

RESEARCH ARTICLE

Identification of an NAC Transcription Factor Family by Deep Transcriptome Sequencing in Onion (*Allium cepa* L.)

Xia Zheng[☉], Shouwei Tang[☉], Siyuan Zhu, Qiuzhong Dai, Touming Liu^{*}

Institute of Bast Fiber Crops and Center of Southern Economic Crops, Chinese Academy of Agricultural Sciences, Changsha, China

☉ These authors contributed equally to this work.

* liutouming@gmail.com



OPEN ACCESS

Citation: Zheng X, Tang S, Zhu S, Dai Q, Liu T (2016) Identification of an NAC Transcription Factor Family by Deep Transcriptome Sequencing in Onion (*Allium cepa* L.). PLoS ONE 11(6): e0157871. doi:10.1371/journal.pone.0157871

Editor: Xiang Jia Min, Youngstown State University, UNITED STATES

Received: March 23, 2016

Accepted: June 6, 2016

Published: June 22, 2016

Copyright: © 2016 Zheng et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: Sequencing reads and transcripts assembled are deposited in the NCBI Sequence Read Archive (SRA) and Transcriptome Shotgun Assembly (TSA) Database, under accession number SRA347079 and GEOY00000000, respectively. The other relevant data are within the paper and its Supporting Information files.

Funding: This work was supported by Technology Innovation Program of China (grant number: CAAS-ASTIP-IBFC), http://www.caas.cn/en/research/research_program/index.shtml. The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Abstract

Although onion has been used extensively in the past for cytogenetic studies, molecular analysis has been lacking because the availability of genetic resources is limited. NAM, ATAF, and CUC (NAC) transcription factors (TFs) are plant-specific proteins, and they play key roles in plant growth, development, and stress tolerance. However, none of the onion NAC (CepNAC) genes had been identified thus far. In this study, the transcriptome of onion leaves was analyzed by Illumina paired-end sequencing. Approximately 102.9 million clean sequence reads were produced and used for *de novo* assembly, which generated 117,189 non-redundant transcripts. Of these transcripts, 39,472 were annotated for their function. In order to mine the CepNAC TFs, CepNAC genes were searched from the transcripts assembled, resulting in the identification of all 39 CepNAC genes. These 39 CepNAC proteins were subjected to phylogenetic analysis together with 47 NAC proteins of known function that were previously identified in other species. The results showed that they can be divided into five groups (NAC-I–V). Interestingly, the NAC-IV and -V groups were found to be likely related to the processes of secondary wall synthesis and stress response, respectively. The transcriptome analysis generated a substantial amount of transcripts, which will aid immensely in identifying important genes and accelerating our understanding of onion growth and development. Moreover, the discovery of 39 CepNAC TFs and the identification of the sequence conservation between them and NAC proteins published will provide a basis for further characterization and validation of their functions in the future.

Introduction

Transcription factors (TFs) are proteins that regulate gene expression by binding specific DNA sequences in the regulatory region of target genes, thereby modulating RNA polymerase activity for the initiation of transcription. NAM, ATAF, and CUC (NAC) TFs are one kind of proteins that only exist in plant. The name of this family was originally derived from the names of three proteins: no apical meristem (NAM), ATAF1-2, and CUC2 (cup-shaped cotyledon) [1,

Competing Interests: The authors have declared that no competing interests exist.

2]. NAC TFs have a highly conserved DNA-binding domain in the N-terminal region and a variable functional domain in the C-terminal region [3, 4]. Generally, this variable C-terminal domain can function as transcriptional activator or repressor [5–7]. The NAC family is a large gene family and comprises a large number of protein members in plants, and at least 151, 117, 163, and 101 non-redundant NAC TF protein members have been identified so far in rice, *Arabidopsis*, *Populus*, and soybean genomes, respectively [8–10].

At present, there are at least 50 NAC genes that have been functionally characterized [11, 12, 13], and they are involved in many aspects of plant growth and development, including embryo development, seed germination, lateral root development, maintenance of the shoot apical meristem, flower formation, plant immune response, cell division and growth, senescence, and formation of secondary walls [3, 11]. Moreover, NAC TFs were also found to participate in regulating of plant tolerance to abiotic/biotic stress [14, 15]. Therefore, NAC family is multifunctional proteins family with various roles in regulating plant growth and development.

