



## Screening and analysis on the differentially expression genes between diploid and autotetraploid watermelon by using of digital gene expression profile

Y. L. Long<sup>a,b</sup>, F. Qiao<sup>c</sup>, X. F. Jiang<sup>a\*</sup>, H. Q. Cong<sup>c</sup>, M. L. Sun<sup>a</sup> and Z. J. Xu<sup>a</sup>

<sup>a</sup>Hainan Key Laboratory of Sustainable Utilization of Tropical Bioresources, Institute of Tropical Agriculture and Forestry, Hainan University – HNU, Haikou 570228, China

<sup>b</sup>Technology Information Institute, Chinese Academy of Tropical Agricultural Sciences – CATAS, Danzhou, Hainan 571737, China

<sup>c</sup>Key Laboratory of Tropical Crops Germplasm Resources Utilization, Ministry of Agriculture, Institute of Tropical Crop Genetic Resources – ITCGR, Chinese Academy of Tropical Agricultural Sciences – CATAS, Danzhou, Hainan 571737, China

\*e-mail: [hjiangxuefei@gmail.com](mailto:hjiangxuefei@gmail.com)

Received: January 16, 2017 – Accepted: December 27, 2017 – Distributed: May 31, 2019  
(With 4 figures)

### Abstract

Synthetic polyploids are key breeding materials for watermelon. Compared with diploid watermelon, the tetraploid watermelon often exhibit wide phenotypic differences and differential gene expression. Digital gene expression (DGE) profile technique was performed in this study to present gene expression patterns in an autotetraploid and its progenitor diploid watermelon, and differentially expressed genes (DEGs) related to the abiotic and biotic stress were also addressed. Altogether, 4,985 DEGs were obtained in the autotetraploid against its progenitor diploid, and 66.02% DEGs is up-regulated. GO analysis shows that these DEGs mainly distributed in ‘metabolic process’, ‘cell’ and ‘catalytic activity’. KEGG analysis revealed that these DEGs mainly cover ‘metabolic pathways’, ‘secondary metabolites’ and ‘ribosome’. Moreover, 134 tolerance related DEGs were identified which cover osmotic adjustment substance, protective enzymes/protein, signaling proteins and pathogenesis-related proteins. This study present the differential expression of stress related genes and global gene expression patterns at background level in autotetraploid watermelons. These new evidences could supplement the molecular theoretical basis for the better resistance after the genome doubling in the gourd family.

*Keywords:* diploidy, autotetraploidy, watermelon, DGE profile, DEGs.

### Triagem e análise da expressão diferencial dos genes entre melancia diplóide e autotetraplóide através do perfil de expressão gênica digital

### Resumo

Poliploides sintéticos são materias fundamentais para melhoramento genético da melancia. Comparativamente ao seu homólogo diploide, a melancia tetraploide apresenta amplas diferenças genotípica e fenotípica e diferença de expressão gênica. A expressão gênica digital ou DGE (digital gene expression) foi utilizada neste estudo para representar o perfil de expressão gênica da melancia autotetraploide e seu progenitor diploide e a expressão diferencial de genes relacionados ao estresse biótico e abiótico. Os resultados mostraram que 4.985 DEGs foram observados no organismo autotetraploide, sendo que, deste total, 66.02% foram supra-regulados. A análise de ontologia gênica (GO) mostrou que estes DEGs estão relacionados principalmente com processos metabólicas, célula e atividade catalítica, abrangendo de acordo com a análise de genes e genoma (KEGG) rotas metabólicas, metabolismo secundário e ribossomos. Além disso, 134 genes de defesa foram identificados, abrangendo substâncias de ajuste osmótico, enzimas/proteínas de proteção, proteínas sinalizadoras e proteínas relacionadas à patogênese. Este estudo mostrou a expressão diferencial de genes relacionados ao estresse e o perfil global de expressão gênica de melancia autotetraploide, estes resultados podem complementar, a nível molecular, o entendimento do fator resistência após a duplicação do genoma em cucurbitáceas.

*Palavras-chave:* diploid, autotetraploide, melancia, perfil DGE, DEGs.

## 1. Introduction

In animal and plant kingdom, many creatures have undergone polyploidization during evolution (Comai, 2005; Benya et al., 2017). Recent evidence indicate that 15% and 31% of angiosperm and fern speciation respectively have experience this event characterized by double or multiply complete sets of chromosomes (Wood et al., 2009). Naturally, polyploidization can be occurred inside a single species or via interspecific hybridization and which is named as autopolyploid or allopolyploid (del Pozo and Ramirez-Parra, 2014). Resynthesize genome duplication of plant cells can be traced back to 1904 by treatment with chloral hydrate and other narcotics (Blakeslee and Avery, 1937). Since then, various crop polyploidys have been created to explore novel phenotypes and improve economic traits such as in cotton (Wendel and Cronn, 2003), wheat (Dubcovsky and Dvorak, 2007), rice (Song et al., 2014) and so on. These successful examples have triggered on great interests to reveal the evolution mechanism between tetraploid and diploid progenitor (Otto and Whitton, 2000; Chantret et al., 2005; Saminathan et al., 2015).

Since plant often show novel phenotypes after polyploidization of chromosome which are not present in their diploid progenitors (Ramsey and Schemske, 1998), then resynthesize polyploidization provide a new way for screening resistant germ-plasma for breeding (Carputo et al., 2000; Vining et al., 2015). Up to now, various polyploidy with drought tolerance (del Pozo and Ramirez-Parra, 2014), cold resistance (Deng et al., 2012), salt tolerance (Chao et al., 2013), pathogen resistance (Oswald and Nuismer, 2007) have been reported, and these phenomena always accompanied by up expression of resistance gene (Wu et al., 2013), protective enzymes (Caverzan et al., 2012), or resistance related small molecules (Willett and Burton, 2002). However, these traits are not always consistent with gene dosage effects since there are more complex mechanism behind it, such as mutation buffering, allelic diversity and heterozygosis, and sub-functionalization of duplicated genes (Osborn et al., 2003; Comai, 2005; Dubcovsky and Dvorak, 2007).

The plant polyploidization could effects on its development and resistance to abiotic stresses, and this effects have been studied through the differential gene expression patterns (Chen and Ni, 2006; Doyle et al., 2008). Gene expression changes could elaborate the molecular basis of polyploid evolutionary event, which have been exemplified in Arabidopsis (Wang et al., 2006), watermelon (Saminathan et al., 2015) and Brassica (Gaeta et al., 2009). In watermelon, RNA sequencing (RNA-seq) revealed that 5,362 (80.63%) differentially expressed genes (DEGs) were up-regulated in autotetraploid, and which involved in arginine biosynthesis, chlorophyllide synthesis, GDP mannose biosynthesis and so on. Moreover, alternative splicing events between diploid and autotetraploid were also found different in spatial expression (Saminathan et al., 2015).

Since autotetraploid watermelon can be routinely induced by colchicines (Suying et al., 1993), and since these autotetraploids always show attractive resistance when challenged by stress stimuli, questions about the gene expression patterns under normal condition between different

autopolyploids and their diploids in watermelon are still waiting exploring. Based on the genome sequenced data (Guo et al., 2013), digital gene expression (DGE) profile technique has become a convenient and economic way for it. Here, we present a gene expression pattern between an autotetraploid and its progenitor diploid watermelon, and DEGs related to the tolerance were also discussed.

