchara[3], Darush Struss[4], Chotipa Sakulsingharoj[1] and Saengtong Pongjaroenkit[1,*]

[1]Program in Genetics, Faculty of Science, Maejo University, Chiang Mai, Thailand
[2]Research and Development, East-West Seed Ltd, Bulacan, Philippines
[3]Research and Development, East-West Seed Ltd, Suphan Buri, Thailand
[4]Research and Development, East-West Seed Ltd, Chiang Mai, Thailand
[*]Corresponding author: saengton@gmaejo.mju.ac.th

Received 11 November 2021
Revised 19 May 2022
Accepted 1 June 2022

Abstract

Dwarfism is a desirable agronomic trait in watermelon. The dwarf growth habit in watermelon is controlled by a single recessive gene (*dwarf; dw*). Linkage analysis of 309 F_2 individuals derived from inbred watermelon lines KK-6939 (viny plant) and TH-15974 (dwarf plant), positioned the *dw* locus at the terminal region of watermelon chromosome 9. Using the gene annotation data derived from watermelon reference genome "97103", 25 genes were located between two flanking markers, WMSNP-0002750 and WMSNP-0002780. Only 2 out of 25 candidate genes, *Cla015407* and *Cla015408*, are related to growth hormone and encode a protein predicted to be a gibberellin 3-beta-hydroxylase (GA3ox). Only the single nucleotide polymorphism (SNP) (G/A) at the 626th nucleotide of *Cla015407* could distinguish dwarf plants from viny plants. The mutation is an acceptor splice site in intron 1 and leads to an altered splicing site at the 639th nucleotide and a 13 base pair (bp) deletion in complementary deoxyribonucleic acid (cDNA), resulting in a truncated protein in dwarf plants. Although *Cla015408* encoded the same protein as *Cla015407*, the expression level of *Cla015408* couldn't be detected from quantitative and semi-quantitative real time polymerase chain reaction (RT-PCR). Therefore, *Cla015408* couldn't function instead of *Cla015407* in dwarf plants. These results suggested that *Cla015407* is the *dw* gene. A high throughput functional marker, Cla015407-GA, was developed and validated in various inbred lines. The functional marker is segregated with the phenotype, demonstrating that the marker has high accuracy and high value to implement in a watermelon breeding program for marker-assisted selection (MAS).

Keywords: Watermelon, Dwarf locus, Single nucleotide polymorphism marker, Gibberellin 3-beta hydroxylase, *Cla015407*, Acceptor splice site, Functional marker

1. Introduction

Watermelon (*Citrullus lanatus* L.) is an economically important cucurbit crop and is produced worldwide [1]. Common watermelons have long vine and climbing habits, therefore the plant must be grown with wide spacing. The exception is dwarf cultivars, which are natural mutants, and can be grown at a tighter spacing. Dwarf watermelon has short vine plants with shorter internodes that are suitable for high-density planting and can save land resources. The crop may be established in the greenhouse by planting seeds or using containerized transplants and dwarf plants can also decrease plant disease during the production period due to beneficial characteristics such as lodging resistance which was associated with reduced plant height [2]. Therefore, dwarf plants are one of the most desirable and valuable traits in watermelon breeding programs.

Dwarfism inheritance in watermelon has been identified from segregating populations that derived from a cross between normal plants and dwarf lines, and several different locus names exist, including *dw-1* in Bush Desert King cultivars [3], *dw-1s* in Somali local [4], *dw-2* in WB-2 cultivar [5], *dw-3* in Dwarf Male-Sterile Watermelon (DMSW) [6], *dw-4* in d5-6y [7], *dsh* in dsh [8], *Cldw-1* in WM102 [9], *Cldf* in N21 [10], *Si* in

w106 [11] and *Cldw* in dwarf cultivar 812 [12]. Even though the dwarf mutants were derived from different sources, all studies reported the inheritance of the dwarf phenotype was controlled by a single recessive gene. Only five out of ten loci were identified as the position on the watermelon genome. *Cldf*, *Si* and *Cldw* were the same locus, which is different from the *dsh* and *Cldw-1* locus.

The main causes of dwarfism in plants are mutations in hormone biosynthesis pathway-related genes such as gibberellins (GAs) [13-15]. GA is a crucial hormone to promote plant growth and stem elongation [16]. There are seven key catalyzing enzymes in the GA biosynthesis pathway, including copalyl diphosphate synthase (CPS), kaurene synthase (KS), kaurene oxidase (KO), kaurenoic acid oxidase (KAO), GA 20-oxidase (GA20ox), GA 2-oxidase (GA2ox) and GA 3-oxidase (GA3ox) [17]. The GA20ox and GA3ox catalyze the final two steps of GA biosynthesis pathway [18,19]. The pumpkin (*Cucurbita moschata* Duchesne) bush mutant plant showed partial response to exogenous GAs hormones but not to indole acetic acid (IAA) and brassinosteroids (BR) [14]. A deletion in the promoter region of the *Cma_004516* gene, that encodes the GA 20-oxidase in the GA biosynthesis pathway causes dwarfness in pumpkin (*Cucurbita maxima* Duch.) [15]. Two single nucleotide polymorphisms (SNPs) in exons together with several SNPs and insertion-deletions (InDels) in the promoter of *Cp4.1LG10g05910.1* gene, which encodes GA2-oxidase, were identified in the dwarf *Cucurbita pepo* L. line X10 [13]. Hypocotyl and root elongation in a dwarf and a normal type of watermelon were investigated. Gibberellins were effective in stimulating hypocotyl elongation of dwarf type, but normal type responded only slightly to the gibberellin employed. To compare gibberellin specific, plants of the dwarf type are most sensitive to GA4 together with GA7 [20].

Dwarfness-related genes in watermelon have been identified and reported from various dwarf lines. The candidate gene, *Cla010726* on chromosome 7 was identified as a dwarfism gene in the *dsh* watermelon line by using Bulked Segregant Analysis (BSA)-seq and linkage analysis. The SNP was detected in the promoter region of *Cla010726* gene, which encoded GA20ox and led to the expression level being significantly lower in the dwarf plants than in viny plants [8]. The deletion of a single nucleotide in an ABC transporter B subfamily gene (*Cla010337*) on chromosome 9 was identified in dwarf line WM102 using BSA-Seq and linkage analysis. The regulation of plant height mechanism of adenosine triphosphate-binding cassette (ABC) transporter genes was unclear but likely due to crosstalk between auxins and other hormone pathways [9]. The 13 base pair (bp) deletion of *Cla015407* coding region (GA3ox) on chromosome 9 was detected in dwarf Duan 125 by molecular mapping [21]. A SNP at the 3′AG splice site of intron 1 and 13 bp deletion of exon 2 of *Cla015407* gibberellin 3-beta-hydroxylase (GA3ox) on chromosome 9 was detected in dwarf N21 line and led to truncated protein in dwarf watermelon [10]. The SNP at the 3′AG splice site of intron 1 of *Cla015407* (GA3ox) generated two splicing isoforms, one splicing isoform retained the intron sequence while the other had a 13-bp deletion in exon 2. Both isoforms lead to truncated protein in dwarf w106 line [11]. A splice site mutation at intron 1 was revealed in *Cla015407* (GA3ox) of dwarf cultivar 812 and resulted in the expression of *Cla015407* gene in dwarf cultivar 812 internode that was significantly lower than viny cultivar W1-1 internode at reproductive growth stage [12]. Therefore, dwarfism of watermelon is mostly mutations in *GA3ox* or *GA20ox* genes. However, many genes encoded *GA3ox* or *GA20ox* enzyme in watermelon genome.