Onion (*Allium cepa*) is one of the most important vegetable crops of the *Allium* family; it is known for its nutritional and medicinal properties and has been widely planted around the world with more than 5000 years [16]. Onion is a diploid ($2n = 16$) plant and has a large nuclear genome of approximately 16.4 Gb per 1C, which is about 38 and 130 times larger than the rice and *Arabidopsis* genomes, respectively [17, 18]. Molecular analysis in onion has been hindered by the limited availability of genetic resources, as a result of the onion's large nuclear genome, which presents challenges for gene identification. Therefore, while the onion has been used extensively in previous studies for cytogenetic analysis, the molecular analysis of onion genes has been lacking. In recent years, the rise of next generation sequencing (NGS) technologies has offered a powerful and cost-efficient tool for high-throughput sequence determination. Transcriptome analysis based on this technology has led to the discovery of new genes in hundreds of species [19–23]. However, no studies to date have identified NAC family genes in onion.

In the present study, we have *de novo* assembled and characterized the transcriptome of onion leaves using Illumina paired-end sequencing technology and searched onion NAC (CepNAC) genes from the assembled transcripts. Furthermore, we have performed phylogenetic analysis of the CepNAC TFs identified in this study along with 47 NAC proteins of known function that were previously identified in other species (S1 Table). The discovery of these CepNAC genes will provide a basis for further understanding of their function in onion growth and development.

Materials and Methods

Plant material and RNA extraction

The study included a local variety of onion (Chalinghuangpi) that produces yellow-skin bulbs and was collected from Chaling, Hunan, China. The onion was planted in the experimental field of the Institute of Bast Fiber Crops, Chinese Academy of Agricultural Sciences, Changsha, China, on Sept. 20, 2014. On Mar. 10, 2015, leaves tissues of three individuals were sampled. The tissue sample was immediately frozen in liquid nitrogen, and then stored at -80°C until use. E.Z.N.A. Plant RNA Kit (OMEGA Bio-Tek, USA) was used to extract the total RNA according to the manufacturer's protocol.

cDNA library construction, sequencing, and assembly

Illumina sequencing was performed at Novogene Bioinformatics Technology Co., Ltd, Beijing, China (www.novogene.cn). RNA of onion leaves was used to construct the cDNA libraries with fragment lengths of approximately 250 bp (± 25 bp). Paired-end sequencing was then

performed using the Illumina sequencing platform (HiSeq™ 2500) according to the manufacturer's instructions (Illumina, San Diego, CA). After trimming adapter sequences and filtering low-quality reads, the clean reads were used to search against NCBI non-redundant protein sequences (NR) database by BLAST (Basic Local Alignment Search Tool) with an E-value threshold of 10^{-1} , and the reads that showed significant similarity with other species might be the sequences derived from other species, thereby being removed. Thereafter, the remaining clean reads were used to *de novo* assemble the onion transcriptome using Trinity [24]. Sequencing reads and transcripts assembled are deposited in the NCBI Sequence Read Archive (SRA) and Transcriptome Shotgun Assembly (TSA) Database, under accession number SRA347079 and GEOY00000000, respectively.

Gene annotation and classification

To annotate the transcripts of onion transcriptome using bioinformatics approaches, each of the assembled transcripts was searched against seven public databases, including KEGG Orthology (KO) database, protein family (PFAM) database, eukaryotic orthologous groups (KOG) database, NR database, NCBI nucleotide sequences (NT) database, Swiss-Prot protein database, and Gene Ontology (GO) database. WEGO software was used to perform GO functional classification of all transcripts to visualize gene functions distribution. Transcripts annotated in the KOG database were classified according to the COG group that each transcript was assigned to, whereas transcripts annotated from the KO database were grouped according to the KEGG pathway that gene was assigned to.

