

1 **Genome wide identification of NAC transcription factors and their role in abiotic**
2 **stress tolerance in *Chenopodium quinoa***

3 Nouf Owdah Alshareef¹, Elodie Rey¹, Holly Khoury¹, Mark Tester¹, Sandra M.
4 Schmöckel^{1,2,*}

5
6 ¹ Division of Biological and Environmental Sciences and Engineering (BESE), King
7 Abdullah University of Science & Technology (KAUST), Thuwal, Saudi Arabia

8 ² Institute of Crop Science, Faculty of Agriculture, University of Hohenheim, Stuttgart,
9 Germany

10

11 * Corresponding author

12 Sandra.schmoeckel@uni-hohenheim.de

13

14 **Short Title:** NAC transcription factors in quinoa

15

16

17 **Author Contributions**

18 **Nouf Owdah Alshareef**

19 Roles: Conceptualization, Data Curation, Formal Analysis, Investigation, Validation,
20 Visualization, Writing – Original Draft Preparation, Writing – review & editing

21 **Elodie Rey**

22 Formal Analysis, Validation

23 **Holly Khoury**

24 Investigation

25 **Mark Tester**

26 Conceptualization, Resources Supervision, Writing – Review & Editing, Funding acquisition

27 **Sandra M. Schmöckel**

28 Conceptualization, Resources, Supervision, Investigation, Formal Analysis, Validation, Writing –
29 original draft, Writing – Review & Editing

30 **Abstract**

31 *Chenopodium quinoa* Willd. (quinoa) is a pseudocereal with high nutritional value and
32 relatively high tolerance to several abiotic stresses, including water deficiency and
33 salt stress, making it a suitable plant for the study of mechanisms of abiotic stress
34 tolerance. NAC (NAM, ATAF and CUC) transcription factors are involved in a
35 range of plant developmental processes and in the response of plants to biotic and
36 abiotic stresses. In the present study, we perform a genome-wide comprehensive
37 analysis of the *NAC* transcription factor gene family in quinoa. In total, we identified
38 107 quinoa *NAC* transcription factor genes, distributed equally between sub-
39 genomes A and B. They are phylogenetically clustered into two major groups and
40 18 subgroups. Almost 75% of the identified *CqNAC* genes were duplicated two to
41 seven times and the remaining 25% of the *CqNAC* genes were found as a single
42 copy. We analysed the transcriptional responses of the identified quinoa *NAC* TF
43 genes in response to various abiotic stresses. The transcriptomic data revealed 28
44 stress responsive *CqNAC* genes, where their expression significantly changed in
45 response to one or more abiotic stresses, including salt, water deficiency, heat and
46 phosphate starvation. Among these stress responsive NACs, some were
47 previously known to be stress responsive in other species, indicating their
48 potentially conserved function in response to abiotic stress across plant species.
49 Six genes were differentially expressed specifically in response to phosphate
50 starvation but not to other stresses, and these genes may play a role in controlling
51 plant responses to phosphate deficiency. These results provide insights into

52 quinoa NACs that could be used in the future for genetic engineering or molecular
53 breeding.

54 **1. Introduction**

55 The NAC transcription factor (TF) family is plant specific and is one of the largest families
56 of TFs in plants. The acronym for NAC is derived from three TFs: NAM, ATAF and CUC,
57 where NAM is an acronym for No Apical Meristem (1), ATAF stands for Arabidopsis
58 Transcription Activator Factor (2), and CUC is a Cup Shaped Cotyledon (1). These three
59 genes share a conserved N-terminal DNA-binding NAC domain. *Arabidopsis thaliana*
60 (Arabidopsis) has 117 NACs, also called ANACs (3), rice has 151 NACs called ONACs
61 (3), and soybean has 152 NACs (4). NAC proteins regulate a wide range of physiological
62 and developmental processes – for example, petunia NAM and Arabidopsis CUC1-2
63 proteins are involved in shoot meristem development (1, 2). Other NAC members are
64 involved in floral morphology (1, 5), plant senescence (6, 7), cell division (8), cell wall
65 synthesis (9) and lateral root development (10).

66 Typical NAC TF proteins are characterized by a highly conserved DNA binding NAC
67 domain at the N-terminus region. The NAC domain spans approximately 150 amino acids
68 and consists of five conserved subdomains (A-E) that make up motifs for DNA binding,
69 protein-protein interaction or transcription factor dimerization (11). Subdomains C and D
70 are conserved and bind to DNA, subdomain A is involved in NAC dimerization, and
71 subdomains B and E are highly divergent and may contribute to the functional diversity
72 of NAC TFs (11, 12). The C-terminal region of NAC TFs is diversified and contains the

73 transcription regulatory domain, which contributes to either transcription activation or
74 repression (11, 13, 14).

75 Some NAC TF proteins possess transmembrane motifs in the C-terminal region that
76 anchor NAC proteins to the intracellular membranes and make it inactive (15). When the
77 NAC protein is activated, it undergoes proteolytic cleavage to release the NAC protein
78 from the membrane to enable its TF function in the nucleus. These membrane-associated
79 NAC TFs are designated as NAC with transmembrane motif1-like (NTL); most of them
80 are associated with plasma membranes and a few are anchored to the endoplasmic
81 reticulum membrane (15). The size of these NTL TFs is larger (from 335-652 amino acids)
82 than the non-membrane associated NAC TFs, which are usually around 320 residues
83 (15). More than 13 NAC TFs in Arabidopsis and six in rice have been described as NTL
84 (15).

85 Some variation in the structure of NAC proteins has been reported. These NACs are
86 called atypical NAC proteins or NAC-like proteins. These variant proteins include some
87 proteins that have only a NAC domain without a C-terminal region (16, 17) and other
88 proteins that have a tandem repeat of the NAC domains (12). Another two variants of
89 NAC proteins are: suppressor of gamma response 1 (SOG1) proteins that have an extra
90 sequence preceding the conserved NAC domain (17, 18), and vascular plant one-zinc-
91 finger (VOZ) proteins that have a DNA-binding zinc finger, transcriptional regulation
92 domain (TRD) at the N-terminal and a NAC domain at the C-terminus (12-14, 19).

93 Quinoa (*Chenopodium quinoa Willd.*, $2n=4x=36$) is a dicotyledonous allotetraploid
94 pseudocereal plant with 18 chromosomes (20). It belongs to the Amaranthaceae family,
95 which also includes other economically important crops such as *Beta*

96 *vulgaris* (beet), *Spinacia oleracea* (spinach) and *Amaranthus hypochondriacus*
97 (amaranth). Quinoa has recently gained much attention due to its high nutritional value
98 and high tolerance to environmental stresses. The grain of quinoa has a balanced ratio
99 of carbohydrates, lipids and protein, a higher content of essential amino acids and is rich
100 in iron and vitamins such as vitamin B1, B6 and E (21). In addition to quinoa's nutritional
101 value, it has a high tolerance to different abiotic stresses, including low temperature,
102 drought and salinity (22, 23). Quinoa maintains the highest biomass when grown at 100
103 mM NaCl and the biomass is reduced by up to 50% when it grows under 500 mM NaCl
104 (24). These traits make quinoa a good model for understanding the mechanisms of stress
105 tolerance.

106 The recent completion of the genome of quinoa has allowed to study NAC genes at the
107 whole genome level. To date, only the HSP17 and WRKY gene families have been
108 systematically analysed in quinoa (25, 26). Here, we identify 107 NAC TF genes in the
109 quinoa genome and investigate their transcriptional responses to different stresses
110 including salt, drought and heat. We perform a systematic analysis of NAC TFs in quinoa
111 using the available high-quality reference genome sequence of quinoa (27). This study
112 provides a basis for future functional characterization of NAC TFs in quinoa that could be
113 used in quinoa stress-tolerance research.

114

115

116

117 **2. Results and discussion**

118 **2.1 Identification of the CqNAC transcription factors family in quinoa**

119 In order to identify NAC TFs in the genome of quinoa, we used a combination of several
120 search methods that have been used previously to identify NACs in different plant
121 species. We employed especially those methods that identified NACs in species with
122 duplicated genomes, such as *Panicum virgatum* (Switchgrass), *Populus trichocarpa*
123 (black cottonwood) and *Gossypium raimondii* (cotton) (28-31).

124 We used the reference genome of quinoa accession QQ74 (27) to identify NAC TFs in
125 quinoa. First, we used the Hidden Markov Model (HMM) profile of the protein family
126 (Pfam) “NAC domain” (PF02365) as a query to search for proteins containing NAC
127 domains in the genome of quinoa using the phytozome database version12 (phytozome,
128 <http://www.phytozome.net>, version 12). This search identified 104 putative NACs
129 proteins. Then, we employed a basic local alignment tool for proteins (BLASTp) against
130 QQ74 peptides using the NAC domain sequences of the 110 Arabidopsis NACs. This
131 search identified 109 putative NACs (six more than the Pfam search using the phytozome
132 database). To find further NACs in the genome of quinoa, we used the TF prediction tool
133 from the Plant Transcription Factor Database (PTFDB, <http://plantfdb.cbi.pku.edu.cn/>,
134 version 4.0) (32) to predict all TFs in the peptide sequences of quinoa as an input (44,776
135 peptide sequences). We were able to identify 2093 TFs belonging to different families (S1
136 Table); among those, 107 CqTFs belong to the NAC family. As a quality control and to
137 verify the reliability of our search, we ran the same search method for Arabidopsis and
138 verified that we received the expected number of Arabidopsis NACs.

139 In order to identify the total number of NAC TFs in quinoa, we generated a combined list
140 of the three search approaches (Pfam, BLASTp and PTFDB) (Fig 1, S2 Table). This list
141 consisted of 112 putative CqNACs. All of the three search methods resulted in a similar
142 number of CqNACs (Fig 1). There were 103 CqNACs common to all of the search
143 methods, with two CqNACs identified only in the PTFDB search and five CqNACs
144 identified in the BLASTp search only. To further confirm these putative CqNAC genes,
145 we scanned the protein sequence of all 112 putative CqNAC proteins for the presence of
146 the NAC domain using the InterProScan program ([http://www.ebi.ac.uk/Tools/
147 InterProScan/](http://www.ebi.ac.uk/Tools/InterProScan/)), the hmmscan function of the HMMER web server
148 (<https://www.ebi.ac.uk/Tools/hmmer/search/hmmscan>, HmmerWeb version 2.28.0) (33)
149 and conserved domain search (CDS) of the NCBI database. The presence of the NAC
150 domain was confirmed in 107 CqNACs sequences, whereas the NAC domain was not
151 present in the remaining five putative CqNAC proteins (all had been identified from the
152 BLASTp search) (Fig 1, S2 Table). All of these search methods suggest that only 107
153 CqNACs genes exist in the reference genome of quinoa. The confirmed 107 putative
154 quinoa NACs are listed in Table 1 and are used for further analyses.

155

156 **Fig 1. Scheme representing the bioinformatic approaches used to identify NAC**
157 **genes in the genome of quinoa.**

158 A combined list of three search approaches was generated (using Pfam, BLASTp and the
159 TF prediction tool of the Plant Transcription Factor Database (PTFDB)). Scanned of all
160 112 putative CqNAC proteins for the presence of the NAC domain using, the hmmscan
161 function of the HMMER web server and conserved domain search (CDS) of the NCBI
162 database. A total of 107 CqNAC proteins were identified.

163

164

165

166

167 **Table 1. NAC transcription factor family in quinoa.**
 168 Illustrating the sub-genome location, start and end of each gene, gene length, number of
 169 exons and protein length. CqNACs in bold with (*) sign are membrane associated CqNAC
 170 predicted by TMHMM server v.2.0. (<http://www.cbs.dtu.dk/services/TMHMM/>).
 171

| Nb. | Gene ID | Gene Name | Brief Description | chromosomal location_subgenome | START | END | gene length (bp) | Nb. exons | protein size (aa) |
|-----|--------------------|----------------|---|--------------------------------|------------------|------------------|------------------|-----------|-------------------|
| 1 | AUR62017896 | CqNAC02 | NAC002:NACdomain-containingprotein2 (ATAF-1) | cq15_A | 20376810 | 20378712 | 1903 | 3 | 224 |
| 2 | AUR62003583 | CqNAC02 | NAC002:NACdomain-containingprotein2 (ATAF-1) | cq9_B | 3900006 | 3901944 | 1939 | 3 | 224 |
| 3 | AUR62000536 | CqNAC02 | NAC002:NACdomain-containingprotein2 (ATAF-1) | cq12_A | 6131839 | 6134224 | 2386 | 4 | 248 |
| 4 | AUR62015520 | CqNAC07 | NAC007: NAC domain-containing protein 7 | cq11_B | 21032546 | 21034478 | 1933 | 4 | 248 |
| 5 | AUR62002191 | CqNAC07 | NAC007: NAC domain-containing protein 7 | cq7_A | 65990389 | 65992323 | 1935 | 4 | 249 |
| 6 | AUR62007469 | CqNAC07 | NAC007: NAC domain-containing protein 7 | cq13_A | 728734 | 730378 | 1645 | 3 | 370 |
| 7 | AUR62039927 | CqNAC07 | NAC007: NAC domain-containing protein 7 | cq16_B | 78627158 | 78632662 | 5505 | 4 | 370 |
| 8 | AUR62004572 | CqNAC08 | NAC008: NAC domain-containing protein 8 | cq1_B | 120553744 | 120556805 | 3062 | 6 | 388 |
| 9 | AUR62022685 | CqNAC08 | NAC008: NAC domain-containing protein 8 | cq10_A | 5677119 | 5680213 | 3095 | 6 | 374 |
| 10 | AUR62002922 | CqNAC08 | NAC008: NAC domain-containing protein 8 | cq6_B | 15831647 | 15834922 | 3276 | 7 | 339 |
| 11 | AUR62035283 | CqNAC08 | NAC008: NAC domain-containing protein 8 | cq14_A | 9028304 | 9032559 | 4256 | 7 | 365 |
| 12 | AUR62029330 | CqNAC10 | NAC010: NAC domain-containing protein 10 | cq12_A | 51513524 | 51515901 | 2378 | 5 | 328 |
| 13 | AUR62031305 | CqONAC010 | ONAC010: NAC transcription factor ONAC010 | cq1_B | 25489514 | 25491226 | 1713 | 2 | 158 |
| 14 | AUR62027390 | CqONAC010 | ONAC010: NAC transcription factor ONAC010 | cq2_A | 40025813 | 40027566 | 1754 | 2 | 168 |
| 15 | AUR62016213 | CqONAC010 | ONAC010: NAC transcription factor ONAC010 | cq7_A | 75752855 | 75754429 | 1575 | 3 | 389 |
| 16 | AUR62011638 | CqONAC010 | ONAC010: NAC transcription factor ONAC010 | cq15_A | 13499168 | 13500708 | 1541 | 3 | 414 |
| 17 | AUR62033952 | CqNAC17 | NAC017: NAC domain-containing protein 17 * | cq0_ | 184747668 | 184752783 | 5116 | 5 | 530 |
| 18 | AUR62029057 | CqNAC17 | NAC017: NAC domain-containing protein 17 * | cq2_A | 37120934 | 37126563 | 5630 | 4 | 524 |
| 19 | AUR62035514 | CqNAC21 | NAC021: NAC domain-containing protein 21/22 | cq5_B | 51399005 | 51402623 | 3619 | 2 | 153 |
| 20 | AUR62043497 | CqNAC21 | NAC021: NAC domain-containing protein 21/22 | cq12_A | 29606703 | 29610871 | 4169 | 2 | 155 |
| 21 | AUR62011853 | CqNAC21 | NAC021: NAC domain-containing protein 21/22 | cq8_A | 1858490 | 1862276 | 3787 | 2 | 145 |
| 22 | AUR62024698 | CqNAC21 | NAC021: NAC domain-containing protein 21/22 | cq9_B | 11061255 | 11065302 | 4048 | 3 | 149 |
| 23 | AUR62001365 | CqNAC25 | NAC025: NAC transcription factor 25 | cq7_A | 99573268 | 99576243 | 2976 | 3 | 307 |
| 24 | AUR62014768 | CqNAC25 | NAC025: NAC transcription factor 25 | cq1_B | 30397116 | 30398602 | 1487 | 3 | 343 |
| 25 | AUR62016954 | CqNAC25 | NAC025: NAC transcription factor 25 | cq2_A | 20121444 | 20123003 | 1560 | 3 | 327 |
| 26 | AUR62005821 | CqNAC29 | NAC029: NAC transcription factor 29 | cq14_A | 57750521 | 57756348 | 5828 | 5 | 288 |
| 27 | AUR62032816 | CqNAC29 | NAC029: NAC transcription factor 29 | cq6_B | 5793401 | 5794833 | 1433 | 3 | 272 |
| 28 | AUR62022690 | CqNAC30 | NAC030: NAC domain-containing protein 30 | cq10_A | 5723986 | 5732908 | 8923 | 5 | 571 |
| 29 | AUR62004648 | CqNAC30 | NAC030: NAC domain-containing protein 30 | cq1_B | 121459828 | 121461574 | 1747 | 3 | 334 |
| 30 | AUR62024101 | CqNAC31 | NAC031: Protein CUP-SHAPED COTYLEDON 3 | cq10_A | 59698378 | 59699514 | 1137 | 4 | 248 |
| 31 | AUR62035165 | CqNAC31 | NAC031: Protein CUP-SHAPED COTYLEDON 3 | cq3_B | 1657868 | 1658173 | 306 | 1 | 101 |
| 32 | AUR62043713 | CqNAC37 | NAC037: NAC domain-containing protein 37 | cq0_ | 14884377 | 14886152 | 1776 | 3 | 265 |
| 33 | AUR62043235 | CqNAC37 | NAC037: NAC domain-containing protein 37 | cq9_B | 41526 | 43301 | 1776 | 3 | 265 |
| 34 | AUR62034813 | CqNAC41 | NAC041: NAC domain-containing protein 41 | cq16_B | 11304423 | 11306000 | 1578 | 3 | 228 |
| 35 | AUR62029543 | CqNAC41 | NAC041: NAC domain-containing protein 41 | cq8_A | 26693056 | 26699989 | 6934 | 5 | 365 |

