Genome-wide identification of candidate aquaporins involved in water accumulation of pomegranate outer seed coat

Jianjian Liu [Equal first author, 1, 2] , Gaihua Qin [Equal first author, 2, 3] , Chunyan Liu [2, 3] , Xiuli Liu [4] , Jie Zhou [4] , Jiyu Li [2, 3] , Bingxin Lu [1] , Jianrong Zhao [Corresp. 1]

1 College of Resource and Environment, Anhui Science and Technology University, Fengyang, China

2 Institute of Horticultural Research(Key Laboratory of Genetic Improvement and Ecophysiology of Horticultural Crop, Anhui Province), Anhui Academy of Agricultural Sciences, Hefei, China

3 Key Laboratory of Fruit Quality and Developmental Biology, Anhui Academy of Agricultural Sciences, Hefei, China

4 State Key Laboratory for Managing Biotic and Chemical Threats to the Quality and Safety of Agro-products, Zhejiang Academy of Agricultural Sciences, Hangzhou, China

Corresponding Author: Jianrong Zhao
Email address: zhaojr@ahstu.edu.cn

Aquaporins (AQPs) are a class of highly conserved integral membrane proteins that facilitate the uptake and transport of water and other small molecules across cell membranes. However, little is known about AQP genes in pomegranate (Punica granatum L.) and their potential role in water accumulation of the outer seed coat. We identified 38 PgrAQP genes in the pomegranate genome and divided them into five subfamilies based on a comparative analysis. Purifying selection played a role in the evolution of PgrAQP genes and a whole-genome duplication event in Myrtales may have contributed to the expansion of PgrTIP, PgrSIP, and PgrXIP genes. Transcriptome data analysis revealed that the PgrAQP genes exhibited different tissue-specific expression patterns. Among them, the transcript abundance of *PgrPIPs* were significantly higher than that of other subfamilies. The mRNA transcription levels of *PgrPIP1.3*, *PgrPIP2.8*, and *PgrSIP1.2* showed a significant linear relationship with water accumulation in seed coats, indicating that PgrPIP1.3/PgrPIP2.8 located in the plasma membrane and PgrSIP1.2 proteins located on the tonoplast may be involved in water accumulation and contribute to the cell expansion of the outer seed coat, which then develops into juicy edible flesh. Overall, our results provided not only information on the characteristics and evolution of PgrAQPs, but also insights on the genetic improvement of outer seed coats.

1 **Title:** Genome-wide identification of candidate aquaporins involved in water

2 accumulation of pomegranate outer seed coat

3 Jianjian Liu[1,2 †], Gaihua Qin[2,3, †], Chunyan Liu[2,3], Xiuli Liu[4], Jie Zhou[4], Jiyu Li[2,3],Bingxin Lu[1], Jianrong

4 Zhao[1, *]

5 1. College of Resource and Environment, Anhui Science and Technology University, Fengyang, China.

6 2. Institute of Horticultural Research, Anhui Academy of Agricultural Sciences (Key Laboratory of

7 Genetic Improvement and Ecophysiology of Horticultural Crop, Anhui Province), Hefei, China.

8 3. Key Laboratory of Fruit Quality and Developmental Biology, Anhui Academy of Agricultural Sciences,

9 Hefei, China.

10 4. State Key Laboratory for Managing Biotic and Chemical Threats to the Quality and Safety of Agro-

11 products, Zhejiang Academy of Agricultural Sciences, Hangzhou, China.

12 [†] These authors have contributed equally to this work and share first authorship.

13 [*] **Corresponding author:**

14 Jianrong Zhao, College of Resource and Environment, Anhui Science and Technology University,

15 Fengyang, China.

16 E-mail: zhaojr@ahstu.edu.cn

Abstract

19 Aquaporins (AQPs) are a class of highly conserved integral membrane proteins that

20 facilitate the uptake and transport of water and other small molecules across cell

21 membranes. However, little is known about AQP genes in pomegranate (*Punica*

22 *granatum* L.) and their potential role in water accumulation of the outer seed coat. We

23   identified 38 *PgrAQP* genes in the pomegranate genome and divided them into five

24   subfamilies based on a comparative analysis. Purifying selection played a role in the

25   evolution of *PgrAQP* genes and a whole-genome duplication event in Myrtales may

26   have contributed to the expansion of *PgrTIP*, *PgrSIP*, and *PgrXIP* genes. Transcriptome

27   data analysis revealed that the *PgrAQP* genes exhibited different tissue-specific

28   expression patterns. Among them, the transcript abundance of *PgrPIPs* were

29   significantly higher than that of other subfamilies. The mRNA transcription levels of

30   *PgrPIP1.3*, *PgrPIP2.8*, and *PgrSIP1.2* showed a significant linear relationship with water

31   accumulation in seed coats, indicating that PgrPIP1.3/PgrPIP2.8 located in the plasma

32   membrane and PgrSIP1.2 proteins located on the tonoplast may be involved in water

33   accumulation and contribute to the cell expansion of the outer seed coat, which then

34   develops into juicy edible flesh. Overall, our results provided not only information on

35   the characteristics and evolution of *PgrAQPs*, but also insights on the genetic

36   improvement of outer seed coats.

38   **Introduction**

39   Pomegranate (*Punica granatum* L.) is an important economic fruit tree species due

40   to its functional and nutraceutical properties, and it is widely consumed as a fruit, juice,

41   wine, and medicine (Johanningsmeier & Harris 2011; Patel et al. 2008). Pomegranate is

42   native to Iran, the northern India side of the Himalayan Mountains, and is widely

43   planted in Mediterranean-like climates around the world, including Tunisia, Turkey,

44   Spain, Egypt, Iran, Morocco, the USA, China, India, Argentina, Israel, and South Africa

45   (Qin et al. 2017). The size of the juicy outer seed coat determines the edible quality of

46   the fruit, so it has become an important fruit characteristic. Notably, the morphological

47   characteristics of the pomegranate seed showed a compressed inner seed coat and an

48   expanded outer seed coat, making it an appealing model for studying development of

49   seed coats (Luo et al. 2020; Niu et al. 2018; Qin et al. 2020).

50   The development of the seed coat is accompanied by the import of organic and

51   inorganic nutrients, including sugars, organic acids, cellulose, and water in developing

52   seeds (Qin et al. 2020; Uçar & Karagöz 2009; Zarei et al. 2016; ZHOU et al. 2007b). The

53   symplastic pathway is one of the main pathways by which water moves, which is

54   mediated by integral membrane proteins called aquaporins (AQPs), a major intrinsic

55   protein family (Adams & Wendel 2005; Lian et al. 2004; Suga et al. 2002). AQPs have

56   been shown to transport small molecules such as carbon dioxide, glycerol, ammonia,

57   urea, hydrogen peroxide, and selenite (Ivanov et al. 2007; Yasui et al. 1999; Zwiazek et

58   al. 2017a).

59   The general AQP structure is highly conserved in plants, and it is predicted to

60   consist of six transmembrane (TM) α-helices (H1 to H6) and two half-helices

61   (Deshmukh et al. 2015; Lee et al. 2005; Tornroth-Horsefield et al. 2006). These

62   transmembrane α-helices are linked by five short loops (Loops A to E), Loops B and E

63   contain the signature sequence Asparagine-Proline-Alanine (NPA) motif, which has the

64   primary function of forming water-selective channels (Chaumont et al. 2001; Gupta &

65 Sankararamakrishnan 2009; Wallace & Roberts 2004). Another important secondary

66 structure known as an aromatic/arginine (ar/R) selectivity filter is formed by four

67 residues contributed by transmembrane helices H2/H5 and the loop LE (Azad et al.

68 2016; Deshmukh et al. 2015; Hove & Bhave 2011; Tornroth-Horsefield et al. 2006). These

69 two structures predominantly determine the specificity of solute transport and

70 transport rate (Deshmukh et al. 2015; Lee et al. 2005; Tornroth-Horsefield et al. 2006).

71 According to sequence similarity and protein subcellular localization, AQPs in higher

72 plant can be classified into five distinct subfamilies: plasma membrane intrinsic proteins

73 (PIPs), tonoplast intrinsic proteins (TIPs), nodulin 26-like intrinsic proteins (NIPs)

74 (Pommerrenig et al. 2015), small basic intrinsic proteins (SIPs), and unrecognized (or X)

75 intrinsic proteins (XIPs)(Chaumont et al. 2001; Danielson & Johanson 2010; Kaldenhoff

76 & Fischer 2006; Khabudaev et al. 2014; Quigley et al. 2002).

77   The AQP gene family has been widely studied in numerous plant species, such as

78 *Arabidopsis thaliana* (Quigley et al. 2002), *Oryza sativa* (Sakurai et al. 2005), *Glycine max*

79 (Zhang et al. 2013), cotton (Li et al. 2019), *Populus trichocarpa* (Gupta &

80 Sankararamakrishnan 2009), and grape (Fouquet et al. 2008), by blasting whole genome

81 sequences. Previous studies have demonstrated that AQPs in higher plants play

82 important roles in various physiological and development processes, such as cell and

83 tissue expansion, fiber development, flower pollination, and seed development (Azad et

84 al. 2004; Eisenbarth & Weig 2005; Gattolin et al. 2011; Soto et al. 2008; Soto et al. 2010;

85 Vander Willigen et al. 2006; Wudick et al. 2014). For instance, in French bean (*Phaseolus*

86   *vulgaris*), *PvPIPs* played a role in the import of water and nutrients by phloem-mediated

87   transport and water recycling in the xylem in developing seeds (Terashima & Ono 2002).

88   In rice, OsPIP1;1 and OsPIP1;3 functions as water channels. Over-expression of

89   *OsPIP1;1* could increase rice yield and seed germination. Similarly, the overexpression

90   of *OsPIP1;3* promoted the speed of seed germination under drought conditions (Liu et

91   al. 2013; Liu et al. 2007). In *Pisum sativum* L., *PsNIP1* showed high expression level in

92   developing pea seed coats, and the overexpression of *PsNIP1* increased the uptake of

93   water and glycerol (Schuurmans et al. 2003; Zhou et al. 2007a). Interestingly, the highly

94   correlations between the expression of *AtTIP3*/*AtTIP1* and seed germination stages (a

95   rapid imbibition of desiccated tissues and embryo growth, respectively) was observed,

96   which also provided insight into the influence of AQPs on the seed developmental

97   process (Gattolin et al. 2011). However, little is known about the role of AQPs on seed

98   coat development.

99   In this study, we identified 38 *PgrAQP* genes in the pomegranate genome, then

100 conducted gene structure, phylogenetics, and evolutionary divergence analyses. The

101 expression patterns of *PgrAQP* genes in different tissues and seed developmental stages

102 were analyzed,and the potential function of PgrPIP genes in response to drought stress

103 was also identified. Furthermore, the candidate genes contributing to the water

104 accumulation in the seed coat were selected based on a correlation analysis of gene

105 expression and water accumulation in seed coats. Our findings provided fundamental

106 information about the gene structure, phylogenetics, and evolutionary divergence of

107 *PgrAQP*. In addition, this study also provided useful information for further

108 investigation of the molecular mechanism by which PgrAQP functions in seed coat

109 development.

110 **2. Materials and Methods**

111 **2.1 Identification of the *PgrAQP* genes in the pomegranate genome**

112   The whole genome sequences of pomegranate were downloaded from NCBI

113 Genome (https://www.ncbi.nlm.nih.gov/genome/?term=Punica+granatum+L). Predicted

114 protein sequences were obtained using TBtools with the following sets: Sequence

115 Toolkit: Batch translate CDS to protein (Chen et al. 2020). The amino acid sequences of

116 the AQPs in *Arabidopsis* and *Populus trichocarpa* were employed as queries to blast

117 searches against the whole-genome sequences in pomegranate using TBtools with a

118 query over 50% and e-value less than 0.01 (Quigley et al. 2002). The candidate

119 sequences were submitted to NCBI for EST blast searches. Finally, the AQP protein

120 sequences of pomegranate were examined to verify the presence of the characteristic

121 MIP and transmembrane helical domains using the SMART program

122 (http://smart.embl-heidelberg.de/) and TMHMM

123 (http://www.cbs.dtu.dk/services/TMHMM/) (Finn et al. 2014; Letunic et al. 2020). The

124 information of PgrAQP gene family (protein length, molecular weight, and isoelectric

125 point) were determined using Expasy (https://web.expasy.org/compute_pi/) (Artimo et

126 al. 2012). The subcellular localization of the PgrAQP proteins was predicted by WoLF

127   PSORT (http://www.genscript.com/psort/wolf_ psort.html) Cell-PLoc 2.0

128   (http://www.csbio.sjtu.edu.cn/bioinf/Cell-PLoc-2/)

129   **2.2 Sequence alignment of *PgrAQP* genes, phylogenetic analysis, and classification**

130   The predicted plant AQP protein sequences were aligned using the ClustalW tool

131   in MEGA7.0 (Kumar et al. 2016). *Arabidopsis thaliana*, grape (*Vitis vinifera*), *Populus*

132   *trichocarpa* and eucalyptus (*Eucalyptus grandis*) AQP protein sequences were collected

133   from NCBI (https://www.ncbi.nlm.nih.gov/). A phylogenic tree was constructed with

134   MEGA7.0 using the neighbor-joining method and bootstrap parameter with 1,000

135   replicates. For this study, Pomegranate AQPs were named according to the sequence

136   homology and phylogenetic relationships with *Populus trichocarpa* . According to the

137   name of the best hit gene in *Populus trichocarpa*, the subfamily classification (PIP, NIP,

138   SIP, TIP, and XIP) and corresponding names of AQPs are shown in Fig. 1. TBtools

139   software was used to determine the localization of *PgrAQP* genes on pomegranate

140   chromosomes.