Genetic mapping or linkage map analysis is one of the earliest methods used to develop molecular markers and map genes on chromosomes. Genetic mapping assigns a specific gene to a particular region of a chromosome and relative distances between genes on the chromosome. Genetic mapping methods utilize marker technology and statistical methodology to localize the gene on chromosomes or identify tightly linked markers [22]. The purposes of this study were to identify and characterize the gene that causes dwarf watermelon by a quantitative trait locus (QTL) mapping method and develop a high throughput genotyping marker for marker-assisted selection (MAS).

2. Materials and methods

2.1 Plant materials

Two watermelon inbred lines, KK-6939 (Viny parent; P_1) and TH-15974 (Dwarf parent; P_2) were used for the developing of three segregating populations, including F_2, BC_1P_1, BC_1P_2 for gene inheritance study and QTL mapping. Five inbred lines, including three viny lines (KK-6939, KK-28334, and KK-28469) and two dwarf lines (TH-15974 and KK-28413) were used for dwarf candidate genes sequencing and gene expression. Moreover, 31 various inbred lines that were derived from different backgrounds were used for functional marker validation.

2.2 Phenotypic evaluation and dwarf gene inheritance study

Three phenotypes were evaluated, including vine length, internode length, and seedling morphology. Vine length and internode length were measured and recorded as numeric data, while seedling morphology was

visually evaluated and then classified as dwarf or viny. All parents, F_1, F_2 individuals, BC_1P_1, BC_1P_2, and inbred lines were evaluated. Using the star software, Duncan's test was used to evaluate the significance of statistical data [23].

The mode of gene inheritance that controls the dwarf growth habit in watermelon was observed and investigated in the F_1 together with three segregating populations (F_2, BC_1P_1, and BC_1P_2). Chi-squared test was used to determine the suitability of observed data with expected segregation ratios.

2.3 Linkage analysis and QTL mapping to localize dw locus

Genomic deoxyribonucleic acid (DNA) was extracted from young leaves following the method described by Doyle and Doyle [24]. A total of 346 watermelon SNP markers which were distributed across the watermelon genome were selected and called from transcriptome re-sequencing of two inbred lines, PH-02714 and PH-02812, by East-West Seed company. Two parental lines were screened to obtain polymorphic markers using these watermelon SNP markers. Then, identified polymorphic markers were used to genotype individual plants of the F_2 population (n=309). Genotyping of each plant was detected by allele-specific polymerase chain reaction (PCR) using Kompetitive Allele Specific PCR (KASP). KASP assays were performed in 384 array tape following the protocol of the manufacturer (LGC Genomics, Berlin, Germany). KASP assays were performed in a thermocycler (Soellex® High Throughput PCR Thermal Cycler, Douglasscientific, USA) with cycling conditions: denaturation at 95°C for 15 min, ten cycles of 95°C for 20 sec, touchdown starting at 65°C for 60 sec (decreasing 0.8°C per cycle), followed by 30 cycles of amplification (95°C for 20 sec; 57°C for 60 sec). End-point fluorescence data were visualized with a light scanner (Araya machine, Douglasscientific, USA) and analyzed using Clustering Caller software (LGC, Middlesex, U.K.).

An initial linkage map was constructed by using genotypes of a segregating population (F_2 population) and seedling morphology, which represented *dw* locus using JoinMap Version 4.0 software [25]. Parameters for map construction were the exclusion of markers with obvious segregation distortion from the expected Mendelian segregation ratios of 1:2:1. Markers were grouped based on limit of detection (LOD) threshold values of 6.0-10.0. Markers order assignment was accomplished using the regression mapping algorithm. Recombination fractions between markers were calculated to map distances in cm using the Kosambi mapping function. Vine length and internode length traits were assumed as QTL site and then used QTL analysis to locate vine length and internode length QTL region. QTLs for two phenotypes were identified using the Interval Mapping program by MapQTL 6.0 [26].

2.4 Dwarf gene identification and sequence analysis

Fine mapping was conducted by calling more SNP markers surrounding significant markers from initial maps. New SNP markers were screened for polymorphism between two parental lines, and then those informative SNP markers were used for genotyping individual plants of the F_2 population and to re-construct the linkage map to generate high-resolution linkage and re-localized *dw* locus. Subsequently, to minimize the genetic interval of the *dw* locus region, candidate genes were annotated by using the 97103 watermelon reference genome [1,27].

The sequence of target genes was obtained using the specific primers, which were designed from the watermelon reference genome (Table 1, Figure 1). Full-length sequences of target genes were amplified by using genomic DNA while coding region (CDs) sequences were obtained from total ribonucleic acid (RNA). Genomic DNA and total RNA were extracted from the whole seedling plant using cetyltrimethylammonium bromide (CTAB) and TRIzol® Reagent (Ambion, Carlsbad, CA, USA), respectively. First-strand complementary deoxyribonucleic acid (cDNA) was synthesized using a RevertAid First Strand cDNA Synthesis Kit (Thermo Scientific, Lithuania). The PCR was carried out in a reaction containing 1 × GoTaq® Master Mixes (Promega, USA), 0.67 μM of each primer (07F with 07R1 and 08F1 with 08R1) and 50 ng genomic DNA or cDNA. Amplifications of candidate genes were performed under the following conditions: 95°C for 5 min; 35 cycles of 30 sec at 95°C, 45 sec at 55°C and 1.30 min at 72°C, followed by a final extension step at 72°C for 10 min. PCR products were analyzed on 1.5% agarose gel electrophoresis and sent to 1st base company for sequencing. The complete sequences *Cla015407* and *Cla015408* of five inbred lines were submitted to GenBank.

Table 1 Specific primer sequence of target genes and expected fragment sizes.

Genes	Primer sequence (5′ – 3′)		Expected size (bp)
	Forward	Reverse	
Cla015407	07F; ATGGGAAGCATCAAAATAACCGAA	07R1; TTAACCTACTTTAACCTGGCTGTG	1,257
		07R2; AAGGCATTGCCAATGAGATGA	235
Cla015408	08F1; ATGATTGGATTCATCCCTTCCTTCA	08R1; TTAGCCTACTTTCACCTGGCTA	1,623
	08F2; CCTCAAGACTACACCAAGCAC	08R2; ACACTGCCGAACTCTTTTCC	192
Cla007792 (actin)	F; CCATGTATGTTGCCATCCAG	R; GGATAGCATGGGGTAGAGCA	220

Figure 1 Illustration of *Cla015408* and *Cla015407* gene map.