## 2. Material and Methods

### 2.1. Materials preparation and RNA isolation

The diploidy and autotetraploidy watermelon (voucher FR-32-1B-2n and FR-32-1B-4n) used in the experiment are kept in Institute of Tropical Crop Genetic Resources (ITCGR), Chinese Academy of Tropical Agricultural Sciences (CATAS). The 3~5 matured leaves were harvested between 15~20 node in the fruit-setting stage watermelons which were cultivated in greenhouse. The harvested leaves were transferred to liquid nitrogen immediately for quick-frozen then grated and mixed into two pools per sample, and kept at -80°C refrigerator for following experiments.

Total RNA was extracted and its quantity and quality was examined according to Qiao et al. (2014).

### 2.2. Illumina RNA-sequencing and data processing

To investigating the expression of genes in diploidy and autotetraploidy watermelons, two cDNA libraries, constructed from diploidy and autotetraploidy watermelon, named as 2n and 4n libraries, were sequenced by an Illumina technology. Specific operations are as follows, firstly, the mRNA was enriched, and then be enzymatic hydrolysis into short sequences. Secondly, taking these short sequences as moldboard, the cDNA and the complementary strand was synthesized. Finally, the double strand cDNA were used for setting up the sample libraries for PCR amplification. Its quality was tested with Agilent 2100 Bioanalyzer and quantity was tested with ABI StepOnePlus Real-Time PCR System. All things being in readiness, the library sequencing tasks were perform using Illumina HiSeq 2000. Millions of raw reads were generated, and then removing the low quality reads to obtain clean tags.

### 2.3. Reads mapping and Differentially Expressed Genes (DEGs) selection

To assume the gene expression characteristics in the diploidy and autotetraploidy watermelons, all clean tags were mapped into watermelon genome database (Cucurbit Genomics Database, 2018) by using Bowtie2 (Langmead et al., 2009) and BWA (Li and Durbin, 2009). Gene expression level was quantified by software RSEM (Li and Dewey, 2011), and the gene expression is represented by the fragments per kilobase per million (FPKM) value. Gene expression level is judged according to the criteria as following: the FPKM threshold 0.1 was set to determine whether the gene was expressed; FPKM in 0.1 to 10 indicates low level; FPKM in 10 to 1,000 indicates medium level, and FPKM > 1,000 indicates high level.

To find differentially expressed genes (DEGs) and perform further function analysis on the diploidy and

autotetraploidy watermelons, DEGs were screened by using Poisson distribution method (Benjamini and Yekutieli, 2011), namely genes had  $p$ -value  $\leq 0.05$  with false discovery rate (FDR)  $\leq 0.001$  and estimated absolute  $|\log_2 \text{ratio}| \geq 1$  in the comparisons of 4n and 2n were selected.

2.4. DEGs GO Annotation and KEGG Pathway Enrichment

To identify and analysis the function and pathway the DEGs involved from the macro-level, The WEGO software (Ye et al., 2006) and the KEGG database was adopted. Taking corrected  $p$ -value  $\leq 0.05$  or  $Q$ -value  $\leq 0.05$  to judge whether a function term or a pathway is significant. This analysis could recognize the main biological functions and pathway of DEGs (Kanehisa et al., 2008).

3. Results

3.1. Analysis of Digital Gene Expression (DGE) profile

According to the sequencing results, in total, approximately 19.4 and 20.6 million raw reads were obtained, and removing low quality reads, a total of 18.6 and 20.1 million clean reads were successfully generated in 2n and 4n libraries,

respectively (shown in Table 1). Furthermore, these tags were mapped to the watermelon reference genome to get its annotation. In our study, 13,206,108 (70.92%) and 13,683,906 (68.19%) total mapped reads were favorably acquired in 2n and 4n libraries, respectively (Table 1). In which 11,442,762 (61.45%) and 11,743,881 (58.52%) clean reads were perfectly matched to the reference genomes, respectively. Further analysis revealed that 12,652,595 clean reads (67.95%) in the 2n library and 13,260,965 clean reads (66.0%) in the 4n library were uniquely mapped (Table 1).

The sequencing depth in the two libraries was analyzed, to ensure our sequence data is sufficient for the DEGs coverage. The number of identified genes is increased followed with the amount of sequenced reads increasing. While, when the amount of sequenced reads reach  $150 \times 100$  k, the growth curve of identified genes tend to be flatten (show in Figure 1), this result means that the number of identified genes is enough for the next step analysis.

3.2. Global analysis of gene expression

The normalcy of the whole DGE profile data was assessed by the dispersion of clean tags expression (Table 2). Obviously, in the diploidy and autotetraploidy

Table 1. Alignment statistics of clean reads align to reference gene.

Summary	2n	4n
Raw reads	19,424,529	20,642,525
Clean reads	18,621,230 (95.86%)	20,067,184 (97.21%)
Total bases	912,440,270	983,292,016
Total mapped reads*	13,206,108 (70.92%)	13,683,906 (68.19%)
Perfect match	11,442,762 (61.45%)	11,743,881 (58.52%)
Unique match	12,652,595 (67.95%)	13,260,965 (66.0%)
Multi-position match	553,513 (2.97%)	422,941 (2.11%)
Mismatch	1,763,346 (9.47%)	1,940,025 (9.67%)
Total unmapped reads	5,415,121 (29.08%)	6,383,277 (31.81%)

\*Total Mapped Reads (%) = Perfect Match (%) + Mismatch (%) = Unique Match (%) + Multi-position Match (%).

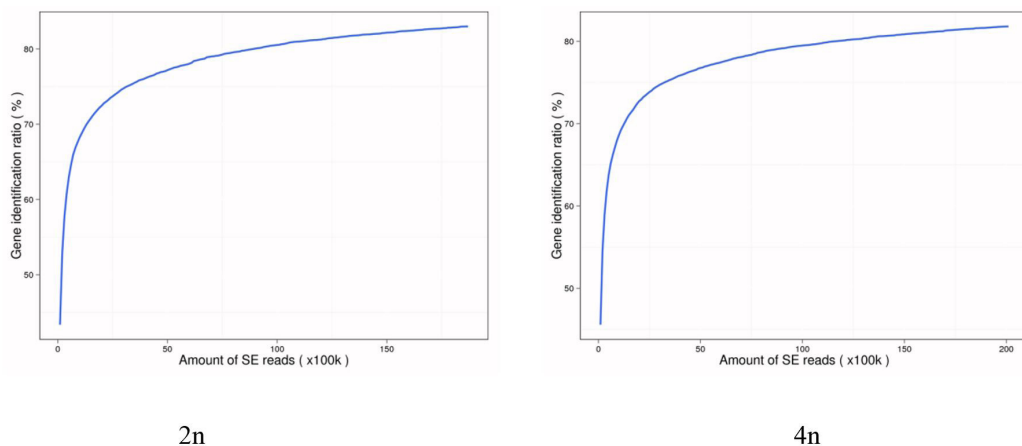


Figure 1. Curve of sequencing saturation. X-axis shows the number of clean reads, units is 100 k -- extreme value is currently the volume of sequencing. Y-axis shows the ratio of identified gene number to number of total gene reported in database. 2n and 4n represent diploidy watermelon and autotetraploidy watermelon, respectively.

