Comparative transcriptome analysis of Powdery mildew Resistance between two Melon (*Cucumis melo* L) with Different Thickness Peel

CHENG Hong^{1,*}, KONG Wei-ping¹, Lü Jun-Feng²

(¹ Vegetable Research Institute, Gansu Academy of Agricultural Sciences, Lanzhou 730070, China; ² College of
Veterinary Medicine, Gansu Agricultural University, Lanzhou, 730070, China)

Abstract

Melon (*Cucumis melo* L.) is widely planted in the world and China is a major producer of muskmelon. Powdery mildew is one of the most common fungal diseases in the world and this disease frequently affects melon (*Cucumis melo* L.) and due to the reduction of melon yield. In this study, one material GanTianmi with thin peel and another material XueLianHua with thick peel were selected. After inoculating the powdery mildew, both materials were used to do the RNA-Seq. In total two RNA-seq libraries were constructed and sequenced separately. The reads per kilobase per Million mapped reads (RPKM) values of all the genes in the two materials were calculated and there were 13828 genes were expressed in the material G and 13944 genes were expressed in the material S (RPKM>1). The differentially expression gene (DEG) analysis result suggested that total 769 the DEGs between the two materials were identified. All the DEGs were annotated with several database and the transcript factors (TFs) that related to disease resistance such as *MYB*, *ERF* and *WRKY* among the DEGs were also identified. This research could not only provide the information about understanding the mechanism of powdery mildew infection but also help researchers breed the varieties with powdery mildew resistance.

Key words Melon (*Cucumis melo* L.). powdery mildew. Transcript sequencing. Candidate genes.
Transcript factors

26 Introduction

27 Melon (*Cucumis melo* L.) is one of the most valuable cash crops that grown all over the word and hina
 28 also is a major producer of muskmelon (Liu, 2003; Chen et al, 2015). However, powdery mildew often
 29 infects muskmelon, thereby resulting in significant losses in its production (Cheng et al., 2006). In
 30 previous studies, there were lots of research that identified candidate genes/function genes about the
 31 powdery mildew resistance (Fukino et al, 2008; Chen et al, 2012, 2013 and 2015; Natarajan et al, 2016;
 32 Wang et al, 2016)
 33 Some function genes were identified through forward genetics such as genetic map construction and
 34 quantitative trait locus (QTL) identification. The genes that located on the confidence intervals of the
 35 QTLs for powdery mildew resistance were considered as candidate genes/function genes (Fukino et al,
 36 2008; Natarajan et al, 2016; Wang et al, 2016). There were also some that identified candidate
 37 genes/function genes for powdery mildew resistance through reverse genetics such as using mutant
 38 (Chen et al, 2012, 2013 and 2015). From these researches, there were also some that suggested the
 39 melon that with thin peel often manifested as disease resistance while the melon that with thick peel
 40 often manifested as disease susceptible.
 41 In this study, one material GanTianmi with thin peel and another material XueLianHua with thick peel
 42 were selected. After inoculating the powdery mildew, both materials were used to do the RNA-Seq. In
 43 total, two RNA-seq libraries were constructed and sequenced separately; generating 47,462,015 mean
 44 clean reads with mean GC content of 49.32 and mean Q30 of 85.89%. Among all the genes in the
 45 genome of melon, 13828 genes were expressed in the material G and 13944 genes were expressed in
 46 the material S (RPKM>1). Through multiple comparisons between libraries, 769 DEGs were identified,
 47 of which the 726 were up-regulated DEGs and 43 were down-regulated ones. Based on the annotation
 48 information in gene ontology (GO), kyoto encyclopedia of genes and genomes (KEGG) and eukaryotic
 49 orthologous groups (KOG) database, most of the DEGs were related to “nucleus”, “chloroplast” and
 50 “cytoplasm” in GO database; related to “Metabolic pathways”, “Biosynthesis of secondary metabolites”
 51 and “Ribosome” in KEGG database and related to “Signal transduction mechanisms” “Posttranslational
 52 modification, protein turnover, chaperones” and “Lipid transport and metabolism” in KOG database.
 53 There were six genes that have annotation information of “Plant-pathogen interaction “. There were
 54 also 46 genes were TFs, among them, four were the *MYB* (Li et al, 2013; Wang et al, 2013; Yang et al,
 55 2019), six were the *ERF* (Xiang et al, 2018; Debbarma et al, 2019) and two were the *WRKY* (Line et al,