2.5 Expression analysis of dwarf candidate genes by quantitative RT-PCR

The expression level of dwarf candidate genes was investigated by quantitative real time polymerase chain reaction (qRT-PCR) in five inbred lines using the dwarf candidate gene-specific primers which were 07F with 07R2 and 08F2 with 08R2 together with *Cla007792* actin gene primer (Table 1). Real time RT-PCR was used to analyze transcript levels with the SensiFAST™ SYBR® No-ROX Kit (Bioline GmbH, GERMANY) and the QIAquant® real-time cyclers (QIAGEN, USA). The reactions were performed for 40 cycles, with 5 sec at 95°C, 10 sec at 60°C and 20 sec at 72°C. The melting curve analysis and 1.5% gel electrophoresis were performed to verify the specific transcript amplification. All reactions were performed in triplicate and were normalized using the Ct values corresponding to the actin gene. A relative expression was calculated using the $2^{-\Delta\Delta CT}$ method [28] against viny parent (KK-6939).

2.6 High throughput functional marker conferring dwarf in watermelon

A high throughput functional marker for a dwarf in watermelon, utilizing the KASP™ marker genotyping method, was developed based on mutation point detected from sequence alignment of dwarf candidate gene between viny and dwarf plants. To validate the functional marker, 31 various inbred lines of viny plant that were derived from different backgrounds were genotyped to confirm the accuracy of the functional markers.

3. Results and discussion

3.1 Phenotypic evaluation and dwarf gene inheritance

The phenotypes of F_1 plants were completely viny (Table 2). The seedling morphology of individual plants in the F_2 population (n = 309) was classified into two groups, consisting of 239 viny and 70 dwarf plants. The average internode length and vine length in the viny group were 69.16 and 3,460.73 mm, respectively. While the internode length and vine length in the dwarf group were 35.19 and 1,295.78 mm, respectively (Table 2). Chi-square tests were performed in the F_2 population to determine the goodness of fit to the single gene model which is a 3:1 (viny:dwarf) ratio. The χ^2 value was 0.907, which was less than 3.84 at the corresponding confidence level of 95%. Thus, the null hypothesis was accepted, indicating no significant deviation between expected and observed values (Table 3). Results indicated that the ratio of viny plants to dwarf plants fit the 3:1 ratio. The single recessive inheritance results obtained from the F_2 population were confirmed by testing in a BC_1P_1 population, derived from a cross between F_1 and KK-6939, which showed all viny and a BC_1P_2 population, derived from F_1 and TH-15974. The phenotype in the BC_1P_2 was consistent with a 1:1 expected ratio (χ^2 = 0.214, p = 0.6439) as shown in Table 3. These results suggested that a single recessive gene controls the dwarf trait, which was confirmed and supported in previous studies [3-7,29,30]. The genes controlling dwarfism of watermelon were similar to other Cucurbitaceae crops such as cucumber and melon, which are controlled by a single recessive gene [31-34], but differ from the dwarf trait in pumpkins is controlled by a single dominant gene [13,35].

Table 2 The phenotypic value summarization and pairwise (mean±SD) comparison of two parental lines and their progenies.

Sample	Observed (plants)	Phenotype		
		Seedling morphology	Internode length (mm) (Mean*)	Vine length (mm) (Mean*)
Parent 1 (KK-6939)	10	Viny (V)	80.33 ± 8.34[a]	3,023.33 ± 1,048[a]
Parent 2 (TH-15974)	10	Dwarf (dw)	34.9 ± 4.23[b]	1,033.3 ± 260.73[b]
F_1	20	Viny (V)	75.00 ± 6.56[a]	3,100.52 ± 1,020[a]
F_2	239	Viny (V)	69.16 ± 11.01[a]	3,460.73 ± 934.46[a]
	70	Dwarf (dw)	35.19 ± 6.82[b]	1,295.78 ± 346.22[b]

*Means followed by the same letter in each column are not significantly different in Duncan's multiple range test (DMRT) at alpha 0.05.

Table 3 Genetic analysis of dwarf growth habit crosses between KK-6939 (viny growth habit) and TH-15974 (dwarf growth habit).

Generations	Number of plants	Observed		Expected ratio	Chi-square (χ^2)	*p*- value
		Viny	Dwarf			
Parent 1 (KK-6939)	10	10	-	-	-	-
Parent 2 (TH-15974)	10	-	10	-	-	-
F_1	20	20	-	-	-	-
F_2	309	239	70	3:1	0.907	0.3409
BC_1P_1 (F_1 x KK-6939)	114	114	-	1:0	-	-
BC_1P_2 (F_1 x TH-15974)	117	56	61	1:1	0.214	0.6439

Significance limit of Chi-square (χ^2) (p = 0.05, df = 1) = 3.84.
$\chi^2 \geq 3.84$ and p value ≤ 0.05 are considered significant.

3.2 QTL mapping and candidate dwarf gene identification

Of a total of 346 SNP markers screened, 65 SNP markers were polymorphic between two parental lines and were used for initial linkage analysis, of which nine failed to be assigned to any linkage group (LG). Thus, 56 SNP markers were mapped to eleven LGs and the *dw* locus was localized on the top of chromosome 9. For QTL analysis, only one major QTL on the top of chromosome 9 was detected for internode length and vine length. The SNP marker "WMSNP-9-88" showed high significant co-segregation with phenotypes and was located 24.4 cm away from the *dw* locus. Therefore, fine mapping was performed to minimize the *dw* locus region. A total of 104 new SNP markers surrounding the WMSNP-9-88 were designed, but only 20 SNP markers showed polymorphism between two parental lines. Those markers were genotyped in the F_2 population and used to reconstruct the genetic map. Finally, two flanking markers were detected; WMSNP-0002750 and WMSNP-0002780, with an interval map of approximately 3.3 cm or 250 kb. One major QTL, underlying the *dw* locus, was detected for both internode length and vine length with LOD score values of 62.36 (R^2 = 62%) and 51.02 (R^2 = 53.8%), respectively (Figure 2). This result confirmed the location of dwarf-related gene on the top of chromosome 9, which is similar to the result from other dwarf lines [10,11,21] but differs from the location of the *dw* locus reported by Zhu et al. [9] which was about 29 Mb downstream of our *dw* locus. The other chromosomes were not significant at all based on the linkage and QTL analysis, including chromosome 7, which was previously reported as containing a dwarf -related gene [8].