141   **2.3 Gene structure and conserved motif analysis of *PgrAQPs***

142   The conserved motifs in the proteins were identified using Multiple Expectation

143   Maximization for Motif Elucidation (MEME v4.12.0, http://meme-suite.org/tools/meme)

144   with the following parameters: maximum number of motifs, 10; width of optimum

145   motif, ≥15 and ≤50 (Bailey et al. 2009). The gene exon–intron structures of AQPs were

146   confirmed using the GSDS database by aligning the open reading frame (ORF) with

147   their corresponding genomic sequences (Hu et al. 2015a). Sequences representing NPA

148   motifs, ar/R filter, and Froger's residue positions were manually identified based on

149   multiple sequence alignments of pomegranate AQPs with heterologous AQPs of

150   *Arabidopsis* (Kumar et al. 2016). TBtools software was used to construct a chromosome

151   localization map of AQP family genes.

152   **2.4 Gene duplication and synteny analysis of the *PgrAQP* gene family**

153   Gene duplication events of *PgrAQP* genes were analyzed using TBtools with the

154   following parameters: Blast Compare Two Seqs, Outfml: Table, NumofThreads: 2, E-

155   value:1e-5, NumofHits: 5, NumofAligns: 5; File Merger For MCScanX, Merge Mode:

156   GtfGff2 SimGxf (Chen et al. 2020). Duplication types were divided into whole-genome

157   (WGD), segmental, and tandem duplications (Qiao et al. 2015). Tandem duplicated

158   genes were defined as two homologous genes within a range of 100-kb and sequence

159   alignment coverage over 75% (Gu et al. 2002; Wang et al. 2010; Yang et al. 2008). The

160   nonsynonymous substitution ratios (*Ka*), synonymous substitution ratios (*Ks*), and

161   *Ka*/*Ks* ratios of the *PgrAQP* family genes were calculated using the TBtools simple *Ka*/*Ks*

162   calculator program, with the following parameters: Simple Ka/Ks Calculator (NG). The

163   *Ks* value was used to calculate the divergence time of duplication events (T=*Ks*/2R Mya,

164   Millions of years), where R is the rate of synonymous substitutions, R =1.5 × 10^{-8}

165   substitutions per synonymous site per year for dicotyledonous plants and T refers to

166   divergence time (Edlund et al. 2004). The microsyntenic relationship of AQP genes

167   between pomegranate, grape, *Arabidopsis*, and eucalyptus was determined using

168   TBtools Dual Systeny Plot for MCscanX program (Chen et al. 2020).

169   **2.5 Plant materials and analysis of transcript profiles of *PgrAQP* genes**

170   Two major pomegranate cultivars, 'Dabenzi' (a hard-seeded cultivar) and 'Tunisia'

171   (a soft-seeded cultivar), were planted in Anhui Province (Hefei, 31°51′9.05″N,

172   117°06′34.33″E), China and grown under the same fertilization and irrigation conditions.

173   Seeds from five fruits of the pomegranate cultivar 'Dabenzi' and 'Tunisia' were

174   collected at 50, 95, and 140 days after pollination (DAP). For outer seed coats, the seeds

175   from fruits collected 95 DAP and 140 DAP was used because it was difficult to visually

176   distinguish the inner and outer seed coats of seeds were collected at 50 DAP. Three

177   biological replicates were taken for RNA sequencing. For each treatment, the fresh

178   weights of the total inner and outer seed coats were recorded and immediately frozen in

179   liquid nitrogen. Part of the inner and outer seed coats were stored at -80°C until they

180   were used for transcriptome sequencing. RNA extraction was performed as described

181   previously (Qin et al. 2020). Briefly, the total RNA was isolated using a Plant RNeasy

182   Mini Elute Cleanup kit (Qiagen, Beijing, China) and the sequencing libraries were

183   prepared using the NEBNext Ultra™ RNA Library Prep Kit for Illumina (New England

184   Biolabs, USA) following the manufacturer's protocols. Transcriptome sequencing was

185   conducted using an Illumina HiSeq 2000 platform.

186   The remaining samples were used for the measurement of water accumulation. The

187   water accumulation (%) in the different seed coats was determined by the fresh weight

188   (FW) and drought weight (DW) using the formula (FW-DW)/FW×100. The freeze-

189   drying analysis was performed using GOLD-SIM FD83 (SIM International group co. ltd,

OFF
OFF

190   USA). The experiments were repeated three times.

191   The abundances of pomegranate AQP transcripts, in the root, flower, leaf, and three

192   developmental stages of the peel and seed coat (inner and outer seed coat), were

193   collected from the microarray data previously published by Qin et al., (Qin et al. 2017).

194   Transcriptional abundances of AQP genes were estimated using the fragments per

195   kilobase of exon per million mapped reads (FRKM) method, and a heat map was

196   generated based on the $\log_2$ FRKM transformation value using TBtools (Hu et al. 2018),

197   with the following protocol: Starting from the "Graphics", click "Heatmap lllustrator",

198   select the Heatmap button, then set input files in each filed, click "Start" and graph will

199   be generated.

200   2.6 **Verification of PgrAQP genes function responding to water deficit in**

201   **pomegranate root**

202   For the analysis of PgrAQP genes expression patterns in response to water deficit,

203   the PEG treatment experiment was performed. For the hydroponic culture experiment,

204   the full Hoagland nutrient solution contained 20% (v/v) polyethylene glycol 6000

205   (PEG6000). The solution pH was adjusted to 5.5. The roots samples were collected after

206   0 h, 1 h, 6 h, 12 h and 24 h of treatment, and were quickly frozen in liquid nitrogen and

207   stored at -80 for subsequent RNA isolation.

208   For performing qRT-PCR analysis, approximately 2 ug of NDA-eliminated total

209   RNA from pomegranate roots were used to synthesize cDNA using a reverse

210   transcription kit (TaKaRa). qRT-PCR was performed on the Applied Biosystems (ABI)

211   StepOne Plus PCR system (Applied Biosystems) using the LightCylcer 96 SYB GREEN I

212   Master (Roche, Indianapolis, IN, USA) in a 20 ul reaction solution. The PgrAQP genes

213   that showed higher expressed in the roots from the RNA-Seq data were selected for

214   qRT-PCR analysis. The relative transcript abundance of each gene was normalized to

215   the pomegranate actin (OWM91407) with the cycle threshold (Ct) $2^{(-\Delta\Delta Ct)}$ method. Three

216   biological replicates and technical replicates were used for each gene. All the specific

217   primers used for each target gene are listed in Table. S1.

218   **2.7 Statistical analysis**

219   The data were analyzed by ANOVA (SPSS 16.0; SPSS Inc., Chicago, IL, USA),

220   followed by Turkey's test (P<0.05) to determine differences of inner and outer seed coats.

221   The data represent the mean $\pm$ SE of three independent biological replicates.

222   **3. Results**

223   **3.1 Genome-wide identification of PgrAQP genes in pomegranate reveal each**

224   **Aquaporin subfamily presents particular physicochemical characteristics**

225   A total of 38 *PgrAQP* members were identified in the pomegranate genome via a

226   genome-wide search using the AQP protein sequences in *Arabidopsis* and *Populus*

227   *trichocarpa* as queries (Quigley et al. 2002). A subsequent conserved domain analysis

228   also confirmed all of the predicted AQPs (**Fig. S1**). The characteristics of *PgrAQP* family

229   genes are shown in **Table. S2**, including the gene ID, protein length, relative molecular

230   weight, transmembrane domains (TMDs), subcellular localization, and calculated

231   isoelectric point (pI). The 38 predicted PgrAQP genes encoded proteins that varied in

232   length from 245 to 359 amino acid residues, with a relative molecular weight of 22.86 to

233   35.18 kDa, and a calculated pI ranging from 5.06 to 10.16. The average PI value of

234   PgrTIPs was less than other AQPs due to the loss of basic residues in the C-terminal

235   domain. The grand average of hydrophobicity index (GRAVY) was used to evaluated

236   protein hydrophobicity and hydrophilicity. The results showed the GRAVY of PgrAQP

237   proteins were all positive, ranging from 0.203 to 0.967, which indicated that all of

238   PgrAQP proteins were hydrophobic. Furthermore, the lowest average of GRAVY value

239   (0.44) was found in the PgrPIP subfamily, suggesting that this subfamily has better

240   interaction with water molecules.

241   The predicted transmembrane domains (TMDs) showed that most *PgrAQP* genes

242   (28 of 38, 73.7%) contained six TMDs, 2.7%, 15.8%, and 7.8% of PgrAQP genes contained

243   four, five, and seven TMDs, respectively (**Fig. S2**). Based on subcellular localization

244   predicted by WoLF PSORT, most PgrPIP and PgrNIP proteins were predicted to

245   localize in plasma membranes, while only PgrPIP2.2 was found in the chloroplast. All

246   PgrTIP proteins were predicted to localized to vacuoles. For PgrSIP proteins, PgrSIP1.1

247   and PgrSIP1.2 were found in the chloroplast and vacuole, respectively. PgrXIP proteins

248   were predicted to localized in plasma membranes.

249   **3.2 Phylogenetic characterization of the pomegranate *PgrAQP* gene family**

250   To investigate the evolutionary relationship of pomegranate AQP family genes, a

251   total of 193 AQP protein sequences from the four species studied (38 in pomegranate, 35

252   in *Arabidopsis*, 33 in grape, 55 in *Populus trichocarpa* and 40 in eucalyptus) were

253   identified. An unrooted phylogenetic tree was constructed based on the alignments of

254   their amino acid sequences in MEGA 7 using the neighbor-joining method (Kumar et al.

255   2016). By comparing amino acid sequences of PgrAQPs with APQs from three other

256   plant species, 38 PgrAQPs were divided into five different subfamilies, that is, 14

257   PgrTIPs, 13 *PgrPIPs*, eight *PgrNIPs*, two *PgrSIPs*, and one *PgrXIP* (**Fig. 1**). The *PgrPIPs*

258   divided into two major subgroups, *PgrPIP1s* and *PgrPIP2s*, which comprised five and

259   eight members, respectively. Furthermore, PgrNIPs formed six subgroups (*PgrNIP1,*

260   *PgrNIP1, PgrNIP4, PgrNIP5, PgrNIP7,* and *PgrNIP8*) in pomegranate. As the largest sub-

261   family, the TIPs members were classified into five subgroups, consisting of seven

262   PgTIP1s, two PgTIP2s, two PgTIP3s, one PgTIP4, and two PgTIP5s. SIPs and XIPs

263   formed one group, containing two and one members, respectively. In the phylogenetic

264   tree, *PgrAQP* genes were more closely related to AQPs in eucalyptus than that in

265   *Arabidopsis* and grape, which is in accordance with the evolutionary relationships

266   among these species (Qin et al. 2017).

267   **3.3 Gene structure and conserved motif analysis of PgrAQP genes confirm the**

268   **phylogenetic classification**

269   Gene structure and conserved motifs can provide information for exploring the

270   evolutionary relationships among a gene family. The number of introns of *PgrAQP*

271   genes ranged from zero to four, while the length of exons was highly similar for each

272   subfamily (**Fig. S3**). Among them, most of the *PgrTIPs* genes had two introns, except for

273   *PgrTIP1.6* and *PgrTIP1.8*, which had only one intron. For PgrPIP genes, most of

274   members had three introns, whereas *PgPIP2.8* contained two introns. The numbers of

275   introns in *PgrNIPs* ranged from one to four introns. Four out of 10 members had four

276   introns (*PgrNIP2.1*, *PgrNIP4.1*, *PgrNIP4.2*, and *PgrNIP3.1*), three members had three

277   introns (*PgrNIP1.2*, *PgrNIP5.1*, and *PgrNIP7.1*), and two members had two introns

278   (*PgrNIP1.1* and *PgrNIP1.3*). *PgrXIP2.1* had three introns. The SIP family formed a small

279   subfamily, among which, two members had two introns and one member had no

280   introns.

281   By analyzing the intron–exon structure of *PgrAQP* genes, we found that the

282   number of introns of each AQPs subfamily (PIP, NIP, TIP, SIP, and XIP) are highly

283   conserved and similar when compared with plant species, such as banana, watermelon,

284   chickpea, and sweet orange, suggesting similar intron loss or gain events were

285   experienced in AQP subfamily over the course of evolution (Deokar & Tar'an 2016; Hu

286   et al. 2015b; Martins Cde et al. 2015; Zhou et al. 2019). Multiple sequence alignments

287   showed that PgrNIPs and PgrPIPs were the most diverse (38.5%) and conserved (72.7%)

288   subfamily at the amino acid level, respectively (**Table. S3**). This finding is consistent

289   with the AQPs from *Nicotiana tabacum*, suggesting that the function and regulatory

290   mechanism of *PgrPIP* subfamily genes showed more conserved than *PgrNIP* subfamily

291   genes in pomegranate (De Rosa et al. 2020; Deokar & Tar'an 2016). Furthermore, the

292   relative conservation of protein sequences and ar/R selectivity filter among the *PgrPIP*

293   and *PgrTIP* subfamily genes suggested that these proteins may share a conserved

294   function in transporting water and other small neutral solutes (**Table. S3** and **S4**) (Zhu

295   et al. 2019).