2011; Jiao et al, 2018; Yang et al, 2018);. This TFs were all related to disease resistance (Ling et al, 2010; Li et al, 2013; Wang et al, 2013; Jiao et al, 2018; Xiang et al, 2018; Yang et al, 2018; Debbarma et al, 2019; Yang et al, 2019). All this research could provide information for breeding muskmelon cultivars that were resistant to powdery mildew.

Materials and methods

Plant materials

There were two materials were used in this research. One material was named GanTianmi (G). This material was manifested as disease resistance with thin peel. The other material was named XueLianHua (H) was manifested as disease susceptible with thick peel.

RNA sample collecting and extraction

The cultivars were sown in a plastic nutrition pot after accelerated germination, and then place in a box under artificial climate, with a night temperature of 18 °C and a day temperature of 25 °C. When the seedlings reached the Three-leaf stage, the powdery mildew germ Race1 was collected using a single lorica (*Podosphaera xanthii*), which is widely distributed in the Gansu Province. The seedlings were then infected with the plant pathogen in a bioclean room by using the leaf inoculation method. Three days after inoculation, spots of powdery mildew on the blades of the seedlings were collected and used for RNA extraction. Total RNA was extracted by using the TRIzol reagent (Beijing Tiangen), following the manufacturer's instructions.

RNA library preparation, sequencing, and quality control

The RNA samples were used to generate sequencing libraries according to the manufacturer of Illumina TruSeq™ RNA Sample Preparation Kits (Illumina, San Diego, CA, USA) Transcriptome sequencing was carried out on an Illumina HiSeq 2500 platform that produced 150 base pair (bp) paired-end (PE) raw reads (Personalbio, Shanghai, China). Raw data (raw reads) in FASTQ format were processed used the trimmomatic software (Bolger *et al.*, 2014). In this step, clean data were obtained as follows: the first, removing reads with $\geq 10\%$ unidentified nucleotides (N); the second, Removing reads with $> 50\%$ bases having phred quality < 5 ; the third, Removing reads with > 10 nt aligned to the adapter, allowing $\leq 10\%$ mismatches; the fourth, removing putative PCR duplicates generated by PCR amplification in the library construction process. After trimmomatic, the quality of the clean data was checked by the software FASTQC

85 (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>).

86 **Read mapping and expression level analysis**

87 The clean sequence tags were mapped to the *Cucumis melo* (<http://cucurbitgenomics.org/organism/3>).

88 An index of the reference genome was built using Bowtie v2.0.6 (Langmead et al, 2012) and PE clean

89 reads were aligned to the reference genome using TopHat v2.0.9 (Kim et al, 2013). The reads per

90 RPKM values of genes were calculated using the software of cufflinks program (Trapnell et al., 2010).

91 After got the RPKM of the genes, the expression heatmap of the genes were drawn with the R package

92 Pheatmap with the value of log2 (RPKM+ 1).

93 **Gene expression pattern analysis**

94 The gene expression level was estimated by counting the number of reads that could be map to genes

95 or exons in the reference genome. To make gene expression data coincident and could be comparable

96 across different genes and experiments, the RPKM was used. The gene expression levels were

97 analyzed with software HTSeq (https://htseq.readthedocs.io/en/release_0.10.0/), model “union”.

98 Differential expression analysis of two genes was analyzed with the software DESeq in the R package.

99 This software could provide statistical routines to determine differential expression in digital gene

100 expression data using a model based on the negative binomial distribution The genes that with q_value

101 less than 0.05 and the log2 (Fold_change) (log2 (FC)) more than one or less than -1 could be

102 considered as DEGs. The genes that with log2 (FC) more than one were up-regulated genes while ones

103 with log2 (FC) less than negative one were down-regulated genes.