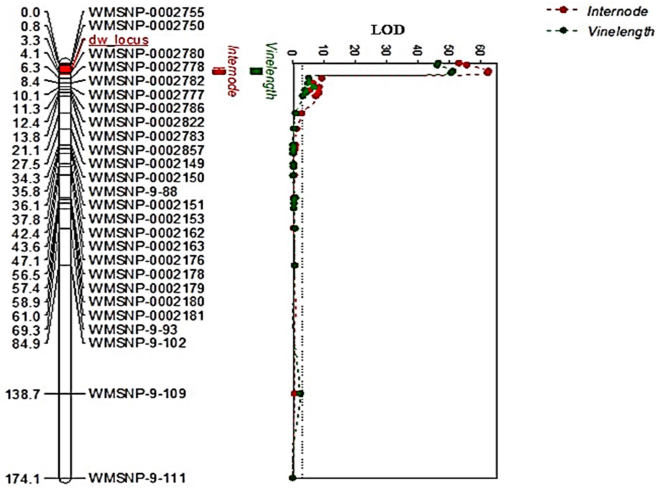


Figure 2 The genetic map of *dw* locus in watermelon chromosome 9, *dw* locus was mapped between WMSNP-0002750 and WMSNP-0002780 SNP markers. Composite interval mapping analysis of quantitative trait loci associated with internode length (red line) and vine length (green line) on chromosome 9.

Two flanking SNP markers were combined with the gene annotation within the major QTL regions to identify candidate genes associated with dwarf growth habits. Candidate genes were annotated by using the "97103" watermelon reference genome database [1,27]. According to the annotation of the watermelon reference genome, there were 25 predicted genes in the 250 kb major QTL region between WMSNP-0002750 and WMSNP-0002780 SNP markers (Figure 3). Of these candidate genes, 20 genes had annotation information (Appendix 1). As dwarf growth habit is mainly determined by stem growth, two genes, *Cla015407* and *Cla015408,* were selected that encoded or were related to growth hormone. Both genes encode a protein predicted to be a gibberellin 3-beta-hydroxylase enzyme (GA3ox), which is a catalyst in the gibberellin biosynthesis III pathway that affects stem elongation in plants. Moreover, many previous studies identified the mutation in GAs-related genes that affect dwarfness in watermelon such as *Cla010726* (GA20ox), *Cla015407* (GA3ox) [8,10,11,21]. *Cla015407* and *Cla015408* are paralogous gene, but *Cla015408* has not been well studied in dwarfism. Therefore, in this study *Cla015407* and *Cla015408* were potential candidate genes underlying the *dw* locus.

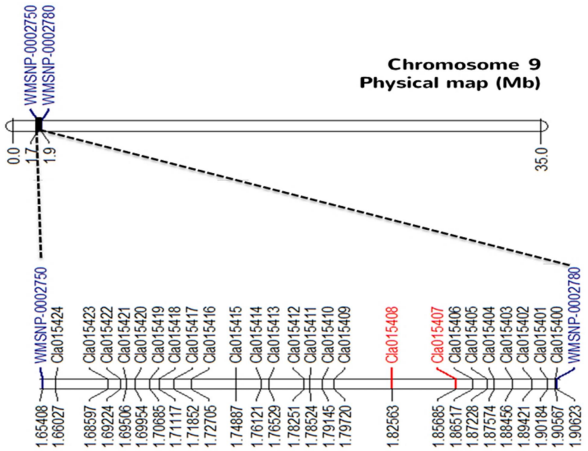


Figure 3 Physical map of watermelon chromosome 9. A total of 25 genes located between WMSNP-0002750 and WMSNP-0002780 SNP markers. Only two candidate genes, *Cla015407* and *Cla015408* (red labeled) are involved in the gibberellin biosynthesis pathway.

3.3 Gene isolation and sequence comparison

Cla015408 was located about 32.47 kb upstream of *Cla015407*. The gene sizes of *Cla015407* and *Cla015408* were 1,257 bp and 1,623 bp, respectively. *Cla015407* and *Cla015408* both contained two exons. However, exon 1 size of *Cla015408* (377 bp) was shorter than *Cla015407* (503 bp) while exon 2 of both genes were of equal size (631 bp) (Figure 1). Both DNA and mRNA were investigated to verify the sequence. A complete DNA sequence of the *Cla015407* gene from dwarf plants (GenBank accession number MZ568772 and MZ568773) were aligned with viny plants (GenBank accession number MZ568769, MZ568770, and MZ568771) is shown in Appendix 3. Whereas in the *Cla015408* gene only full-length DNA was obtained, then a complete DNA sequence of dwarf plants (GenBank accession number MZ568774 and MZ568775) were aligned with viny plants (GenBank accession number MZ568776, MZ568777, and MZ568778) as shown in Appendix 4. At the DNA sequence level of the *Cla015407* gene, two SNP variants were detected among five accessions but only one SNP at the 626th nucleotide could distinguish dwarf plants from viny plants completely perfectly. The viny plants contained the "G" allele while dwarf plants contained the "A" allele at the 626th nucleotide. This G allele is an acceptor splice site in intron 1 of viny plant or normal type, whereas the 3′AG splice site of dwarf plants moved to the 639th nucleotide because of the SNP mutation event in dwarf plants (Figure 4A). Additionally, a 13 bp deletion was detected in cDNA sequences of the *Cla015407* gene obtained from dwarf plants (Figure 4B). This one alternative splicing form was reported in dwarf N21, dwarf cultivar 812 and dwarf Duan 125 [10,21], whereas two splicing isoforms were found in dwarf w106 [11]. This 13 bp deletion was located at the beginning of the second exon of *Cla015407* gene and led to frameshift translation and the premature stop codon "TGA" at 656th-658th nucleotide resulting in a truncated protein in dwarf plants. A truncated protein of *Cla015407* in dwarf plants had a high impact on GA3ox function as shown in previous studies [10,11,21]. The DNA sequence of *Cla015408* obtained from five inbred lines showed one SNP at intron

1 that could not be classified dwarf plants from viny plants. Moreover, this detected SNP was not in the coding region and had no effect on the amino acid sequence.

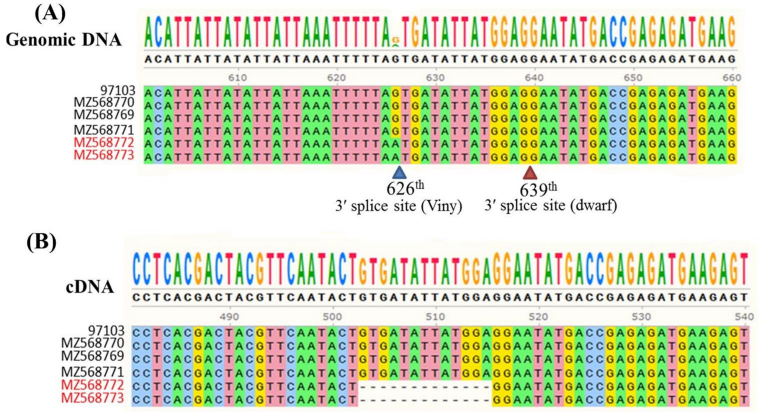


Figure 4 (A) Partial alignment of genomic DNA sequence, blue arrow indicated 3′AG splice site at 626th nucleotide of viny plants and red arrow indicated 3′AG splice site at 639th nucleotide of dwarf plants, and (B) Partial alignment of cDNA sequences that were obtained from dwarf and viny plants.