**Table 2.** Number of genes at different FPKM interval after significance level correction.

FPKM interval	2n	4n	Remarks
0~0.1	141 (0.78%)	128 (0.71%)	Genes unexpressed
0.1~10	8,129 (44.82%)	7,378 (40.84%)	Low level of gene expression
10~1,000	9,742 (53.71%)	10,432 (57.75%)	Medium level of gene expression
>1,000	125 (0.69%)	127 (0.70%)	High level of gene expression

watermelon, the dispersion of the distinct clean tags expression between the 2n and 4n libraries showed seriously similar tendencies that a majority of genes are in a low or medium expression level, nevertheless, a few genes unexpressed or highly expressed, which imply that the whole DGE data between 2n and 4n libraries are normally distributed. Actually, some researches also revealed that a majority genes were low expressed (Wei et al., 2013; Lu et al., 2012; 't Hoen et al., 2008).

### 3.3. Different Expression Genes (DEGs) analysis

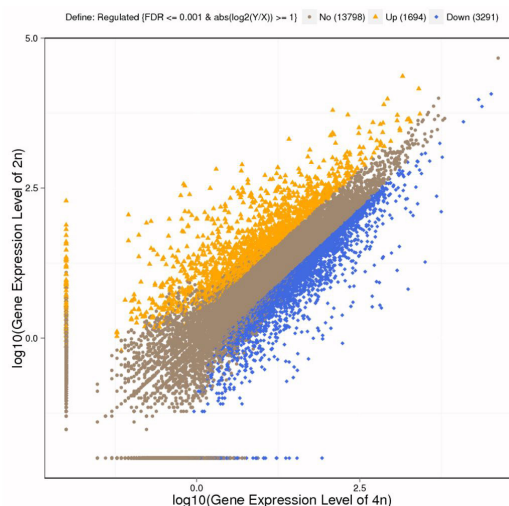
Altogether, there are 4,985 DEGs between the diploidy and autotetraploidy watermelons (Table S1 in Supplementary Materials). The fold changes of DEG expression by using of log 2 ratio to represent, interesting, the majority (95%) of these DEGs varied in log 2 ratio changes are lower than 5 but higher than 1, while only a few (5%) of them exhibited log 2 ratio higher than 5 or lower than 1 in the 2n and 4n libraries (Table S1 in Supplementary Materials), these results conform to the normal distribution of gene expression.

In the Figure 2, blue imply up-regulation genes in 4n library and orange imply up-regulation gene in 2n library. So compared with 4n library, there were only 33.98% DEGs up regulated genes (orange dots) in 2n library, that means 66.02% DEGs up regulated (blue dots) in 4n library (Figure 2).

### 3.4. Functional annotation of DEGs

A wide variety of regulatory and metabolic processes related to altered gene's expression throughout the life of the plant (DeRisi et al., 1997). Based on the allocated Gene ontology (GO) terms, the DEGs in 2n and 4n libraries were classified into several categories using WEGO software. All DEGs were categorized into 3 categories, namely biological process (19 subcategories), cellular component (12 subcategories) and molecular function (11 subcategories) (Figure 3). The 'metabolic process', 'cellular process' and 'single-organism process' subcategories were most prominent in the biological process group. In the cellular component group, 'cell', 'cell part' and 'organelle' were the outstanding terms. For molecular function group, the major subcategories were 'catalytic activity' and 'binding' (Figure 3).

A pathway analysis was carried out to appraise the metabolic network of the DEG's expression changes after chromosome doubling in the watermelons. The front 20 significantly concentrated pathways of the DEGs are



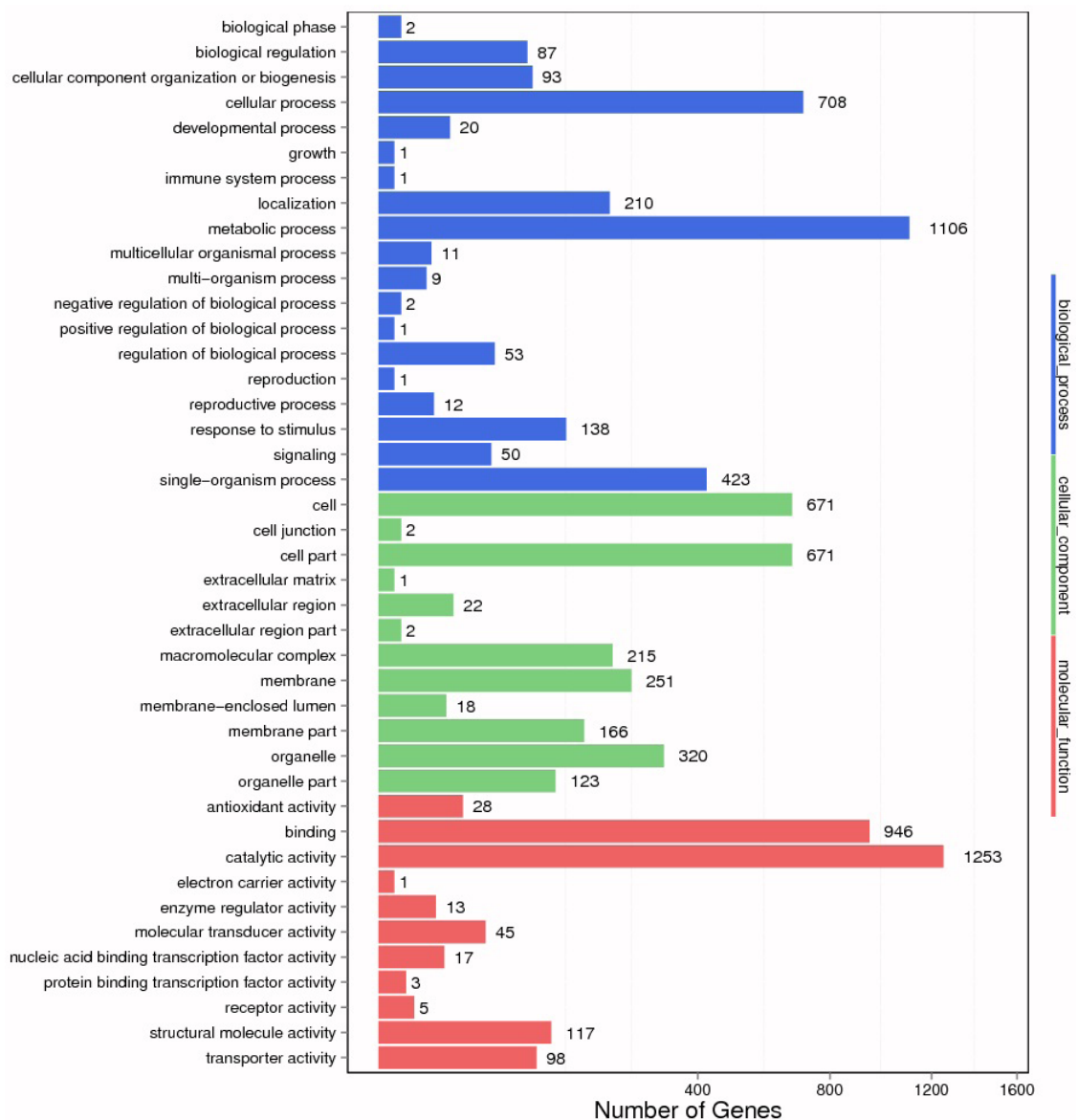
**Figure 2.** Scatter plots of all expressed genes in 2n or 4n library. X-axis means log2 value of gene expression in 4n library, and Y-axis means log2 value of gene expression in 2n library. Blue imply down regulated genes, orange means up regulated genes in 2n library, and brown means non-regulation gene whether in 2n or in 4n library. If a gene expressed just in one sample, its expression value in another sample will be replaced by the minimum value of all expressed genes in control and case samples. Screening threshold is on top legend.

presented in Figure 4. By comparing 2n with 4n library, the pathway analysis results of DEGs indicate that the top 3 pathway of DEGs were 'metabolic pathways', 'biosynthesis of secondary metabolites' and 'ribosome' (Figure 4).