104 **Different expressed genes annotation and transcription factor analysis**

105 The GO annotation information of the candidate genes was got from the GO databases

106 (<http://archive.geneontology.org/latest-lite/>) and (<ftp://ftp.ncbi.nlm.nih.gov/gene/DATA/>). The software

107 BLAST was used to compare the sequence of the genes and the sequence in the GO database. The

108 KEGG annotation information was got through the software KOBAS 3.0 (Xie et al, 2010). The KOG

109 annotation information of candidate genes was got from the KOG database

110 (<ftp://ftp.ncbi.nih.gov/pub/COG/KOG>). The rules of comparison were same as it in GO analysis. All

111 the transcripts of melon were downloaded in the database of Plant TFDB

112 (<http://plantTFDB.cbi.pku.edu.cn/>) and then the DEGs were compared with these TFs with the NCBI

113 Blast. The sequence results with *E_value* less then 1e-10 could be considered as TFs.

114 **Results**

115 **Transcriptome sequencing and alignment to the reference genome**

116 With the purpose of systematic identification of the candidate genes that related to the difference in
117 disease resistance between the materials G and H, RNA-seq libraries of fiber samples were constructed
118 and analyzed. For the materials G, there were 56,645,955 raw reads, total 16,540,359,199bp, 49.24%
119 GC content and the Q30 was 86.02%; after trimmomatic, there were 48,482,165 clean reads, total
120 137692342bp. For the materials H, there were 54343040 raw reads, total 15,812,206,404bp, 49.4% GC
121 content and the Q30 was 49.40%; after trimmomatic, there were 46441864 clean reads, total
122 13,113,637,322bp (Table S1).

123 In the alignment analysis, the proportion of clean data was mapped to the genome of *Cucumis melo*.
124 For the materials G, there were 87,830,047 (90.58%) reads that could be uniquely mapped to the
125 genome of *Cucumis melo* and 87001260 of them could be mapped to genes (99.60% in exon). For the
126 materials H, there were 82,313,164 (88.62%) reads that could be uniquely mapped to the genome of
127 *Cucumis melo* and 80435970 of them could be mapped to genes (99.59% in exon) (Table S2).

128 **The expression level of the genes**

129 From the RNA sequencing, all the expression levels of the genes in the *Cucumis melo* genome were
130 calculated. For the materials G, total 7222 genes were not expressed as the RPKM of these genes were
131 zero; total 6377 genes were low expressed as the RPKM of these genes were between zero and one;
132 total 8944 genes were medium expressed as the RPKM of these genes were between one and ten; total
133 4884 genes were high expressed as the RPKM of these genes were more than 10; For the materials H,
134 total 7275 genes were not expressed; total 6237 genes were low expressed; total 9103 genes were
135 medium expressed a; total 4841 genes were high expressed (Table S3, Table 1, Figure 1).

136 For the no expressed genes there were 6549 genes (82.7%) that could be detected in both materials, 673
137 could be detected in only G and 696 could be detected in only H; for the weak expression genes, there
138 were 5083 (67.5%) could be detected in both materials, 1294 (17.2%) could be detected in only G and
139 1154 (15.3) could be detected in only H; for the medium expressed genes, there were 7932 could be
140 detected in both materials, 1012 (10.0%) could be detected in only G and 1171 (11.6%) could be
141 detected in only H; for the high expressed genes, there were 4391 (82.3%) could be detect in both
142 materials, 493 could be detected in only G and 451 could be only detected in H (Table S3, Table 1,
143 Figure 1).

144 **The different expression genes**

145 To find out the different expression genes between the two materials G and H, the software
146 Deseq2 were used and the genes with the absolute value of log2FC more than one and Q_value
147 less than 0.05 could be considered as different expression genes. In this research, there were 769
148 different expression genes. Among them, 726 were up-regulation genes and the other 43 were
149 down-regulation genes (Table S4, Figure 2).