3.4 Expression analysis of dwarf candidate genes: Cla015407 and Cla015408 gene

There were two candidate genes, *Cla015407* and *Cla015408* that encoded the same protein GA3ox. Based on the sequence analysis, *Cla015407* in dwarf plants had the point mutation at an acceptor splice site in intron 1 and resulting in a nonfunctional protein. *Cla015408* only one SNP was detected at intron 1 region and could not affect the translation (Appendix 4). Therefore, *Cla015408* was supposed to translate as normal amino acid and produced functional protein. The expression analysis was investigated to prove the assumption why *Cla015408* could not function instead of *Cla015407* mutant. The qRT-PCR was performed in whole seedling plants at seedling stage. Consider the expression of *Cla015408* which could not be detected in all samples at seedling (Figure 5A). The relative of *Cla015407* expression was not significantly different among viny and dwarf plants (Figure 5B). Therefore the expression level of *Cla015407* in this study did not relate to plant types at the seedling stage while the expression of *Cla015407* was previously reported as significantly different between viny and dwarf plants in internodes at the reproductive growth stage [12]. The different result was because the plants were studied in different parts and stages. Although the expression of *Cla015407* did not affect the plant type in this study, the mutation of *Cla015407* led to a non-functional protein and resulted in the dwarf phenotype. This result strongly supported *Cla015407* gene played an important role to control watermelon growth habits and suggested that *Cla015408* could not replace *Cla015407* in the dwarf plant at seedling stage. Therefore, the *Cla015407* gene should be dwarf gene in the TH-15974 line which is similar to other watermelon dwarf lines [10,11,21].

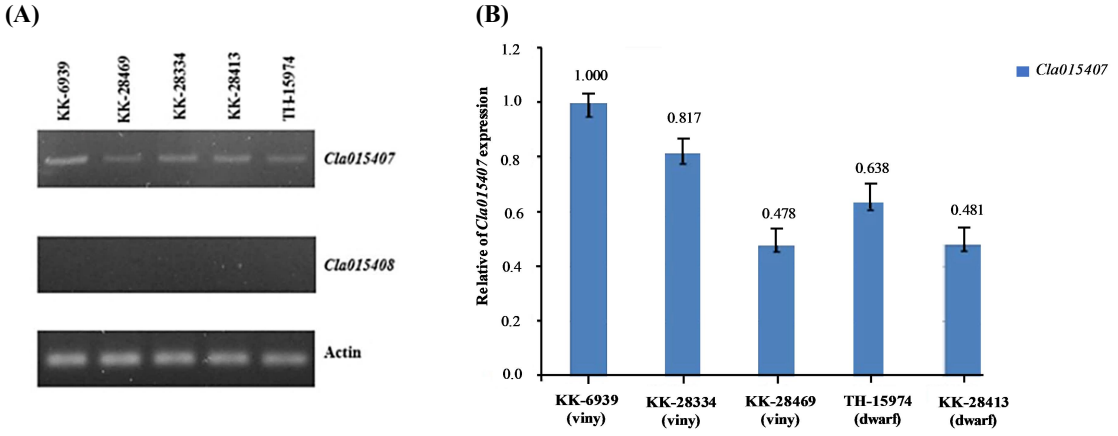


Figure 5 Gene expression analysis of GA3ox (*Cla015407* and *Cla015408*) transcripts by qRT-PCR from five independent watermelon inbred lines; (A) 1.5% gel electrophoresis of qRT-PCR products, and (B) relative of *Cla015407* expression. The error bar indicates the SD of the triplicate.

3.5 GA3ox amino acid sequence analysis

GA3ox protein had two crucial domains, which effect GA3ox function, including non-haem dioxygenase N-terminal domain (DIOX_N domain; IPR026992) and oxoglutarate/iron-dependent dioxygenase domain (FE2OG_OXY domain; IPR005123). Full-length amino acid of Cla015407_dwarf and Cla015407_viny were translated from cDNA sequences which were obtained from five inbred lines. *Cla015408* amino acid sequence was obtained from the watermelon genome database because mRNA fragments could not be obtained from five inbred lines. Full-length amino acids of Cla015407_dwarf, Cla015407_viny, and Cla015408 were conducted in multiple alignments (Figure 6). In the normal plant, Cla015407_viny translated 377 amino acids while in the mutant plant, Cla015407_dwarf had a premature stop codon, leading to the translation of only 173 amino acids and the loss of FE20G dioxygenase domain, which is a crucial domain of GA3ox, and finally leading to *Cla015407* gene producing a nonfunctional enzyme which is the same as previous studies [10,11,21]. Whereas the *Cla015408* amino acid was expected to produce the functional enzyme but the *Cla015408* gene could not recover the dwarf plant because the transcript could not be detected at the mRNA level.

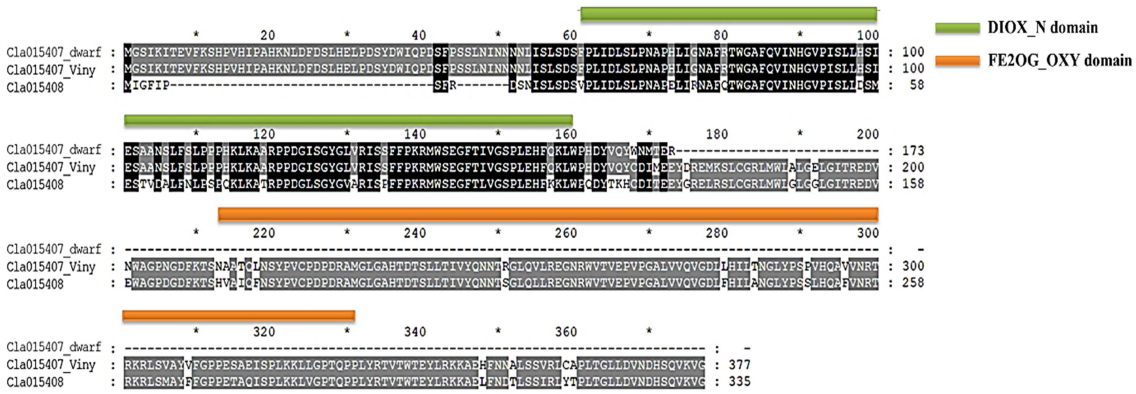


Figure 6 Illustration of two important domains of GA3ox protein and multiple alignment of GA3ox protein of Cla015407_dwarf, Cla015407_viny, and Cla015408 amino acid sequences.