Identification of NAC genes and phylogenetic analysis

In order to identify the NAC genes, the keywords “NAC,” “no apical meristem,” or “NAM” were used as the query to search against annotations of the onion transcripts assembled. ORFs of identified NAC genes were screened using the findorf software. All non-redundant genes identified were further validated as NAC genes by manually checking for their conserved NAM domain. In addition, information regarding 47 functionally characterized NAC genes was collected from previous studies conducted in other plant species, and can be found in [S1 Table](#). We performed the phylogenetic analysis on these 47 NAC proteins of known function, together with the CepNAC proteins identified in this study. Clustal X (version 1.83) program [25] was used to perform multiple sequence alignments of full-length protein sequences, including the highly conserved N-terminal NAM domain and the more divergent C-terminal activation domain. Neighbor-joining (NJ) method was used to construct an unrooted phylogenetic tree by using MEGA 4.0 [26], and the bootstrap analysis was performed with 1000 replicates. NAC proteins of known function that had closest evolutionary relationship to each CepNAC protein were ascertained based on the phylogenetic tree, and their percentage similarity analyzed using the BLASTP program.

Results

Illumina paired-end sequencing and *de novo* assembly

RNA extracted from onion leaves was used to construct a library of cDNA sequences. In all 102.9 million clean sequence reads with lengths of 125 bp were sequenced from the library. Next, these sequence reads were used for the *de novo* assembly using Trinity software. Sequences that were not extended on either end were defined as transcripts. Eventually, a total of 117,189 non-redundant transcripts were assembled. The average length and total length of these *de novo* assembled transcripts were 632 bp and 74.03 Mb, respectively ([Table 1](#)). The

Table 1. Summary of transcriptomes assembled by Illumina sequencing.

| | |
|-----------------------------------|-----------------|
| Clear reads | 102,874,740 |
| Clean bases (Gb) | 12.86 |
| Transcript number | 117,189 |
| Total length of transcript (Mb) | 74.03 |
| Average length of transcript (bp) | 632 |
| Transcript number annotated | 39,472 (33.68%) |

doi:10.1371/journal.pone.0157871.t001

length was between 500 and 1000 for 21,559 (18.4%) transcripts, between 1001 and 2000 bp for 13,322 (11.4%) transcripts, more than 2000 bp for 6,008 (5.1%) transcripts, and no more than 500 bp for all remaining transcripts (Fig 1).

Functional annotation

To validate and annotate these assembled transcripts, sequence similarity searches were performed against in seven public databases. Of 117,189 transcripts, 32,919 (28.1%) showed significant similarity to known proteins in the NCBI non-redundant protein sequence (NR) database, 11,141 (9.5%) in the NCBI nucleotide sequence (NT) database, 10,741 (9.2%) in the KEGG orthology (KO) database, 22,543 (19.2%) in the SwissProt database, 24,578 (21.0%) in the protein family (PFAM) database, 24,672 (21.1%) in the Gene Ontology (GO) database, and 11,175 (9.5%) in the eukaryotic orthologous groups (KOG) database (Fig 2). Finally, 39,472 (33.7%) transcripts showed similarity to known proteins in at least one of these seven databases, and were achieved for annotating function.

Functional classification

Functional classification of assembled transcripts was performed by BLAST searching against GO, KOG and KEGG databases. The results of the GO analysis showed matches to known proteins of 56 GO classes for all 24,672 transcripts, with a total of 125,547 functional terms. The majority of these GO functional terms consisted of assignments to the biological process

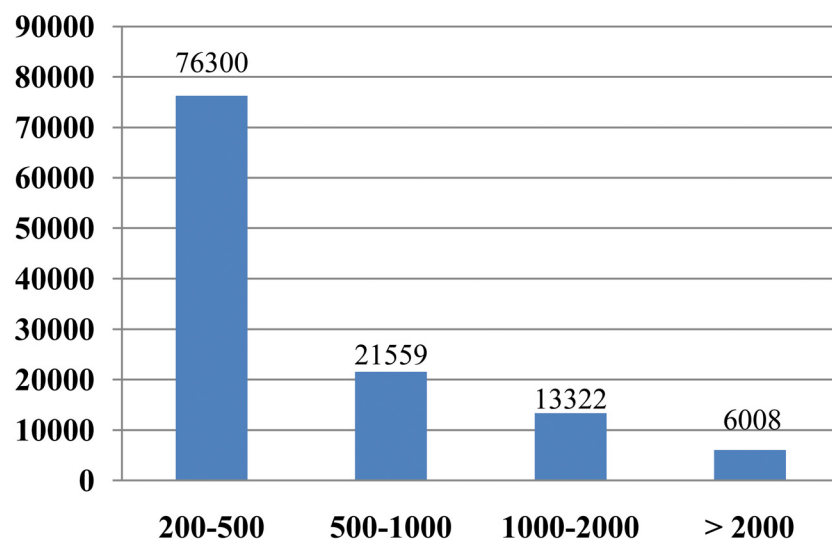


Fig 1. Length distribution of assembled transcripts.