150 **The annotation with different expression gene**

151 All the different expression genes were annotated with GO, KEEG and KOG. For the GO
152 annotation, all the DEGs could be annotated into 968 GO terms as one DEG maybe annotated into
153 more than one GO term. Among the 968 GO terms, 326 belonged to the category of molecular
154 function; 146 belonged to the category of cellular component and 496 belonged to the category of
155 biological process. The GO term nucleus harbored the most number of DEGs (143), and followed
156 with the GO term molecular function, biological process and cytoplasm, harbored 125 116 and
157 114 DEGs respectively (Table S5, Figure 2). For the KEEG analysis, all the DEGs were annotated
158 into 76 KEEG pathways. Among them, “Metabolic pathways”, “Biosynthesis of secondary
159 metabolite” and “Ribosome” harbored the most DEGs (76, 53 and 14) respectively (Table S6,
160 Figure 2); For the KOG analysis, all the DEGs were located into 21 KOG baskets. Among them,
161 the basket “Signal transduction mechanisms”, “Lipid transport and metabolism” and “Energy
162 production and conversion”harbored the most DEGs (67, 50 and 42) respectively. There were also
163 153 DEGs that with unclear KOG annotation information as 109 of them located on the KOG
164 basket of “General function prediction only” and the other 44 located on the KOG basket of
165 “Function unknown” (Table S7, Figure 2).

166 **Transcription factor analysis**

167 ALL The DEGs were compared with the TFs in the database of (<http://planttfdb.cbi.pku.edu.cn/>).
168 Totally, 46 DEGs could be considered as TFs. Among them, there were six DEGs that were *ERF*
169 TF, four were *MYB* TF and two were *WARK* TF. The detail information about TF analysis was in
170 the Figure 2 and Table S8. The TFs *ERF*, *MYB*, and *WARK* were related to disease resistance
171 (Table S8, Figure 2).

172 **Discussion**

173 From the annotation information of the DEGs, there were six genes that annotated with the KEEG

174 pathway “Plant-pathogen interaction”. The six DEGs were *MELO3C016746*, *MELO3C015295*,
175 *MELO3C021706*, *MELO3C018539*, *MELO3C020725* and *MELO3C000209*. Among them, the
176 DEG *MELO3C018539* have been published that it played the role of against *cucurbit chlorotic*
177 *yellow virus* in melon (Takeshita et al, 2013), but the other five DEGs have not been published in
178 melon. In the *Arabidopsis thaliana* annotation database, the homologous gene of *MELO3C016746*
179 in *Arabidopsis thaliana* (*AT1G29150*) was specifically interacted with FUS6/COP11 via the
180 C-terminal domain of *FUS6/COP11* and associates with an ATPase subunit of the 19S proteasome
181 regulatory complex, AtS6A. The mRNA is cell-to-cell mobile (Cho et al, 2006 and 2015;); the
182 homologous gene of *MELO3C015295* in *Arabidopsis thaliana* (*AT5G56010*) was a member of
183 heat shock protein 90 (HSP90) gene family and expressed in all tissues and abundant in root apical
184 meristem, pollen and tapetum (Cha et al, 2013; Bologna et al, 2018); the homologous gene of
185 *MELO3C021706* in *Arabidopsis thaliana* (*AT3G18990*) was related to regulation of flower
186 development and vernalization response (Kinjo et al, 2012; King et al, 2013); the homologous
187 gene of *MELO3C020725* in *Arabidopsis thaliana* (*AT2G37650*) was belong to GRAS family
188 transcription factor and it related to biological processes in regulating root and shoot development,
189 plant growth, development and stress responses (Liu et al, 2018; Zhang et al, 2018); the homologous
190 gene of *MELO3C000209* in *Arabidopsis thaliana* (*AT5G56030*) was also a member of heat shock
191 protein 90 (HSP90) gene family and the function was similar with the gene *MELO3C015295* (Cha
192 et al, 2013; Bologna et al, 2018).