296   To detect the structural diversity and provide further support of the grouping of

297   PgrAQPs, a total of 10 conserved motifs were identified (**Fig. S3**). Generally, motif

298   compositions were conserved within each subfamily. For PgrPIP, eight motifs were

299   found in all family members. Motifs 1, 2, 3, 6, 8, and 10 were common for TIP and NIP

300   subfamily members, except for PgrNIP7.1. Interestingly, motifs 4 and 5 were only

301   identified in the PIP subfamily and similar motifs were found in bread wheat,

302   indicating the PIP subfamily may have unique functions (Madrid-Espinoza et al. 2018).

303   **3.4 Comparison of substrate-specific residues in *PgrAQP* proteins**

304   The NPA motifs, ar/R selectively filter, and Froger's positions were identified by

305   multiple sequence alignment between the *PgrAQPs* and *AtAQPs* using MEGA 7.0.

306   These highly conserved motifs and positions were critical for the substrate selectively of

307   AQPs (Tornroth-Horsefield et al. 2006). Conserved domain searches using CDD tool

308   from NCBI confirmed all the predicted AQP genes in Pomegranate encoded MIP

309   domains (Fig.S7). As shown in **Table. S4**, all of PgrTIPs, PIPs, and XIP proteins

310   harbored two conserved NPA domains in both loop B (LB) and loop E (LE). For the

311   PgrNIPs subfamily, except for PgrNIP5.1, the rest of numbers showed the third residue

312   of the first NPA motifs was serine rather than alanine. In addition, PgrSIP1.2 showed

313   the substitution of alanine by threonine in their first NPA motif.

314   The residues of the ar/R selectivity filter and Froger's position displayed conserved

315   regions within each subfamily, but regions were more variable across different

316   subfamilies (Kayum et al. 2017). For example, all of the members of the PgrPIPs showed

317   conservative residues at the ar/R selectively filter with phenylalanine-histidine-

318   threonine-arginine, which is typical of aquaporin protein structure. The residues of

319   Froger's position were conserved in PgrPIPs, including S at P2, A at P3, Y at P4, and W

320   at P5, while the P1 position was variable with the Q/M residues. Different residues were

321   observed at the ar/R selectively filter and Froger's position in PgrNIPs, which had

322   W/G/A/V-V/S/I-A/G-R and F/L/Y-S/T-A-Y/F-I/V/LM residue compositions, respectively.

323   For PgrTIPs members, the residues of P3, P4, and P5 positions were highly conserved,

324   while the P1 and P2 positions were variable residues. The ar/R selectivity filters were

325   variable with H/S/N-I/V-A/P-V/R/C residues. The residues of Froger's position were

326   conserved in PgrSIP1.1 and PgrSIP1.2, but the ar/R filter showed distinct difference.

327   The pore diameter and hydrophobicity of AQP proteins determines their substrate

328   specificity (Almasalmeh et al. 2014; Hove & Bhave 2011). The highly conserved amino

329   acid features of AQPs included six transmembrane domains, the NPA domain, and the

330   ar/R selectivity filter(Froger et al. 1998). For example, all of the PgrPIP subfamily

331   members showed a highly conserved ar/R filter structure (T-H-R-T) (**Table. S4**), which

332   was observed in PIP family genes from other plant species, such as watermelon,

333   *Arabidopsis*, *Brassica rapa*, soybean, and chickpea, indicating that the substrate specificity

334   of this subfamily may be more specific than others subfamilies (Deokar & Tar'an 2016;

335   Kayum et al. 2017; Quigley et al. 2002; Zhang et al. 2013). Among the different TIP

336   subgroups, the highly conserved NPA motif, ar/R H2, H5, and Forger's P3 to P5 were

337 observed. Furthermore, the highly conserved ar/R filter (H-I-A-V) and Froger's

338 positions (T-S-A-Y-W) of the *PgrTIP1* subfamily were reported to function as urea and

339 H_2O_2 transporter (Hove & Bhave 2011). In addition, the conserved ar/R filter G-S-G-R

340 residues were found in *PgrNIP2.1*, and this characteristic was identified as the indicator

341 of Si transporters, indicating that *PgrNIP2.1* may be involved in the transport of Si

342 (Deshmukh et al. 2013; Deshmukh et al. 2015; Zhou et al. 2019). In pomegranate, valine

343 was present at position H5 in the ar/R selectivity filters of *PgrXIP2.1*, suggesting that the

344 hydrophobicity of *PgrXIP2.1* is greater than other PgrAQPs subfamily members

345 (Danielson & Johanson 2008; Gupta & Sankararamakrishnan 2009).

346 **3.5 Segmental duplication events have contributed to the expansion of the *PgrAQP***

347 **family genes**

348 To investigate the localization of *PgrAQP* genes and duplication events in

349 pomegranate, we anchored the *PgrAQPs* on chromosomes and conducted a duplication

350 analysis. The physical position of the *PgrAQP* genes were found to be unevenly

351 distributed across all pomegranate chromosomes (**Fig. 2**). The PIP subfamily genes were

352 randomly anchored on chromosomes, except for Chromosome 7, and TIP subfamily

353 genes were found in all chromosomes except Chromosome 4. Aside from Chromosomes

354 2 and 8, NIP group genes were located in each chromosome. Genes in the SIP subfamily

355 were present only on Chromosome 4.

356 We further analyzed the gene duplication modes of *PgrAQP* genes in pomegranate.

357 As shown in **Fig. 2**, 47% of the *PgrAQP* genes had been duplicated by

358   tandem/segmental duplication events. We found one tandem duplication event in

359   Chromosome 3 (*PgrTIP1.2/PgrTIP1.7*). There were eight pairs of segmental duplications

360   detected among six chromosomes. As shown in **Table 1**, the proportion of segmental

361   *PgrAQP* gene duplications was 88%, indicating that segmental duplication events have

362   played a key role in the expansion of the *PgrAQP* gene family. To access the selection

363   pressure and the date at which such duplication events occurred, estimation of the *Ka*

364   and *Ks* substitution rates of these duplication *PgrAQP* gene pairs were calculated. A

365   *Ka*/*Ks* ratio >1 and <1 indicate positive Darwinian selection or purifying selection,

366   whereas a value of 1 indicates neutral selection. The *Ka*/*Ks* ratios of *PgrAQ*P duplication

367   gene pairs showed a *Ka*/*Ks* ratio of <1, indicating that these *PgrAQP* genes have

368   experienced purifying selection during the course of evolution. According to the

369   mathematical formula (T=*Ks*/2λ) used to calculate the evolutionary date, we assessed

370   the divergence time of these duplication events and found that the gene duplication

371   events occurred approximately 1.74–6.97 million years ago.

372   Further, we constructed the comparative synteny maps of three plants species

373   (pomegranate *vs*. *Arabidopsis*, pomegranate vs. grape, and pomegranate vs. *Eucalyptus*)

374   to explore the evolutionary process of *PgrAQP* genes (**Fig. 3, Table. S5**) and found that

375   24, 28, and 41 orthologous AQP gene pairs were identified, respectively. Remarkably,

376   the numbers of orthologous TIP gene pairs between pomegranate and grape/*Eucalyptus*

377   were significantly higher than that in *Arabidopsis*. Nevertheless, the PIP genes were

378   highly conserved within the species. The syntenic relationship detected in TIP genes

379   indicates that the expansion of *PgrTIPs* and *VvTIPs/EucTIPs* genes may have occurred

380   after that of *Arabidopsis*, while the PIP genes have been evolutionarily conserved.

381   **3.6 Analysis of *PgrAQP* gene expression profiles of different pomegranate tissues**

382   **and water deficit**

383   Identifying tissue-specific genes is a basic strategy to select candidate genes involved in

384   biological processes. To explore the possible functions of *PgrAQP* genes in various

385   developmental stages of different organs of pomegranate, a heat map of *Pgr*AQP

386   expression profiles was conducted. The heat map showed various expression patterns

387   of the 38 *PgrAQP* genes analyzed (**Fig. 4A, Table. S6**). Most of TIPs, such as *PgrTIP1.1*,

388   *PgrTIP1.4*, *PgrTIP1.6*, *PgrTIP1.8*, and *PgrTIP2.1*, showed higher expression in the roots,

389   leaves, and flowers, whereas the transcripts of *PgrTIP2.3, PgrTIP3.1,* and *PgrTIP3.2*

390   showed extremely low expression levels. For the NIPs subfamily, *PgrNIP4.1* and

391   *PgrNIP5.1* were highly expressed in the leaves and roots, respectively. Interestingly, the

392   transcripts of *PgrNIP1.3* could only be detected in flowers, indicating that *PgrNIP1*

393   might be involved in the development of pomegranate flowers, while other members

394   were transcribed at extremely low levels. The accumulation of transcripts of two

395   *PgrSIPs* was detected in all tissues analyzed, whereas *PgrSIP1.2* had higher relative

396   expression levels than that of *PgrSIP1.1*. In the PIPs subfamily, *PgrPIP3*, *PgrPIP2.3*, and

397   *PgrPIP2.4* had a low level of transcripts in all tissues and stages analyzed, whereas the

398   remaining members had higher expression levels in all tissues analyzed. Notably,

399   *PgrPIP1.3, PgrPIP1.5, PgrPIP2.1,* and *PgrPIP2.8* had higher expression levels in the peel

400   and seed coats at all three experimental stages, indicating that these genes may play

401   important roles in seed development.

402   To reveal the potential function of PgrAQPs in response to water deficit in

403   pomegranate root, the transcript levels of seven selected PgrAQPs from PIP subfamily

404   for seedling under 20% PEG supply condition were tested using qRT-PCR. According to

405   the qRT-PCR results, except the *PgrPIP1.5*, of which the transcript was enhanced in the

406   roots after the 24 h of 20% PEG supply conditions, the expression of the other five

407   paralogues, *PgrPIP1.1, PgrPIP1.2, PgrPIP1.3, PgrPIP2.1* and *PgrPIP1.4,* were

408   significantly repressed under the 20% PEG supply condition (Figure 5). It should be

409   noted that the expression levels of the *PgrPIP1.2* was strongly and rapidly decreased

410   under the 20% PEG treatment. In contrast, the expression pattern of *PgrPIP2.4* showed

411   barely changed until 12 h after 20% PEG treatment, the down-regulation was observed

412   at 24 h.

413   **3.7 Identification of candidate *PgrAQPs* involved in water accumulation in the outer**

414   **seed coat of pomegranate**

415   To understand water transport and accumulation in pomegranate seed coats,

416   especially of the juicy outer seed coat, we detected water accumulation in the inner and

417   outer seed coats at different developmental stages, as well as the relative transcript

418   levels of *PgrAQP* genes in corresponding samples. We found that the water content was

419   significantly increased in outer seed coats during seed development in 'Dabenzi' and

420   'Tunisia,' and the water that accumulated in the outer seed coats was higher than that in

421 the inner seed coats (**Fig. 6**). Genes, including *PgrPIP1.3, PgrPIP2.8, PgrPIP1.5,*

422 *PgrPIP2.6, PgrPIP2.1, PgrPIP2.2, PgrPIP2.5,* and *PgSIP1.2,* had relatively high levels of

423 transcript accumulation in inner and outer seed coats. Among the abovementioned

424 genes, *PgrPIP1.3, PgrPIP2.8,* and *PgSIP1.2,* had high levels of accumulation of

425 transcripts only in the outer seed coat at the later developmental stages (**Fig. 5**), which

426 indicated that these genes may be involved in the accumulation of water in outer seed

427 coats at the later developmental stages.

428   In Arabidopsis, a plasma membrane aquaporin, AtPIP1;2 is involved in aquaporin-

429 mediated leaf water transport, lateral root formation, and water uptake in root. To

430 further understand the function of *PgrPIP1.3, PgrPIP2.8,* and *PgrSIP1.2* in the water

431 accumulation in the seed coat, the linear relationship between the water content in the

432 seed coat and differences in transcript levels for *PrgPIP1.3* and *PgrSIP1.2* in seed coat

433 development was determined. Significant positive correlations in both *PgrPIP1.3* and

434 *PgrSIP1.2* and water content in outer seed coats were observed (**Fig. 7**). Therefore, it was

435 concluded that *PgrPIP1.3* and *PgrSIP1.2* might be involved in the accumulation of water

436 in pomegranate seed coats.

437 **4. Discussion**

438 **4.1** *PgrAQP* **identification and structure analysis**

439   AQP proteins play an important role in various physiological and developmental

440 processes of different organs and tissues, and function as a transporter for water and/or

441 small neutral solutes in plants. Numerous studies have been conducted on AQPs in

442   plants, such as in *Arabidopsis*, grape, soybean, rice, and *Populus trichocarpa* (Fouquet et al.