| | | | | | | | | | |
|----|--------------------|----------------|---|---------------|------------------|------------------|-------------|----------|------------|
| 36 | AUR62026904 | CqNAC42 | NAC042: NAC domain-containing protein 42 | cq4_A | 44976454 | 44984863 | 8410 | 3 | 344 |
| 37 | AUR62040889 | CqNAC42 | NAC042: NAC domain-containing protein 42 | cq1_B | 72702858 | 72708102 | 5245 | 3 | 213 |
| 38 | AUR62004554 | CqNAC43 | NAC043: NAC domain-containing protein 43 | cq1_B | 120299216 | 120301237 | 2022 | 3 | 334 |
| 39 | AUR62019936 | CqNAC43 | NAC043: NAC domain-containing protein 43 | cq16_B | 72986058 | 72987726 | 1669 | 3 | 341 |
| 40 | AUR62022669 | CqNAC43 | NAC043: NAC domain-containing protein 43 | cq10_A | 5477634 | 5479368 | 1735 | 3 | 250 |
| 41 | AUR62000492 | CqNAC45 | NAC045: NAC domain-containing protein 45 | cq12_A | 5500539 | 5501585 | 1047 | 2 | 327 |
| 42 | AUR62004778 | CqNAC45 | NAC045: NAC domain-containing protein 45 | cq12_A | 25145721 | 25148970 | 4056 | 6 | 576 |
| 43 | AUR62003400 | CqNAC45 | NAC045: NAC domain-containing protein 45 | cq9_B | 1759783 | 1763912 | 4130 | 6 | 573 |
| 44 | AUR62037478 | CqNAC45 | NAC045: NAC domain-containing protein 45 | cq5_B | 62910115 | 62914170 | 3250 | 3 | 241 |
| 45 | AUR62015629 | CqNAC56 | NAC056: NAC transcription factor 56 | cq5_B | 3821692 | 3822921 | 1230 | 3 | 302 |
| 46 | AUR62029331 | CqNAC56 | NAC056: NAC transcription factor 56 | cq12_A | 51490439 | 51492813 | 2375 | 4 | 327 |
| 47 | AUR62001531 | CqNAC60 | NAC60: NAC domain-containing protein 60 * | cq7_A | 101732643 | 101734168 | 1526 | 3 | 330 |
| 48 | AUR62007071 | CqNAC66 | NAC066: NAC domain-containing protein 66 | cq13_A | 5441599 | 5443027 | 1429 | 3 | 279 |
| 49 | AUR62016200 | CqNAC67 | NAC67: NAC domain-containing protein 67 | cq7_A | 75912843 | 75914874 | 2032 | 3 | 257 |
| 50 | AUR62036626 | CqNAC68 | NAC68: NAC domain-containing protein 68 | cq3_B | 56560424 | 56570523 | 10100 | 4 | 279 |
| 51 | AUR62025580 | CqNAC68 | NAC68: NAC domain-containing protein 68 | cq18_B | 24671246 | 24673006 | 1761 | 3 | 342 |
| 52 | AUR62025636 | CqNAC68 | NAC68: NAC domain-containing protein 68 | cq18_B | 23977517 | 23979442 | 1926 | 3 | 354 |
| 53 | AUR62015630 | CqNAC72 | NAC072: NAC domain-containing protein 72 | cq5_B | 3807048 | 3808199 | 1152 | 4 | 278 |
| 54 | AUR62022486 | CqNAC72 | NAC072: NAC domain-containing protein 72 | cq1_B | 96255823 | 96265494 | 9672 | 6 | 386 |
| 55 | AUR62029344 | CqNAC72 | NAC072: NAC domain-containing protein 72 | cq12_A | 51286101 | 51287415 | 1315 | 4 | 298 |
| 56 | AUR62021364 | CqNAC73 | NAC073: NAC domain-containing protein 73 | cq1_B | 16997945 | 17004953 | 7009 | 6 | 167 |
| 57 | AUR62033276 | CqNAC73 | NAC073: NAC domain-containing protein 73 | cq2_A | 52375299 | 52382203 | 6905 | 5 | 319 |
| 58 | AUR62031086 | CqNAC73 | NAC073: NAC domain-containing protein 73 | cq2_A | 5673254 | 5676404 | 3151 | 2 | 251 |
| 59 | AUR62014879 | CqNAC73 | NAC073: NAC domain-containing protein 73 | cq1_B | 32318489 | 32327115 | 8627 | 2 | 252 |
| 60 | AUR62026259 | CqNAC73 | NAC073: NAC domain-containing protein 73 | cq14_A | 24108385 | 24115692 | 7308 | 4 | 191 |
| 61 | AUR62043434 | CqNAC73 | NAC073: NAC domain-containing protein 73 | cq0_ | 19574734 | 19575424 | 691 | 3 | 138 |
| 62 | AUR62036063 | CqNAC73 | NAC073: NAC domain-containing protein 73 | cq15_A | 45801852 | 45808337 | 6486 | 3 | 204 |
| 63 | AUR62041525 | CqNAC78 | NAC078: NAC domain-containing protein 78 * | cq5_B | 57361157 | 57364321 | 3165 | 6 | 550 |
| 64 | AUR62043491 | CqNAC78 | NAC078: NAC domain-containing protein 78 | cq12_A | 28893343 | 28901372 | 8030 | 4 | 393 |
| 65 | AUR62041539 | CqNAC78 | NAC078: NAC domain-containing protein 78 * | cq12_A | 20501797 | 20505411 | 3615 | 6 | 556 |
| 66 | AUR62043741 | CqNAC78 | NAC078: NAC domain-containing protein 78 | cq1_B | 84648679 | 84662099 | 13421 | 4 | 355 |
| 67 | AUR62006885 | CqNAC81 | NAC081: Protein ATAF2 | cq5_B | 74997825 | 75001903 | 4079 | 4 | 329 |
| 68 | AUR62006407 | CqNAC82 | NAC082: NAC domain-containing protein 82 | cq7_A | 90788443 | 90789561 | 1119 | 1 | 372 |
| 69 | AUR62029755 | CqNAC82 | NAC082: NAC domain-containing protein 82 | cq17_B | 59895805 | 59896911 | 1107 | 1 | 368 |
| 70 | AUR62006837 | CqNAC82 | NAC082: NAC domain-containing protein 82 | cq5_B | 75576113 | 75577156 | 1044 | 2 | 319 |
| 71 | AUR62000493 | CqNAC82 | NAC082: NAC domain-containing protein 82 | cq12_A | 5522447 | 5523436 | 990 | 1 | 329 |
| 72 | AUR62028872 | CqNAC83 | NAC083: NAC domain-containing protein 83 | cq6_B | 3320845 | 3322156 | 1312 | 4 | 257 |
| 73 | AUR62005686 | CqNAC83 | NAC083: NAC domain-containing protein 83 | cq14_A | 55484398 | 55485672 | 1275 | 4 | 256 |
| 74 | AUR62018752 | CqNAC86 | NAC086: NAC domain-containing protein 86 | cq16_B | 75482814 | 75484845 | 2032 | 4 | 325 |
| 75 | AUR62007247 | CqNAC86 | NAC086: NAC domain-containing protein 86 | cq13_A | 3165461 | 3167486 | 2026 | 4 | 327 |
| 76 | AUR62010080 | CqNAC86 | NAC086: NAC domain-containing protein 86 | cq15_A | 5122099 | 5126510 | 4412 | 4 | 352 |
| 77 | AUR62013074 | CqNAC86 | NAC086: NAC domain-containing protein 86 | cq17_B | 79021049 | 79025517 | 4469 | 4 | 352 |
| 78 | AUR62020251 | CqNAC89 | NAC089: NAC domain-containing protein 89 * | cq18_B | 32521438 | 32523891 | 2454 | 5 | 430 |

| | | | | | | | | | |
|-----------|--------------------|----------------|---|---------------|-----------------|-----------------|-------------|----------|------------|
| 79 | AUR62001522 | CqNAC90 | NAC090: NAC domain-containing protein 90 | cq7_A | 101656423 | 101660122 | 3700 | 4 | 268 |
| 80 | AUR62020261 | CqNAC90 | NAC090: NAC domain-containing protein 90 | cq18_B | 32661219 | 32663097 | 1879 | 3 | 258 |
| 81 | AUR62006223 | CqNAC90 | NAC090: NAC domain-containing protein 90 | cq15_A | 2810155 | 2814517 | 4363 | 3 | 276 |
| 82 | AUR62036991 | CqNAC90 | NAC090: NAC domain-containing protein 90 | cq17_B | 56071227 | 56073064 | 1838 | 2 | 151 |
| 83 | AUR62039815 | CqNAC94 | ANAC094: Putative NAC domain-containing protein 94 | cq4_A | 38381888 | 38383418 | 1531 | 3 | 385 |
| 84 | AUR62034400 | CqNAC94 | ANAC094: Putative NAC domain-containing protein 94 | cq1_B | 55324061 | 55325551 | 1491 | 3 | 385 |
| 85 | AUR62027152 | CqNAC94 | ANAC094: Putative NAC domain-containing protein 94 | cq15_A | 1809369 | 1810776 | 1408 | 3 | 222 |
| 86 | AUR62012948 | CqNAC94 | ANAC094: Putative NAC domain-containing protein 94 | cq17_B | 80277875 | 80278498 | 624 | 2 | 146 |
| 87 | AUR62040120 | CqNAC94 | ANAC094: Putative NAC domain-containing protein 94 | cq6_B | 34103882 | 34105819 | 1938 | 5 | 390 |
| 88 | AUR62032488 | CqNAC94 | ANAC094: Putative NAC domain-containing protein 94 * | cq14_A | 41611790 | 41618469 | 6680 | 8 | 516 |
| 89 | AUR62038748 | CqNAC98 | NAC098: Protein CUP-SHAPED COTYLEDON 2 | cq1_B | 106129053 | 106135097 | 6045 | 4 | 294 |
| 90 | AUR62020634 | CqNAC98 | NAC098: Protein CUP-SHAPED COTYLEDON 2 | cq4_A | 16178718 | 16185086 | 6369 | 4 | 297 |
| 91 | AUR62019477 | CqNAC98 | NAC098: Protein CUP-SHAPED COTYLEDON 2 | cq12_A | 48919545 | 48921791 | 2247 | 3 | 400 |
| 92 | AUR62036158 | CqNAC98 | NAC098: Protein CUP-SHAPED COTYLEDON 2 | cq5_B | 11161825 | 11164198 | 2374 | 5 | 324 |
| 93 | AUR62034435 | CqNAC100 | NAC100: NAC domain-containing protein 100 | cq1_B | 56919532 | 56920931 | 1400 | 3 | 341 |
| 94 | AUR62018102 | CqNAC100 | NAC100: NAC domain-containing protein 100 | cq4_A | 35692016 | 35693362 | 1347 | 3 | 333 |
| 95 | AUR62021652 | CqNAC100 | NAC100: NAC domain-containing protein 100 | cq8_A | 9672471 | 9674125 | 1655 | 3 | 361 |
| 96 | AUR62008438 | CqNAC100 | NAC100: NAC domain-containing protein 100 | cq16_B | 4681091 | 4683124 | 2034 | 3 | 364 |
| 97 | AUR62021621 | CqNAC100 | NAC100: NAC domain-containing protein 100 | cq8_A | 10126416 | 10133113 | 6698 | 4 | 198 |
| 98 | AUR62028103 | CqNTL9 | NTL9: Protein NTM1-like 9 * | cq11_B | 68788278 | 68792879 | 4602 | 6 | 630 |
| 99 | AUR62005966 | CqNTL9 | NTL9: Protein NTM1-like 9 * | cq7_A | 70111368 | 70115912 | 4545 | 6 | 631 |
| 100 | AUR62013541 | CqSMB | SMB: Protein SOMBRERO | cq3_B | 65117611 | 65120613 | 3003 | 3 | 346 |
| 101 | AUR62026741 | CqSMB | SMB: Protein SOMBRERO | cq10_A | 20749557 | 20757214 | 7658 | 4 | 283 |
| 102 | AUR62005772 | CqFEZ | FEZ: Protein FEZ | cq14_A | 56770446 | 56773311 | 2866 | 3 | 446 |
| 103 | AUR62044373 | CqFEZ | FEZ: Protein FEZ | cq6_B | 5175069 | 5177844 | 2776 | 3 | 454 |
| 104 | AUR62017681 | CqBRN2 | BRN2: Protein BEARSKIN2 | cq0_ | 137754423 | 137760607 | 6185 | 4 | 418 |
| 105 | AUR62000898 | CqNAM-B1 | NAM-B1: NAC transcription factor NAM-B1 | cq12_A | 10832642 | 10835485 | 2844 | 3 | 418 |
| 106 | AUR62006836 | CqNAC103 | NAC103: NAC domain containing protein 103 | cq5_B | 75580825 | 75581836 | 1011 | 2 | 315 |
| 107 | AUR62004773 | CqNAC28 | NAC domain containing protein 28 | cq5_B | 62725181 | 62726259 | 1078 | 2 | 140 |

173 Quinoa NAC genes encode proteins ranging in size from 101 to 631 amino acids (aa) in
174 length, with an average of 322 aa.