### 3.5. Differentially expressed genes related to stress resistance

A large number of studies demonstrated that polyploids exhibited better stress resistance, for instance, drought tolerance (del Pozo and Ramirez-Parra, 2014), cold resistance (Deng et al., 2012), salt tolerance (Chao et al., 2013), pathogen resistance (Oswald and Nuismer, 2007).

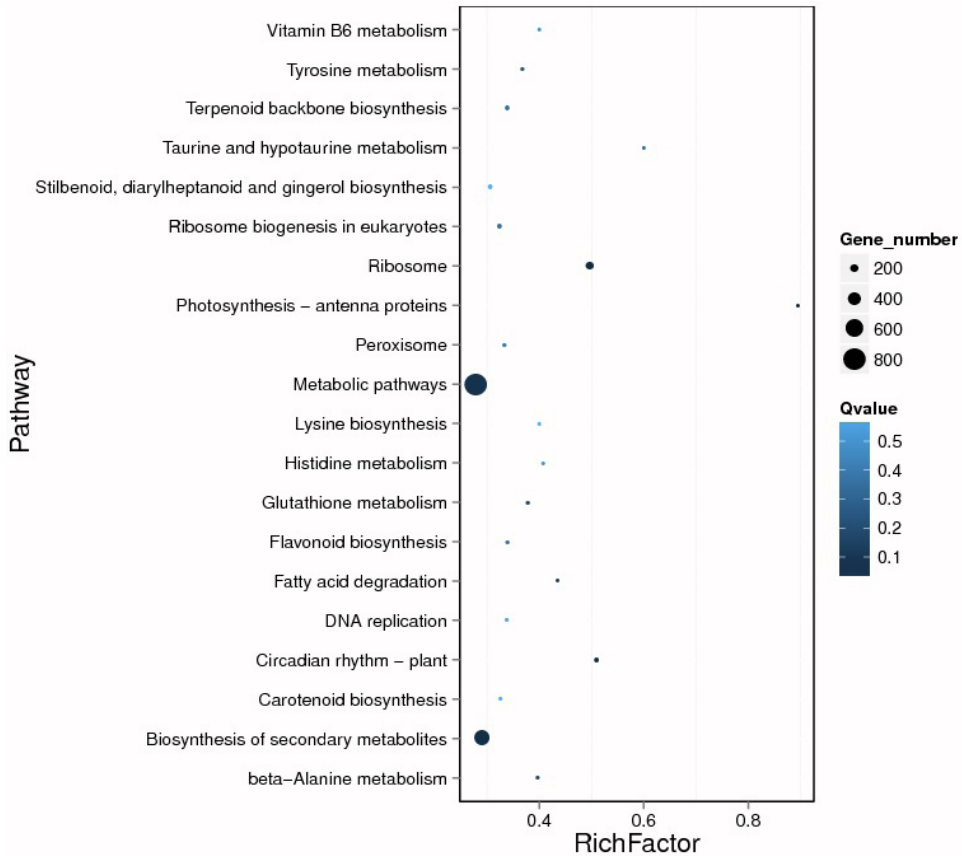
In the present research, candidate genes with known roles in stress resistance were selected on the basis of DGEs between the 2n and 4n. 134 candidate genes were identified with 34 up regulated genes in 2n library and 100 up regulated genes in 4n library (Table S2 in Supplementary Materials).



**Figure 3.** GO functional classification on DEGs for 4n/2n. X axis means number of DEGs (the number is presented by its square root value). Y axis represents GO terms. All GO terms are grouped in to three ontologies: blue is for biological process, green is for cellular component and red is for molecular function.

According to gene’s function, these genes can be classified into (1) osmotic adjustment substance related genes (13 genes), which are mainly closely involved the synthesis and metabolism of carbohydrate, proline and betaine; (2) protective enzymes/protein related genes (39 genes), including SOD、CAT、GSTs, etc.; (3) signaling proteins related genes (71 genes), including transcription factor dominated by WRKY (25 genes), hormone-related and protein kinase (46 genes); (4) pathogen interaction related genes (11 genes) (Table 3).

In view of metabolism, these genes cover six pathways, including glutathione metabolism (12 genes), peroxisome (13 genes), starch and sucrose metabolism (7 genes), plant hormone signal transduction (11 genes), plant-pathogen interaction (71 genes) and proline metabolism (5 genes). Coincidentally, 11 genes in plant hormone signal transduction pathway also belong to plant-pathogen interaction pathway. In addition, there are 26 genes belong to two or more than two metabolic pathways which imply these genes have more critical roles in the stress tolerance.



**Figure 4.** Statistics of pathway enrichment of DEGs in 4n/2n. Rich Factor Refers to the ratio of DGEs and total genes located in this pathway. The larger the rich factor, the greater the degree of DEGs enrichment. Q-value is P-value after correction by a multiple hypothesis test ranging from 0~1, the closer to zero, indicating that the more significant enrichment of this pathway.

**Table 3.** Differentially expressed genes related to stress resistance between diploid (2n) and tetraploid (4n) of watermelon.

Classification	Gene ID
<b>Osmotic adjustment substance related gene</b>	Cla018637, Cla009709, Cla014481, Cla019627, Cla010566, Cla023306, Cla014322, Cla015539, Cla006489, Cla014007, Cla004999, Cla019569, Cla019160, Cla023447, Cla011759, Cla007739, Cla011457, Cla010856, Cla006938, Cla017447, Cla011409, Cla017196, Cla021932, Cla009381, Cla010691, Cla008101, Cla016768, Cla003261, Cla011317, Cla017735, Cla017039, Cla005224, Cla018159, Cla018150, Cla018613, Cla018162, Cla021338, Cla022390, Cla017854, Cla014638, Cla005410, Cla009135, Cla018163, Cla018160, Cla018942, Cla016490, Cla010932, Cla001193, Cla000612, Cla005204, Cla005206, Cla005202.
<b>Protective enzymes/protein related genes</b>	Cla017851, Cla015154, Cla006772, Cla014818, Cla020642, Cla014433, Cla018026, Cla009969, Cla010918, Cla004233, Cla021806, Cla006015, Cla007656, Cla021984, Cla007306, Cla013402, Cla008346, Cla006599, Cla013532, Cla008805, Cla021148, Cla003232, Cla017537, Cla012699, Cla012266.
<b>Transcription factor</b>	Cla011143, Cla002086, Cla006701, Cla009668, Cla011091, Cla020951, Cla021582, Cla008973, Cla014170, Cla002091, Cla012165, Cla010842, Cla020738, Cla016575, Cla010229, Cla013340.
<b>signaling proteins related genes</b>	Cla014185, Cla014725, Cla022981, Cla006067, Cla010434, Cla011861, Cla017994, Cla003893, Cla004486, Cla017316, Cla016798, Cla006308, Cla000205, Cla001497, Cla006519, Cla012567, Cla017256, Cla005545, Cla011936, Cla005313, Cla009754, Cla006764, Cla022876, Cla004866, Cla010135, Cla020836, Cla010883, Cla013089, Cla021418, Cla008563.
<b>Plant pathogen Interaction related genes</b>	Cla018800, Cla003318, Cla010001, Cla007904, Cla014936, Cla020173, Cla018801, Cla003651, Cla003654, Cla001623, Cla006334.