doi:10.1371/journal.pone.0157871.g001

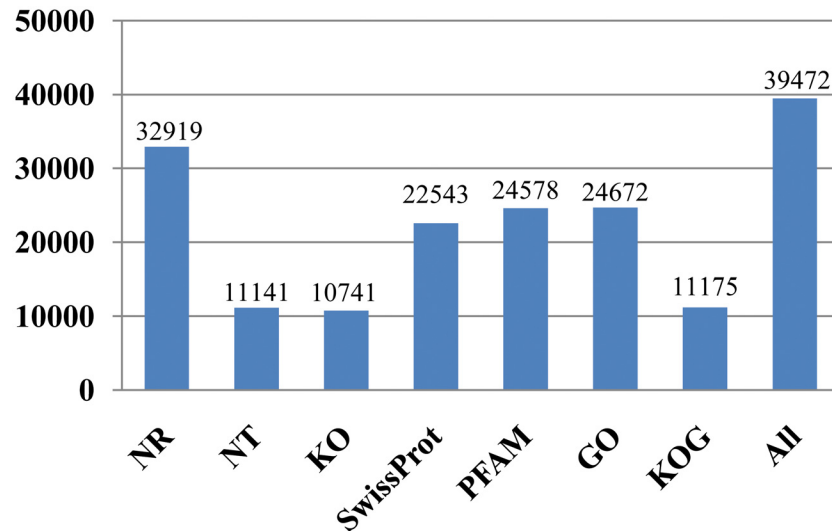


Fig 2. Transcript numbers annotated in the seven public databases searched.

doi:10.1371/journal.pone.0157871.g002

ontology (59,019; 47.0%), followed by cellular component (37,756; 30.1%) and molecular function (28,772; 22.9%) ontologies (Fig 3). KOG functional classification resulted in the assignment of 11,175 to all 25 KOG categories. Some transcripts were assigned to several KOG categories, which led to a total of 12,517 sequences being classified into 25 KOG categories (Fig 4). In addition, on using BLASTx with an E-value threshold of 10^{-5} against the KEGG database, 10,741 transcript sequences matched a total of 12,035 functional terms in the database and were assigned to 32 KEGG categories (Fig 5).

Identification of the CepNAC genes

The NAC genes of onion were discovered by searching annotated sequences among the assembled transcripts. A total of 39 *CepNAC* genes were identified, and were named *CepNAC01* through *CepNAC39* (Table 2). Among these 39 genes, 35 contained the full-length open reading frame (ORF), and the other 4 genes were partial (Table 2). The proteins encoded by these 4 genes with partial sequence contained the complete NAM domain. The DNA sequences of *CepNAC* genes and their predicted protein sequences are shown in S1 File.

Phylogenetic analysis of CepNAC proteins and functionally characterized NAC proteins identified in other species

All 39 *CepNAC* protein sequences identified in this study, together with 47 functionally characterized NAC proteins of other species, were subjected to phylogenetic analysis. The results suggested that the 86 NAC proteins can be classified into five groups (I–V, Fig 6). A total of 6, 7, 13, 1, and 7 *CepNAC* members were classified into NAC-I, II, III, IV, and V, respectively (Table 2, Fig 6). Moreover, a total of 8, 9, 9, and 16 NAC proteins of known function were assigned to the NAC-II, III, IV and V subfamilies, respectively, whereas none of the NAC TFs reported fell into the NAC-I group (Fig 6). In addition, five *CepNAC* members and five NAC proteins of known function were not assigned to any phylogenetic group (Fig 6). Interestingly, all 9 NAC TFs reported to be in the NAC-IV subfamily function as secondary wall synthesis regulators, and all 16 NAC TFs reported to be in the NAC-V group function as stress tolerance modulators (Fig 6; S1 Table). This suggests that the NAC-IV and V subfamilies might be