175 Several NAC TFs are present as membrane-bound transcription factors (MTFs) (15). To
176 identify membrane associated NAC TFs in quinoa, we scanned all of the full length
177 CqNAC peptide sequences for the presence of α -helical transmembrane (TM) motives
178 using the TMHMM server v.2.0. (<http://www.cbs.dtu.dk/services/TMHMM/>). We identified
179 nine putative membrane-bound CqNACs proteins containing α -helical TM motives in the
180 C-terminal region (Table 1).

181 The number of CqNACs we identified in this study is less than the number of NACs in
182 Arabidopsis and rice. We expected at least double the number of NACs due to the
183 tetraploid nature of the quinoa genome. However, we found that the total number of
184 predicted TFs in the genome of quinoa is in fact even slightly less than the total number
185 of TFs in Arabidopsis (2093 CqTFs in quinoa compared with 2296 TFs in Arabidopsis, S1
186 Fig, S1 Table). The number of TFs in each family is almost 0.8-fold (median) compared
187 with Arabidopsis, with only two families showing 3 and 5 times more TFs in quinoa
188 compared with Arabidopsis i.e., the FAR1 and LFY families, respectively. Moreover, fewer
189 sequences belonging to the heat shock proteins family 70 (HSP70) were identified in
190 quinoa compared with Arabidopsis (quinoa has 16 HSP70 members, while Arabidopsis
191 has 18 members) (26). However, by comparing the number of CqNAC genes in quinoa
192 with its close relatives, quinoa has more NACs than spinach (*Spinacia oleracea*) and
193 sugar beet (*Beta vulgaris*), which have 80 and 59 NACs genes, respectively.

194 **2.2 Chromosomal location and gene duplication of quinoa CqNAC genes**

195 In total, 103 CqNACs are localized to the 18 chromosomes of quinoa and only four
196 CqNACs could not be mapped to any chromosome and are therefore assigned to
197 chromosome zero (Table 1). CqNACs genes appear to be equally distributed between
198 sub-genome A and B (S2a Fig); however, CqNACs are unequally distributed across the
199 chromosomes (S2b Fig). The largest number of CqNACs genes are localized to
200 chromosome 1 (13 CqNACs, ~12.38 %), followed by chromosome 12 (12 CqNACs,
201 ~11.43 %), while chromosome 11 and chromosome 13 have the smallest number of
202 CqNACs (only two CqNACs ~1.9 % are localized to chromosome 11 and three CqNACs
203 ~2.86 % are localized to chromosome 13) (S2b Fig).

204 We noticed that in most cases, we found two or more quinoa *CqNAC* genes for every
205 Arabidopsis orthologue and thus we grouped quinoa CqNACs as duplicated and un-
206 duplicated genes. The duplicated group consists of 30 duplicated CqNAC genes (each
207 gene has from 2 to 7 copies) giving a total number of 97 genes distributed between
208 subgenome A and subgenome B. The un-duplicated CqNACs consist of 10 CqNACs
209 genes (Fig 2a,b). In total, only 40 Arabidopsis NAC genes have orthologues in the
210 genome of quinoa (36.5% of Arabidopsis NACs).

211

212 **Fig 2. Duplicated and non-duplicated quinoa *CqNAC* genes.**

213 a) Copy number of each quinoa *CqNAC* gene. b) Pie chart shows the percentage of
214 duplicated and unduplicated *CqNAC*s. The percentage of genes are calculated based on
215 the total number of *CqNAC*s orthologues to Arabidopsis *NAC*s (40 *NAC* genes).

216

217 We performed synteny analysis using MCScanX to reveal the relationship between the
218 duplicated CqNACs, and to determine if these duplicated genes form homologous pairs
219 and if these homologs are located between subgenomes A and B (to form homoeologous

220 pairs). Synteny analysis revealed that different types of homology occurs between the
221 duplicated genes, which we classify as: (1) genes that form a homoeologous pair located
222 in the same phylogenetic cluster (S3a Table) – in quinoa, this applies to 52 genes (26
223 pairs, 48.6% of all of the genes); (2) genes that appear to form a homoeologous pair but
224 are located in different phylogenetic clusters - 18 genes in quinoa (16.8%), (S3b Table);
225 (3) genes that have more than one homolog located in different clusters - 10 genes in
226 quinoa (9.3%) (S3c Table); and (4) genes that have no homolog according to our analyses
227 - 27 genes in quinoa (25.2%) (S3d Table). Although some of these genes are duplicated,
228 such as the orthologue of *AtJUB1*, none of the duplicated copies have a homoeolog. The
229 absence of some homoeolog copies in some of the *CqNAC* genes suggests that gene
230 loss occurred during the evolution of quinoa, causing some gene loss in the NAC family.
231 To confirm this, the sequences of the two progenitor genomes needs to be investigated.
232 Similar findings have been observed in the WRKY TF family in quinoa, where some of
233 the homologous copies of some of the WRKY genes have been lost (25).

234 **2.3 Phylogenetic analysis of the NAC TFs family in quinoa and Arabidopsis**

235 We constructed a Neighbor-Joining (N-J) phylogenetic tree to study the relationship
236 between quinoa CqNACs. The NAC domain of 107 quinoa NAC proteins was used in the
237 phylogenetic construction. From the phylogenetic tree (and according to the bootstrap
238 values), quinoa NAC proteins can be classified into two major groups (Fig 3a). Group I
239 has the largest number of CqNACs (96 CqNACs) and is subdivided into 15 subgroups
240 (Fig 3b). Group II contains 11 CqNACs.

241

242 **Fig 3. Phylogenetic tree of quinoa CqNACs.**

243 a) The phylogenetic tree was constructed in MEGA 7.0 using the Neighbour-Joining (NJ)
244 method with 1000 bootstrap interactions based on the multiple sequence alignment of
245 107 amino acid sequences of NAC domains belonging to the NAC proteins from
246 *Chenopodium quinoa*. Nodes that have a bootstrap value of less than 50 are collapsed.
247 The tree divided CqNACs into two major groups: Group I (with blue background) and
248 Group II (with pink background). Group I is further subdivided into 15 subgroups that have
249 different background colours. b) Details of the subgroups from the phylogenetic tree in
250 (a).

251

252 Generally, most of the quinoa CqNACs form sister pairs (there are 49 pairs) except for
253 nine CqNACs that occur as a single CqNAC (Fig 3b). These results indicate that genes
254 in each subgroup might originate from the same duplication event. Similar phylogenetic
255 topologies have been also found in other plants with duplicated genomes, such as in
256 switchgrass (28). For most NACs, closely related members in the same phylogenetic
257 subgroup share a similar exon/intron structure (number and length), with few exceptions
258 (Fig 4).

259

260 **Fig 4. Exon/ intron structure of quinoa CqNAC genes belonging to different**
261 **phylogenetic subgroups.**

262

263 To find the phylogenetic relationship between quinoa and Arabidopsis, we constructed
264 another phylogenetic tree, in this case from the alignments of the full-length sequences
265 of NAC proteins from Arabidopsis (110 seq.) and quinoa (107 seq.). (Fig 5). The
266 phylogenetic tree divided NACs into different sub-groups, consistent with the previous
267 classification (34). The bootstrap values were sometimes low; however, this was also
268 reported in previous studies (35-39). NAC TFs from the same phylogenetic group are
269 likely to have a similar function. For instance, the NAM subfamily includes NACs that

270 function in shoot boundary formation (CUC1/2/3) (40, 41), while NACs belonging to
271 subfamily OsNAC7 are involved in secondary wall formation (Vascular related NAC-
272 Domain proteins “VNDs” (42) and in root development (SMBs (43)). The OsNAC8 family
273 are another example, which includes NACs involved in stress response (ANAC019,
274 ANAC055 and ANAC072) (44). Interestingly, there is a subgroup of Arabidopsis NACs
275 that are absent in quinoa (subgroups 8, 16 and 18), might have been lost during quinoa
276 evolution. Subgroups 8, 16 and 18 contain mostly unknown NACs, where the function
277 has not been described yet. There are only four genes previously described, ANAC005
278 involved in xylem formation (45), NTM1 involved in cell division (46) and two genes
279 involved in abiotic stress, NTM2 and NAC67 (47, 48). Quinoa has a high tolerance to
280 salinity and water deficit; it is possible, that the absence of these genes contributes to
281 quinoa’s tolerance. There are only two subgroups of NACs for which quinoa has a larger
282 number of members than Arabidopsis, and there are three groups that quinoa is missing
283 (S4 Table). This may explain why quinoa has less NACs than Arabidopsis.

284

285 **Fig 5. Phylogenetic relationship of NAC proteins from Arabidopsis and Quinoa.**

286 The phylogenetic tree was constructed in MEGA 7.0 using the Neighbor-Joining (NJ)
287 method with 1000 bootstrap interactions based on multiple sequence alignment of 107
288 amino acid sequences of full-length peptide sequence from *Chenopodium quinoa* and
289 110 sequences from Arabidopsis thaliana. The tree was divided into different subgroups
290 and named according to the classification of OoKa et al. (2013). Red circles represent
291 NAC peptides from Arabidopsis and green triangles represent peptides from quinoa. Blue
292 boxes represent subgroups that have members from both Arabidopsis and quinoa, yellow
293 boxes represents subgroups that have larger number of quinoa NACs than Arabidopsis
294 NACs and pink boxes represents subgroups that have members from Arabidopsis only.

295

296 **2.4 Transcriptional analyses of quinoa *CqNAC* genes**

297 **2.4.1 Expression of duplicated genes (functionalization of duplicated genes)**

298 The high proportion of gene duplication among quinoa *CqNAC* genes raises the question
299 of their functional redundancy. During the evolutionary process, duplicated genes may
300 have obtained non-functionalization, neo-functionalization or sub-functional redundancy
301 (49), which could be indicated by differences in their patterns of expression. Thus, we
302 investigated the expression of duplicated *CqNACs* in the shoots and roots of quinoa
303 plants grown in soil and in a hydroponic system under normal growth conditions. We used
304 the available RNAseq data for quinoa accession QQ74 assessed under different stress
305 conditions, such as heat, drought, salinity and low-phosphate. First, to determine if all of
306 the *CqNACs* that we identified in this study were expressed, we calculated their
307 cumulative expression in all of the samples and all of the conditions. We found that five
308 genes were not expressed at all under any of the conditions used in this work, and four
309 genes were expressed only at very low levels (Fragments Per Kilobase of transcript per
310 Million mapped reads, FPKM<1) (S5 Table).

311 Expression analysis of the duplicated genes showed a different expression pattern
312 between the shoots and the roots among the duplicated copies (Fig 6). Some of the
313 duplicated *CqNAC* genes have a similar expression pattern, which could suggest a similar
314 or redundant function. Other duplicated *CqNAC* genes have different expression patterns,
315 suggesting that some of the duplicated copies may have developed different functions.
316 Similar expression patterns were also observed in diploid cotton *Gossypium raimondii*
317 (30) and Tartary buckwheat (*Fagopyrum tataricum*) (50). We also noticed that not all of
318 the duplicated copies are expressed. For example, in two of the genes that are triplicated

319 and one of the genes that are found in four copies, only two copies are expressed and
320 the other two copies are not expressed.

321

322 **Fig 6. Heat map of the expression of duplicated *CqNAC* in the shoots and roots of**
323 **quinoa.**

324 Quinoa plants were grown in hydroponic system under normal growth conditions.
325 a) Genes duplicated two times, b) Genes duplicated three times, c) Genes duplicated four
326 times, d) Genes duplicated more than four times, and e) Genes that have a single copy.
327 The corresponding expression values are in S6 Table.
328

329 NAC TFs have long been known as an important TF family that are involved in regulating
330 plant growth, development and responses to stress (1, 2, 51-53). To detect NACs that
331 are potentially involved in root and shoot development, we identified the expression of all
332 of the 107 putative quinoa NAC TFs genes in the shoots and roots of quinoa plants grown
333 under growth conditions in soil as well as in hydroponic conditions (S6 Table). To find the
334 genes that are potentially involved in root or shoot development, the expression of
335 *CqNACs* in the shoots was compared to their expression in the roots. We chose genes
336 that were significantly differentially expressed (DE) between shoot and root (FDR=0.05),
337 as these might be involved in regulating root or shoot development. Those *CqNACs*
338 comprise 24 genes (19 genes being more expressed in roots “root specific” and five genes
339 being more induced in the shoots “shoot specific”) (Fig 7 and S3 Fig and S6 Table). The
340 expression of *CqNACs* in the shoots and roots of plants grown in soil and in hydroponic
341 showed a consistent expression pattern. However, there are a few differences in the
342 number of significantly differentially expressed (DE) genes. Among those NAC genes that
343 are significantly more expressed in the roots than the shoots (in both soil and hydroponic)
344 is the *AUR62013541* gene, which is the ortholog of the well-characterized *SOMBRERO*

345 (*SMB*) gene. *SMB* has been shown to be involved in controlling the reorientation and
346 timing of cell division in root stem cells (43). *AUR62000536* is a second example, the
347 orthologue of the Arabidopsis *AtNAC2* gene. *AtNAC2* is specifically expressed in the roots
348 and significantly increases the number of lateral root formation in response to ethylene
349 and auxin signalling in Arabidopsis (54). In addition, *AUR62032816* (the orthologue of
350 *ANAC029*) is strongly and significantly induced in the roots compared to the shoots.
351 *ANAC029* was found to be involved in root morphogenesis in *Medicago truncatula* (55).
352 *AUR62004554* is another example, which is the orthologue of Arabidopsis *ANAC043*. In
353 Arabidopsis, *ANAC043* belongs to secondary wall NAC genes (SWN), which are also
354 known as *secondary wall thickening promoting factor1 (NST1)*; however, its role in root
355 growth and development has not yet been identified (56). With respect to shoot related
356 *CqNACs*, we identified only five *CqNACs* to be significantly more expressed in the shoots
357 than the roots, i.e., the two orthologues of *ANAC072* (*AUR62015630*, *AUR62029344*),
358 and the two orthologs of *ANA0C94* (*AUR62032488*, *AUR62040120*) and the orthologue
359 of *NAM-B* (*AUR62000898*). The function of these genes in shoot formation has not been
360 identified before. Thus, future studies should test if those genes are indeed involved in
361 shoot formation or shoot related processes in quinoa.

362

363 **Fig 7. Heat map of the expression of *CqNAC* significantly differentially expressed**
364 **(DE) between roots and shoots.**

365 Expression level (in FPKM) of significantly DE genes between shoot and root of plants
366 grown in a hydroponic system or soil. Significant difference is calculated at FDR=0.05.