193 There were also some DEGs that were TFs that related to disease resistance like *MYB* (Katlyar et
194 al, 2012; Li et al, 2013; Wang et al, 2013; Yang et al, 2019), *ERF* (Xiang et al, 2018; Debbarma et al,
195 2019) and *WRKY* (Line et al, 2011; Jiao et al, 2018; Yang et al, 2018). In plants, MYB TFs were
196 related to plant development, secondary metabolism, hormone signal transduction, disease
197 resistance and abiotic stress tolerance (Katlyar et al, 2012); the ERF TFs plays an important role
198 in plant development processes and stress responses such as salt, drought, heat, and cold with
199 well-conserved DNA-binding domain (Xiang et al, 2018; Debbarma et al, 2019); The *WRKY* TFs
200 were widely implicated in defense responses and various other physiological processes (Yang et al,
201 2018); it could evolve diverse functions with different biochemical properties and implicated to
202 modulate seed development, flowering, fruit ripening, senescence and various metabolic processes
203 (Jiao et al, 2018).

204 In our research, these DEGs were related to disease resistance and could be considered as
205 candidate genes. But whether these DEGs could make some contributions to increase the disease
206 resistance of melon, it still need further studies like transgene.

207 Acknowledgements

208 The National Natural Science Foundation of China (31471896 , 31360434) 、 Technology Support
209 Project of the GAAS (2017GAAS48) and The Major Project of Gansu Science & Technology
210 (17ZD2NA015-04) supported this study.

211

212 Conflict of interest

213 The authors declare no competing interests.

214

215

216

217

218 Reference

219 Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence
220 data. *Bioinformatics*. 30(15):2114-20

221 Bologna N, Lselin R, Abriata LA, Sarazin A, Pumplin N, Jay F, Grentzinger T, Peraro MD,
222 Voinnet O (2018). Nucleo-cytosolic Shuttling of ARGONAUTE1 Prompts a Revised Model of the
223 Plant MicroRNA Pathway. *Molecular Cell*. 69(4), 709-719

224 Cha JY, Ahn G, Kim JY, Kang SB, Kim MR, Su'udi M, Kim WY, Son DY (2013) Structural and
225 functional differences of cytosolic 90-kDa heat-shock proteins (Hsp90s) in *Arabidopsis thaliana*.
226 *Plant Physiology and Biochemistry*. 70: 368-373

227 Chen H, Kun WP, Liu DS, Su, YQ, He QW (2012). Molecular cloning and expression analysis of
228 CmMlo1 in melon. *Mol Biol Rep*. 39:1903–1907

229 Chen H, Kong WP, Hou D, Lv JF, Tao XL (2013) Isolation, characterization, and expression
230 analysis of CmMLO2 in muskmelon. *Mol Biol Rep*. 40:2609–2615

231 Chen H, Kong WP, Lv JF, Li JP (2015). Analysis of Powdery Mildew Resistance in Wild Melon
232 MLO Mutants. *Horticultural Plant Journal*, 1 (3): 174–180

233 Cheng, Zh.J., Wang, H.S., Zhang, Zh.B., Guo, Sh.R., 2006. Research on the genetic mechanisms
234 of resistance to powdery mildew melon. Jiangsu Agricultural Sciences, (6): 224–225.

235 Cho SK, Chuang SH, Ryu MY, Park MJ, Lee MM, Bahk Y, Kim J, Pai HS, Kim TW (2006)
236 Heterologous Expression and Molecular and Cellular Characterization of CaPUB1 Encoding a
237 Hot Pepper U-Box E3 Ubiquitin Ligase Homolog. Plant Physiology. 142: 1664–1682

238 Cho SK, Bae H, Ryu MY, Yang SW, Kim WT (2015). PUB22 and PUB23 U-BOX E3 ligases
239 directly ubiquitinate RPN6, a 26S proteasome lid subunit, for subsequent degradation in
240 *Arabidopsis thaliana*. Biochemical and Biophysical Research Communications 464: 994-999

241 Debbarma J, Sarki YN, Saikia B, Boruah HPD, Singha DL, Chikkaputtaiah C (2019) Ethylene
242 Response Factor (ERF) Family Proteins in Abiotic Stresses and CRISPR-Cas9 Genome Editing of
243 ERFs for Multiple Abiotic Stress Tolerance in Crop Plants: A Review. Mol Biotechnol. doi:
244 10.1007/s12033-018-0144-x.