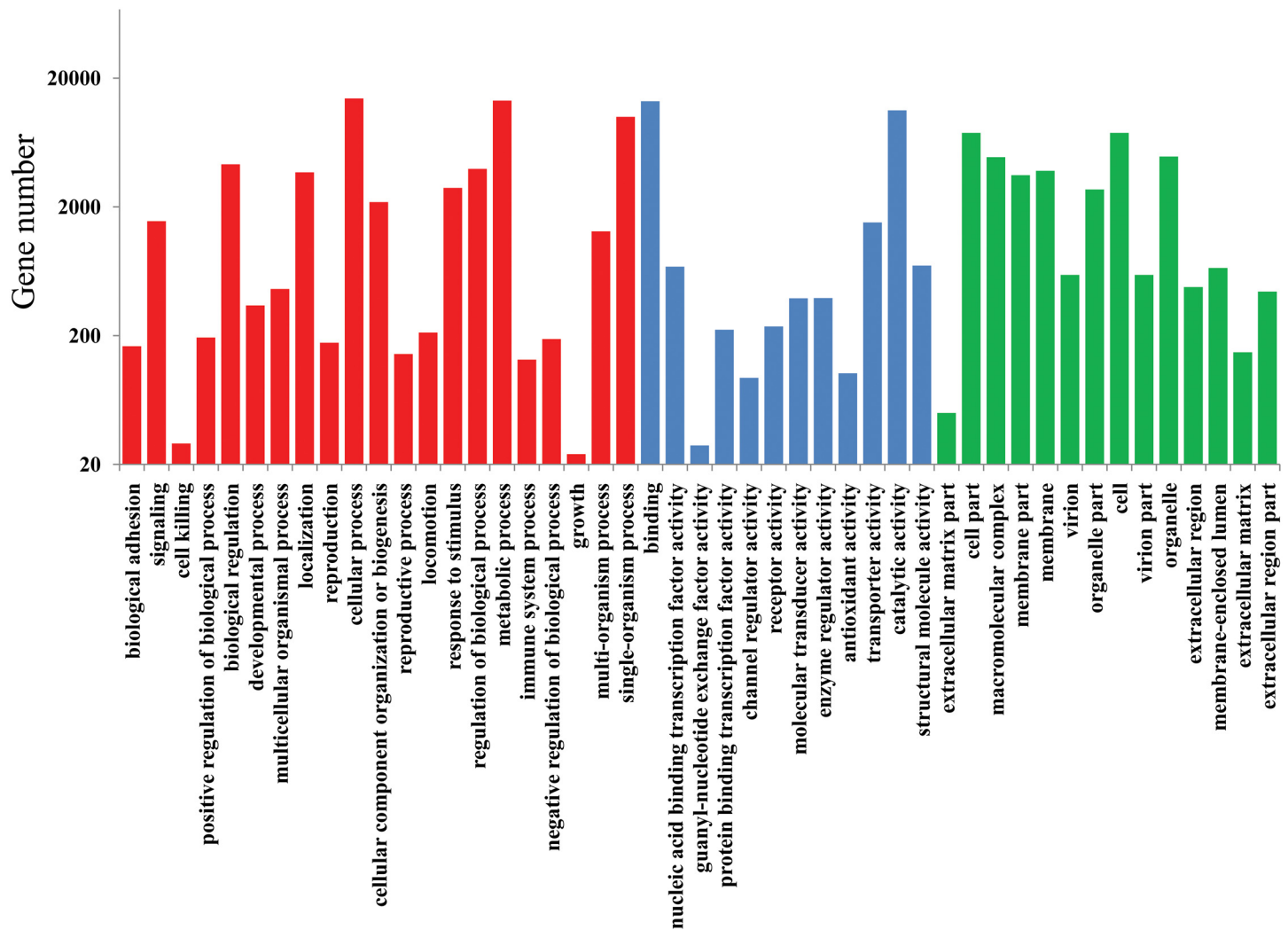


Fig 3. Gene ontology classifications of assembled transcripts; the red, blue, and green histogram bar represented three main categories: biological process, molecular function, and cellular component.

doi:10.1371/journal.pone.0157871.g003

functionally related to the processes of secondary wall synthesis and stress response, respectively. On the other hand, the function of NAC TFs reported in the NAC-II and III groups was very diverse (S1 Table). Seventeen CepNAC TFs showed highest homology to the functionally characterized NAC proteins reported in other species, with a sequence similarity ratio ranging from 32% to 82% (Table 2). For 7 of those CepNAC proteins, the similarity ratio with the NAC proteins of known function was more than 60%.