367

368 **2.4.2 Expression of quinoa *CqNAC* TFs in response to abiotic stresses**

369 NAC TFs have been identified as an important regulator of stress responses, and given
370 that quinoa has a high stress tolerance to abiotic stresses, we hypothesize that NAC TFs
371 are involved in quinoa's adaptation to stress tolerance. Here, we examine the expression
372 of all of the putative quinoa NACs in response to different abiotic stresses (salt, drought,
373 heat and phosphate starvation) using RNAseq previously generated (27). Around 35-48
374 *CqNACs* are DE in the roots or shoots of quinoa plants exposed to one of the abiotic
375 stresses examined in this study. However, the largest number of genes being DE were in
376 the roots of the plants that were exposed to phosphate starvation (Fig 8a).

377

378 **Fig 8. Quinoa NACs differentially expressed in response to different abiotic** 379 **stresses.**

380 a) Number of genes differentially expressed in the shoots and roots of quinoa plants
381 exposed to different abiotic stresses, including salt, drought, heat and phosphate
382 starvation. b) Heat map of *CqNACs* DE after salt stress. c) Heat map of *CqNACs* DE after
383 drought stress. d) Heat map of *CqNACs* DE after heat stress. e) Heat map of *CqNACs*
384 DE after phosphate starvation. The significance between stress and control is calculated
385 at FDR=0.05. f) Venn diagram illustrating the overlap between different stress-responsive
386 *CqNAC* genes (only significant DE genes).

387

388 NAC genes are known to be involved in salt stress response and tolerance in Arabidopsis
389 and many other plant species (57-60). Here, we identified nine genes in quinoa that were
390 significantly upregulated in the shoots, roots or both in response to salt stress (5 genes
391 in shoot, 1 gene in root and 3 genes in both shoot and root) (Fig 8b). Interestingly, some
392 of these genes are orthologues of NACs genes, which have been previously identified as
393 salt responsive genes and are known to play a role in salinity tolerance when they are

394 overexpressed in plants. For example, *AUR62028872* and *AUR62005686* are two
395 orthologues of the Arabidopsis *ANAC083* (*VND-INTERACTING 2* or *VIN2*), a negative
396 regulator of xylem vessel formation. In Arabidopsis, *ANAC083* was induced by salt
397 treatment and its overexpression enhanced salinity tolerance by directly binding to the
398 promoters of two important genes involved in salt stress response (61). Our data showed
399 a similar trend as the two quinoa orthologues of *ANAC083* were significantly induced by
400 salt stress. *AtNAP* (*ANAC029*) was identified as another NAC TF gene that is induced by
401 salt stress (62). *AtNAP* functions as a negative regulator of salt stress response via
402 repressing the expression of *AREB1* and thus its overexpression increases the plants'
403 sensitivity to salt stress (62). We found a similar expression trend in quinoa as the quinoa
404 orthologue of *AtNAP*, *AUR62005821*, was also induced by salt stress in both shoot and
405 root. Two other genes that are also significantly upregulated by salt treatment are
406 *AUR62000536*, the *NAC2* orthologue and *AUR62029344*, the *ANAC072* orthologue –
407 those two genes were found to be induced by salt stress, and plants were found to have
408 enhanced stress tolerance when they were overexpressed separately (10, 63). Another
409 example is *AUR62006885*, the *ATAF2* orthologue, which has been previously identified
410 to be induced by salt stress and pathogen attack (64, 65). Here we report a similar
411 expression pattern in quinoa as it is also induced in the shoots and roots of quinoa plants
412 in response to salt stress.

413 Some salt responsive *CqNAC* genes, such as *AUR62036626*, *AUR62041525* and
414 *AUR62043497* (the orthologues of Arabidopsis *ANAC068*, *ANAC078* and *ANAC021*,
415 respectively), are significantly DE in response to salt, but their roles in response to salt

416 stress have not been studied before. These genes have the potential to be involved in
417 salt stress response and adaptation in quinoa.

418 Quinoa is known to have a high drought tolerance and many NACs have been found to
419 be involved in drought tolerance in rice and Arabidopsis (57-60). After drought treatment,
420 expression of sixteen *CqNAC* genes were significantly different between the stress and
421 control in the shoots or roots, or both (Fig 8c). Some of these significantly DE genes are
422 orthologous to well-known drought responsive NAC genes in other plant species. For
423 example, in Arabidopsis, overexpression of *RD26/ANAC072* and *ATAF1/ANAC002*
424 confers drought tolerance (44, 66). We found that in quinoa the *ATAF-1* orthologue
425 *AUR62000536* and the two orthologues of *ANAC072* *AUR62015630* and *AUR62029344*
426 were upregulated in response to drought stress, which is consistent with their role in
427 drought tolerance as for Arabidopsis (44, 66). *AUR62005821* and *AUR62032816*, the
428 two orthologues of Arabidopsis *AtNAP*, present two further examples, which we found to
429 be significantly upregulated in the shoots and downregulated in the roots in response to
430 drought stress. In rice, *OsNAP* functions as a positive regulator of drought tolerance (59).
431 In addition, the wheat *TaNAP29* enhances drought tolerance in transgenic Arabidopsis
432 (60). We identify ten genes that are significantly differentially expressed, but their roles in
433 response to drought stress have not been previously identified.

434 The focus of the role of NAC TFs in heat response has only recently emerged and there
435 are only a few reports focusing on their role in heat tolerance. Thus, their role has not
436 been studied as thoroughly as other stresses (53, 67-70). Here, we identified 13 genes
437 that were significantly differentially expressed in response to heat stress in the shoots or
438 roots or both. Interestingly, *AUR62015630* and *AUR62029344* (the two orthologues of

439 *ANAC072/RD26*) were downregulated in both the shoots and roots. In our previous study
440 of the transcriptomic analysis of NAC TFs in response to heat priming in Arabidopsis, we
441 found that *RD26* and *AtNAP* were responsive to heat priming (Alshareef *et al*,
442 unpublished). This suggests the possible role of NAC TF genes in regulating heat stress
443 response in quinoa.

444 Adequate supply of inorganic phosphate (P_i) is important for plant growth and
445 development, whereby low levels of P_i in soil adversely affect plant growth and
446 development. Thus, plants have evolved several regulatory mechanisms to adapt to
447 phosphorus deficiency and optimize P_i uptake and assimilation (57). Here, we found that
448 the largest number of quinoa NAC TF genes was induced by phosphate deficiency
449 treatment. Nineteen genes were significantly upregulated in the roots or shoots or both
450 (16 *CqNAC* genes were upregulated in the roots and 10 *CqNACs* were upregulated in in
451 the shoots in response to phosphate deficiency (Fig 8e). In Arabidopsis, it has been
452 reported that five NAC genes are induced by more than a 2 fold change in response to
453 P-starvation, namely *At3g15500/AtNAC3*, *At3g04070/ANAC047*, *At2g43000/JUB1*,
454 *At1g52890/ANAC019*, *At1g02220/ANAC003* (71-73) and (Hammond and Bennet, (AT-
455 O61 Genevestigator). However, none of these gene orthologues in quinoa are
456 significantly differentially expressed in our data.

457 From the above results on the transcriptional responses of quinoa NACs to different
458 abiotic stresses, we found there are three *CqNACs* that are differentially expressed in
459 response to all of the stresses, i.e., *AUR62029344*, *AUR62000536* and *AUR62006885*,
460 which are the orthologues of *ANAC072*, *ANAC002* and *ATAF2*, respectively (Fig 8e). This
461 suggests the presence of a crosstalk pathway that mediates stress tolerance between

462 different stresses. Although these genes are common between stresses, they have
463 different expression patterns in response to different stresses and perhaps they are
464 involved in different gene regulatory networks. For instance, *AUR62029344*, a
465 *ANAC072/RD26* orthologue, was induced in the shoots and roots of quinoa plants after
466 salt, phosphate and drought stresses, and only downregulated in response to heat stress
467 in both shoots and roots.

468 The consistent transcriptional responses between *CqNACs* (in response to abiotic stress)
469 and those that have been reported previously for *Arabidopsis* or other plant species,
470 indicate a conserved function of NACs across different species and the possible sharing
471 of a common stress tolerance mechanism. However, we have also identified some
472 *CqNACs* that are stress responsive, transcriptionally, but their function have not been
473 identified previously. These NACs may participate in a novel stress tolerance in quinoa
474 plants.

475 **2.4.3 qRT-PCR analysis of stress responsive *CqNACs***

476 To verify the response of *CqNACs* to different abiotic stresses, we conducted an
477 additional salinity, drought and heat experiments and investigated the expression of some
478 selected *CqNAC* by qRT-PCR. We identified the selected *CqNACs* to be DE in response
479 to stress according to our RNAseq data and/or are orthologue to NACs functionally
480 annotated as a stress responsive gene in other plant species. We analysed the
481 expression of six salt stress responsive *CqNAC* genes in response to salt stress (300 mM
482 NaCl). These genes are orthologues of *Arabidopsis ATAF1*, *ATAF2*, *ANAC072*,
483 *ANAC078* and *ANAC083* (Fig 9a). The same trend for changes in expression were
484 observed for all of the tested genes when comparing qRT-PCR and RNAseq (S3a Fig),

485 which confirms their response to salt stress. Among these genes, we tested the
486 expression of the two copies of *ANAC083* orthologues (*AUR62037478*, *AUR62043497*)
487 and the results showed a stronger response in one of the copies (*AUR62043497*)
488 compared to the other copy (*AUR62037478*).

489

490 **Fig 9. qRT-PCR analysis of some selected *CqNACs* in response to different abiotic**
491 **stresses.**

492 a) Relative expression of *CqNACs* after salt treatment relative to control. b) Relative
493 expression of *CqNACs* after drought treatment relative to control. a) Relative expression
494 of *CqNACs* after heat treatment relative to control. The expression level represents the
495 mean of three independent biological replicates. The error bar is standard error of the
496 mean.

497

498 In response to drought stress, we analysed the expression of nine genes (Fig 9b). The
499 results showed that the expression pattern of some of the *CqNAC* genes in response to
500 drought stress is consistent between RNAseq and qRT-PCR, except for four genes
501 (*ATAF1*, *ANAC083* and the two copies of *ANAC029*), which showed a different
502 expression pattern between RNAseq and qRT-PCR (S3b Fig). We also analysed the
503 expression pattern between the copies of the duplicated genes, i.e., the orthologues of
504 Arabidopsis *ANAC100*, *ANAC072* and *ANAC029*. The orthologues of Arabidopsis
505 *ANAC100* and *ANAC072* showed a similar expression pattern between the duplicated
506 copies, which could suggest that both of the copies have a similar function. The two
507 copies of *ANAC072* showed an opposite expression pattern in response to drought
508 between the two copies (RNAseq and qRT-PCR) (Fig 9b) raising the possibility that these
509 two genes have a different function in response to drought stress in quinoa.

510 In response to heat stress, we analysed the expression of eight genes. The expression
511 of all of the *CqNAC* genes in response to heat stress was consistent between the RNAseq

512 and qRT-PCR for six of the genes. Two genes appeared to have opposite expression
513 patterns between the RNAseq and qRT-PCR, the orthologues of Arabidopsis *ANAC083*
514 and *ANAC072* (S3c Fig). The experiments for heat stress for RNAseq and qRT-PCR
515 studies were performed under different conditions. The heat stress for RNAseq samples
516 was for much shorter duration, which may explain the different expression pattern of some
517 genes.

518 Among the analysed genes, there are some duplicated genes, namely the orthologues of
519 Arabidopsis *ANAC100*, *ANAC083* and *ANAC072*. In the two duplicated genes (*ANAC100*
520 and *ANAC083*), both copies of each gene showed a similar expression pattern in
521 response to heat stress. This is similar to what we have seen in response to drought
522 stress for these two genes. However, the two copies of *ANAC072* showed a different
523 expression pattern in response to heat stress, suggesting that these two copies of
524 *ANAC072* gene may have a different function (Fig 9c).

525 To conclude, the high consistency of *CqNAC* expression between the RNAseq and qRT-
526 PCR experiments confirmed their response to abiotic stress. Moreover, by analysing the
527 expression among the duplicated genes in response to stress, our results indicate that
528 there is a similarity between the expression pattern and some of the duplicated genes,
529 which indicates similar or redundant functions of these duplicated copies. However, in
530 some duplicated genes, the expression of one copy is weaker or different (opposite) to
531 the expression of the other copy, which may suggest different functions of the duplicated
532 genes in response to stress. We propose that these genes should be further investigated
533 for their role in stress responses and potentially stress tolerance.

534

535 **3. Conclusions**

536 In conclusion, this study provides a comprehensive identification and characterization on
537 the transcriptional level of the NAC TFs family in quinoa. We identified 107 NAC TFs in
538 the genome of quinoa plants. We phylogenetically and functionally classified these NACs
539 into different phylogenetic subgroups, in alignment with previously identified NACs in
540 Arabidopsis. This functional classification was also supported by the global expression
541 analysis that we carried out in two tissues (shoots and roots) in response to abiotic
542 stresses, as several NACs involved in root development were significantly differentially
543 expressed in the root of quinoa plants. Similarly, we found several NACs responded
544 transcriptionally to different abiotic stresses in a similar manner to previously
545 characterized NACs in Arabidopsis and other plants species. This indicates that some of
546 the quinoa NAC TFs have a similar function to previously identified NACs. Our results
547 provide valuable information for further functional research of NAC TF in quinoa and their
548 role in its adaptation to different stresses.

549 **4. Materials and Methods**

550 **4.1 Sequence database search**

551 We employed three approaches to identify putative NAC family genes in quinoa. In the
552 first approach, we searched for peptides that have a protein family (Pfam) (PF02365) in
553 the genome of quinoa in the phytozome database version12 (phytozome,
554 <http://www.phytozome.net>, version 12). In the second approach, we downloaded all
555 protein sequences of the quinoa (44,776 peptide sequences) from phytozome database

556 (phytozome, <http://www.phytozome.net>, version 12), and we used these sequences as a
557 query in the transcription factors prediction server from PTFDB. In the third approach, we
558 used a Basic Local Alignment search tool (BLASTp); the protein sequences of the 110
559 NAC domain of *Arabidopsis thaliana* (*Arabidopsis*) were used to BLASTp against the
560 quinoa QQ74 peptides. We selected the protein sequence with E-value of $< 1e^{-10}$ as a
561 candidate NAC gene, keeping only alignments that cover at least 60% of the query length
562 (*Arabidopsis* sequences). We downloaded sequences of *Arabidopsis* and rice NAC
563 proteins from Phytozome database version12. We considered peptide sequences that
564 had less than 100 residues to be truncated proteins and thus they were removed from the
565 analysis. We confirmed all of the putative NAC proteins by verifying the presence of the
566 NAC domain with the InterProScan program (<http://www.ebi.ac.uk/Tools/InterProScan/>),
567 hmmscan function of the HMMER web server
568 (<https://www.ebi.ac.uk/Tools/hmmer/search/hmmscan>, HmmerWeb version 2.28.0) of
569 NAM domain (PF02365) (33) and a conserved domain search (CDS) of the NCBI
570 database.

571 We performed the inference of homologous copies of genes using a software that
572 searches for multiple collinearity regions (MCScanX) with default parameters. Briefly, we
573 performed between all quinoa QQ74 peptide sequences against themselves with an E-
574 value of $< 1e^{-10}$ and we reported a maximum of 5 best hits. Subsequently, we inferred
575 pairwise collinear genes from collinear blocks defined by a minimum of 5 homologous
576 sequences in conserved order. We inferred homologous genes from collinear genes
577 between A and B sub-genomes.

578 The prediction of membrane-bound NACs proteins was performed using the software
579 TMHMM (<http://www.cbs.dtu.dk/services/TMHMM/>, server v.2.0) (74).