## 4. Discussion

### 4.1. The global gene transcription changes in diploid and tetraploid watermelon

Nowadays, DEGs provide a powerful tool to analysis the gene expression in a global view, this greatly facilitate breeders to decipher the secrets of autopolyploid. Among these secrets, whether various watermelon autopolyploids share some common gene expression patterns are still waiting to be revealed.

This proportional increasement of gene expression to gene dosage have been widely observed (del Pozo and Ramirez-Parra, 2014; Galitski et al., 1999; Guo et al., 1996). Moreover, the disproportional gene expression has been explained as mutation buffering, allelic diversity and heterozygosis, and sub-functionalization (Comai, 2005; Dubcovsky and Dvorak, 2007).

In the present study, two DEG profiles of autopolyploid watermelons were profiled to identify genes which are differentially expressed. More than 18 million clean tags per library was reached, including more than 12 million unique mapped tags (Table 1), suggesting that the database selected is relatively completed. The global gene transcription approximately significantly altered 27.5% due to the chromosome doubling in diploid watermelon, and 66.02% in these DEGs up-regulated in autotetraploid watermelon.

### 4.2. General gene expression characters in the autopolyploid watermelon

Autopolyploidy can greatly alter the cytological process (Sharma and Gohil, 2013), biochemical process (Pall et al., 2015), genetic process (Lloyd and Bomblies, 2016) and physiological process (del Pozo and Ramirez-Parra, 2015) in various organisms. The autotetraploid watermelon used in our study results in larger petals and pedicel than diploid, and pollen grain of autotetraploid is larger with the increasing number of abnormal pollen grains, as a result, the seeds of autotetraploid are less than diploid (Feng, 2012).

GO analysis of DEGs in our DGE profile revealed that in cellular component, 'cell', 'organelle' and 'cell part' are dominant, which may contribute to the above mentioned morphological changes. Besides, higher level of secondary metabolites accumulation was also found in autopolyploidy (Lavana et al., 2012), such as alkaloids, terpene, scopolamine, etc (Madani et al., 2015; Dehghan et al., 2012; Lavana et al., 2012). Here, in our tetraploid, massive up-regulated genes are also fell into 'biosynthesis of secondary metabolites' pathway.

### 4.3. Gene expression divergence caused by tetraploidization is involved with stress response

Evolutionarily, polyploidization has occurred in many plants naturally and in turn some of which may be with better change to be survival against extreme environments (Murray, 1995). Under harsh environment, the plants with duplicated genome always show a high level of expression of the tolerant related genes which contributed to their adaptability (Bertrand et al., 2015; Tamayo-Ordoñez et al., 2016). However, whether the polyploidized plant has to keep higher level of gene expression even under suitable condition or whether only some specific pathways are taking

the surveillance job are still not well documented. New clues in this study showed that there are 134 defense related genes which were differentially expressed in the tetraploid and the diploid (show in Table 3). These genes are mainly involved in signal transduction, osmotic adjustment, ROS related protective enzymes and disease resistance genes.

Efficiency of transform extracellular stimuli into intracellular physiological signals and transduction speed of intracellular signals play vital roles in stress tolerance. And these signals can be various substance, including sugars, transcription factors (TFs), protein kinases and calmodulin, (Kunz et al., 2014; Latz et al., 2013; Nakashima et al., 2012; Yoshida et al., 2014). In this research, 52.99% tolerance related DEGs encoding signaling substance, such as WRKY TFs, protein kinases and calmodulin, and 71.83% of them up-regulated in 4n library. A recent report revealed that allopolyploid *Coffea arabica* is with higher content of sucrose than its two diploid parental species under cold temperature, whereas few differences were found at related gene expression level, supports that allopolyploid is more efficient in signaling (Bertrand et al., 2015).

Besides powerful signaling, polyploid plants can also up-regulate resistance genes involved in biosynthesis such as proline, betaine, ROS related protective enzymes and pathogenesis-related proteins (del Pozo and Ramirez-Parra, 2014; Bertrand et al., 2015; Tamayo-Ordoñez et al., 2016). Here, in our non-stressed 4n library, DEGs of proline synthesis, GB synthesis, protective enzymes and pathogenesis-related protein are strongly up-regulated.

## 5. Conclusion

A total of 18,783 DEGs were identified from diploidy (2n) and autotetraploidy (4n) watermelon. In which, 4,985 genes expressed significantly different and 3,291 DEGs were up-regulated in the autotetraploidy. GO analysis show that these DEGs mainly distributed in 'metabolic process', 'cell' and 'catalytic activity'. KEGG analysis revealed that these DEGs mainly cover 'metabolic pathways', 'secondary metabolites' and 'ribosome'. And 134 genes closely related to the stress resistance including osmotic adjustment substance, protective enzymes/protein, signaling proteins and pathogenesis-related protein in these significant DEGs.

These results suggest that there already exist a lot of differentially expressed genes in our diploidy and autotetraploidy watermelon under non-stress condition. And these DEGs can be related to various plant metabolic regulations, which contribute to the adaption ability and tolerances in autotetraploidy. These results supplement the molecular regulation system in autopolyploidy watermelon and also provides evidences of the basis of the resistance in autopolyploids.

## Acknowledgements

The authors thank the National Natural Science Foundation of China (31760595), Key Technologies Program of Haikou (2017045) and Science Research Foundation for High Education by Educational Commission of Hainan Province (Hnky2017-6).