3.6 High throughput functional marker conferring dwarf in watermelon

KASP^{TM} genotyping marker was designed based on SNP at an acceptor splice site on intron 1 of *Cla015407* and named as “Cla015407-GA " marker. This functional marker was validated in 31 inbred lines to confirm the accuracy of the marker. The “G/G " genotype was present in all 29 viny plants whereas “A/A” genotype was present in 2 dwarf plants. The genotyping result of this marker was perfectly consistent with the phenotype (Appendix 2). Therefore, the functional marker (Cla015407-GA) was the value marker for MAS in the watermelon dwarf breeding program. This finding leads to understanding the mechanism of the root cause of dwarfism in watermelon. In the case that prefers to recover plant height, we need to apply the gibberellin exogenously [20].

4. Conclusion

Two candidate genes of the *dw* locus, *Cla015407*, and *Cla015408*, encoded a GA3ox enzyme. Only a SNP at 626th nucleotide of intron 1 *Cla015407* distinguished dwarf plants and viny plants. This SNP was a 3′AG splice site and led to altered splicing which produced a truncated protein in dwarf plants. The relative expression of *Cla015407* was not significantly related to plant type while *Cla015408* could not be detected in all samples. Due to this, the *Cla015408* could not function instead of *Cla015407* in the dwarf plant. The mutation at the acceptor splice site was used to develop the “Cla015407-GA” functional marker.

5. Acknowledgements

We would like to thank East-West Seed company for providing plant materials and high throughput SNP genotyping. We thank Caleb Orchard for editing the English text of a draft of this manuscript.

6. References

[1] Guo S, Zhang J, Sun H, Salse J, Lucas WJ, Zhang H, et al. The draft genome of watermelon (*Citrullus lanatus*) and resequencing of 20 diverse accessions. Nat Genet. 2013;45(1):51-58.

[2] Smitchger JA. Quantitative trait loci associated with lodging, stem strength, yield, and other important agronomic traits in dry field peas [dissertation]. Bozeman, Montana: Montana State University; 2017.

[3] Mohr H, Sandhu M. Inheritance and morphological traits of a double recessive dwarf in watermelon, *Citrullus lanatus* (Thunb.) Mansf J Am Soc Hortic Sci . 1975.

[4] Dyutin K, Afanas'eva E. Inheritance of the short vine trait in watermelon. Cytol Genet. 1987;21(3):71-73.

[5] Liu P, Loy J. Inheritance and morphology of two dwarf mutants in watermelon. J Am Soc Hortic Sci. 1972:745-748.

[6] Hexun H, Xiaoqi Z, Zhencheng W, Qinghuai L, Xi L. Inheritance of male-sterility and dwarfism in watermelon [*Citrullus lanatus* (Thunb.) Matsum. and Nakai]. Sci Hortic. 1998;74(3):175-181.

[7] Yang H, LI YG, Yang DX, Yang J. A new dwarf mutant dw-4 in watermelon. CGC Report. 2009;31(32):11-14.

[8] Dong W, Wu D, Li G, Wu D, Wang Z. Next-generation sequencing from bulked segregant analysis identifies a dwarfism gene in watermelon. Sci Rep. 2018;8(1):1-7.

[9] Zhu H, Zhang M, Sun S, Yang S, Li J, Li H, et al. A single nucleotide deletion in an ABC transporter gene leads to a dwarf phenotype in watermelon. Front Plant Sci. 2019;10:1399.

[10] Wei C, Zhu C, Yang L, Zhao W, Ma R, Li H, et al. A point mutation resulting in a 13 bp deletion in the coding sequence of *Cldf* leads to a GA-deficient dwarf phenotype in watermelon. Hortic Res. 2019;6(1):1-12.

[11] Sun Y, Zhang H, Fan M, He Y, Guo P. A mutation in the intron splice acceptor site of a *GA3ox* gene confers dwarf architecture in watermelon (*Citrullus lanatus* L.). Sci Rep. 2020;10(1):1-15.

[12] Zhang T, Liu J, Amanullah S, Ding Z, Cui H, Luan F, et al. Fine mapping of *Cla015407* controlling plant height in watermelon. J Am Soc Hortic Sci . 2021;1(aop):1-26.

[13] Ding W, Wang Y, Qi C, Luo Y, Wang C, Xu W, et al. Fine mapping identified the gibberellin 2-oxidase gene *CpDw* leading to a dwarf phenotype in squash (*Cucurbita pepo* L.). Plant Sci. 2021;306:110857.

[14] Wu T, Cao J, Qin Z, Du Y. Identification of a novel ga-related bush mutant in pumpkin (*cucurbita moschata* duchesne). Pak J Bot. 2015;47(4):1359-1366.

[15] Zhang G, Ren Y, Sun H, Guo S, Zhang F, Zhang J, et al. A high-density genetic map for anchoring genome sequences and identifying QTLs associated with dwarf vine in pumpkin (*Cucurbita maxima* Duch.). BMC Genom. 2015;16(1):1-13.

[16] Sun TP. Gibberellin-GID1-DELLA: a pivotal regulatory module for plant growth and development. Plant Physiol. 2010;154(2):567-570.

[17] Sakamoto T, Miura K, Itoh H, Tatsumi T, Tanaka UM, Ishiyama K, et al. An overview of gibberellin metabolism enzyme genes and their related mutants in rice. Plant Physiol. 2004;134(4):1642-1653.

[18] Chen Y, Hou M, Liu L, Wu S, Shen Y, Ishiyama K, et al. The maize *DWARF1* encodes a gibberellin 3-oxidase and is dual localized to the nucleus and cytosol. Plant Physiol. 2014;166(4):2028-2039.

[19] Qin X, Liu JH, Zhao WS, Chen XJ, Guo ZJ, Peng YL. Gibberellin 20-oxidase gene *OsGA20ox3* regulates plant stature and disease development in rice. Mol Plant Microbe Interact. 2013;26(2):227-239.

[20] Loy JB, Liu PB. Response of seedlings of a dwarf and a normal strain of watermelon to gibberellins. Plant Physiol. 1974;53(3):325-330.

[21] Gebremeskel H, Dou J, Li B, Zhao S, Muhammad U, Lu X, et al. Molecular mapping and candidate gene analysis for GA3 responsive short internode in watermelon (*Citrullus lanatus*). Int J Mol Sci. 2020;21(1):290.

[22] Collard BC, Jahufer M, Brouwer J, Pang E. An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: the basic concepts. Euphytica. 2005;142(1):169-196.

[23] International Rice Research Institute. STAR-Statistical Tool for Agricultural Research v. 2.0.1. http://bbi.irri.org/products [accessed 15 Jun 2019].

[24] Doyle JJ, Doyle JL. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull. 1987;19:11-15.

[25] Van Ooijen J. JoinMap® 4, Software for the calculation of genetic linkage maps in experimental populations. Wageningen, Netherlands: Kyazma B.V.; 2006.

[26] Van Ooijen J. MapQTL® 6, Software for the mapping of quantitative trait loci in experimental populations of diploid species. Wageningen, Netherlands: Kyazma B.V.; 2009.