580 **4.2 Phylogenetic tree analysis**

581 Multiple sequence alignment of the full length protein sequences or NAC domain
582 sequences was performed using the tool MUSCLE (75) built in the software MEGA 7.0
583 (76). We constructed the unrooted phylogenetic tree using the software MEGA 7.0 (76)
584 using the Neighbor-Joining (NJ) method with a bootstrap test value of 1000.

585 **4.3 Gene structure and conserved motif**

586 We used the gene structure display server (GSDS) program (<http://gsds.cbi.pku.edu.cn/>,
587 version 2.0) (77) to display the organization of exons and introns of NAC genes by
588 comparing cDNA with the DNA sequences of individual gene.

589 **4.4 Plant material, growth conditions and stress treatment and transcriptomic** 590 **analysis**

591 We used the gene expression (RNAseq) data previously generated (27). Plant growth
592 conditions, samples collections and stress application have been previously described
593 (27). Briefly, RNA was isolated from roots and shoots of quinoa plants exposed to different
594 stresses including drought, heat, salinity and low phosphate conditions. For drought and
595 heat treatments, plants were grown in soil, for salinity and low phosphate treatments,
596 plants were grown in hydroponics (along with their respective controls without treatment):

597 Soil grown plants: Plants were grown in soil in a growth chamber under well-watered
598 conditions at 20°C and 12 h daily light for three weeks and then either left without water
599 for one week (as drought treatment) or transferred to another growth chamber with 37°C
600 day and 32°C night temperatures (as heat treatment).

601 Hydroponically grown plants: After germination on agar for two weeks, seedlings were
602 transferred to aerated tanks containing basal nutrient solution for another week before
603 either being moved to tanks containing basal nutrient solution lacking KH_2PO_4 and
604 supplemented with a compensatory amount of KCl (as low phosphate treatment) or to
605 tanks containing basal nutrient solution supplemented with 150 mM NaCl and then
606 increase to 300 mM NaCl after 24 h (as salinity treatment). Paired-end sequencing of
607 100-bp libraries was performed using an Illumina HiSeq2000. Three replicates for each
608 treatment and each tissue were sequenced, one sample for salt treated roots was
609 excluded as it did not pass quality control.

610 **4.5 Expression analysis of *CqNAC* genes using qRT-PCR**

611 Quinoa accession QQ74 was grown in 10 cm diameter pots with standard potting mix,
612 under controlled conditions (day/night, 12/12 h, 22/18, 50% humidity, photon flux density
613 $300\text{-}350 \mu\text{mol m}^{-2} \text{s}^{-1}$) in plant growth Biochambers until the 11th leaf growth stage (~26
614 days) before they were subjected to treatment.

615 Plants in all treatment groups, other than drought, were watered on the first day of
616 treatment by filling their trays with water (or saline water in case of salt stress). Plants
617 were also watered throughout the course of experiment to maintain pots at a water holding
618 capacity of 60%, except for salt and drought treatments.

619 For heat stress, plants were moved to a chamber with the same settings as control, only
620 the temperature was set to day/night 35/30°C. For drought stress, plants were watered
621 up to 20-30% water holding capacity WHC (For drought plants). For salt treatment, pots
622 were soaked in 300 mM NaCl solution for four hours.

623 For gene expression analysis, leaf 10 was sampled six days after the treatment, frozen
624 in liquid nitrogen and then RNA was extracted using the Direct-Zol Plant RNA extraction
625 kit according to the manufacturer's instruction (Zymo Research, Germany). Quality of the
626 RNA was assessed using gel electrophoresis and quantity using a NanoDrop. For cDNA
627 synthesis, 1 µg of high-quality RNA was used to prepare cDNA using the SuperScript III
628 kit according to the manufacturer's instruction (Life Technologies).

629 We prepared the real time qRT-PCR reaction using the SYBR Green master mix (Applied
630 Biosystems): cDNA 4 µL, SYBR Green master mix (2X) 5 µL, 1 µL of forward and reverse
631 primer mix (final concentration of each primer is 0.5 µM). We incubated the reactions in
632 the ABI PRISM 7900 HT sequence detection system (Applied Biosystems). We set the
633 reaction using the following conditions: initial step of 95°C for 10 min followed by 40 cycles
634 of these two steps (95°C for 15 sec, 60°C for 1 min). We used polyubiquitin 10 gene
635 (*AUR62015654*) as the reference gene for data analysis. The primer sequences of all of
636 the primers is provided in S7 Table. After 40 reaction cycles, we generated a melting
637 curve by heating from 60°C to 95°C with a ramp speed of 1.9°C min⁻¹ to verify
638 amplification of the desired product.

639 We analysed data using the comparative Ct method as described in Kamranfar, Xue (78).
640 Briefly, we calculated the delta Ct value (ΔCt) by normalizing each Ct value with the Ct
641 value of the reference gene *UBQ10*. Then, the level of gene expression was expressed

642 as the difference between an arbitrary value of 40 and the ΔCt value (a high $40-\Delta\text{Ct}$ value
643 indicates a high gene expression level). We used a threshold of 1.5-fold change to select
644 differentially expressed genes.

645 **Acknowledgements**

646 We thank Noha Saber (KAUST, Saudi Arabia) for practical support.

647 5. References

- 648 1. Aida M, Ishida T, Fukaki H, Fujisawa H, Tasaka M. Genes involved in organ
649 separation in Arabidopsis: an analysis of the cup-shaped cotyledon mutant. *The Plant*
650 *Cell Online*. 1997;9(6):841-57.
651
- 652 2. Souer E, van Houwelingen A, Kloos D, Mol J, Koes R. The no apical meristem
653 gene of *Petunia* is required for pattern formation in embryos and flowers and is expressed
654 at meristem and primordia boundaries. *Cell*. 1996;85(2):159-70.
655
- 656 3. Nuruzzaman M, Manimekalai R, Sharoni AM, Satoh K, Kondoh H, Ooka H, et al.
657 Genome-wide analysis of NAC transcription factor family in rice. *Gene*. 2010;465(1-2):30-
658 44.
659
- 660 4. Le DT, Nishiyama R, Watanabe Y, Mochida K, Yamaguchi-Shinozaki K, Shinozaki
661 K, et al. Genome-wide survey and expression analysis of the plant-specific NAC
662 transcription factor family in soybean during development and dehydration stress. *DNA*
663 *research*. 2011;18(4):263-76.
664
- 665 5. Sablowski RW, Meyerowitz EM. A homolog of NO APICAL MERISTEM is an
666 immediate target of the floral homeotic genes APETALA3/PISTILLATA. *Cell*.
667 1998;92(1):93-103.
668
- 669 6. Uauy C, Distelfeld A, Fahima T, Blechl A, Dubcovsky J. A NAC gene regulating
670 senescence improves grain protein, zinc, and iron content in wheat. *Science*.
671 2006;314(5803):1298-301.
672
- 673 7. Balazadeh S, Riaño-Pachón D, Mueller-Roeber B. Transcription factors regulating
674 leaf senescence in *Arabidopsis thaliana*. *Plant Biology*. 2008;10(s1):63-75.
675
- 676 8. Kim YS, Kim SG, Park JE, Park HY, Lim MH, Chua NH, et al. A membrane-bound
677 NAC transcription factor regulates cell division in *Arabidopsis*. *Plant Cell*.
678 2006;18(11):3132-44.
679
- 680 9. Zhong R, Richardson EA, Ye Z-H. Two NAC domain transcription factors, SND1
681 and NST1, function redundantly in regulation of secondary wall synthesis in fibers of
682 *Arabidopsis*. *Planta*. 2007;225(6):1603-11.
683
- 684 10. He XJ, Mu RL, Cao WH, Zhang ZG, Zhang JS, Chen SY. AtNAC2, a transcription
685 factor downstream of ethylene and auxin signaling pathways, is involved in salt stress
686 response and lateral root development. *The Plant Journal*. 2005;44(6):903-16.
687
- 688 11. Ernst HA, Nina Olsen A, Skriver K, Larsen S, Lo Leggio L. Structure of the
689 conserved domain of ANAC, a member of the NAC family of transcription factors. *EMBO*
690 *reports*. 2004;5(3):297-303.

- 691
692 12. Jensen MK, Kjaersgaard T, Nielsen MM, Galberg P, Petersen K, O'Shea C, et al.
693 The Arabidopsis thaliana NAC transcription factor family: structure-function relationships
694 and determinants of ANAC019 stress signalling. *The Biochemical journal*.
695 2010;426(2):183-96.
696
697 13. Olsen AN, Ernst HA, Leggio LL, Skriver K. NAC transcription factors: structurally
698 distinct, functionally diverse. *Trends in plant science*. 2005;10(2):79-87.
699
700 14. Olsen AN, Ernst HA, Leggio LL, Skriver K. DNA-binding specificity and molecular
701 functions of NAC transcription factors. *Plant Science*. 2005;169(4):785-97.
702
703 15. Kim SY, Kim SG, Kim YS, Seo PJ, Bae M, Yoon HK, et al. Exploring membrane-
704 associated NAC transcription factors in Arabidopsis: implications for membrane biology
705 in genome regulation. *Nucleic Acids Research*. 2007;35(1):203-13.
706
707 16. Shen H, Yin Y, Chen F, Xu Y, Dixon RA. A Bioinformatic Analysis of NAC Genes
708 for Plant Cell Wall Development in Relation to Lignocellulosic Bioenergy Production.
709 *BioEnergy Research*. 2009;2(4):217-32.
710
711 17. Christiansen MW, Holm PB, Gregersen PL. Characterization of barley (*Hordeum*
712 *vulgare* L.) NAC transcription factors suggests conserved functions compared to both
713 monocots and dicots. *BMC Res Notes*. 2011;4:302.
714
715 18. Yoshiyama K, Conklin PA, Huefner ND, Britt AB. Suppressor of gamma response
716 1 (SOG1) encodes a putative transcription factor governing multiple responses to DNA
717 damage. *Proceedings of the National Academy of Sciences*. 2009;106(31):12843-8.
718
719 19. Mitsuda N, Hisabori T, Takeyasu K, Sato MH. VOZ; isolation and characterization
720 of novel vascular plant transcription factors with a one-zinc finger from Arabidopsis
721 thaliana. *Plant and cell physiology*. 2004;45(7):845-54.
722
723 20. Kolano B, McCann J, Orzechowska M, Siwinska D, Temsch E, Weiss-
724 Schneeweiss H. Molecular and cytogenetic evidence for an allotetraploid origin of
725 *Chenopodium quinoa* and *C. berlandieri* (Amaranthaceae). *Molecular phylogenetics and*
726 *evolution*. 2016;100:109-23.
727
728 21. Repo-Carrasco-Valencia R, Espinoza C, Jacobsen S-E. Nutritional Value and Use
729 of the Andean Crops Quinoa (*Chenopodium quinoa*) and Kañiwa (*Chenopodium*
730 *pallidicaule*)2003. 179-89 p.
731
732 22. Isabelle Adolf V, Jacobsen S-E, Shabala S. Salt tolerance mechanisms in quinoa
733 (*Chenopodium quinoa* Willd.)2013. 43–54 p.
734

- 735 23. Jacobsen SE, Mujica A, Jensen CR. The Resistance of Quinoa (*Chenopodium*
736 *quinoa* Willd.) to Adverse Abiotic Factors. Food Reviews International. 2003;19(1-2):99-
737 109.
738
- 739 24. Hariadi Y, Marandon K, Tian Y, Jacobsen SE, Shabala S. Ionic and osmotic
740 relations in quinoa (*Chenopodium quinoa* Willd.) plants grown at various salinity levels.
741 Journal of experimental botany. 2011;62(1):185-93.
742
- 743 25. Yue H, Chang X, Zhi Y, Wang L, Xing G, Song W, et al. Evolution and Identification
744 of the WRKY Gene Family in Quinoa (*Chenopodium quinoa*). Genes. 2019;10(2).
745
- 746 26. Liu J, Wang R, Liu W, Zhang H, Guo Y, Wen R. Genome-Wide Characterization
747 of Heat-Shock Protein 70s from *Chenopodium quinoa* and Expression Analyses of
748 Cqhsp70s in Response to Drought Stress. Genes. 2018;9(2).
749
- 750 27. Jarvis DE, Ho YS, Lightfoot DJ, Schmöckel SM, Li B, Borm TJA, et al. The genome
751 of *Chenopodium quinoa*. Nature. 2017;542:307.
752
- 753 28. Yan H, Zhang A, Ye Y, Xu B, Chen J, He X, et al. Genome-wide survey of
754 switchgrass NACs family provides new insights into motif and structure arrangements and
755 reveals stress-related and tissue-specific NACs. Scientific reports. 2017;7(1):3056.
756
- 757 29. Hu R, Qi G, Kong Y, Kong D, Gao Q, Zhou G. Comprehensive Analysis of NAC
758 Domain Transcription Factor Gene Family in *Populus trichocarpa*. BMC Plant Biology.
759 2010;10(1):145.
760
- 761 30. Shang H, Li W, Zou C, Yuan Y. Analyses of the NAC Transcription Factor Gene
762 Family in *Gossypium raimondii* Ulbr.: Chromosomal Location, Structure, Phylogeny, and
763 Expression Patterns. Journal of Integrative Plant Biology. 2013;55(7):663-76.
764
- 765 31. Sun H, Hu M, Li J, Chen L, Li M, Zhang S, et al. Comprehensive analysis of NAC
766 transcription factors uncovers their roles during fiber development and stress response in
767 cotton. BMC Plant Biology. 2018;18(1):150.
768
- 769 32. Jin J, Tian F, Yang D-C, Meng Y-Q, Kong L, Luo J, et al. PlantTFDB 4.0: toward a
770 central hub for transcription factors and regulatory interactions in plants. Nucleic Acids
771 Research. 2017;45(D1):D1040-D5.
772
- 773 33. Potter SC, Luciani A, Eddy SR, Park Y, Lopez R, Finn RD. HMMER web server:
774 2018 update. Nucleic Acids Research. 2018;46(W1):W200-W4.
775
- 776 34. Ooka H, Satoh K, Doi K, Nagata T, Otomo Y, Murakami K, et al. Comprehensive
777 analysis of NAC family genes in *Oryza sativa* and *Arabidopsis thaliana*. DNA Research
778 2003;10:239-47.
779