443   2008; Gupta & Sankararamakrishnan 2009; Quigley et al. 2002; Sakurai et al. 2005;

444   Zhang et al. 2013). However, the genome-wide identification of the AQP gene family in

445   pomegranate was absent due to the limitations of an available genome sequence (Luo et

446   al. 2020; Qin et al. 2017). In this study, 38 *PgrAQP* genes were identified and

447   characterized in pomegranate. The characteristics of *PgrAQP* were comparable with

448   other plants species. For instance, the numbers of exons and structures of intron/exons

449   in the *PgrAQP* genes were highly conserved in different species, such as *Arabidopsis*,

450   olive, chickpea, *Arachis hypogea*, and banana (Deokar & Tar'an 2016; Faize et al. 2020; Hu

451   et al. 2015b; Quigley et al. 2002; Shivaraj et al. 2019). Similar to the organizations of

452   intron/exons, *PgrAQP* proteins of each subfamily possess the same conserved motifs

453   (**Fig. 2**). These results suggest that the gene structures of PgrAQPs are closely related to

454   homologous genes.

455   Furthermore, the TMDs showed that some of the PgrAQPs were lack one

456   (PgrSIP1.1) or two (PgrNIP5.1, PgrPIP2.8) TMDs (Fig. S3). Variation in the number of

457   TMDs to different plant species has been reported (Ayadi et al. 2011; Zhu et al. 2019).

458   For instance, a truncated form of wheat TdPIP2;1 aquaporin, showed no water transport

459   activity. Interestingly, the truncated tdpip2;1 could reach the plasma membrane by

460   interact with the functional TdPIP2;1, and then may affect the functional form and

461   reduce the water transport activity of aquaporin (Ayadi et al. 2011). Therefore, the

462   absence of TMDs may affect the PgrAQPs subcellular localization and water transport

activity.

4.2 The evolutionary relationships among the *PgrAQP* gene family

Gene duplication is considered as a major driving force for the evolution of gene families, and several duplication events have been identified over the course of evolution of some plant species. At approximately 117 million years ago (Mya), all core eudicots experienced a genome triplication event (the γ event), including Arabidopsis, *Eucalyptus*, grape, and pomegranate (Jiao et al. 2012; Jiao et al. 2014). Then, Arabidopsis experienced two recent WGDs (α and β), whereas pomegranate and *Eucalyptus* underwent a Myrtales lineage-specific WGD event (109.9Mya, M), but grape did not undergo any additional WGDs (Myburg et al. 2014; Qin et al. 2017). In *Arabidopsis*, a total of 35 AQP genes were identified and further evolution analysis revealed that *AtAQPs* result from different types of gene duplication, such as γ WGD (1), β WGD (2), α WGD (8), tandem (2), and transposed (4) (Bowers et al. 2003; Zwiazek et al. 2017b). For poplar, tandem duplication (4) and the recent WGD (20) were the major driving forces for 55 AQP genes (Zou & Yang 2019).

In this study, the 38-member PgrAQP family is comparable to *Arabidopsis* (33) and *Eucalyptus* (40), but had a greater number of genes than grape (28). Furthermore, all the PgrNIP, PgrSIP, and PgrXIP subfamilies were found to have a close relationship with the corresponding genes in *Eucalyptus* AQPs, which is consistent with the evolutionary relationships among the species, suggesting that these AQP subfamilies might have functional conservation in Myrtales (Qin et al. 2017). It is reasonable to deduce that

484 WGD events may function as a trigger of *PgrAQP* family genes expansion. In additional,

485 we identified eight gene duplication events in the *PgrAQPs* subfamily, including seven

486 segmental duplications and one tandem duplication event (**Fig. 2, Table 1**).

487 Interestingly, the *Ka/Ks* ratio of the eight duplications was < 1, indicating that the

488 evolution of the *PgrAQP* genes is mediated by large-scale purifying selection, similar to

489 the AQP family in *B. rapa* and wheat (Kayum et al. 2017; Madrid-Espinoza et al. 2018).

490 The synonymous substitution rate was also used to estimate the evolutionary timescale,

491 and the divergence time of duplicated *PgrAQP* genes occurred 1.77 to 6.97 million years

492 ago, which is in accordance with the divergence time of *BrAQPs*. This indicates that

493 duplicated divergence of the PgrAQP genes occurred after the triplication events and

494 Myrtales WGD duplication events (Kayum et al. 2017; Qin et al. 2017).

495   According to a genome-wide analysis of different organisms, it was hypothesized

496 that the frequency of gene duplication events was important to the evolution of a

497 species (Flagel & Wendel 2009; Lynch & Conery 2000). In cotton and sesame, tandem

498 duplicated genes showed functional differentiation, although they shared structural

499 conservation (Li et al. 2019; WU Xiang-yang 2016). In our study, *PgrTIP1.2* and

500 *PgrTIP1.7* were identified as tandem duplicated genes, and they were predominantly

501 expressed in roots and outer seed coats in pomegranate, respectively. Our results

502 provided more information for understanding the evolution of plant AQPs.

503 **4.3 Potential functions of *PgrAQP* genes**

504   As the largest AQP subfamily in most plant species, PIPs play a crucial role in

505   water absorption of roots and leaves. Furthermore, PIPs can affect photosynthesis by

506   enhancing the diffusion of CO_2 in mesophyll tissue of rice and *N. tabacum* (Flexas et al.

507   2006; Xu et al. 2019). In this study, *PgrPIP* subfamily genes had higher expression levels

508   in all analyzed samples compared with other analyzed AQPs. Interestingly, among the

509   *RhPIPs*, *RhPIP1s* and *RhPIP2s* were involved in the expansion of rose petals via an

510   ethylene-dependent pathway (Chen et al. 2013; Ma et al. 2008). The expression analysis

511   showed that *PgrPIP1.3* and *PgrPIP2.1* had similar expression patterns in pomegranate

512   flowers, suggesting a similar role of PIPs during pomegranate flower development (**Fig.**

513   **4**). In the higher plants, TIPs are widely used as markers for vacuolar compartments

514   and function as transporter for small solutes in various tissue (Bienert et al. 2007; Holm

515   et al. 2005; Liu et al. 2003; Porcel et al. 2018). In *Arabidopsis*, at least six TIP subfamily

516   numbers, including *TIP1;1*, *TIP1;2*, *TIP2;1*, *TIP2;2*, *TIP2;3*, and *TIP4;1*, showed specific

517   expression patterns in roots (Gattolin et al. 2011). In addition, the rice *OsTIP2;1* was only

518   detected in roots (Nguyen et al. 2013). In pomegranate, *PgrTIP1.2* and *PgrTIP2.3* show

519   higher expression patterns in the roots than other organs, indicated that these gene may

520   specifically participate in the absorption and transport of small solutes, such as NH_4^+,

521   H_2O_2, and urea, in the roots of pomegranate (Bienert et al. 2007; Holm et al. 2005; Liu et

522   al. 2003). In our study, the XIP and majority of NIPs showed lower expression levels

523   than PIPs and TIPs. Interestingly, we found that *PgrNIP5.1* was highly expressed in

524   roots. Such specific root expression of NIPs was reported for *AtNIP5;1* and *HvNIP2;1*

525   (Schnurbusch et al. 2010; Takano et al. 2006), which were involved in boron (B)

526 homeostasis, indicating that *PgNIP5;1* may participate in B absorption and translocation

527 in pomegranate roots.

528   Drought causes tissue dehydration due to an imbalance between plant water

529 uptake and transpiration. Evidence shows that AQPs play an important role in drought

530 tolerance in plants. Ectopically expression of *MpPIP2;1* in *Arabidopsis* has been shown to

531 enhance drought and salinity tolerance. The decrease or increase of the *AQPs* transcript

532 levels could prevent water losses or helps plants to direct water flow to specific organ

533 under drought stress. In this study, we found most members of the *PgrPIP* genes were

534 suppressed by drought stress (Fig 5), including *PgrPIP1.1, PgrPIP1.2, PgrPIP1.3,*

535 *PgrPIP2.1* and *PgrPIP2.4*, suggested their involvement in reduced water losses in

536 Pomegranate plants. Furthermore, the upregulation of the transcript levels of *PgrPIP1.5*

537 was observed under 20% PEG supply condition, indicated that PgrPIP1.5 might play a

538 critical for tolerance to drought in pomegranate.

539 **4.4 Identification of candidate *PgrAQP* genes involved in pomegranate seed coat**

540 **development**

541   Plant cell expansion is primarily driven by turgor and requires steady water intake,

542 the rate of tissue growth is primarily restricted due to decreasing of turgor (Peret et al.

543 2012; Picaud et al. 2003; Reuscher et al. 2013). For example, in higher plants, the

544 development of a seed coat was primarily initiated by fertilization and driven by cell

545 expansion and growth (Figueiredo et al. 2016). Accordingly, most of the AQPs were

546 reported to be strongly expressed in tissues that can be hydraulically limited during

547 growth. In *Arabidopsis*, the specific expression pattern and regulatory mechanism

548 showed that *AtTIP1;1* plays a critical role in cell expansion (Beebo et al. 2009; Ludevid et

549 al. 1992). Overexpression of ginseng TIP in *Arabidopsis* resulted in a significant increase

550 of leaf cell sizes compared with the wild type plants (Lin et al. 2007). In rice, *OsPIP1;1* is

551 highly expressed in leaves and roots, and overexpression of *OsPIP1;1* exhibited a higher

552 germination rate than the control plants (Liu et al. 2013).

553   In pomegranate, the expanded outer seed coats had higher water accumulation

554 than the rigid inner seed coats in both the hard-seeded cultivar 'Dabenzi' and the soft-

555 seeded cultivar 'Tunisia' (**Fig. 6**). Accordingly, significant positive correlations were

556 found between the expression level of *PgrSIP1.2, PgrPIP1.3*, and *PgrPIP2.8* and water

557 content in the seed coats (**Fig. 7**). PgrSIP1.2 was predicted to be localized in the vacuole.

558 Considering that promotion of cell expansion and maintenance of turgor requires the

559 transfer of substantial amounts of water to cells, it is reasonable to proposed that

560 *PgrPIP1.3, PgrPIP2.8*, and *PgrSIP1.2* might be involved in mediating the water

561 accumulation in the inner and outer seed coats of pomegranate.

562 **Conclusions**

563 In this study, a total of 38 AQP genes were identified and their characteristics, including

564 protein physicochemical properties, gene structure, phyletic evolutionary, and

565 expression patterns were studied. These PgrAQP genes are distributed across 9

566 pomegranate chromosomes and divided into five subfamilies. Purifying selection were

567 undergone during the evolution of *PgrAQP* family genes basing on the syntenic

568   relationships and duplication events analysis, and a whole-genome duplication event in

569   Myrtales may contribute to the expansion of *PgrTIP*, *PgrSIP*, and *PgrXIP* genes.

570   Furthermore, the high expression of *PgrPIP1.3, PgrPIP2.8,* and *PgrSIP1.2* in seed coats

571   and the positive correlation between transcript levels of *PgrAQP* genes and the water

572   content revealed these three genes may be the potential candidate genes involving in

573   outer seed coat development. Hence, further studies on functions of this three AQP

574   genes are needed for genetic improvement of outer seed coats in pomegranate.

575

References

577   Adams KL, and Wendel JF. 2005. Polyploidy and genome evolution in plants. *Curr Opin Plant Biol* 8:135-141.
578   10.1016/j.pbi.2005.01.001
579   Almasalmeh A, Krenc D, Wu B, and Beitz E. 2014. Structural determinants of the hydrogen peroxide permeability of
580   aquaporins. *FEBS J* 281:647-656. 10.1111/febs.12653
581   Artimo P, Jonnalagedda M, Arnold K, Baratin D, Csardi G, de Castro E, Duvaud S, Flegel V, Fortier A, Gasteiger E,
582   Grosdidier A, Hernandez C, Ioannidis V, Kuznetsov D, Liechti R, Moretti S, Mostaguir K, Redaschi N, Rossier
583   G, Xenarios I, and Stockinger H. 2012. ExPASy: SIB bioinformatics resource portal. *Nucleic Acids Res*
584   40:W597-603. 10.1093/nar/gks400
585   Ayadi M, Cavez D, Miled N, Chaumont F, Masmoudi K. Identification and characterization of two plasma membran
586   e aquaporins in durum wheat (Triticum turgidum L. subsp. durum) and their role in abiotic stress toleranc
587   e. Plant Physiol Biochem. 2011 Sep;49(9):1029-
588   39. doi: 10.1016/j.plaphy.2011.06.002. Epub 2011 Jun 17. PMID: 21723739.
589   Azad AK, Ahmed J, Alum MA, Hasan MM, Ishikawa T, Sawa Y, and Katsuhara M. 2016. Genome-Wide
590   Characterization of Major Intrinsic Proteins in Four Grass Plants and Their Non-Aqua Transport Selectivity
591   Profiles with Comparative Perspective. *PLoS One* 11:e0157735. 10.1371/journal.pone.0157735
592   Azad AK, Sawa Y, Ishikawa T, and Shibata H. 2004. Phosphorylation of plasma membrane aquaporin regulates
593   temperature-dependent opening of tulip petals. *Plant Cell Physiol* 45:608-617. 10.1093/pcp/pch069
594   Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, and Noble WS. 2009. MEME SUITE:
595   tools for motif discovery and searching. *Nucleic Acids Res* 37:W202-208. 10.1093/nar/gkp335
596   Beebo A, Thomas D, Der C, Sanchez L, Leborgne-Castel N, Marty F, Schoefs B, and Bouhidel K. 2009. Life with and
597   without AtTIP1;1, an Arabidopsis aquaporin preferentially localized in the apposing tonoplasts of adjacent
598   vacuoles. *Plant Mol Biol* 70:193-209. 10.1007/s11103-009-9465-2
599   Bienert GP, Moller AL, Kristiansen KA, Schulz A, Moller IM, Schjoerring JK, and Jahn TP. 2007. Specific aquaporins