[27] Cucurbit Coordinated Agricultural Project (CucCAP). Watermelon (97103) v1 Genome, http://cucurb itgenomics.org/organism/1 [accessed 15 Jun 2019].

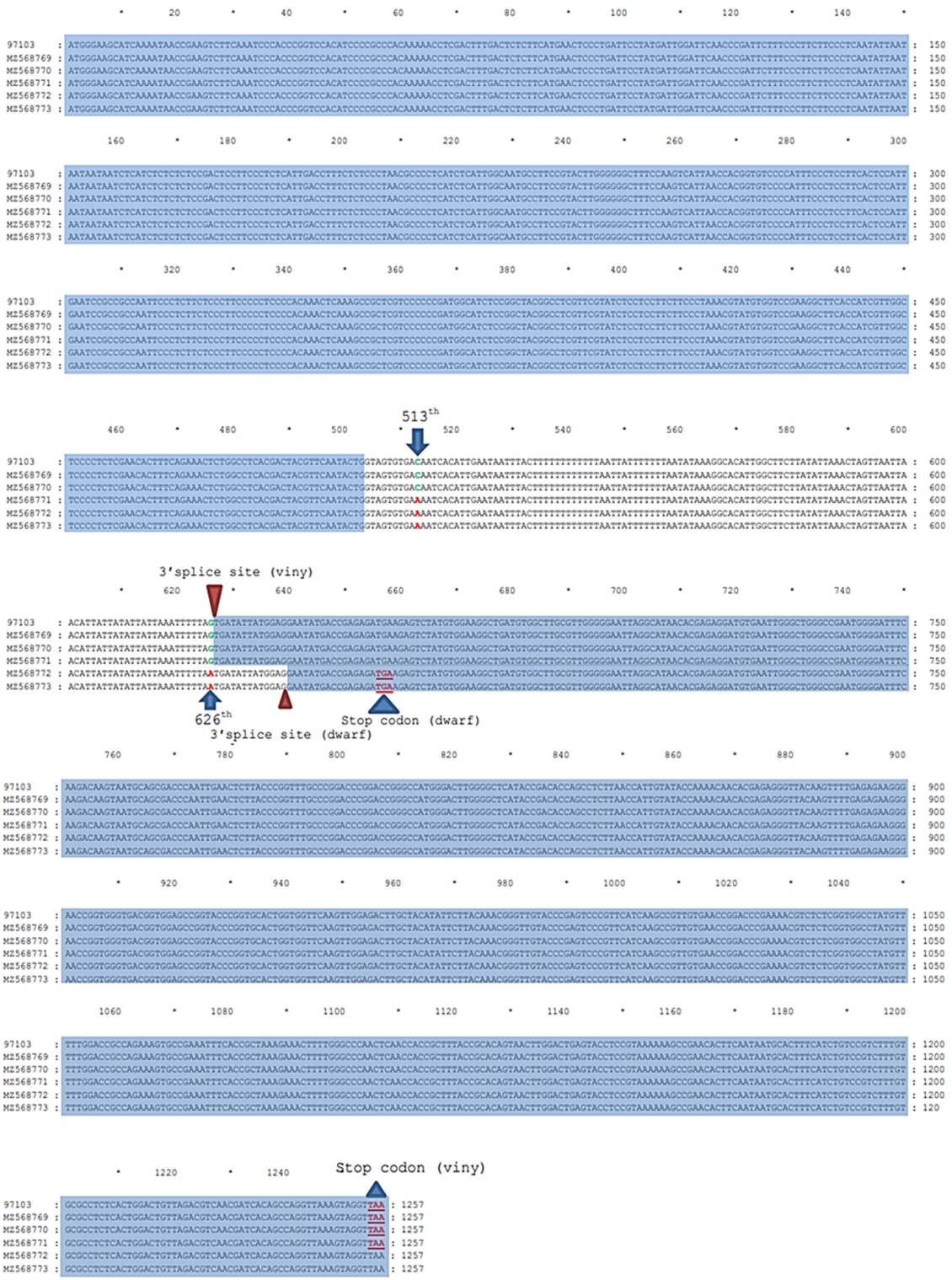
[28] Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2− ΔΔCT method. Methods. 2001;25(4):402-408.
[29] Guner N, Wehner TC. The genes of watermelon. Hortic Sci. 2004;39(6):1175-1182.
[30] Mohr HC. Mode of inheritance of the bushy growth characteristics in watermelon. Proc Assn Southern Agr Workers; 1956;53:174
[31] Hwang J, Oh J, Kim Z, Staub JE, Chung SM, Park Y. Fine genetic mapping of a locus controlling short internode length in melon (*Cucumis melo* L.). Mol Breed. 2014;34(3):949-961.
[32] Li Y, Yang L, Pathak M, Li D, He X, Weng Y. Fine genetic mapping of cp: a recessive gene for compact (dwarf) plant architecture in cucumber, *Cucumis sativus* L. Theor Appl Genet. 2011;123(6):973.
[33] Xin M, Qin Z, Wang L, Wu T, Zhou X. Genetic identification of a dwarf mutant in cucumber (*Cucumis sativus* L.). Afr J Biotechnol. 2012;11(20):4493-4498.
[34] Zhang T, Liu J, Liu S, Ding Z, Luan F, Gao P. Bulked-segregant analysis identified a putative region related to short internode length in melon. Sci Hortic. 2019;54(8):1293-1298.
[35] Wu T, Zhou J, Zhang Y, Cao J. Characterization and inheritance of a bush-type in tropical pumpkin (*Cucurbita moschata* Duchesne). Sci Hortic. 2007;114(1):1-4.

Appendix 1 The list of candidate genes located between two flanking markers (WMSNP-0002750 - WMSNP-0002780) orange box indicates growth hormone related genes.