- 780 35. Zhang Y, Wang L. The WRKY transcription factor superfamily: its origin in
781 eukaryotes and expansion in plants. *BMC evolutionary biology*. 2005;5:1.
782
- 783 36. Wilkins O, Nahal H, Foong J, Provart NJ, Campbell MM. Expansion and
784 diversification of the *Populus* R2R3-MYB family of transcription factors. *Plant physiology*.
785 2009;149(2):981-93.
786
- 787 37. Li X, Duan X, Jiang H, Sun Y, Tang Y, Yuan Z, et al. Genome-wide analysis of
788 basic/helix-loop-helix transcription factor family in rice and *Arabidopsis*. *Plant physiology*.
789 2006;141(4):1167-84.
790
- 791 38. Lijavetzky D, Carbonero P, Vicente-Carbajosa J. Genome-wide comparative
792 phylogenetic analysis of the rice and *Arabidopsis* Dof gene families. *BMC evolutionary*
793 *biology*. 2003;3:17.
794
- 795 39. Wang D, Guo Y, Wu C, Yang G, Li Y, Zheng C. Genome-wide analysis of CCCH
796 zinc finger family in *Arabidopsis* and rice. *BMC Genomics*. 2008;9(1):44.
797
- 798 40. Hibara K, Karim MR, Takada S, Taoka K, Furutani M, Aida M, et al. *Arabidopsis*
799 CUP-SHAPED COTYLEDON3 regulates postembryonic shoot meristem and organ
800 boundary formation. *Plant Cell*. 2006;18(11):2946-57.
801
- 802 41. Vroemen CW, Mordhorst AP, Albrecht C, Kwaaitaal MACJ, de Vries SC. The CUP-
803 SHAPED COTYLEDON3 Gene Is Required for Boundary and Shoot Meristem Formation
804 in *Arabidopsis*. *The Plant Cell*. 2003;15(7):1563-77.
805
- 806 42. Zhou J, Zhong R, Ye ZH. *Arabidopsis* NAC domain proteins, VND1 to VND5, are
807 transcriptional regulators of secondary wall biosynthesis in vessels. *PLoS One*.
808 2014;9(8):e105726.
809
- 810 43. Willemsen V, Bauch M, Bennett T, Campilho A, Wolkenfelt H, Xu J, et al. The NAC
811 domain transcription factors FEZ and SOMBRERO control the orientation of cell division
812 plane in *Arabidopsis* root stem cells. *Developmental cell*. 2008;15(6):913-22.
813
- 814 44. Tran L-SP, Nakashima K, Sakuma Y, Simpson SD, Fujita Y, Maruyama K, et al.
815 Isolation and Functional Analysis of *Arabidopsis* Stress-Inducible NAC Transcription
816 Factors That Bind to a Drought-Responsive cis-Element in the early responsive to
817 dehydration stress 1 Promoter. *The Plant Cell*. 2004;16(9):2481-98.
818
- 819 45. Zhao J, Liu JS, Meng FN, Zhang ZZ, Long H, Lin WH, et al. ANAC005 is a
820 membrane-associated transcription factor and regulates vascular development in
821 *Arabidopsis*. *Journal of integrative plant biology*. 2016;58(5):442-51.
822
- 823 46. Kim YS, Park CM. Membrane regulation of cytokinin-mediated cell division in
824 *Arabidopsis*. *Plant signaling & behavior*. 2007;2(1):15-6.
825

- 826 47. Rahman H, Ramanathan V, Nallathambi J, Duraialagaraja S, Muthurajan R. Over-
827 expression of a NAC 67 transcription factor from finger millet (*Eleusine coracana L.*)
828 confers tolerance against salinity and drought stress in rice. BMC biotechnology. 2016;16
829 Suppl 1:35.
830
- 831 48. Jung JH, Park CM. Auxin modulation of salt stress signaling in Arabidopsis seed
832 germination. Plant signaling & behavior. 2011;6(8):1198-200.
833
- 834 49. Prince VE, Pickett FB. Splitting pairs: the diverging fates of duplicated genes.
835 Nature reviews Genetics. 2002;3(11):827-37.
836
- 837 50. Liu M, Ma Z, Sun W, Huang L, Wu Q, Tang Z, et al. Genome-wide analysis of the
838 NAC transcription factor family in Tartary buckwheat (*Fagopyrum tataricum*). BMC
839 Genomics. 2019;20(1):113.
840
- 841 51. Shahnejat-Bushehri S, Tarkowska D, Sakuraba Y, Balazadeh S. Arabidopsis NAC
842 transcription factor JUB1 regulates GA/BR metabolism and signalling. Nature plants.
843 2016;2.
844
- 845 52. Thirumalaikumar VP, Devkar V, Mehterov N, Ali S, Ozgur R, Turkan I, et al. NAC
846 transcription factor JUNGBRUNNEN1 enhances drought tolerance in tomato. Plant
847 Biotechnology Journal. 2018;16(2):354-66.
848
- 849 53. Lee S, Lee HJ, Huh SU, Paek KH, Ha JH, Park CM. The Arabidopsis NAC
850 transcription factor NTL4 participates in a positive feedback loop that induces
851 programmed cell death under heat stress conditions. Plant science : an international
852 journal of experimental plant biology. 2014;227:76-83.
853
- 854 54. He XJ, Mu RL, Cao WH, Zhang ZG, Zhang JS, Chen SY. AtNAC2, a transcription
855 factor downstream of ethylene and auxin signaling pathways, is involved in salt stress
856 response and lateral root development. The Plant journal : for cell and molecular biology.
857 2005;44(6):903-16.
858
- 859 55. De Zelicourt A, Diet A, Marion J, Laffont C, Ariel F, Moison M, et al. Dual
860 involvement of a *Medicago truncatula* NAC transcription factor in root abiotic stress
861 response and symbiotic nodule senescence. The Plant journal : for cell and molecular
862 biology. 2012;70(2):220-30.
863
- 864 56. Zhong R, Ye Z-H. The Arabidopsis NAC transcription factor NST2 functions
865 together with SND1 and NST1 to regulate secondary wall biosynthesis in fibers of
866 inflorescence stems. Plant signaling & behavior. 2015;10(2):e989746-e.
867
- 868 57. Nuruzzaman M, Sharoni AM, Kikuchi S. Roles of NAC transcription factors in the
869 regulation of biotic and abiotic stress responses in plants. Frontiers in Microbiology.
870 2013;4:248.
871

- 872 58. Golldack D, Lüking I, Yang O. Plant tolerance to drought and salinity: stress
873 regulating transcription factors and their functional significance in the cellular
874 transcriptional network. *Plant cell reports*. 2011;30(8):1383-91.
875
- 876 59. Chen X, Wang Y, Lv B, Li J, Luo L, Lu S, et al. The NAC family transcription factor
877 OsNAP confers abiotic stress response through the ABA pathway. *Plant & cell*
878 *physiology*. 2014;55(3):604-19.
879
- 880 60. Huang Q, Wang Y, Li B, Chang J, Chen M, Li K, et al. TaNAC29, a NAC
881 transcription factor from wheat, enhances salt and drought tolerance in transgenic
882 *Arabidopsis*. *BMC Plant Biol*. 2015;15:268.
883
- 884 61. Yang SD, Seo PJ, Yoon HK, Park CM. The *Arabidopsis* NAC transcription factor
885 VNI2 integrates abscisic acid signals into leaf senescence via the COR/RD genes. *The*
886 *Plant cell*. 2011;23(6):2155-68.
887
- 888 62. Seok HY, Woo DH, Nguyen LV, Tran HT, Tarte VN, Mehdi SM, et al. *Arabidopsis*
889 *AtNAP* functions as a negative regulator via repression of *AREB1* in salt stress response.
890 *Planta*. 2017;245(2):329-41.
891
- 892 63. Tran LS, Nakashima K, Sakuma Y, Simpson SD, Fujita Y, Maruyama K, et al.
893 Isolation and functional analysis of *Arabidopsis* stress-inducible NAC transcription factors
894 that bind to a drought-responsive cis-element in the early responsive to dehydration
895 stress 1 promoter. *Plant Cell*. 2004;16(9):2481-98.
896
- 897 64. Delessert C, Kazan K, Wilson IW, Van Der Straeten D, Manners J, Dennis ES, et
898 al. The transcription factor *ATAF2* represses the expression of pathogenesis-related
899 genes in *Arabidopsis*. *The Plant journal : for cell and molecular biology*. 2005;43(5):745-
900 57.
901
- 902 65. Seki M, Narusaka M, Ishida J, Nanjo T, Fujita M, Oono Y, et al. Monitoring the
903 expression profiles of 7000 *Arabidopsis* genes under drought, cold and high-salinity
904 stresses using a full-length cDNA microarray. *The Plant journal : for cell and molecular*
905 *biology*. 2002;31(3):279-92.
906
- 907 66. Wu Y, Deng Z, Lai J, Zhang Y, Yang C, Yin B, et al. Dual function of *Arabidopsis*
908 *ATAF1* in abiotic and biotic stress responses. *Cell research*. 2009;19(11):1279-90.
909
- 910 67. Shahnejat-Bushehri S, Mueller-Roeber B, Balazadeh S. *Arabidopsis* NAC
911 transcription factor *JUNGBRUNNEN1* affects thermomemory-associated genes and
912 enhances heat stress tolerance in primed and unprimed conditions. *Plant Signaling &*
913 *Behavior*. 2012;7(12):1518-21.
914
- 915 68. Guan Q, Yue X, Zeng H, Zhu J. The protein phosphatase *RCF2* and its interacting
916 partner *NAC019* are critical for heat stress-responsive gene regulation and
917 thermotolerance in *Arabidopsis*. *Plant Cell*. 2014;26(1):438-53.

- 918
919 69. Fang Y, Liao K, Du H, Xu Y, Song H, Li X, et al. A stress-responsive NAC
920 transcription factor SNAC3 confers heat and drought tolerance through modulation of
921 reactive oxygen species in rice. *Journal of experimental botany*. 2015;66(21):6803-17.
922
923 70. Guo W, Zhang J, Zhang N, Xin M, Peng H, Hu Z, et al. The Wheat NAC
924 Transcription Factor TaNAC2L Is Regulated at the Transcriptional and Post-Translational
925 Levels and Promotes Heat Stress Tolerance in Transgenic Arabidopsis. *PLoS One*.
926 2015;10(8):e0135667.
927
928 71. Morcuende R, Bari R, Gibon Y, Zheng W, Pant BD, Blasing O, et al. Genome-wide
929 reprogramming of metabolism and regulatory networks of Arabidopsis in response to
930 phosphorus. *Plant, cell & environment*. 2007;30(1):85-112.
931
932 72. Misson J, Raghothama KG, Jain A, Jouhet J, Block MA, Bligny R, et al. A genome-
933 wide transcriptional analysis using *Arabidopsis thaliana* Affymetrix gene chips determined
934 plant responses to phosphate deprivation. *Proceedings of the National Academy of
935 Sciences of the United States of America*. 2005;102(33):11934-9.
936
937 73. Nilsson L, Müller R, Nielsen TH. Dissecting the plant transcriptome and the
938 regulatory responses to phosphate deprivation. *Physiologia Plantarum*. 2010;139(2):129-
939 43.
940
941 74. Krogh A, Larsson B, von Heijne G, Sonnhammer EL. Predicting transmembrane
942 protein topology with a hidden Markov model: application to complete genomes. *Journal
943 of molecular biology*. 2001;305(3):567-80.
944
945 75. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high
946 throughput. *Nucleic Acids Research*. 2004;32(5):1792-7.
947
948 76. Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics
949 Analysis Version 7.0 for Bigger Datasets. *Molecular biology and evolution*.
950 2016;33(7):1870-4.
951
952 77. Hu B, Jin J, Guo A-Y, Zhang H, Luo J, Gao G. GSDS 2.0: an upgraded gene
953 feature visualization server. *Bioinformatics*. 2015;31(8):1296-7.
954
955 78. Kamranfar I, Xue G-P, Tohge T, Sedaghatmehr M, Fernie AR, Balazadeh S, et al.
956 Transcription factor RD26 is a key regulator of metabolic reprogramming during dark-
957 induced senescence. *New Phytologist*. 2018;218(4):1543-57.

958

959

960 **Supporting Information:**

961 **S1 Fig. Bar graph showing the number of genes in each transcription factor family** 962 **in Arabidopsis and quinoa.**

963 Data generated using the plant transcription factor prediction tool in the plant transcription
964 factor database (PTFDB, <http://planttfdb.cbi.pku.edu.cn/>, version 4.0) (32).

965

966 **S2 Fig. Distribution of NAC genes along the genome.**

967 a) CqNAC distribution across sub-genome A and sub-genome B. b) Number of CqNAC
968 genes per chromosome.

969

970 **S2 Fig. Comparison of RNAseq with qRT-PCR results.**

971 Relative expression of CqNACs after treatment relative to control. a) Expression after salt
972 treatment. b) Expression after drought treatment. c) Expression after heat treatment.

973

974 **S1 Table. TFs in quinoa predicted by PTFDB.**

975 List of all predicted transcription factors belonging to different families in the genome of
976 quinoa as predicted by the plant transcription factor prediction tool in the plant
977 transcription factor database (PTFDB) (32).

978

979 **S2 Table. Predicted NACs in quinoa.**

980 List of predicted NACs obtained from the three search tools including BLASTp, Pfam and
981 PTFDB, and list of CqNACs validated by HHmer and prosite scanning tool.

982

983 **S3 Table: List of CqNACs and their homoeologes according to BLASTp and** 984 **synteny analysis.**

985 a) List of CqNACs that has a single homoeolog in same cluster, 52 genes "26 pairs". b)
986 List of CqNACs that has single homoeolog in different clusters (7 pairs "14genes"). c) List
987 of CqNACs that has more than one homolog in different cluster, (16 genes). d) List of
988 CqNACs that has no homolog according our parameters "25 genes".

989 **S4 Table: Comparison between phylogenetic subgroups of CqNACs and AtNACs.**

990 a) Number of Arabidopsis and quinoa NAC members in each phylogenetic subgroups. b)
991 Function of genes in the phylogenetic subgroups 8, 16 and 18 that are absent in quinoa.

992 **S5 Table: Cumulative expression from RNAseq data of all of the Quinoa CqNACs** 993 **in all of the treatments.**

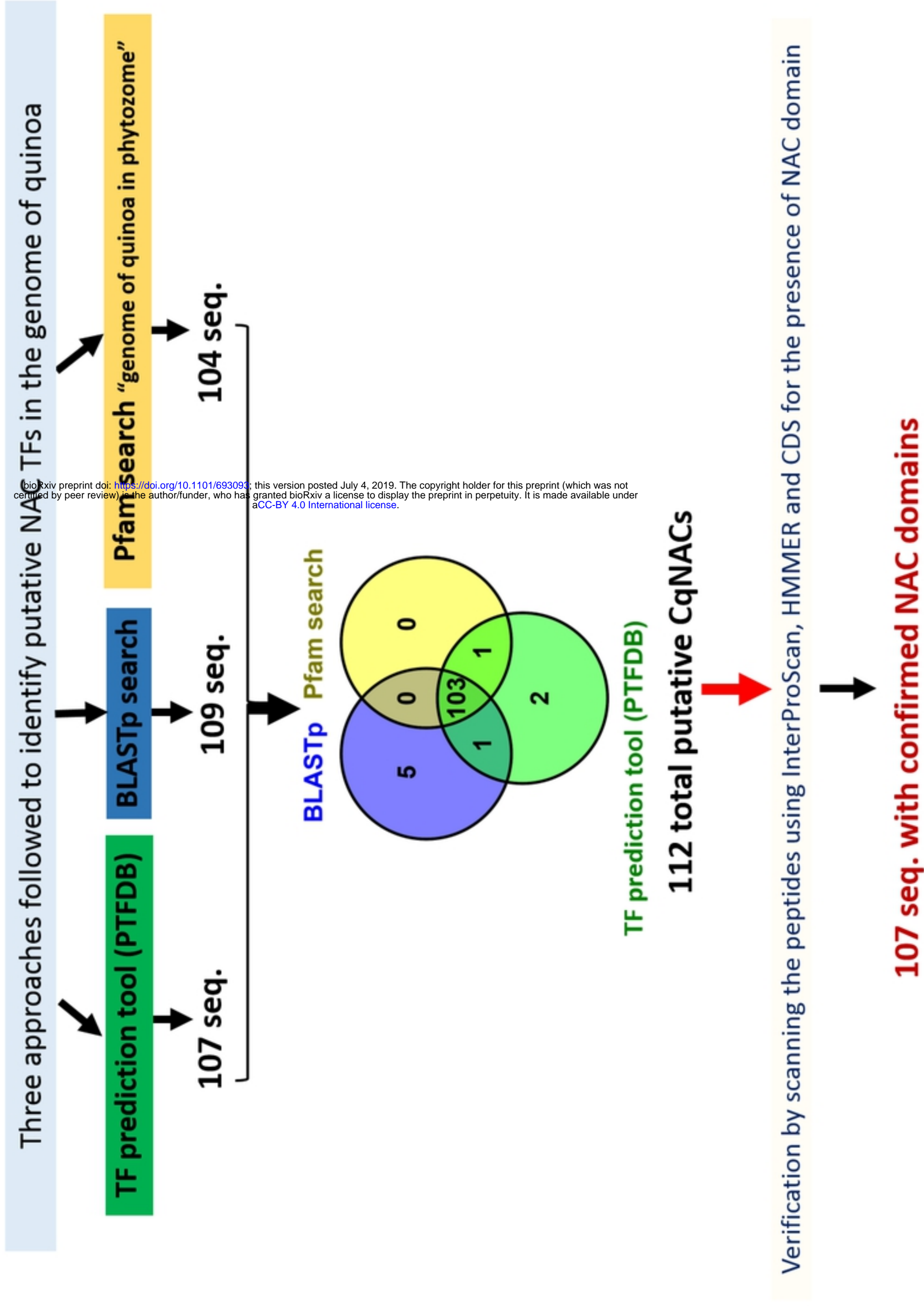
994

995 **S6 Table: Expression values of duplicated CqNACs in control and stress** 996 **conditions (salt, drought, heat and phosphate starvation).**