## References

- 'T HOEN, P.A.C., ARIYUREK, Y., THYGESEN, H.H., VREUGDENHIL, E., VOSSSEN, R.H.A.M., DE MENEZES, R.X., BOER, J.M., VAN OMMEN, G.-J.B. and DEN DUNNEN, J.T., 2008. Deep sequencing-based expression analysis shows major advances in robustness, resolution and inter-lab portability over five microarray platforms. *Nucleic Acids Research*, vol. 36, no. 21, pp. 1-11. <http://dx.doi.org/10.1093/nar/gkn705>.
- BENJAMINI, Y. and YEKUTIELI, D., 2011. The control of the false discovery rate in multiple testing under dependency. *Annals of Statistics*, vol. 29, no. 4, pp. 65-88. <http://dx.doi.org/10.1214/aos/1013699998>.
- BENYA, E.G.F., LEAL-ZANCHET, A.M. and HAUSER, J., 2017. Polyploidy as a chromosomal component of stochastic noise: variable scalar multiples of the diploid chromosome complement in the invertebrate species *Girardia schubarti* from Brazil. *Brazilian Journal of Biology = Revista Brasileira de Biologia*, vol. 77, no. 4, pp. 745-751. <http://dx.doi.org/10.1590/1519-6984.20615>. PMID:28614425.
- BERTRAND, B., BARDIL, A., BARAILLE, H., DUSSERT, S., DOULBEAU, S., DUBOIS, E., SEVERAC, D., DEREPPER, A. and ETIENNE, H., 2015. The greater phenotypic homeostasis of the allopolyploid *Coffea arabica* improved the transcriptional homeostasis over that of both diploid parents. *Plant & Cell Physiology*, vol. 56, no. 10, pp. 2035-2051. <http://dx.doi.org/10.1093/pcp/pcv117>. PMID:26355011.
- BLAKESLEE, A.F. and AVERY, A.G., 1937. Methods of inducing doubling of chromosomes in plants by treatment with colchicine. *The Journal of Heredity*, vol. 28, no. 12, pp. 393-411. <http://dx.doi.org/10.1093/oxfordjournals.jhered.a104294>.
- CARPUTO, D., BARONE, A. and FRUSCIANTE, L., 2000. 2n gametes in the potato: essential ingredients for breeding and germplasm transfer. *Theoretical and Applied Genetics*, vol. 101, no. 5-6, pp. 805-813. <http://dx.doi.org/10.1007/s001220051547>.
- CAVERZAN, A., PASSAIA, G., ROSA, S.B., RIBEIRO, C.W., LAZZAROTTO, F. and MARGIS-PINHEIRO, M., 2012. Plant responses to stresses: role of ascorbate peroxidase in the antioxidant protection. *Genetics and Molecular Biology*, vol. 35, no. 4, suppl. 1, pp. 1011-1019. <http://dx.doi.org/10.1590/S1415-47572012000600016>. PMID:23412747.
- CHANTRET, N., SALSE, J., SABOT, F., RAHMAN, S., BELLEC, A., LAUBIN, B., DUBOIS, I., DOSSAT, C., SOURDILLE, P., JOUDRIER, P., GAUTIER, M.F., CATTOLICO, L., BECKERT, M., AUBOURG, S., WEISSENBACH, J., CABOCHE, M., BERNARD, M., LEROY, P. and CHALHOUB, B., 2005. Molecular basis of evolutionary events that shaped the hardness locus in diploid and polyploid wheat species (*Triticum* and *Aegilops*). *The Plant Cell*, vol. 17, no. 4, pp. 1033-1045. <http://dx.doi.org/10.1105/tpc.104.029181>. PMID:15749759.
- CHAO, D.Y., DILKES, B., LUO, H., DOUGLAS, A., YAKUBOVA, E., LAHNER, B. and SALT, D.E., 2013. Polyploids exhibit higher potassium uptake and salinity tolerance in *Arabidopsis*. *Science*, vol. 341, no. 6146, pp. 658-659. <http://dx.doi.org/10.1126/science.1240561>. PMID:23887874.
- CHEN, Z.J. and NI, Z., 2006. Mechanisms of genomic rearrangements and gene expression changes in plant polyploids. *BioEssays*, vol. 28, no. 3, pp. 240-252. <http://dx.doi.org/10.1002/bies.20374>. PMID:16479580.
- COMAI, L., 2005. The advantages and disadvantages of being polyploid. *Nature Reviews. Genetics*, vol. 6, no. 11, pp. 836-846. <http://dx.doi.org/10.1038/nrg1711>. PMID:16304599.
- Cucurbit Genomics Database, 2018 [viewed 25 September 2015]. *Watermelon (97103) Genome* [online]. Available from: <http://cucurbitgenomics.org/>
- DEHGHAN, E., HÄKKINEN, S.T., OKSMAN-CALDENTEY, K.-M. and SHAHRIARI AHMADI, F., 2012. Production of tropane alkaloids in diploid and tetraploid plants and in vitro hairy root cultures of Egyptian henbane (*Hyoscyamus muticus* L.). *Plant Cell, Tissue and Organ Culture*, vol. 110, no. 1, pp. 35-44. <http://dx.doi.org/10.1007/s11240-012-0127-8>.
- DEL POZO, J.C. and RAMIREZ-PARRA, E., 2014. Deciphering the molecular bases for drought tolerance in *Arabidopsis* autotetraploids. *Plant, Cell & Environment*, vol. 37, no. 12, pp. 2722-2737. <http://dx.doi.org/10.1111/pce.12344>. PMID:24716850.
- DEL POZO, J.C. and RAMIREZ-PARRA, E., 2015. Whole genome duplications in plants: an overview from *Arabidopsis*. *Journal of Experimental Botany*, vol. 66, no. 22, pp. 6991-7003. <http://dx.doi.org/10.1093/jxb/erv432>. PMID:26417017.
- DENG, B., DU, W., LIU, C., SUN, W., TIAN, S. and DONG, H., 2012. Antioxidant response to drought, cold and nutrient stress in two ploidy levels of tobacco plants: low resource requirement confers polytolerance in polyploids? *Plant Growth Regulation*, vol. 66, no. 1, pp. 37-47. <http://dx.doi.org/10.1007/s10725-011-9626-6>.
- DERISI, J.L., IYER, V.R. and BROWN, P.O., 1997. Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science*, vol. 278, no. 5338, pp. 680-686. <http://dx.doi.org/10.1126/science.278.5338.680>. PMID:9381177.
- DOYLE, J.J., FLAGEL, L.E., PATERSON, A.H., RAPP, R.A., SOLTIS, D.E., SOLTIS, P.S. and WENDEL, J.F., 2008. Evolutionary genetics of genome merger and doubling in plants. *Annual Review of Genetics*, vol. 42, no. 1, pp. 443-461. <http://dx.doi.org/10.1146/annurev.genet.42.110807.091524>. PMID:18983261.
- DUBCOVSKY, J. and DVORAK, J., 2007. Genome plasticity a key factor in the success of polyploid wheat under domestication. *Science*, vol. 316, no. 5833, pp. 1862-1866. <http://dx.doi.org/10.1126/science.1143986>. PMID:17600208.
- FENG, Z., 2012. *Study on tetraploid induction of mini-watermelon and fruit characteristics and the embryo development of the different ploidy*. Haikou: Hainan University, pp. 15-43. Master's Dissertation in Crop Genetics and Breeding.
- GAETA, R.T., YOO, S.Y., PIRES, J.C., DOERGE, R.W., CHEN, Z.J. and OSBORN, T.C., 2009. Analysis of gene expression in resynthesized *Brassica napus* allopolyploids using *Arabidopsis* 70mer oligo microarrays. *PLoS One*, vol. 4, no. 3, pp. e4760. <http://dx.doi.org/10.1371/journal.pone.0004760>. PMID:19274085.
- GALITSKI, T., SALDANHA, A.J., STYLES, C.A., LANDER, E.S. and FINK, G.R., 1999. Ploidy regulation of gene expression. *Science*, vol. 285, no. 5425, pp. 251-254. <http://dx.doi.org/10.1126/science.285.5425.251>. PMID:10398601.
- GUO, M., DAVIS, D. and BIRCHLER, J.A., 1996. Dosage effects on gene expression in a maize ploidy series. *Genetics*, vol. 142, no. 4, pp. 1349-1355. PMID:8846910.
- GUO, S., ZHANG, J., SUN, H., SALSE, J., LUCAS, W., ZHANG, H., ZHENG, Y., MAO, L., REN, Y., WANG, Z., MIN, J., GUO, X., MURAT, F., HAM, B.K., ZHANG, Z., GAO, S., HUANG, M., XU, Y., ZHONG, S., BOMBARELY, A., MUELLER, L.A., ZHAO, H., HE, H., ZHANG, Y., ZHANG, Z., HUANG, S., TAN,