600   facilitate the diffusion of hydrogen peroxide across membranes. *J Biol Chem* 282:1183-1192.
601   10.1074/jbc.M603761200
602   Bowers JE, Chapman BA, Rong J, and Paterson AH. 2003. Unravelling angiosperm genome evolution by
603   phylogenetic analysis of chromosomal duplication events. *Nature* 422:433-438. 10.1038/nature01521
604   Chaumont F, Barrieu F, Wojcik E, Chrispeels MJ, and Jung R. 2001. Aquaporins constitute a large and highly
605   divergent protein family in maize. *Plant Physiol* 125:1206-1215. 10.1104/pp.125.3.1206
606   Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, and Xia R. 2020. TBtools: An Integrative Toolkit Developed
607   for Interactive Analyses of Big Biological Data. *Mol Plant* 13:1194-1202. 10.1016/j.molp.2020.06.009
608   Chen W, Yin X, Wang L, Tian J, Yang R, Liu D, Yu Z, Ma N, and Gao J. 2013. Involvement of rose aquaporin RhPIP1;1
609   in ethylene-regulated petal expansion through interaction with RhPIP2;1. *Plant Mol Biol* 83:219-233.
610   10.1007/s11103-013-0084-6
611   Danielson JA, and Johanson U. 2008. Unexpected complexity of the aquaporin gene family in the moss
612   Physcomitrella patens. *BMC Plant Biol* 8:45. 10.1186/1471-2229-8-45
613   Danielson JA, and Johanson U. 2010. Phylogeny of major intrinsic proteins. *Adv Exp Med Biol* 679:19-31.
614   10.1007/978-1-4419-6315-4_2
615   De Rosa A, Watson-Lazowski A, Evans JR, and Groszmann M. 2020. Genome-wide identification and
616   characterisation of Aquaporins in Nicotiana tabacum and their relationships with other Solanaceae
617   species. *BMC Plant Biol* 20:266. 10.1186/s12870-020-02412-5
618   Deokar AA, and Tar'an B. 2016. Genome-Wide Analysis of the Aquaporin Gene Family in Chickpea (Cicer arietinum
619   L.). *Front Plant Sci* 7:1802. 10.3389/fpls.2016.01802
620   Deshmukh RK, Vivancos J, Guérin V, Sonah H, Labbé C, Belzile F, and Bélanger RR. 2013. Identification and
621   functional characterization of silicon transporters in soybean using comparative genomics of major
622   intrinsic proteins in Arabidopsis and rice. *Plant Mol Biol* 83:303-315. 10.1007/s11103-013-0087-3
623   Deshmukh RK, Vivancos J, Ramakrishnan G, Guérin V, Carpentier G, Sonah H, Labbé C, Isenring P, Belzile FJ, and
624   Bélanger RR. 2015. A precise spacing between the NPA domains of aquaporins is essential for silicon
625   permeability in plants. *Plant J* 83:489-500. 10.1111/tpj.12904
626   Edlund AF, Swanson R, and Preuss D. 2004. Pollen and stigma structure and function: the role of diversity in
627   pollination. *Plant Cell* 16 Suppl:S84-97. 10.1105/tpc.015800
628   Eisenbarth DA, and Weig AR. 2005. Dynamics of aquaporins and water relations during hypocotyl elongation in
629   Ricinus communis L. seedlings. *J Exp Bot* 56:1831-1842. 10.1093/jxb/eri173
630   Faize M, Fumanal B, Luque F, Ramirez-Tejero JA, Zou Z, Qiao X, Faize L, Gousset-Dupont A, Roeckel-Drevet P, Label
631   P, and Venisse JS. 2020. Genome Wild Analysis and Molecular Understanding of the Aquaporin Diversity in
632   Olive Trees (Olea Europaea L.). *Int J Mol Sci* 21. 10.3390/ijms21114183
633   Figueiredo DD, Batista RA, Roszak PJ, Hennig L, and Kohler C. 2016. Auxin production in the endosperm drives seed
634   coat development in Arabidopsis. *Elife* 5. 10.7554/eLife.20542
635   Finn RD, Bateman A, Clements J, Coggill P, Eberhardt RY, Eddy SR, Heger A, Hetherington K, Holm L, Mistry J,
636   Sonnhammer EL, Tate J, and Punta M. 2014. Pfam: the protein families database. *Nucleic Acids Res*
637   42:D222-230. 10.1093/nar/gkt1223
638   Flagel LE, and Wendel JF. 2009. Gene duplication and evolutionary novelty in plants. *New Phytol* 183:557-564.
639   10.1111/j.1469-8137.2009.02923.x
640   Flexas J, Ribas-Carbo M, Hanson DT, Bota J, Otto B, Cifre J, McDowell N, Medrano H, and Kaldenhoff R. 2006.

641   Tobacco aquaporin NtAQP1 is involved in mesophyll conductance to CO2 in vivo. *Plant J* 48:427-439.
642   10.1111/j.1365-313X.2006.02879.x

643 Fouquet R, Léon C, Ollat N, and Barrieu F. 2008. Identification of grapevine aquaporins and expression analysis in
644   developing berries. *Plant Cell Reports* 27:1541-1550. 10.1007/s00299-008-0566-1

645 Froger A, Tallur B, Thomas D, and Delamarche C. 1998. Prediction of functional residues in water channels and
646   related proteins. *Protein Sci* 7:1458-1468. 10.1002/pro.5560070623

647 Gattolin S, Sorieul M, and Frigerio L. 2011. Mapping of tonoplast intrinsic proteins in maturing and germinating
648   Arabidopsis seeds reveals dual localization of embryonic TIPs to the tonoplast and plasma membrane. *Mol*
649   *Plant* 4:180-189. 10.1093/mp/ssq051

650 Gu Z, Cavalcanti A, Chen FC, Bouman P, and Li WH. 2002. Extent of gene duplication in the genomes of Drosophila,
651   nematode, and yeast. *Mol Biol Evol* 19:256-262. 10.1093/oxfordjournals.molbev.a004079

652 Gupta AB, and Sankararamakrishnan R. 2009. Genome-wide analysis of major intrinsic proteins in the tree plant
653   Populus trichocarpa: characterization of XIP subfamily of aquaporins from evolutionary perspective. *BMC*
654   *Plant Biol* 9:134. 10.1186/1471-2229-9-134

655 Holm LM, Jahn TP, Moller AL, Schjoerring JK, Ferri D, Klaerke DA, and Zeuthen T. 2005. NH3 and NH4+ permeability
656   in aquaporin-expressing Xenopus oocytes. *Pflugers Arch* 450:415-428. 10.1007/s00424-005-1399-1

657 Hove RM, and Bhave M. 2011. Plant aquaporins with non-aqua functions: deciphering the signature sequences.
658   *Plant Mol Biol* 75:413-430. 10.1007/s11103-011-9737-5

659 Hu B, Jin J, Guo AY, Zhang H, Luo J, and Gao G. 2015a. GSDS 2.0: an upgraded gene feature visualization server.
660   *Bioinformatics* 31:1296-1297. 10.1093/bioinformatics/btu817

661 Hu W, Hou X, Huang C, Yan Y, Tie W, Ding Z, Wei Y, Liu J, Miao H, Lu Z, Li M, Xu B, and Jin Z. 2015b. Genome-Wide
662   Identification and Expression Analyses of Aquaporin Gene Family during Development and Abiotic Stress
663   in Banana. *Int J Mol Sci* 16:19728-19751. 10.3390/ijms160819728

664 Hu W, Hua X, Zhang Q, Wang J, Shen Q, Zhang X, Wang K, Yu Q, Lin YR, Ming R, and Zhang J. 2018. New insights
665   into the evolution and functional divergence of the SWEET family in Saccharum based on comparative
666   genomics. *BMC Plant Biol* 18:270. 10.1186/s12870-018-1495-y

667 Ivanov, II, Loktyushkin AV, Gus'kova RA, Vasil'ev NS, Fedorov GE, and Rubin AB. 2007. Oxygen channels of
668   erythrocyte membrane. *Dokl Biochem Biophys* 414:137-140. 10.1134/s160767290703012x

669 Jiao Y, Leebens-Mack J, Ayyampalayam S, Bowers JE, McKain MR, McNeal J, Rolf M, Ruzicka DR, Wafula E, Wickett
670   NJ, Wu X, Zhang Y, Wang J, Zhang Y, Carpenter EJ, Deyholos MK, Kutchan TM, Chanderbali AS, Soltis PS,
671   Stevenson DW, McCombie R, Pires JC, Wong GK, Soltis DE, and Depamphilis CW. 2012. A genome
672   triplication associated with early diversification of the core eudicots. *Genome Biol* 13:R3. 10.1186/gb-
673   2012-13-1-r3

674 Jiao Y, Li J, Tang H, and Paterson AH. 2014. Integrated syntenic and phylogenomic analyses reveal an ancient
675   genome duplication in monocots. *Plant Cell* 26:2792-2802. 10.1105/tpc.114.127597

676 Johanningsmeier SD, and Harris GK. 2011. Pomegranate as a Functional Food and Nutraceutical Source. *Annual*
677   *Review of Food Science and Technology* 2:181-201. 10.1146/annurev-food-030810-153709

678 Kaldenhoff R, and Fischer M. 2006. Functional aquaporin diversity in plants. *Biochim Biophys Acta* 1758:1134-1141.
679   10.1016/j.bbamem.2006.03.012

680 Kayum MA, Park JI, Nath UK, Biswas MK, Kim HT, and Nou IS. 2017. Genome-wide expression profiling of
681   aquaporin genes confer responses to abiotic and biotic stresses in Brassica rapa. *BMC Plant Biol* 17:23.

682   10.1186/s12870-017-0979-5

683   Khabudaev KV, Petrova DP, Grachev MA, and Likhoshway YV. 2014. A new subfamily LIP of the major intrinsic
684   proteins. *BMC Genomics* 15:173. 10.1186/1471-2164-15-173

685   Kumar S, Stecher G, and Tamura K. 2016. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger
686   Datasets. *Mol Biol Evol* 33:1870-1874. 10.1093/molbev/msw054

687   Lee JK, Kozono D, Remis J, Kitagawa Y, Agre P, and Stroud RM. 2005. Structural basis for conductance by the
688   archaeal aquaporin AqpM at 1.68 A. *Proc Natl Acad Sci U S A* 102:18932-18937.
689   10.1073/pnas.0509469102

690   Letunic I, Khedkar S, and Bork P. 2020. SMART: recent updates, new developments and status in 2020. *Nucleic*
691   *Acids Res*. 10.1093/nar/gkaa937

692   Li W, Zhang D, Zhu G, Mi X, and Guo W. 2019. Combining genome-wide and transcriptome-wide analyses reveal
693   the evolutionary conservation and functional diversity of aquaporins in cotton. *BMC Genomics* 20:538.
694   10.1186/s12864-019-5928-2

695   Lian HL, Yu X, Ye Q, Ding X, Kitagawa Y, Kwak SS, Su WA, and Tang ZC. 2004. The role of aquaporin RWC3 in drought
696   avoidance in rice. *Plant Cell Physiol* 45:481-489. 10.1093/pcp/pch058

697   Lin W, Peng Y, Li G, Arora R, Tang Z, Su W, and Cai W. 2007. Isolation and functional characterization of PgTIP1, a
698   hormone-autotrophic cells-specific tonoplast aquaporin in ginseng. *J Exp Bot* 58:947-956.
699   10.1093/jxb/erl255

700   Liu C, Fukumoto T, Matsumoto T, Gena P, Frascaria D, Kaneko T, Katsuhara M, Zhong S, Sun X, Zhu Y, Iwasaki I, Ding
701   X, Calamita G, and Kitagawa Y. 2013. Aquaporin OsPIP1;1 promotes rice salt resistance and seed
702   germination. *Plant Physiol Biochem* 63:151-158. 10.1016/j.plaphy.2012.11.018

703   Liu HY, Yu X, Cui DY, Sun MH, Sun WN, Tang ZC, Kwak SS, and Su WA. 2007. The role of water channel proteins and
704   nitric oxide signaling in rice seed germination. *Cell Res* 17:638-649. 10.1038/cr.2007.34

705   Liu LH, Ludewig U, Gassert B, Frommer WB, and von Wiren N. 2003. Urea transport by nitrogen-regulated
706   tonoplast intrinsic proteins in Arabidopsis. *Plant Physiol* 133:1220-1228. 10.1104/pp.103.027409

707   Ludevid D, Hofte H, Himelblau E, and Chrispeels MJ. 1992. The Expression Pattern of the Tonoplast Intrinsic Protein
708   gamma-TIP in Arabidopsis thaliana Is Correlated with Cell Enlargement. *Plant Physiol* 100:1633-1639.
709   10.1104/pp.100.4.1633