245 Fukino N, Ohara T, Monforte AJ, Sugiyama M, Sakata Y, Kuniyama M, Matsumoto S
246 (2008) .Identification of QTLs for resistance to powdery mildew and SSR markers diagnostic for
247 powdery mildew resistance genes in melon (*Cucumis melo* L.). Theor Appl Genet. 118(1):165-75.

248 Jiao ZG, Sun JL, Wang CQ, Dong YM, Xiao SH, Gao XL, Cao QW, Li LB, LI WD, Gao C (2018)
249 Genome-wide characterization, evolutionary analysis of WRKY genes in Cucurbitaceae species
250 and assessment of its roles in resisting to powdery mildew disease. Plos One, 13(12): e0199851

251 Katlyar A, Smita S, Lenka SK, Rajwanshi R, Chinnusamy V, Bansal KC (2012) Genome-wide
252 classification and expression analysis of *MYB* transcription factor families in rice and *Arabidopsis*.
253 BMC Genomics. 13:544

254 Kinjo H, Shitsukawa N, Takumi S, Murai K (2012) Diversification of three
255 APETALA1/FRUITFULL-like genes in wheat. Mol Genet Genomics. 287:283–294

256 Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. (2013). TopHat2: accurate
257 alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome
258 Biology. 14: R36.

259 King GJ, Chanson AH, McCallum EJ, Ohme-Takag M, Byriel K, Hill JM, Martin JL, Mylne JS
260 (2013). The Arabidopsis B3 Domain Protein VERNALIZATION1 (VRN1) Is Involved in
261 Processes Essential for Development, with Structural and Mutational Studies Revealing Its
262 DNA-binding Surface. J Biol Chem. 288(5): 3198–3207

263 Langmead B and Salzberg S (2012). Fast gapped-read alignment with Bowtie 2. Nature Methods.
264 9: 357-359.

265 Li YH, Wen CL, Weng YQ (2013) Fine mapping of the pleiotropic locus B for black spine and
266 orange mature fruit color in cucumber identifies a 50 kb region containing a R2R3-MYB
267 transcription factor. Theor Appl Genet. 126:2187–2196

268 Ling J, Jiang WJ, Zhang Y, Yu HJ, Mao ZC, Gu XF, Huang SW, Xie BY (2010) Genome-wide
269 analysis of WRKY gene family in *Cucumis sativus*. BMC Genomics. 12:471

270 Liu B, Sun Y, Xue J, Jia X, Li R. Genome-wide characterization and expression analysis
271 of GRAS gene family in pepper (*Capsicum annuum* L.). Peer J. 29;6:e4796

272 Natarajan S, Kim HT, Thamilarasan SK, Veerappan K, Park JI, Nou IS (2016). Whole Genome
273 Re-Sequencing and Characterization of Powdery Mildew Disease-Associated Allelic Variation in
274 Melon. PLoS One. 16;11(6):e0157524

275 Liu JP. 2003. Suggestions of production status and development on watermelon and melon.
276 Chinese Watermelon and melon, (4): 1–3.

277 Takeshita M, Okuda M, Okuda S, Hyodo A, Hamano K, Furuya N, Tsuchiya K (2013) Induction
278 of Antiviral Responses by Acibenzolar-S-Methyl Against Cucurbit chlorotic yellows virus in
279 Melon. Phytopathology. 103(9): 960-5

280 Trapnell C, Williams B A, Pertea G, et al. Transcript assembly and quantification by RNA-Seq reveals
281 unannotated transcripts and isoform switching during cell differentiation. Nature Biotechnology, 2010,
282 28(5):511-5.