- T., PANG, E., LIN, K., HU, Q., KUANG, H., NI, P., WANG, B., LIU, J., KOU, Q., HOU, W., ZOU, X., JIANG, J., GONG, G., KLEE, K., SCHOOF, H., HUANG, Y., HU, X., DONG, S., LIANG, D., WANG, J., WU, K., XIA, Y., ZHAO, X., ZHENG, Z., XING, M., LIANG, X., HUANG, B., LV, T., WANG, J., YIN, Y., YI, H., LI, R., WU, M., LEVI, A., ZHANG, X., GIOVANNONI, J.J., WANG, J., LI, Y., FEI, Z. and XU, Y., 2013. The draft genome of watermelon (*Citrullus lanatus*) and resequencing of 20 diverse accessions. *Nature Genetics*, vol. 45, no. 1, pp. 51-58. <http://dx.doi.org/10.1038/ng.2470>. PMID:23179023.
- KANEHISA, M., ARAKI, M., GOTO, S., HATTORI, M., HIRAKAWA, M., ITOH, M., KATAYAMA, T., KAWASHIMA, S., OKUDA, S., TOKIMATSU, T. and YAMANISHI, Y., 2008. KEGG for linking genomes to life and the environment. *Nucleic Acids Research*, vol. 36, suppl. 1, pp. D480-D484. <http://dx.doi.org/10.1093/nar/gkm882>. PMID:18077471.
- KUNZ, S., PESQUET, E. and KLECZKOWSKI, L.A., 2014. Functional dissection of sugar signals affecting gene expression in *Arabidopsis thaliana*. *PLoS One*, vol. 9, no. 6, pp. e100312. <http://dx.doi.org/10.1371/journal.pone.0100312>. PMID:24950222.
- LANGMEAD, B., TRAPNELL, C., POP, M. and SALZBERG, S.L., 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology*, vol. 10, no. 3, pp. R25-34. <http://dx.doi.org/10.1186/gb-2009-10-3-r25>. PMID:19261174.
- LATZ, A., MEHLMER, N., ZAPF, S., MUELLER, T.D., WURZINGER, B., PFISTER, B., CSASZAR, E., HEDRICH, R., TEIGE, M. and BECKER, D., 2013. Salt stress triggers phosphorylation of the *Arabidopsis* vacuolar K<sup>+</sup> channel TPK1 by calcium-dependent protein kinases (CDPKs). *Molecular Plant*, vol. 6, no. 4, pp. 1274-1289. <http://dx.doi.org/10.1093/mp/sss158>. PMID:23253603.
- LAVANIA, U.C., SRIVASTAVA, S., LAVANIA, S., BASU, S., MISRA, N.K. and MUKAI, Y., 2012. Autopolyploidy differentially influences body size in plants, but facilitates enhanced accumulation of secondary metabolites, causing increased cytosine methylation. *The Plant Journal*, vol. 71, no. 4, pp. 539-549. <http://dx.doi.org/10.1111/j.1365-3113X.2012.05006.x>. PMID:22449082.
- LI, B. and DEWEY, C.N., 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics*, vol. 12, no. 1, pp. 323. <http://dx.doi.org/10.1186/1471-2105-12-323>. PMID:21816040.
- LI, H. and DURBIN, R., 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, vol. 25, no. 14, pp. 1754-1760. <http://dx.doi.org/10.1093/bioinformatics/btp324>. PMID:19451168.
- LLOYD, A. and BOMBLIES, K., 2016. Meiosis in autopolyploid and allopolyploid *Arabidopsis*. *Current Opinion in Plant Biology*, vol. 30, pp. 116-122. <http://dx.doi.org/10.1016/j.pbi.2016.02.004>. PMID:26950252.
- LU, J., DU, Z.X., KONG, J., CHEN, L.N., QIU, Y.H., LI, G.F., MENG, X.H. and ZHU, S.F., 2012. Transcriptome analysis of *Nicotiana tabacum* infected by *Cucumber mosaic virus* during systemic symptom development. *PLoS One*, vol. 7, no. 8, pp. e43447. <http://dx.doi.org/10.1371/journal.pone.0043447>. PMID:22952684.
- MADANI, H., HOSSEINI, B., DEHGHAN, E. and REZAEI-CHIYANEH, E., 2015. Enhanced production of scopolamine in induced autotetraploid plants of *Hyoscyamus reticulatus* L. *Acta Physiologiae Plantarum*, vol. 37, no. 3, pp. 55. <http://dx.doi.org/10.1007/s11738-015-1795-x>.
- MURRAY, D.F., 1995. Causes of arctic plant diversity: origin and evolution. In: F.S. CHAPIN, C. KÖRNER, eds. *Arctic and Alpine biodiversity: patterns, causes and ecosystem consequences*. Berlin: Springer, vol. 113, pp. 21-32. [http://dx.doi.org/10.1007/978-3-642-78966-3\\_2](http://dx.doi.org/10.1007/978-3-642-78966-3_2).
- NAKASHIMA, K., TAKASAKI, H., MIZOI, J., SHINOZAKI, K. and YAMAGUCHI-SHINOZAKI, K., 2012. NAC transcription factors in plant abiotic stress responses. *Biochimica et Biophysica Acta*, vol. 1819, no. 2, pp. 97-103. <http://dx.doi.org/10.1016/j.bbarm.2011.10.005>.
- OSBORN, T.C., CHRIS PIRES, J., BIRCHLER, J.A., AUGER, D.L., JEFFERY CHEN, Z., LEE, H.-S., COMAI, L., MADLUNG, A., DOERGE, R.W., COLOT, V. and MARTIENSSEN, R.A., 2003. Understanding mechanisms of novel gene expression in polyploids. *Trends in Genetics*, vol. 19, no. 3, pp. 141-147. [http://dx.doi.org/10.1016/S0168-9525\(03\)00015-5](http://dx.doi.org/10.1016/S0168-9525(03)00015-5). PMID:12615008.
- OSWALD, B.P. and NUISMER, S.L., 2007. Neopolyploidy and pathogen resistance. *Proceedings. Biological Sciences*, vol. 274, no. 1624, pp. 2393-2397. <http://dx.doi.org/10.1098/rspb.2007.0692>. PMID:17686733.
- OTTO, S.P. and WHITTON, J., 2000. Polyploid incidence and evolution. *Annual Review of Genetics*, vol. 34, no. 1, pp. 401-437. <http://dx.doi.org/10.1146/annurev.genet.34.1.401>. PMID:11092833.
- PALII, A., BATÏRU, G., ROTARI, A., ROTARI, E. and COMAROV, G., 2015. The influence of ploidy level on some biochemical characteristics of maize grains with OPAQUE-2 endosperm. *Scientific Papers - Series A. Agronomy (Basel)*, vol. 4, pp. 273-276.
- QIAO, F., CONG, H., JIANG, X., WANG, R., YIN, J., QIAN, Q., WANG, Z. and NICK, P., 2014. *De Novo* characterization of a *Cephalotaxus hainanensis* transcriptome and genes related to paclitaxel biosynthesis. *PLoS One*, vol. 9, no. 9, pp. e106900. <http://dx.doi.org/10.1371/journal.pone.0106900>. PMID:25203398.
- RAMSEY, J. and SCHEMSKE, D.W., 1998. Pathways, mechanisms, and rates of polyploid formation in flowering plants. *Annual Review of Ecology and Systematics*, vol. 29, no. 1, pp. 467-501. <http://dx.doi.org/10.1146/annurev.ecolsys.29.1.467>.
- SAMINATHAN, T., NIMMAKAYALA, P., MANOHAR, S., MALKARAM, S., ALMEIDA, A., CANTRELL, R., TOMASON, Y., ABBURI, L., RAHMAN, M.A., VAJJA, V.G., KHACHANE, A., KUMAR, B., RAJASIMHA, H.K., LEVI, A., WEHNER, T. and REDDY, U.K., 2015. Differential gene expression and alternative splicing between diploid and tetraploid watermelon. *Journal of Experimental Botany*, vol. 66, no. 5, pp. 1369-1385. <http://dx.doi.org/10.1093/jxb/eru486>. PMID:25520388.
- SHARMA, G. and GOHIL, R.N., 2013. Origin and cytology of a novel cytotype of *Allium tuberosum* Rottl. Ex Spreng. (2n = 48). *Genetic Resources and Crop Evolution*, vol. 60, no. 2, pp. 503-511. <http://dx.doi.org/10.1007/s10722-012-9852-4>.
- SONG, Z.J., DU, C.Q., ZHANG, X.H., CHEN, D.L., HE, Y.C. and CAI, D.T., 2014. Studies on awns in polyploid rice (*Oryza sativa* L.) and preliminary cross experiments of a special awnless tetraploid rice. *Genetic Resources and Crop Evolution*, vol. 61, no. 4, pp. 797-807. <http://dx.doi.org/10.1007/s10722-013-0074-1>.
- SUYING, T., XIUQIANG, H. and JIWEI, L., 1993. Raising the frequency of inducing tetraploid watermelon by treating of colchicine. *Acta Horticulturae*, no. 402, pp. 18-22. <http://dx.doi.org/10.17660/ActaHortic.1995.402.2>.
- TAMAYO-ORDOÑEZ, M.C., RODRIGUEZ-ZAPATA, L.C., NARVAEZ-ZAPATA, J.A., TAMAYO-ORDONEZ, Y.J., AYIL-