710   Luo X, Li H, Wu Z, Yao W, Zhao P, Cao D, Yu H, Li K, Poudel K, Zhao D, Zhang F, Xia X, Chen L, Wang Q, Jing D, and
711   Cao S. 2020. The pomegranate (Punica granatum L.) draft genome dissects genetic divergence between
712   soft- and hard-seeded cultivars. *Plant Biotechnol J* 18:955-968. 10.1111/pbi.13260

713   Lynch M, and Conery JS. 2000. The evolutionary fate and consequences of duplicate genes. *Science* 290:1151-1155.
714   10.1126/science.290.5494.1151

715   Ma N, Xue J, Li Y, Liu X, Dai F, Jia W, Luo Y, and Gao J. 2008. Rh-PIP2;1, a rose aquaporin gene, is involved in
716   ethylene-regulated petal expansion. *Plant Physiol* 148:894-907. 10.1104/pp.108.120154

717   Madrid-Espinoza J, Brunel-Saldias N, Guerra FP, Gutierrez A, and Del Pozo A. 2018. Genome-Wide Identification
718   and Transcriptional Regulation of Aquaporin Genes in Bread Wheat (Triticum aestivum L.) under Water
719   Stress. *Genes (Basel)* 9. 10.3390/genes9100497

720   Martins Cde P, Pedrosa AM, Du D, Goncalves LP, Yu Q, Gmitter FG, Jr., and Costa MG. 2015. Genome-Wide
721   Characterization and Expression Analysis of Major Intrinsic Proteins during Abiotic and Biotic Stresses in
722   Sweet Orange (Citrus sinensis L. Osb.). *PLoS One* 10:e0138786. 10.1371/journal.pone.0138786

723   Myburg AA, Grattapaglia D, Tuskan GA, Hellsten U, Hayes RD, Grimwood J, Jenkins J, Lindquist E, Tice H, Bauer D,
724   Goodstein DM, Dubchak I, Poliakov A, Mizrachi E, Kullan AR, Hussey SG, Pinard D, van der Merwe K, Singh
725   P, van Jaarsveld I, Silva-Junior OB, Togawa RC, Pappas MR, Faria DA, Sansaloni CP, Petroli CD, Yang X,
726   Ranjan P, Tschaplinski TJ, Ye CY, Li T, Sterck L, Vanneste K, Murat F, Soler M, Clemente HS, Saidi N, Cassan-
727   Wang H, Dunand C, Hefer CA, Bornberg-Bauer E, Kersting AR, Vining K, Amarasinghe V, Ranik M, Naithani
728   S, Elser J, Boyd AE, Liston A, Spatafora JW, Dharmwardhana P, Raja R, Sullivan C, Romanel E, Alves-Ferreira
729   M, Kulheim C, Foley W, Carocha V, Paiva J, Kudrna D, Brommonschenkel SH, Pasquali G, Byrne M, Rigault
730   P, Tibbits J, Spokevicius A, Jones RC, Steane DA, Vaillancourt RE, Potts BM, Joubert F, Barry K, Pappas GJ,
731   Strauss SH, Jaiswal P, Grima-Pettenati J, Salse J, Van de Peer Y, Rokhsar DS, and Schmutz J. 2014. The
732   genome of Eucalyptus grandis. *Nature* 510:356-362. 10.1038/nature13308
733   Nguyen MX, Moon S, and Jung KH. 2013. Genome-wide expression analysis of rice aquaporin genes and
734   development of a functional gene network mediated by aquaporin expression in roots. *Planta* 238:669-
735   681. 10.1007/s00425-013-1918-9
736   Niu J, Cao D, Li H, Xue H, Chen L, Liu B, and Cao S. 2018. Quantitative proteomics of pomegranate varieties with
737   contrasting seed hardness during seed development stages. *Tree Genetics & Genomes* 14:14.
738   10.1007/s11295-018-1229-1
739   Patel C, Dadhaniya P, Hingorani L, and Soni MG. 2008. Safety assessment of pomegranate fruit extract: acute and
740   subchronic toxicity studies. *Food Chem Toxicol* 46:2728-2735. 10.1016/j.fct.2008.04.035
741   Peret B, Li G, Zhao J, Band LR, Voss U, Postaire O, Luu DT, Da Ines O, Casimiro I, Lucas M, Wells DM, Lazzerini L,
742   Nacry P, King JR, Jensen OE, Schaffner AR, Maurel C, and Bennett MJ. 2012. Auxin regulates aquaporin
743   function to facilitate lateral root emergence. *Nat Cell Biol* 14:991-998. 10.1038/ncb2573
744   Picaud S, Becq F, Dedaldechamp F, Ageorges A, and Delrot S. 2003. Cloning and expression of two plasma
745   membrane aquaporins expressed during the ripening of grape berry. *Funct Plant Biol* 30:621-630.
746   10.1071/FP02116
747   Pommerrenig B, Diehn TA, and Bienert GP. 2015. Metalloido-porins: Essentiality of Nodulin 26-like intrinsic
748   proteins in metalloid transport. *Plant Sci* 238:212-227. 10.1016/j.plantsci.2015.06.002
749   Porcel R, Bustamante A, Ros R, Serrano R, and Mulet Salort JM. 2018. BvCOLD1: A novel aquaporin from sugar beet
750   (Beta vulgaris L.) involved in boron homeostasis and abiotic stress. *Plant Cell Environ* 41:2844-2857.
751   10.1111/pce.13416
752   Qiao X, Li M, Li L, Yin H, Wu J, and Zhang S. 2015. Genome-wide identification and comparative analysis of the heat
753   shock transcription factor family in Chinese white pear (Pyrus bretschneideri) and five other Rosaceae
754   species. *BMC Plant Biol* 15:12. 10.1186/s12870-014-0401-5
755   Qin G, Liu C, Li J, Qi Y, Gao Z, Zhang X, Yi X, Pan H, Ming R, and Xu Y. 2020. Diversity of metabolite accumulation
756   patterns in inner and outer seed coats of pomegranate: exploring their relationship with genetic
757   mechanisms of seed coat development. *Hortic Res* 7:10. 10.1038/s41438-019-0233-4
758   Qin G, Xu C, Ming R, Tang H, Guyot R, Kramer EM, Hu Y, Yi X, Qi Y, Xu X, Gao Z, Pan H, Jian J, Tian Y, Yue Z, and Xu Y.
759   2017. The pomegranate (Punica granatum L.) genome and the genomics of punicalagin biosynthesis. *Plant
760   J* 91:1108-1128. 10.1111/tpj.13625
761   Quigley F, Rosenberg JM, Shachar-Hill Y, and Bohnert HJ. 2002. From genome to function: the Arabidopsis
762   aquaporins. *Genome Biol* 3:RESEARCH0001. 10.1186/gb-2001-3-1-research0001
763   Reuscher S, Akiyama M, Mori C, Aoki K, Shibata D, and Shiratake K. 2013. Genome-wide identification and

764    expression analysis of aquaporins in tomato. *PLoS One* 8:e79052. 10.1371/journal.pone.0079052

765 Sakurai J, Ishikawa F, Yamaguchi T, Uemura M, and Maeshima M. 2005. Identification of 33 rice aquaporin genes

766    and analysis of their expression and function. *Plant Cell Physiol* 46:1568-1577. 10.1093/pcp/pci172

767 Schnurbusch T, Hayes J, Hrmova M, Baumann U, Ramesh SA, Tyerman SD, Langridge P, and Sutton T. 2010. Boron

768    Toxicity Tolerance in Barley through Reduced Expression of the Multifunctional Aquaporin HvNIP2;1.

769    153:1706-1715. 10.1104/pp.110.158832 %J Plant Physiology

770 Schuurmans JA, van Dongen JT, Rutjens BP, Boonman A, Pieterse CM, Borstlap AC. Members of the aquaporin famil

771    y in the developing pea seed coat include representatives of the PIP, TIP, and NIP subfamilies. Plant Mol Bi

772    ol. 2003 Nov;53(5):633-45. doi: 10.1023/B:PLAN.0000019070.60954.77. PMID: 15010602.

773 Shivaraj SM, Deshmukh R, Sonah H, and Belanger RR. 2019. Identification and characterization of aquaporin genes

774    in Arachis duranensis and Arachis ipaensis genomes, the diploid progenitors of peanut. *BMC Genomics*

775    20:222. 10.1186/s12864-019-5606-4

776 Soto G, Alleva K, Mazzella MA, Amodeo G, and Muschietti JP. 2008. AtTIP1;3 and AtTIP5;1, the only highly

777    expressed Arabidopsis pollen-specific aquaporins, transport water and urea. *FEBS Lett* 582:4077-4082.

778    10.1016/j.febslet.2008.11.002

779 Soto G, Fox R, Ayub N, Alleva K, Guaimas F, Erijman EJ, Mazzella A, Amodeo G, and Muschietti J. 2010. TIP5;1 is an

780    aquaporin specifically targeted to pollen mitochondria and is probably involved in nitrogen remobilization

781    in Arabidopsis thaliana. *Plant J* 64:1038-1047. 10.1111/j.1365-313X.2010.04395.x

782 Suga N, Takada H, Nomura A, Ohga S, Ishii E, Ihara K, Ohshima K, and Hara T. 2002. Perforin defects of primary

783    haemophagocytic lymphohistiocytosis in Japan. *Br J Haematol* 116:346-349. 10.1046/j.1365-

784    2141.2002.03266.x

785 Takano J, Wada M, Ludewig U, Schaaf G, von Wiren N, and Fujiwara T. 2006. The Arabidopsis major intrinsic

786    protein NIP5;1 is essential for efficient boron uptake and plant development under boron limitation. *Plant*

787    *Cell* 18:1498-1509. 10.1105/tpc.106.041640

788 Terashima I, and Ono K. 2002. Effects of HgCl(2) on CO(2) dependence of leaf photosynthesis: evidence indicating

789    involvement of aquaporins in CO(2) diffusion across the plasma membrane. *Plant Cell Physiol* 43:70-78.

790    10.1093/pcp/pcf001

791 Tornroth-Horsefield S, Wang Y, Hedfalk K, Johanson U, Karlsson M, Tajkhorshid E, Neutze R, and Kjellbom P. 2006.

792    Structural mechanism of plant aquaporin gating. *Nature* 439:688-694. 10.1038/nature04316

793 Uçar S, and Karagöz S. 2009. The slow pyrolysis of pomegranate seeds: The effect of temperature on the product

794    yields and bio-oil properties. *Journal of Analytical and Applied Pyrolysis* 84:151-156.

795    https://doi.org/10.1016/j.jaap.2009.01.005

796 Vander Willigen C, Postaire O, Tournaire-Roux C, Boursiac Y, and Maurel C. 2006. Expression and inhibition of

797    aquaporins in germinating Arabidopsis seeds. *Plant Cell Physiol* 47:1241-1250. 10.1093/pcp/pcj094

798 Wallace IS, and Roberts DM. 2004. Homology modeling of representative subfamilies of Arabidopsis major intrinsic

799    proteins. Classification based on the aromatic/arginine selectivity filter. *Plant Physiol* 135:1059-1068.

800    10.1104/pp.103.033415

801 Wang L, Guo K, Li Y, Tu Y, Hu H, Wang B, Cui X, and Peng L. 2010. Expression profiling and integrative analysis of

802    the CESA/CSL superfamily in rice. *BMC Plant Biol* 10:282. 10.1186/1471-2229-10-282