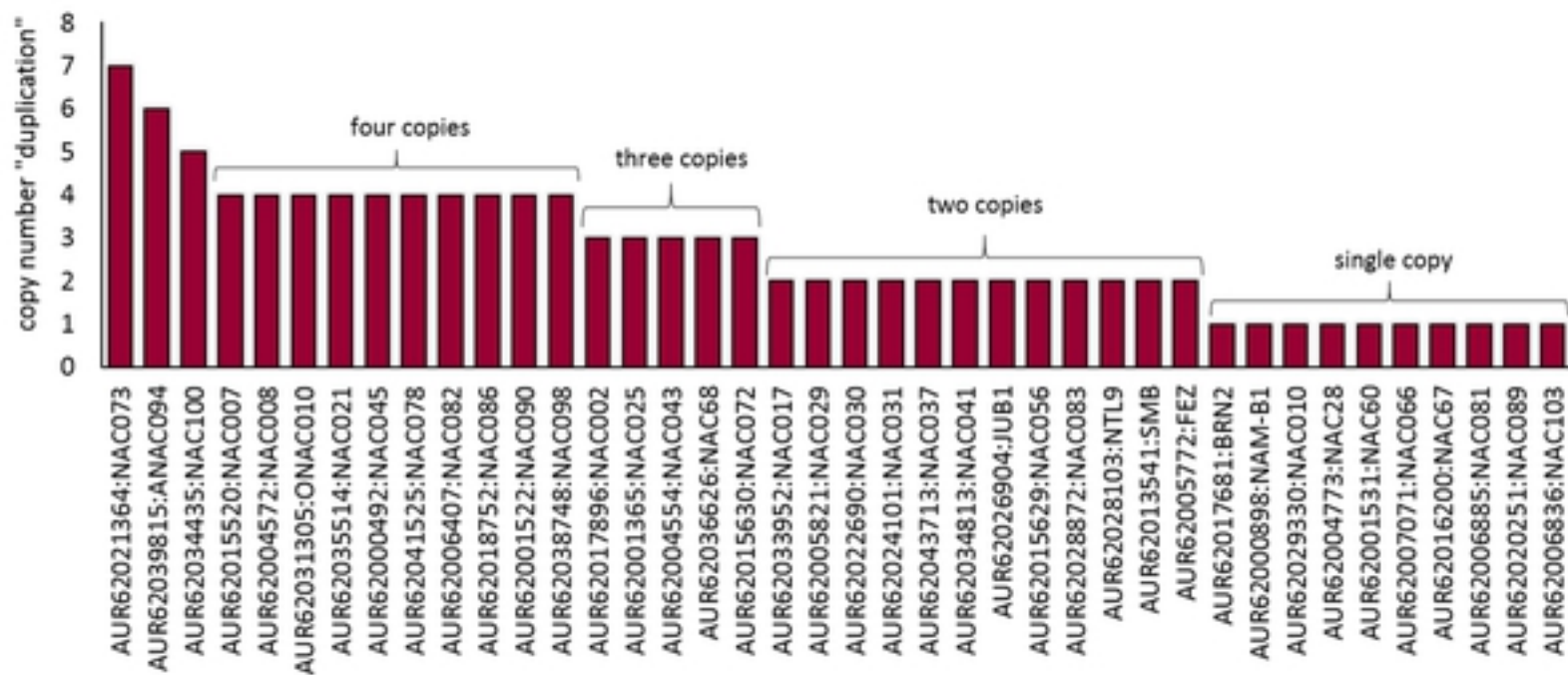
997 Expression values are based on RNAseq of three biological replicates. The significance
998 of gene expression differences was calculated at FDR=0.05.

999 **S7 Table: List of primer sequences used in qRT-PCR reaction.**

Fig 1



a)



b)



Fig 2

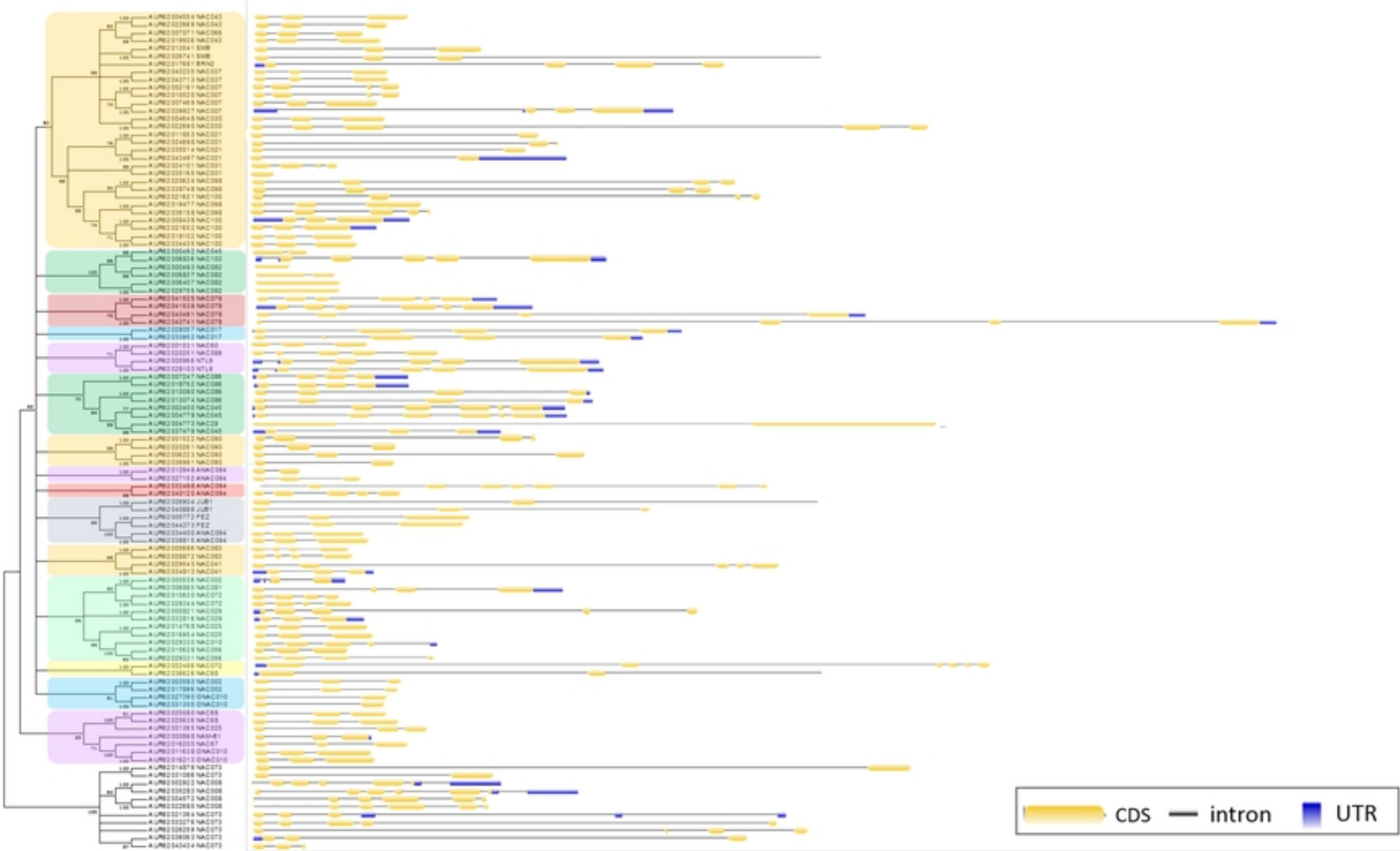


Fig 4

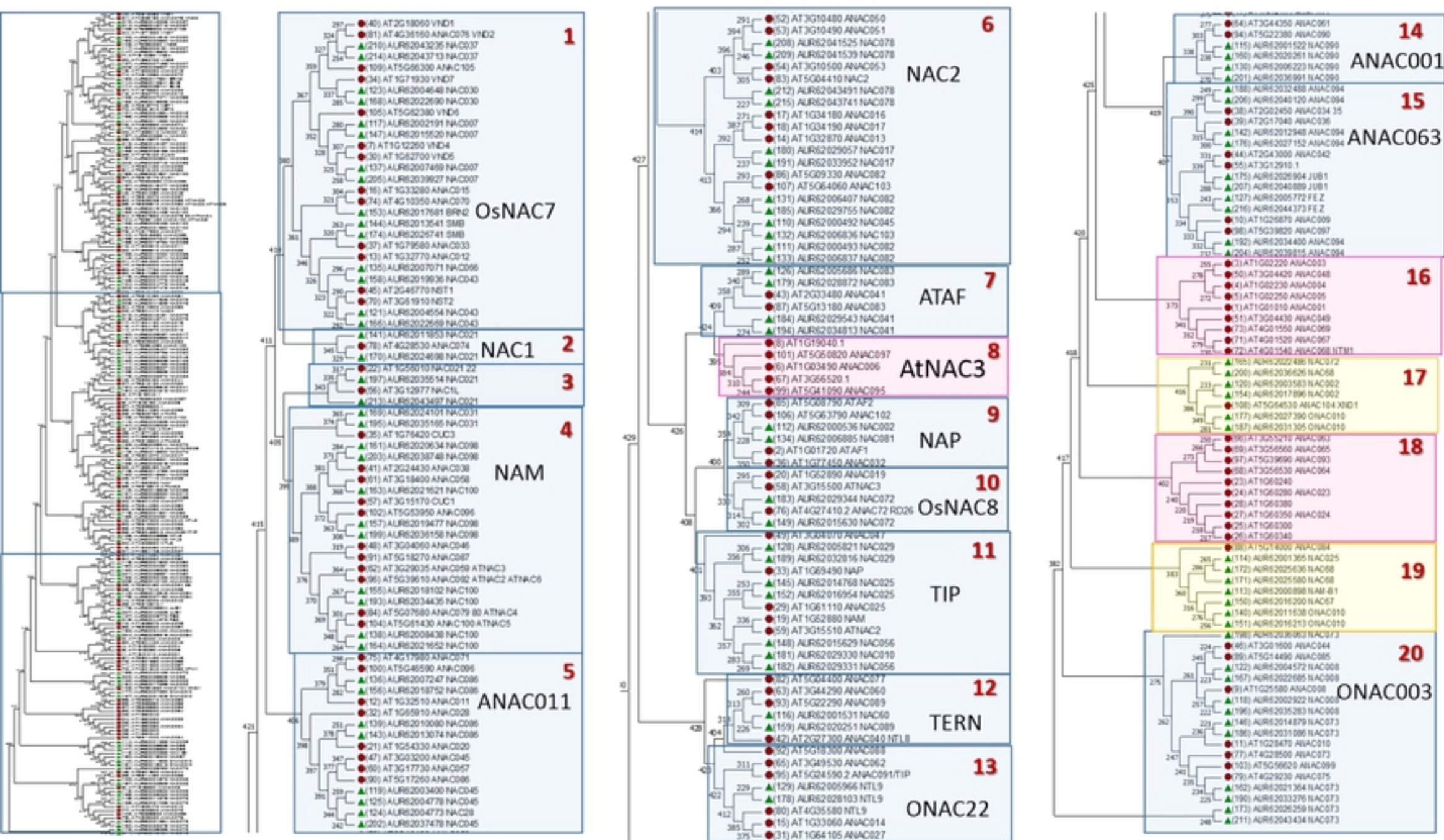
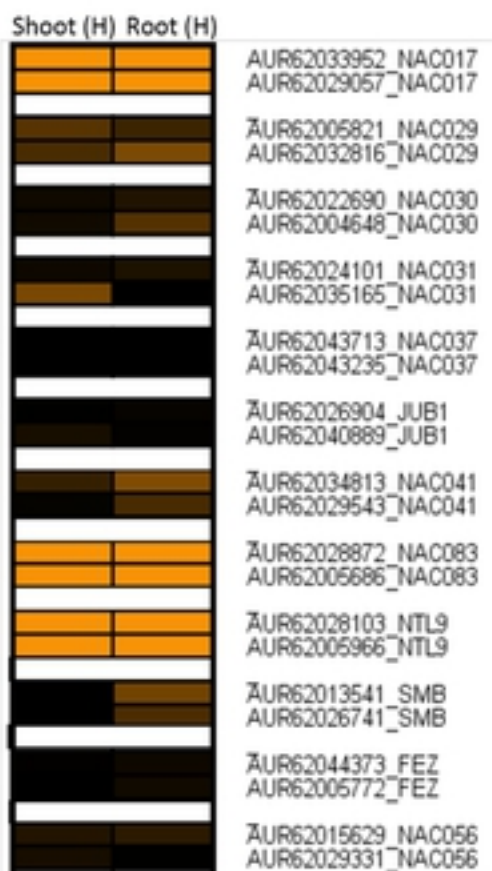
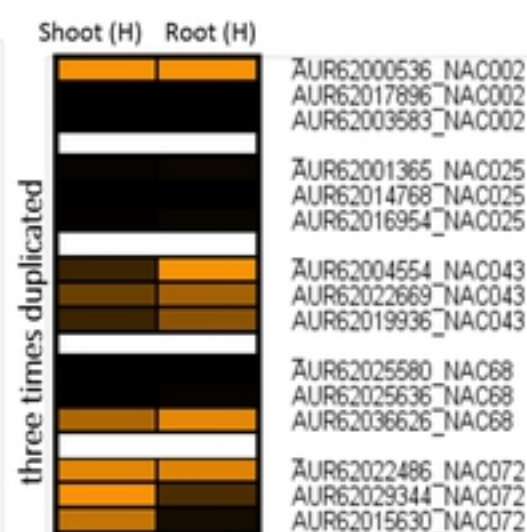