- GUTIERREZ, B.A., BARREDO-POOL, F. and SANCHEZ-TEYER, L.F., 2016. Morphological features of different polyploids for adaptation and molecular characterization of CC-NBS-LRR and LEA gene families in *Agave L.* *Journal of Plant Physiology*, vol. 195, pp. 80-94. <http://dx.doi.org/10.1016/j.jplph.2016.03.009>. PMID:27016883.
- VINING, K.J., DAVIS, T.M., JAMIESON, A.R. and MAHONEY, L.L., 2015. Germplasm resources for verticillium wilt resistance breeding and genetics in strawberry (*Fragaria*). *Journal of Berry Research*, vol. 5, no. 4, pp. 183-195. <http://dx.doi.org/10.3233/JBR-150096>.
- WANG, J., TIAN, L., LEE, H.S., WEI, N.E., JIANG, H., WATSON, B., MADLUNG, A., OSBORN, T.C., DOERGE, R.W., COMAI, L. and CHEN, Z.F., 2006. Genome wide nonadditive gene regulation in *Arabidopsis* allotetraploids. *Genetics*, vol. 172, no. 1, pp. 507-517. <http://dx.doi.org/10.1534/genetics.105.047894>. PMID:16172500.
- WEI, M., SONG, M., FAN, S. and YU, S., 2013. Transcriptomic analysis of differentially expressed genes during anther development in genetic male sterile and wild type cotton by digital gene-expression profiling. *BMC Genomics*, vol. 14, no. 1, pp. 97. <http://dx.doi.org/10.1186/1471-2164-14-97>. PMID:234022Ta79.
- WENDEL, J.F. and CRONN, R.C., 2003. Polyploidy and the evolutionary history of cotton. *Advances in Agronomy*, vol. 78, pp. 139-186. [http://dx.doi.org/10.1016/S0065-2113\(02\)78004-8](http://dx.doi.org/10.1016/S0065-2113(02)78004-8).
- WILLETT, C.S. and BURTON, R.S., 2002. Proline biosynthesis genes and their regulation under salinity stress in the euryhaline copepod *Tigriopus californicus*. *Comparative Biochemistry and Physiology. Part B, Biochemistry & Molecular Biology*, vol. 132, no. 4, pp. 739-750. [http://dx.doi.org/10.1016/S1096-4959\(02\)00091-X](http://dx.doi.org/10.1016/S1096-4959(02)00091-X). PMID:12128060.
- WOOD, T.E., TAKEBAYASHI, N., BARKER, M.S., MAYROSE, I., GREENSPOON, P.B. and RIESEBERG, L.H., 2009. The frequency of polyploid speciation in vascular plants. *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, no. 33, pp. 13875-13879. <http://dx.doi.org/10.1073/pnas.0811575106>. PMID:19667210.
- WU, P., QIN, Z., ZHAO, W., ZHOU, X.Y., WU, T., XIN, M. and GUO, Q.Q., 2013. Transcriptome analysis reveals differentially expressed genes associated with propamocarb response in cucumber (*Cucumis sativus* L.) fruit. *Acta Physiologiae Plantarum*, vol. 35, no. 8, pp. 2393-2406. <http://dx.doi.org/10.1007/s11738-013-1274-1>.
- YE J., FANG, L., ZHENG, H., ZHANG, Y., CHEN, J., ZHANG, Z., WANG, J., LI, S., LI, R., BOLUND, L. and WANG, J., 2006. WEGO: a web tool for plotting GO annotations. *Nucleic Acids Research*, vol. 34, suppl. 2, pp. W293-W297. <http://dx.doi.org/10.1093/nar/gkl031>. PMID: 16845012.
- YOSHIDA, T., MOGAMI, J. and YAMAGUCHI-SHINOZAKI, K., 2014. ABA-dependent and ABA-independent signaling in response to osmotic stress in plants. *Current Opinion in Plant Biology*, vol. 21, pp. 133-139. <http://dx.doi.org/10.1016/j.pbi.2014.07.009>. PMID:25104049.

## Supplementary Material

Supplementary material accompanies this paper.

**Table S1.** Different Expression Genes in the diploidy and autotetraploidy watermelons.

**Table S2.** Up regulated resistance related genes in diploidy and autotetraploidy watermelons.

This material is available as part of the online article from <http://www.scielo.br/bjb>.