803 WU Xiang-yang CC-z, Lü Gao-qiang, WANG Xin-yu. 2016. Identification and Characterization of the AQP Gene

804    Family in Sesame. 49:1844-1858. 10.3864/j.issn.0578-1752.2016.10.002

805 Wudick MM, Luu DT, Tournaire-Roux C, Sakamoto W, and Maurel C. 2014. Vegetative and sperm cell-specific
806   aquaporins of Arabidopsis highlight the vacuolar equipment of pollen and contribute to plant
807   reproduction. *Plant Physiol* 164:1697-1706. 10.1104/pp.113.228700
808 Xu F, Wang K, Yuan W, Xu W, Liu S, Kronzucker HJ, Chen G, Miao R, Zhang M, Ding M, Xiao L, Kai L, Zhang J, and Zhu
809   Y. 2019. Overexpression of rice aquaporin OsPIP1;2 improves yield by enhancing mesophyll CO2
810   conductance and phloem sucrose transport. *Journal of Experimental Botany* 70:671-681.
811   10.1093/jxb/ery386
812 Yang S, Zhang X, Yue JX, Tian D, and Chen JQ. 2008. Recent duplications dominate NBS-encoding gene expansion in
813   two woody species. *Mol Genet Genomics* 280:187-198. 10.1007/s00438-008-0355-0
814 Yasui M, Hazama A, Kwon TH, Nielsen S, Guggino WB, and Agre P. 1999. Rapid gating and anion permeability of an
815   intracellular aquaporin. *Nature* 402:184-187. 10.1038/46045
816 Zarei A, Zamani Z, Fatahi R, Mousavi A, Salami SA, Avila C, and Cánovas FM. 2016. Differential expression of cell
817   wall related genes in the seeds of soft- and hard-seeded pomegranate genotypes. *Scientia Horticulturae*
818   205:7-16. https://doi.org/10.1016/j.scienta.2016.03.043
819 Zhang DY, Ali Z, Wang CB, Xu L, Yi JX, Xu ZL, Liu XQ, He XL, Huang YH, Khan IA, Trethowan RM, and Ma HX. 2013.
820   Genome-wide sequence characterization and expression analysis of major intrinsic proteins in soybean
821   (Glycine max L.). *PLoS One* 8:e56312. 10.1371/journal.pone.0056312
822 Zhou Y, Setz N, Niemietz C, Qu H, Offler CE, Tyerman SD, and Patrick JW. 2007a. Aquaporins and unloading of
823   phloem-imported water in coats of developing bean seeds. *Plant Cell Environ* 30:1566-1577.
824   10.1111/j.1365-3040.2007.01732.x
825 ZHOU Y, SETZ N, NIEMIETZ C, QU H, OFFLER CE, TYERMAN SD, and PATRICK JW. 2007b. Aquaporins and unloading
826   of phloem-imported water in coats of developing bean seeds.   30:1566-1577.
827   https://doi.org/10.1111/j.1365-3040.2007.01732.x
828 Zhou Y, Tao J, Ahammed GJ, Li J, and Yang Y. 2019. Genome-wide identification and expression analysis of
829   aquaporin gene family related to abiotic stress in watermelon. *Genome* 62:643-656. 10.1139/gen-2019-
830   0061
831 Zou Z, and Yang J. 2019. Genome-wide comparison reveals divergence of cassava and rubber aquaporin family
832   genes after the recent whole-genome duplication. *BMC Genomics* 20:380. 10.1186/s12864-019-5780-4
833 Zwiazek J, Xu H, Tan X, Navarro-Ródenas A, and Morte A. 2017a. Significance of oxygen transport through
834   aquaporins. *Scientific Reports* 7:40411. 10.1038/srep40411
835 Zwiazek JJ, Xu H, Tan X, Navarro-Rodenas A, and Morte A. 2017b. Significance of oxygen transport through
836   aquaporins. *Sci Rep* 7:40411. 10.1038/srep40411

837

Figure 1

Phylogenetic analysis of AQP proteins from pomegranate, grape, *Arabidopsis*, *Populus trichocarpa* and *Eucalyptus*.

The tree was generated by the neighbor-joining (NJ) method in MEGA 7.0 with 1,000 bootstrap replicates. Different background colors indicate the different subfamilies of the AQP proteins.

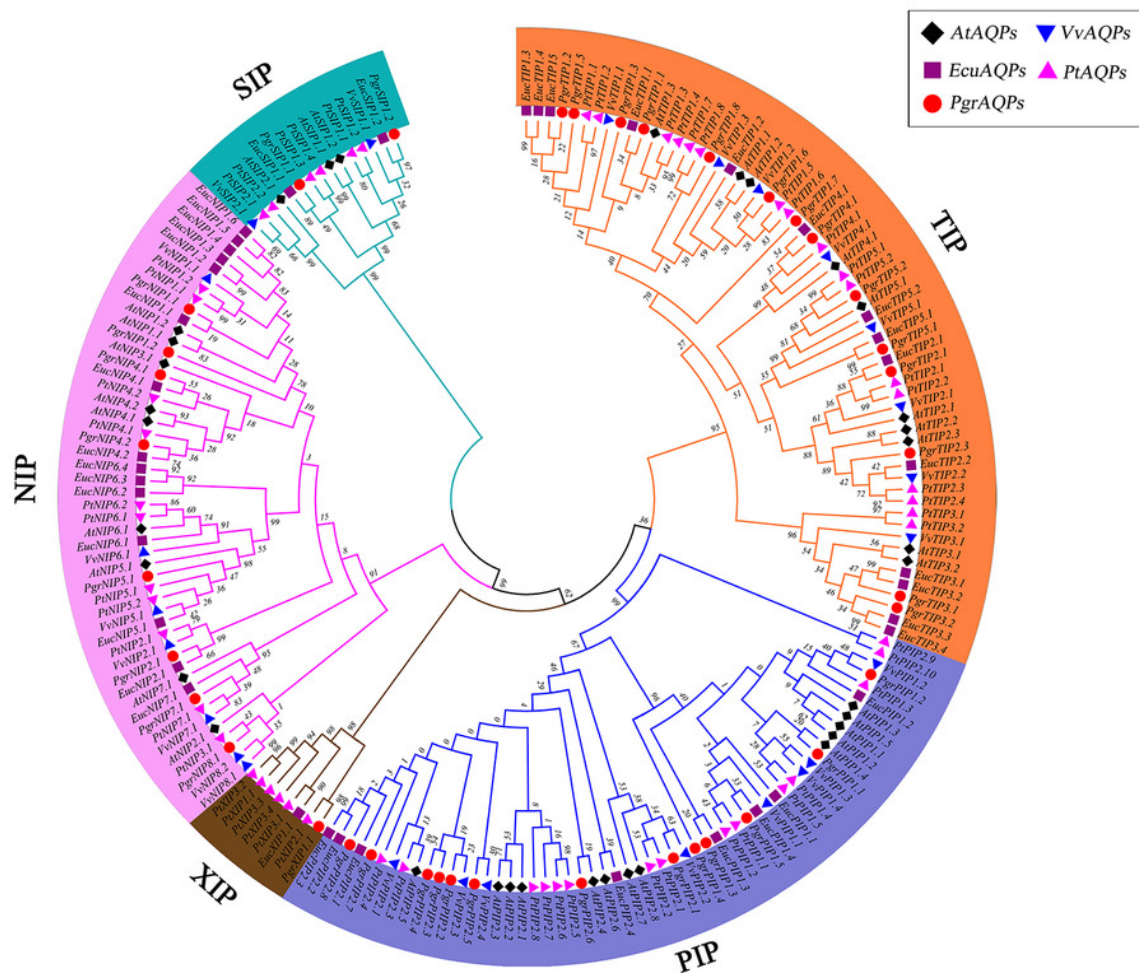



Figure 2

Analysis of chromosomal locations and syntenic relationships of *PgrAQP* genes.

The AQP genes in pomegranate were mapped to different chromosomes using TBtools, and AQP genes in red and blue represent genes with segmental and tandem duplications, respectively.

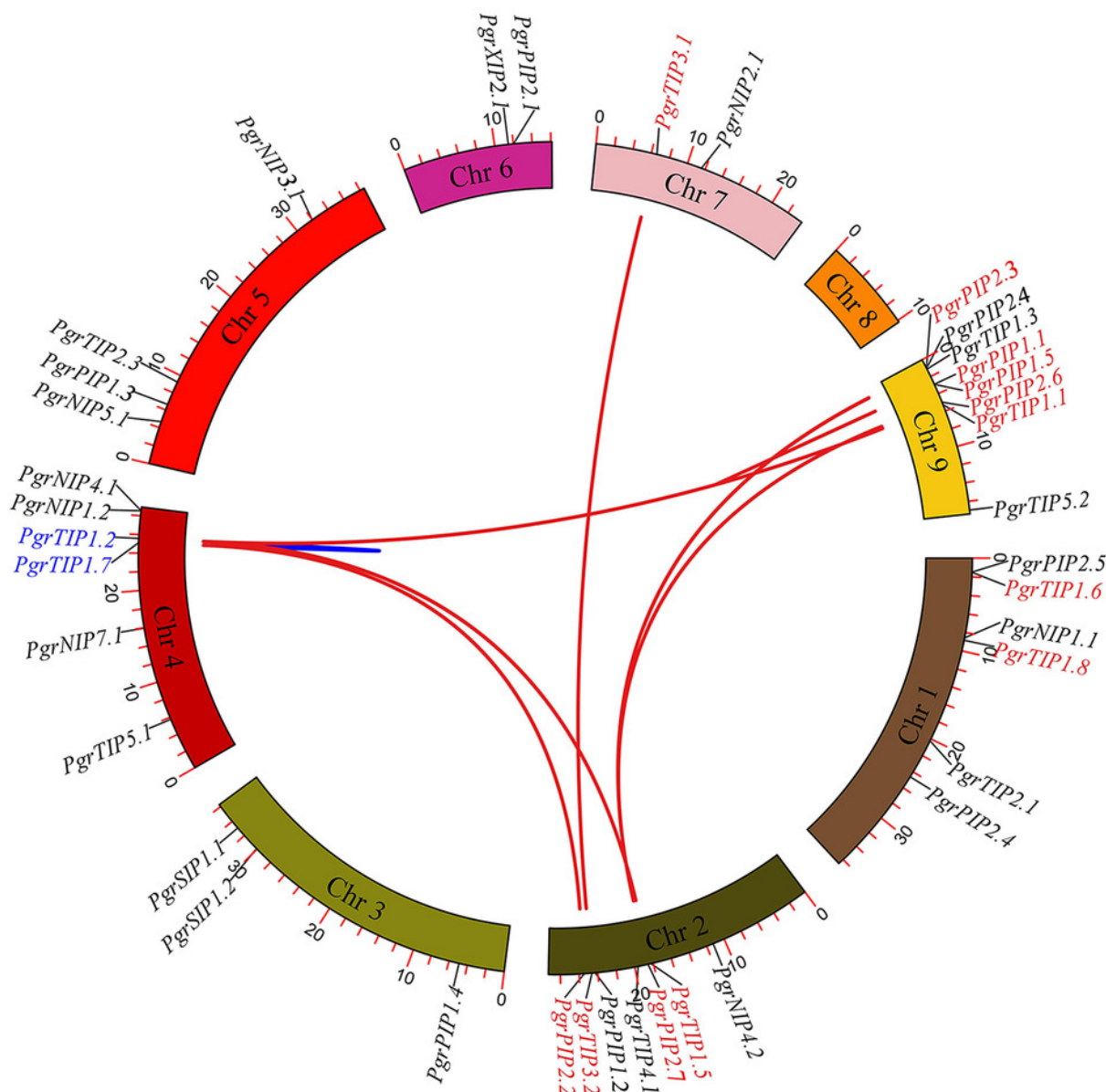



Figure 3

Synteny analysis of *PgrAQP* genes between pomegranate and three plant species.

(A) Pomegranate and *A. thaliana*, **(B)** pomegranate and *Vitis vinifera*, **(C)** pomegranate and *Eucalyptus grandis*. The gray lines indicated collinearity between pomegranate and other species. The red lines highlight the syntenic AQP gene pairs. The chromosome name is indicated at the top of every chromosome.

Figure 4

Expression analysis of the *PgrAQP* genes in pomegranate.

(A) Expression profile of *PgrAQP* genes in the cultivated pomegranate cultivar 'Dabenzi,' including roots, flowers, leaves, and three stages of the peel, inner, and outer seed coats (50, 95, and 140 days after pollination). **(B)** Expression profiling of *PgrAQP* genes at different developmental stages of the seed coats in pomegranate. The abbreviations are as follows: D: *P. granatum* 'Dabenzi,' T: *P. granatum* 'Tunisia,' O: Outer seed coat, I: Inner seed coat. The number represents the number of days after pollination (DAP). The heat map was generated using TBtools. Expression data were calculated with log2 normalization based on FPKM values.

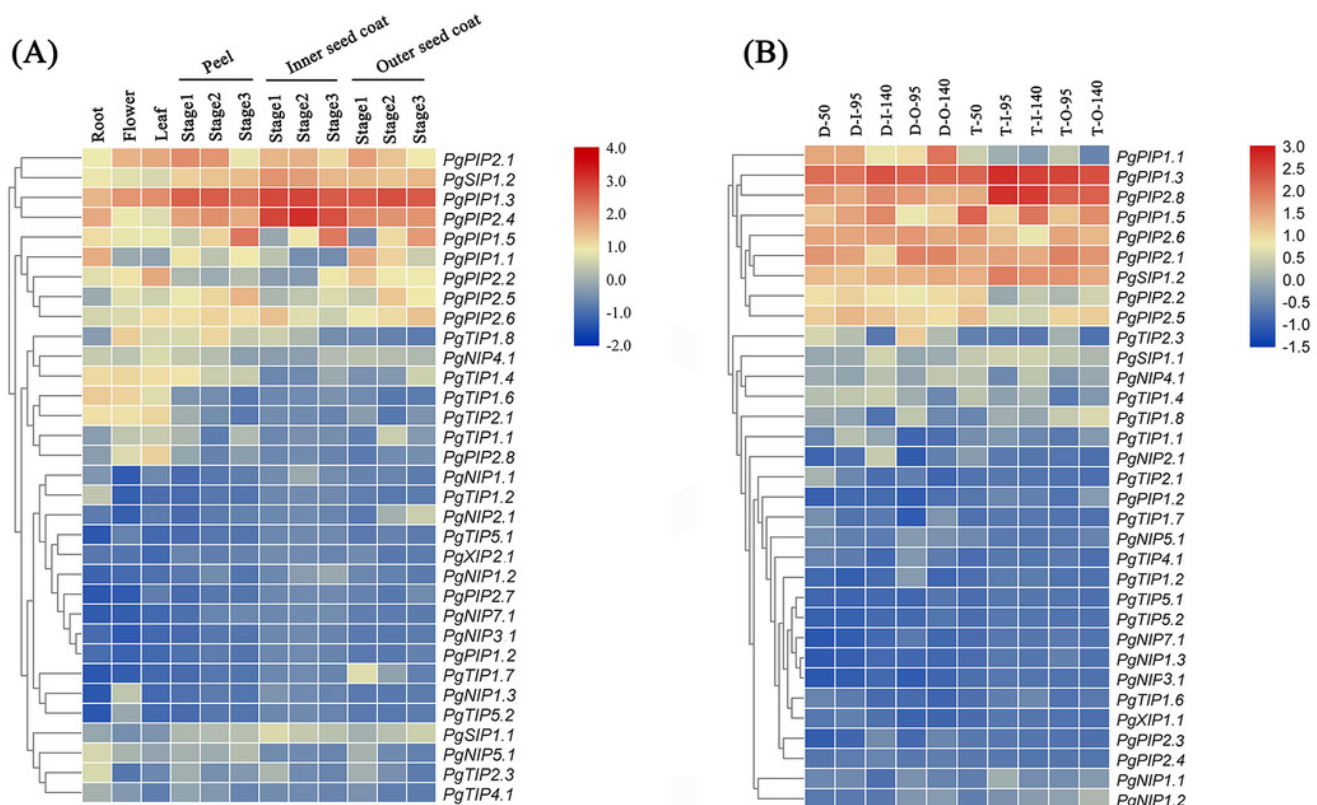



Figure 5

Expression profiles of six Pomegranate PIPs roots in response to water deficit

qRT-PCR was performed to determined the relative transcript level for the six *PgrPIP* genes. Relative expression level was normalized relative to untreated control group (0 h PEG treatment). Error bar represent SE of three independent biological replicated. Asterisk indicate significant differences, *P < 0.05.