Fig 5

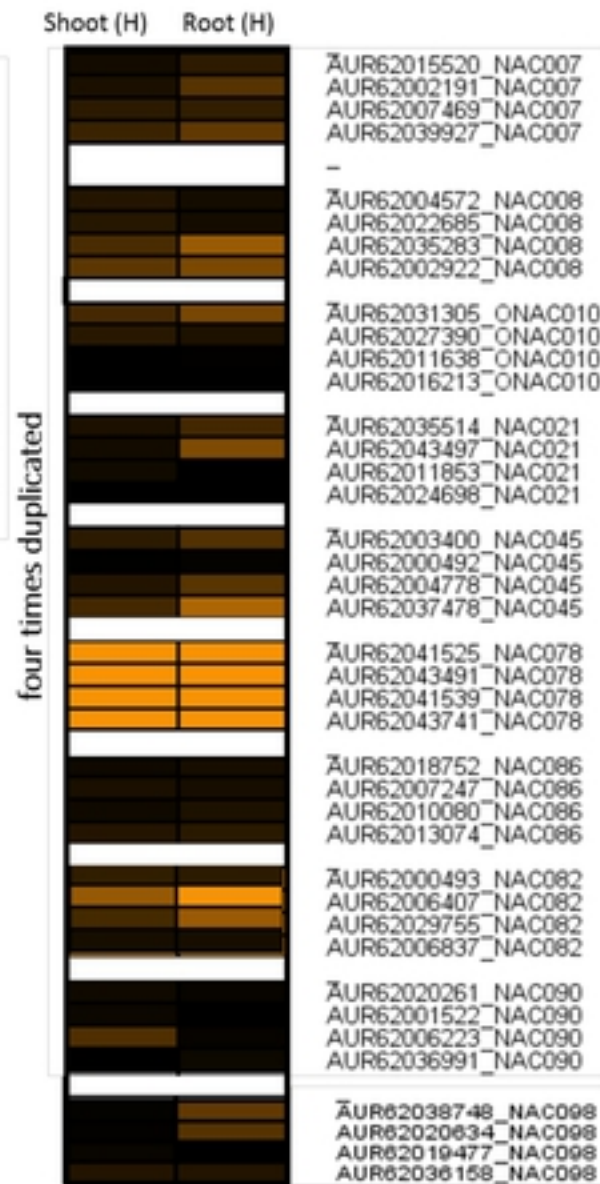
a) Genes two times duplicated:



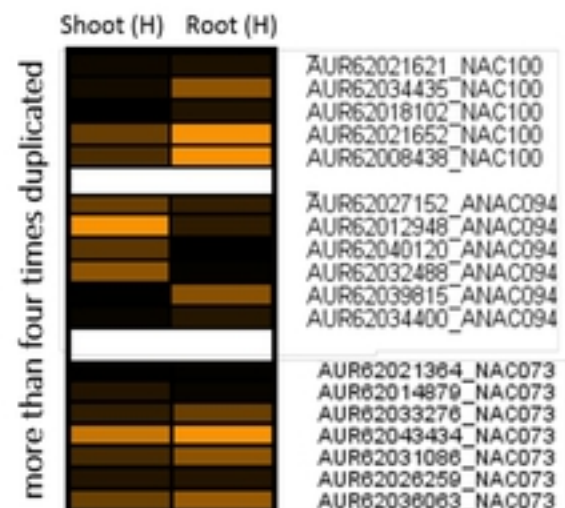
b) Genes three times duplicated:



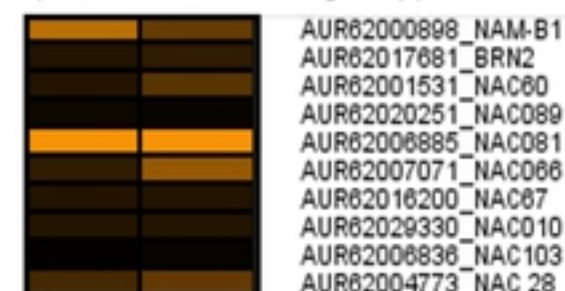
c) Genes four times duplicated:



d) Genes more than four times duplicated



e) Genes that have a single copy



Expression level: FPKM values



Fig 6

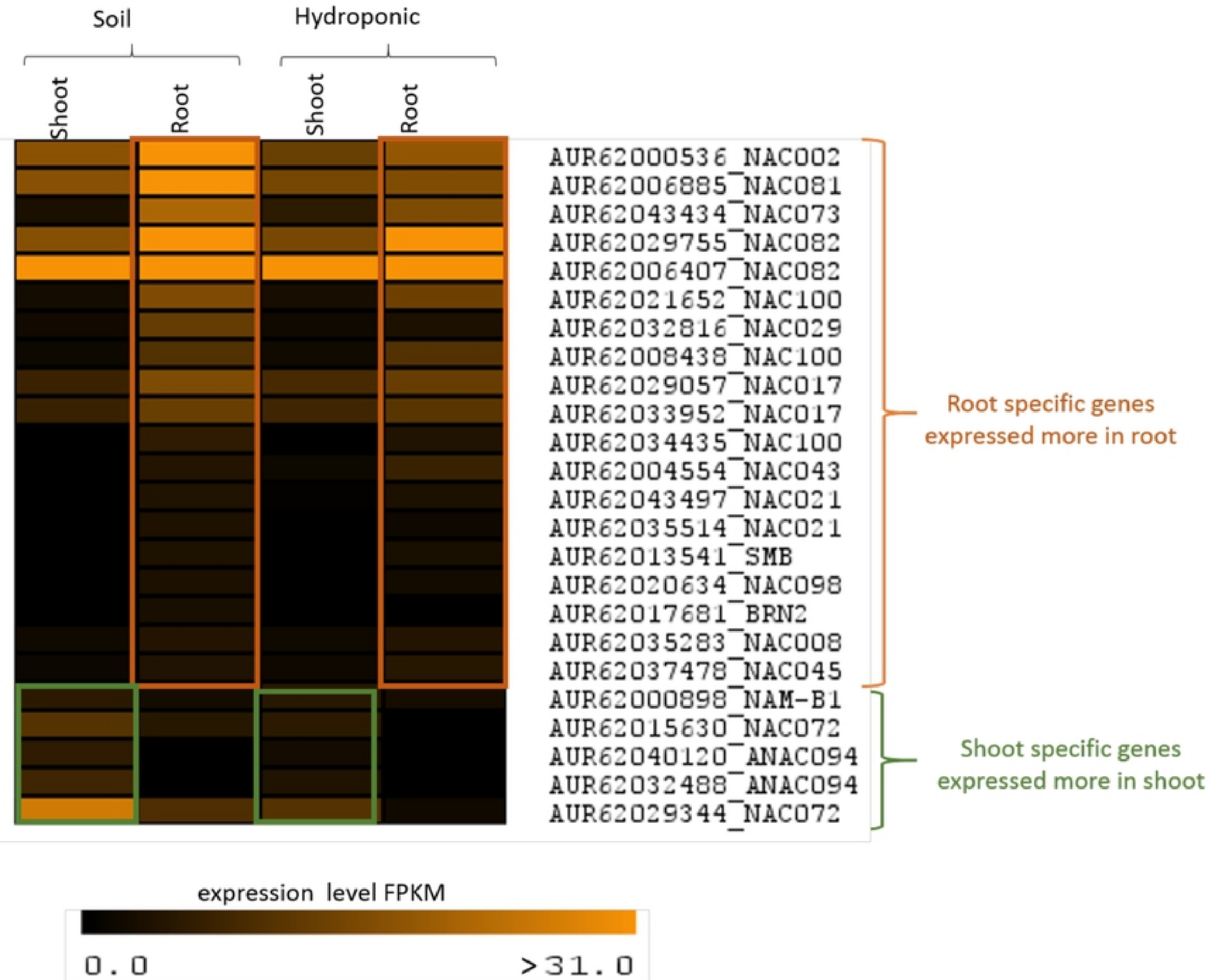


Fig 7

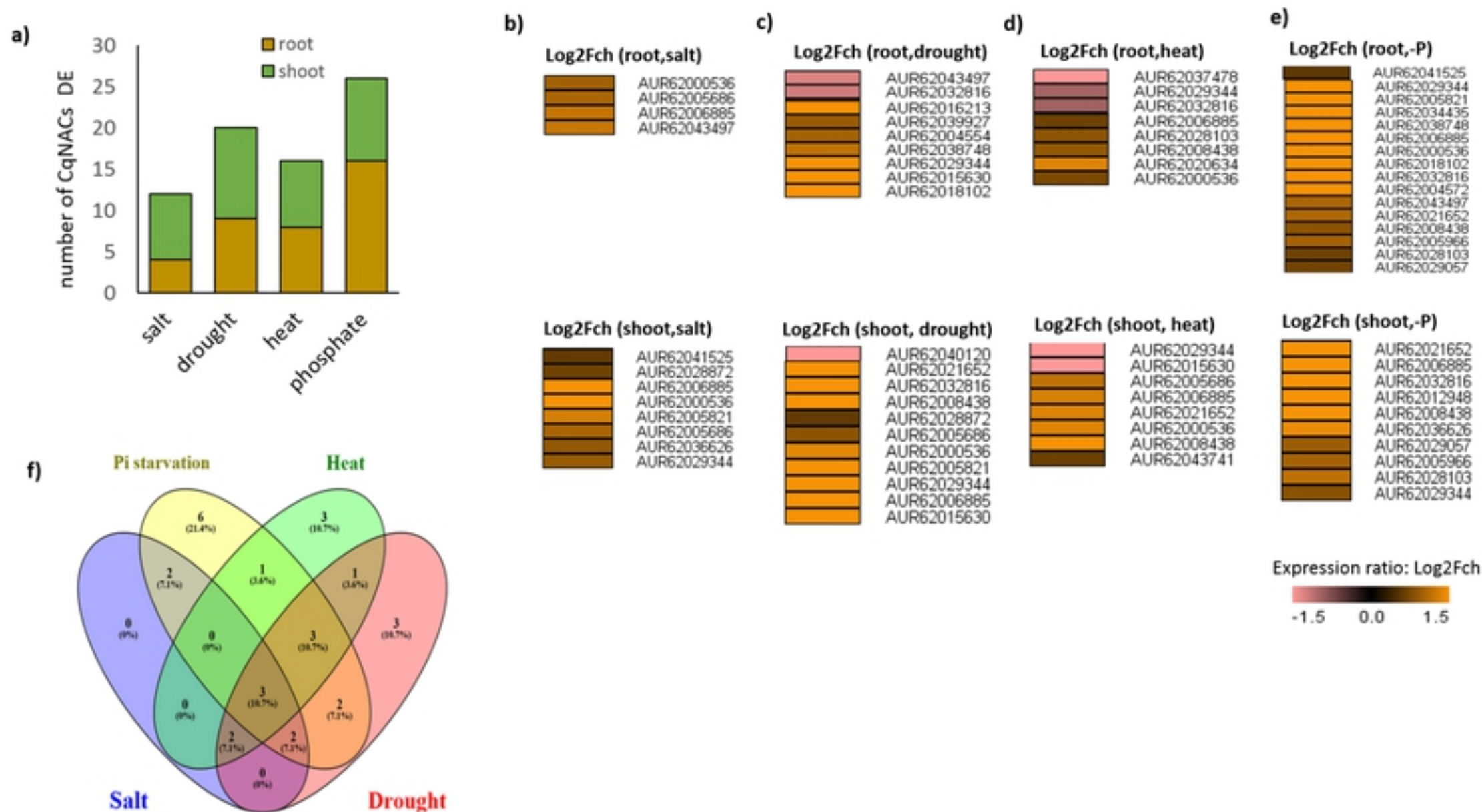


Fig 8

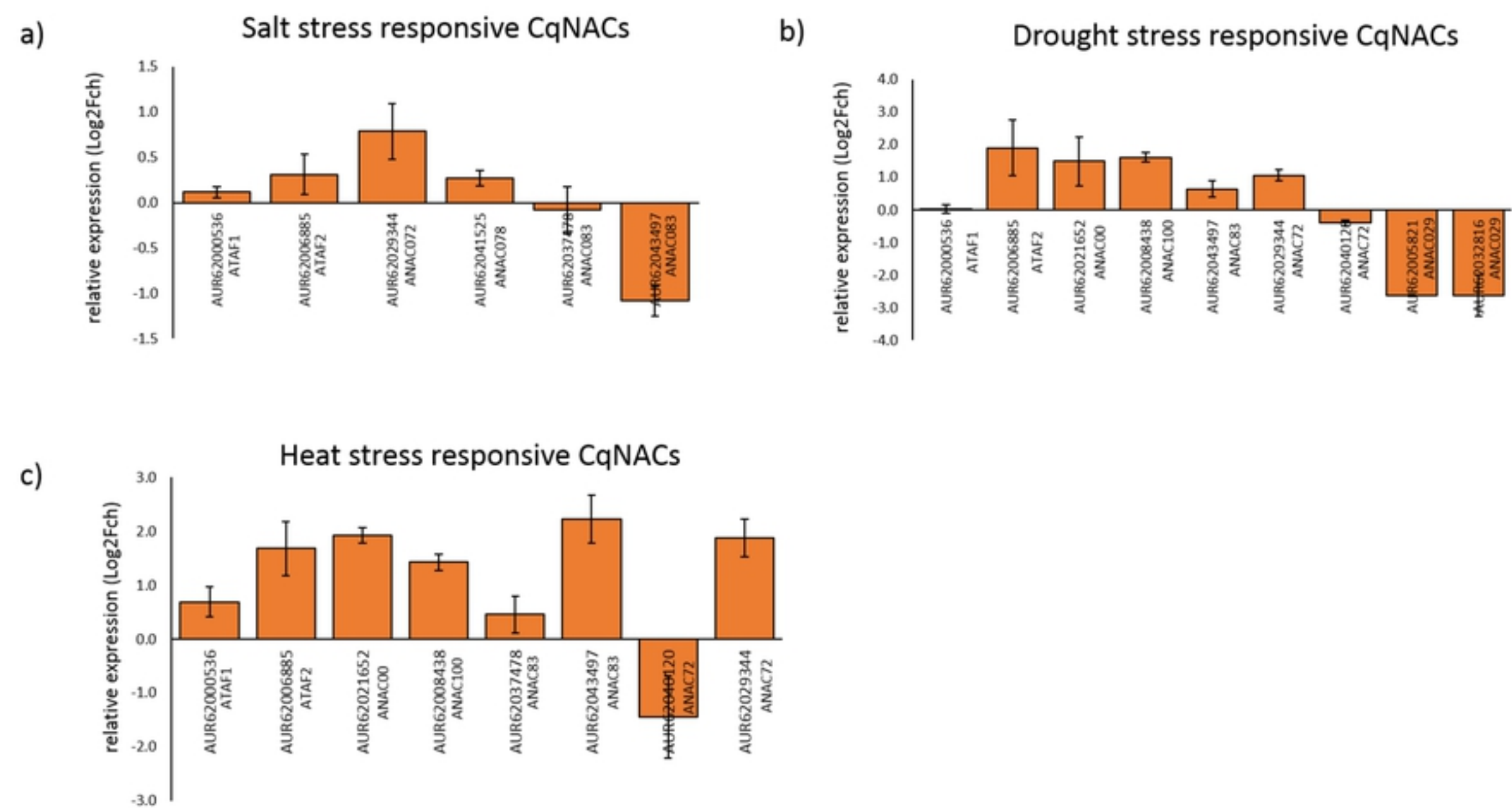


Fig 9