283 Wang XM, Zhou J, Yang Y, Yu FB, Chen J, Yu CL, Wang F, Cheng Y, Yan CQ, Chen JP. (2013)
284 Transcriptome Analysis of a Progeny of Somatic Hybrids of Cultivated Rice (*Oryza sativa* L.) and
285 Wild Rice (*Oryza meyeriana* L.) With High Resistance to Bacterial Blight. J Phytopathol
286 161:324–334

287 Wang YH, Wu DH, Huang JH, Tsao SJ, Hwu KK, Lo HF (2016). Mapping quantitative trait loci
288 for fruit traits and powdery mildew resistance in melon (*Cucumis melo*). Botanical Studies. 57:19

289 Xiang LX, Liu C, Luo JZ, He L, Deng YS, Yuan J, Wu CF, Cai YF (2018) A tuber mustard
290 AP2/ERF transcription factor gene, BjABR1, functioning in abscisic acid and abiotic stress
291 responses, and evolutionary trajectory of the ABR1 homologous genes in Brassica species. Peer J
292 6:e6071

293 Xie, C., Mao, X.Z., Huang, J.J., Ding, Y., Wu, J.M., Dong, S., Kong, L. et al. (2011) KOBAS 2.0:
294 a web server for annotation and identification of enriched pathways and diseases. *Nucleic Acids*
295 *Research*. 39, 316-322

296 Yang KB, Li Y, Wang SN, Xu XR, Sun HY, Zhao HS, Li XP, Gao ZM (2019). Genome-wide
297 identification and expression analysis of the MYB transcription factor in moso bamboo
298 (*Phyllostachys edulis*). *Peer J*. DOI 10.7717/peerj.6242

299 Yang XZ, Li H, Yang YC, Wang YQ, Mo YL, Zhang RM, Zhang Y, Ma JX, Wei CH, Zhang X
300 (2018) Identification and expression analyses of WRKY genes reveal their involvement in growth
301 and abiotic stress response in watermelon (*Citrullus lanatus*). *Plos one*. 13(1): e0191308

302 Zhang B, Liu J, Yang ZE, Chen EY, Zhang CJ, Zhang XY, Li FG (2018) Genome-wide analysis
303 of GRAS transcription factor gene family in *Gossypium hirsutum* L. *BMC Genomics*. 19(1):348
304
305

306 Table 1. The Basic statistics of RPKM_value of the genes in two materials

RPKM Value	G	H
RPKM=0	7222	7245
0<RPKM<1	6377	6237
1<RPKM<10	8944	9103
RPKM>10	1884	4842

307 Table 2. The number of TFs in the DEGs

Transcription factor	No.Gene
B3	2
bHLH	5
bZIP	1
C2H2	5
C3H	1
CO-like	1
DBB	4
ERF	6
G2-like	1
GATA	1
LBD	2
LSD	4
MYB	4
NAC	3
Nin-like	1
SBP	2
TALE	1

308

309

310 **Supporting information**

311 Table S1. The detail information of RNA-Seq quality evaluate

312 Table S2. The detail information of comparison efficiency evaluate

313 Table S3. The RPKM value of all the genes in the material G and H

314 Table S4. The log2FoldChange value and FDR value of the DEGs between the materials G and H

315 Table S5. The detail GO annotation information of DESs between the materials G and H

316 Table S6. The detail KEGG annotation information of DESs between the materials G and H

317 Table S7. The detail KOG annotation information of DESs between the materials G and H

318 Table S8. The detail TF annotation information of DEGs between the materials G and H

319

320

321 **Figure legends**

322 **Figure 1**

323 a. The number of genes that with RPKM intervals of RPKM=0, $0 < \text{RPKM} < 1$, $1 < \text{RPKM} < 10$,

324 $\text{RPKM} > 10$

325 b. The intersection and union of the genes that with RPKM intervals of RPKM=0, $0 < \text{RPKM} < 1$,

326 $1 < \text{RPKM} < 10$, $\text{RPKM} > 10$ between the material G and H

327 c. The log2FoldChange value and FDR value of the positive regulation DEGs and negative

328 regulation DEGs

329 **Figure 2. The annotation information of the DEGs between the materials G and H**

330 a. The GO annotation information of the DEGs between the materials G and H

331 b. The KEGG annotation information of the DEGs between the materials G and H

332 c. The KOG annotation information of the DEGs between the materials G and H

333 d. The TF annotation information of the DEGs between the materials G and H

334

a

b

c

a

b

c

d