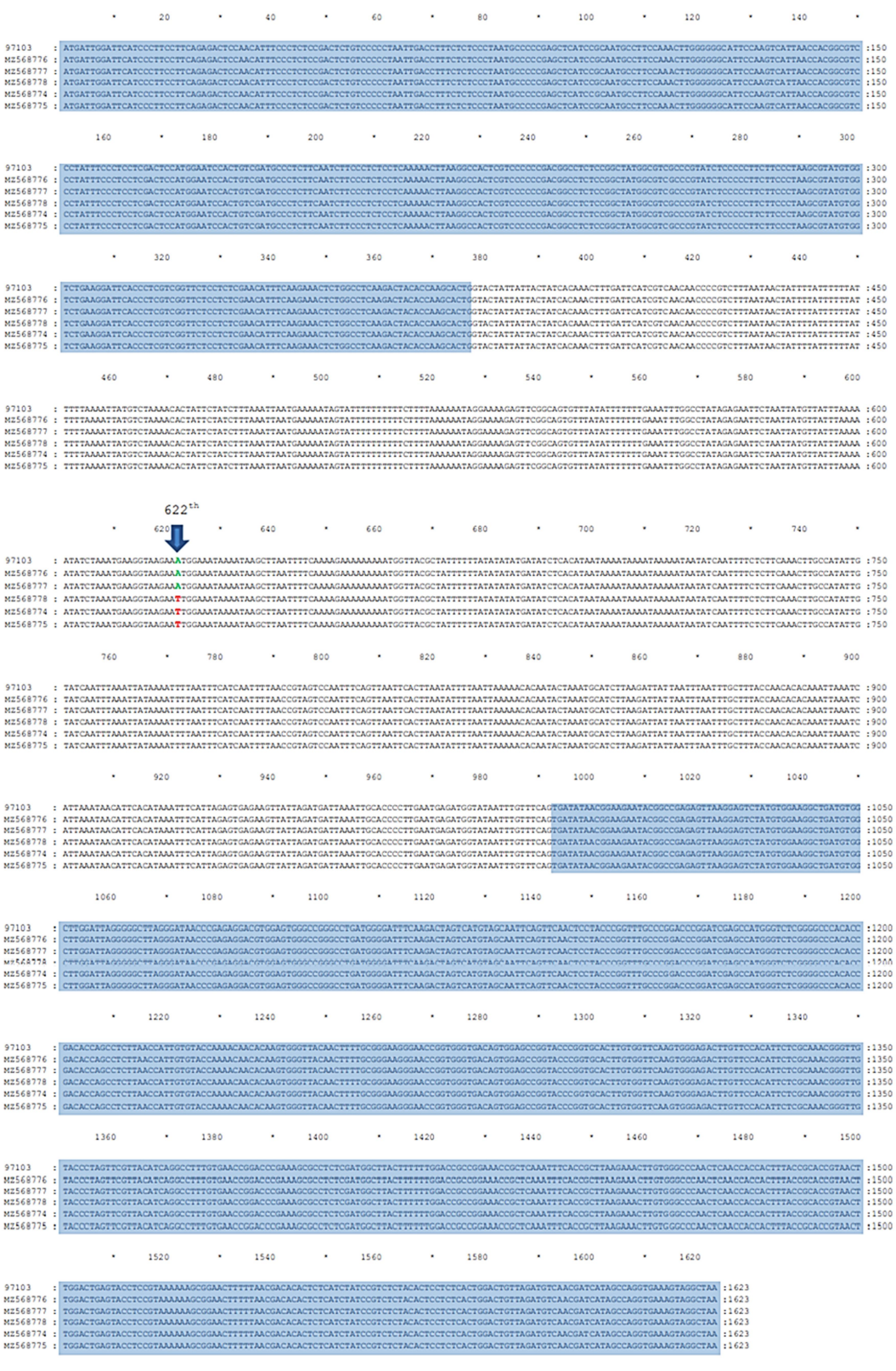
Marker/gene	Position (bp)	Protein encoded
WMSNP-0002750	1.65 Mb	flanking marker
Cla015424	1660268..1660982	Unknown
Cla015423	1685975..1686560	Unknown
Cla015422	1692240..1692870	Unknown
Cla015421	1695059..1695678	Unknown
Cla015420	1699541..1703958	Expressed protein
Cla015419	1706846..1709473	Receptor-like kinase
Cla015418	1711170..1716626	CENP-C
Cla015417	1718518..1722713	DNA repair prtein like XRCC1
Cla015416	1727050..1746431	NMDA receptor-regulated 1-like
Cla015415	1748873..1751565	Acyl-ACP thioesterase
Cla015414	1761209..1761589	Idole-3-acetic acid inducible 19
Cla015413	1765291..1767390	D-glycerate 3-kinase
Cla015412	1782514..1783426	Cyclin-A2-1
Cla015411	1785242..1787309	G2/mitotic-specific cyclin-B
Cla015410	1791454..1796416	Histidinol-phosphate aminotransferase
Cla015409	1797195..1801483	Katanin p60 ATPase-containing subunit
Cla015408	1825632..1827254	Gibberellin 3-beta-hydroxylase
Cla015407	1856847..1858103	Gibberellin 3-beta-hydroxylase
Cla015406	1865167..1867760	2-oxoglutarate-dependent dioxygenase
Cla015405	1872282..1874413	2-oxoglutarate-dependent dioxygenase
Cla015404	1875741..1877969	2-oxoglutarate-dependent dioxygenase
Cla015403	1884562..1891460	ABC transporter G subfamily member 29
Cla015402	1894213..1897458	Unknown Protein (AHRD V1)
Cla015401	1901841..1903265	Sulfate transporter
Cla015400	1905671..1908064	Pentatricopeptide repeat-containing protein
WMSNP-0002780	1.91 Mb	flanking marker

Appendix 2 Genotyping result of functional marker Cla015407-GA in watermelon inbred lines.

Entry	Acc.	Phenotype	Cla015407-GA
1	TH-15974	Dwarf	A/A
2	KK-28413	Dwarf	A/A
3	KK-6939	Viny	G/G
4	KK-28334	Viny	G/G
5	KK-28469	Viny	G/G
6	KK-43359	Viny	G/G
7	KK-13592	Viny	G/G
8	KK-33238	Viny	G/G
9	KK-43389	Viny	G/G
10	KK-44178	Viny	G/G
11	KK-43724	Viny	G/G
12	KK-44208	Viny	G/G
13	KK-43383	Viny	G/G
14	KK-31298	Viny	G/G
15	KK-31711	Viny	G/G
16	KK-43346	Viny	G/G
17	KK-24920	Viny	G/G
18	KK-34799	Viny	G/G
19	KK-34618	Viny	G/G
20	KK-34635	Viny	G/G
21	KK-34805	Viny	G/G
22	KK-35966	Viny	G/G
23	KK-41861	Viny	G/G
24	KK-40383	Viny	G/G
25	KK-34644	Viny	G/G
26	KK-31403	Viny	G/G
27	KK-34724	Viny	G/G
28	KK-33239	Viny	G/G
29	KK-44144	Viny	G/G
30	KK-43344	Viny	G/G
31	KK-44101	Viny	G/G
32	KK-44501	Viny	G/G
33	KK-44105	Viny	G/G
34	KK-44281	Viny	G/G
35	KK-44268	Viny	G/G
36	KK-44154	Viny	G/G



Appendix 3 Multiple alignment of *Cla015407* among viny lines (MZ568769, MZ568770, and MZ568771), dwarf lines (MZ568772 and MZ568773) and reference genome of watermelon (97103). There were two SNPs at 513th and 626th nucleotide within intron 1 region. The TAA at 1255th – 1257th nucleotide was stop codon of viny plants whereas TGA at 656th – 658th nucleotide was the stop codon of dwarf plants.



Appendix 4 Multiple alignment of *Cla015408* among viny lines (MZ568776, MZ568777, and MZ568778), dwarf lines (MZ568774 and MZ568775) and reference genome of watermelon (97103). There was one SNP at 622nd nucleotide within the intron 1 region.