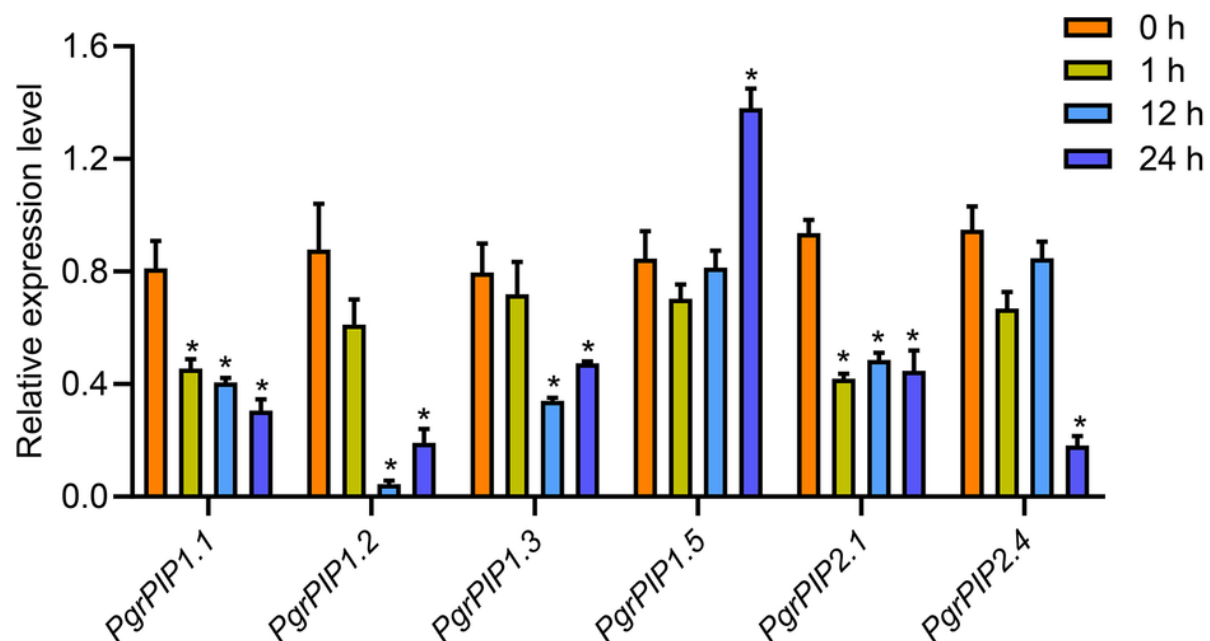



Figure 6

Water accumulation in seed coats of pomegranate.

The abbreviations in the sample designations represent the cultivar names: D for *P. granatum* 'Dabenzi' and T for *P. granatum 'Tunisia*.' The second letter represents tissue: O and I stands for the outer seed coats and the inner seed coat, respectively. The numbers represent the three stages of seed coat development at different days after flowering (DAF). Different letters indicate a significant difference (P<0.05).

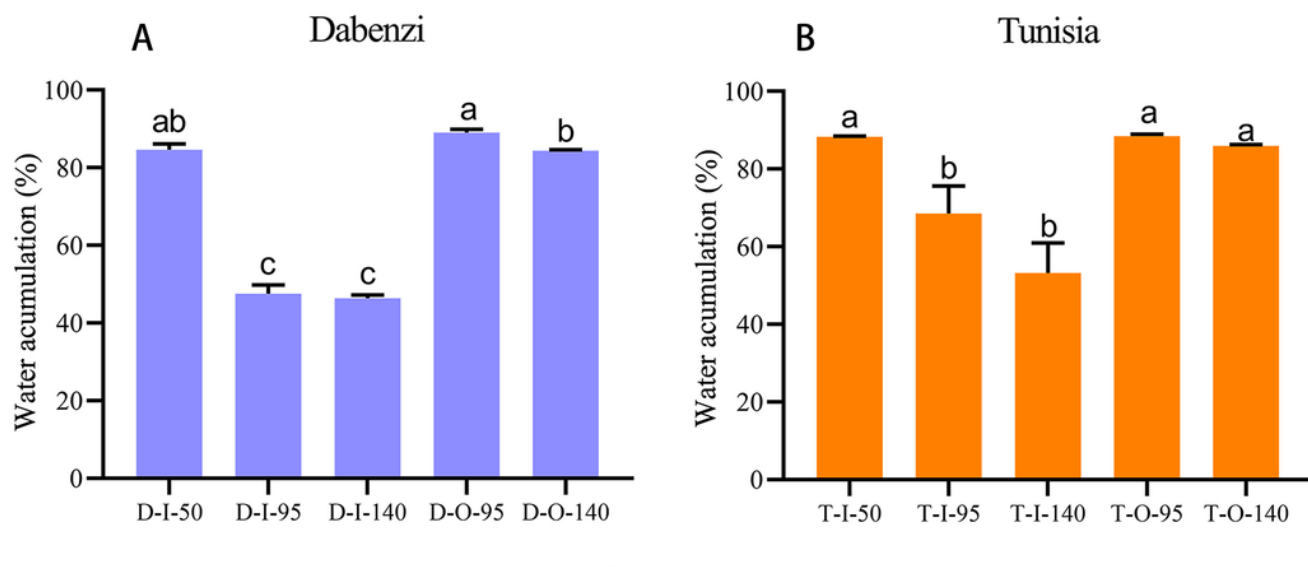



Figure 7

Linear regression between gene expression of PgrAQPs and water accumulation.

Correlation analysis between gene expression of PgrPIP1.3, PgrPIP2.8 and PgrSIP1.2 andwater accumulation in seed coats from P. granatum 'Dabenzi' (A, C, E) and P. granatum'Tunisia' (B, D, F).

'Dabenzi'   'Tunisia'

A   *PgrPIP1.3*

$y=17.65x+718.6$
$r=0.9738$
$P=0.0051$

B   *PgrPIP1.3*

$y= 40.69*x - 956.4$
$R^2=0.8915$
$P=0.0422$

C   *PgrPIP2.8*

$y= 43.01x-1530$
$r=0.9687$
$P=0.0066$

D   *PgrPIP2.8*

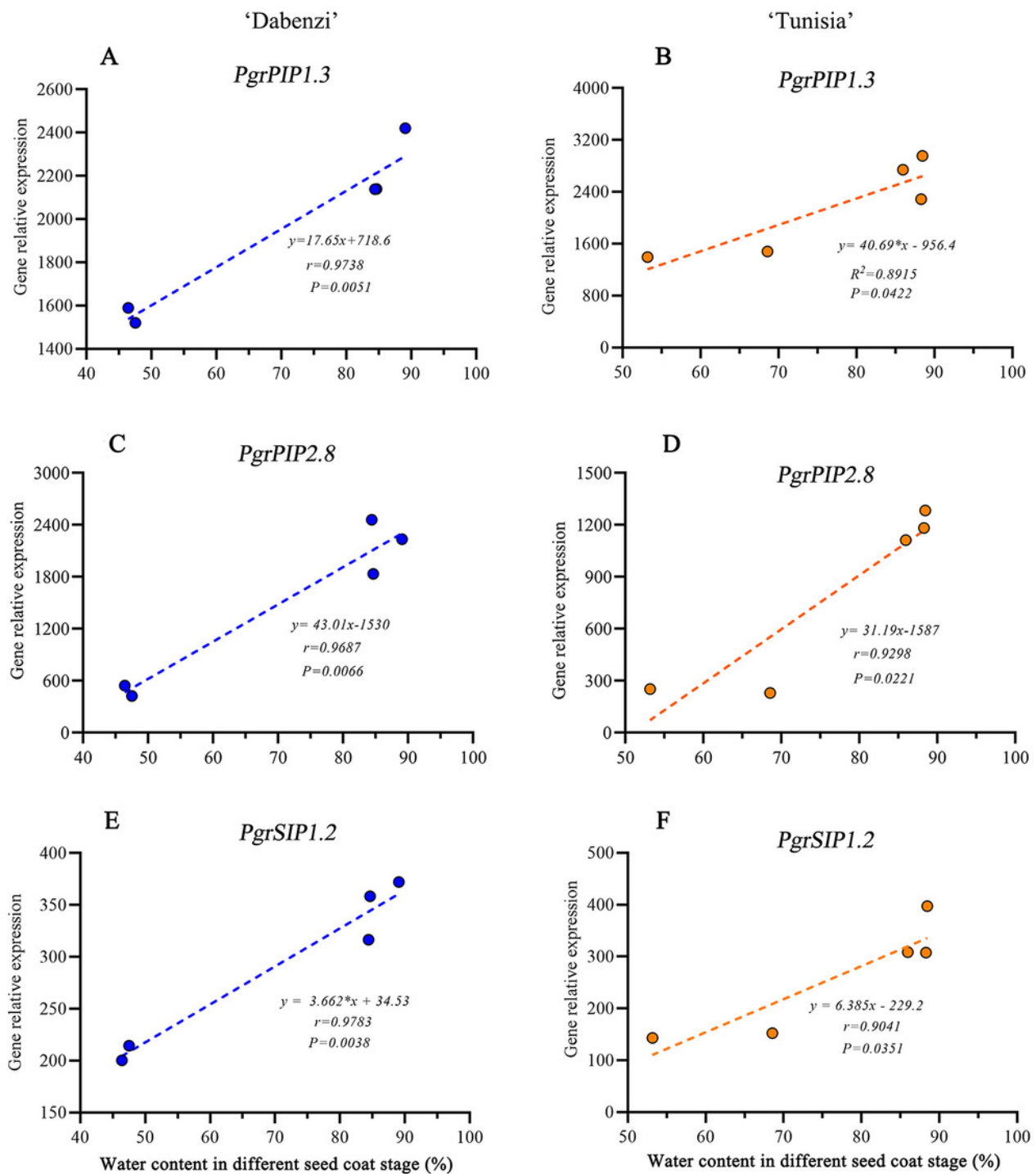
$y= 31.19x-1587$
$r=0.9298$
$P=0.0221$

E   *PgrSIP1.2*

$y = 3.662*x + 34.53$
$r=0.9783$
$P=0.0038$

Water content in different seed coat stage (%)

F   *PgrSIP1.2*

$y = 6.385x - 229.2$
$r=0.9041$
$P=0.0351$

Water content in different seed coat stage (%)

The *Ka* and *Ks* values of duplicated *PgrAQP* gene pairs.

1   **Table 1 The *Ka* and *Ks* values of duplicated *PgrAQP* gene pairs.**

Duplicated gene pairs	Duplicate type	*Ka*	*Ks*	*Ka/Ks*	Time (Mya)	Purify selection
PgrTIP1.2 vs PgrTIP1.5	Segmental	0.074	1.54	0.048	5.12	Yes
PgrTIP3.1 vs PgrTIP3.2	Segmental	0.254	1.19	0.213	3.97	Yes
PgrPIP1.5 vs PgrPIP1.1	Segmental	0.078	0.86	0.090	2.88	Yes
PgrPIP2.5 vs PgrPIP2.2	Segmental	0.123	1.26	0.098	4.21	Yes
PgrPIP2.7 vs PgrPIP2.3	Segmental	0.102	1.38	0.074	4.59	Yes
PgrPIP2.6 vs PgrPIP2.7	Segmental	0.073	1.30	0.057	4.32	Yes
PgrTIP1.2 vs PgrTIP1.1	Segmental	0.096	2.09	0.046	6.97	Yes
PgrTIP1.2 vs PgrTIP1.7	Tandem	0.101	0.52	0.194	1.74	Yes

Note: *Ka*, non-synonymous substitution rate; *Ks*, synonymous substitution rate; Mya, Million years ago.