Discussion

Characterization of the transcriptome of onion leaves

Because of its cost-efficiency, high throughput and speed, NGS technology offers several advantages over traditional Sanger sequencing, which is costly, time-consuming, and laborious. NGS has been proven a powerful tool in several researching areas, including re-sequencing, micro-RNA expression profiling, DNA methylation, and especially *de novo* transcriptome sequencing for non-model organisms [27]. Based on this technology, the transcriptomes of

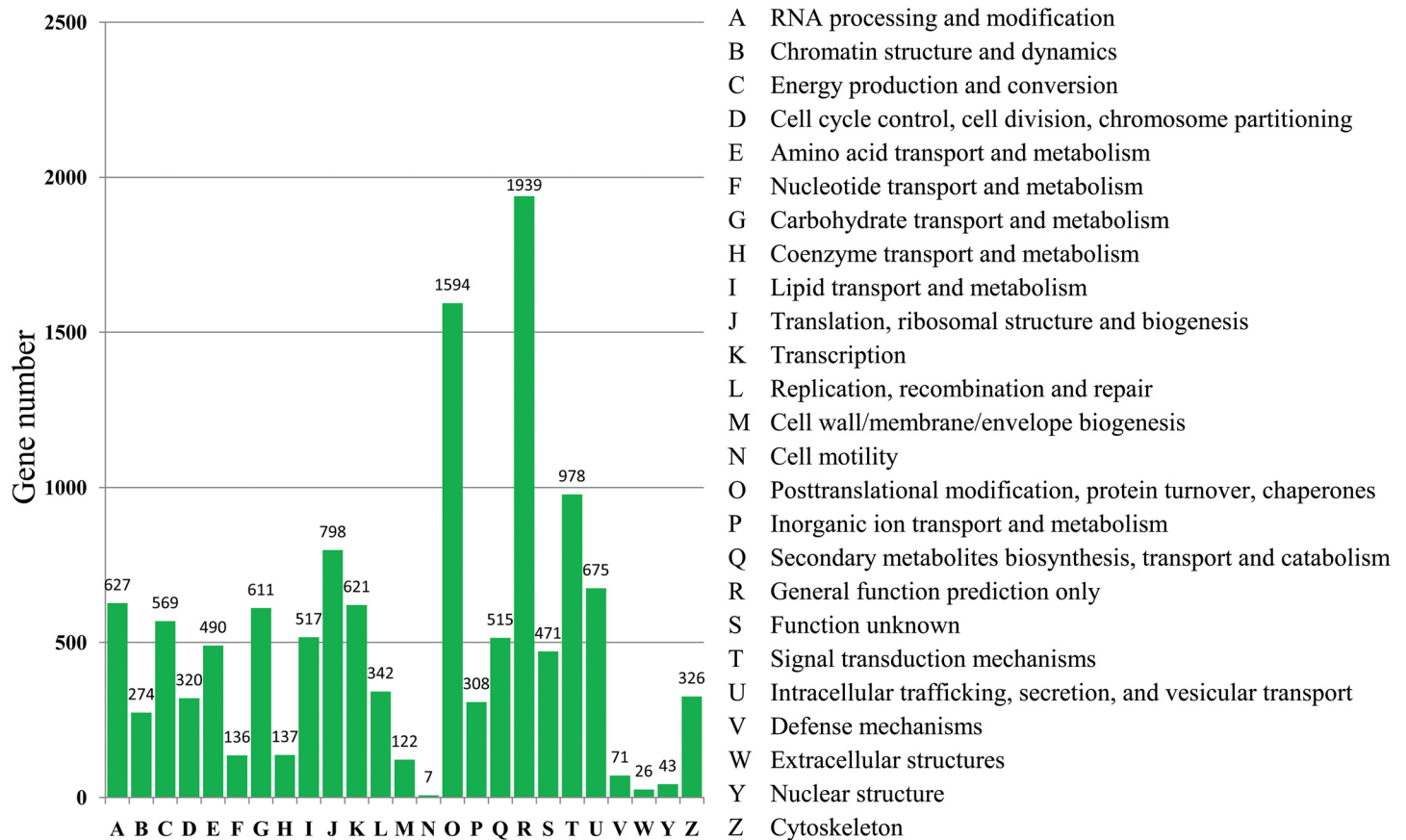


Fig 4. Histogram presentation of clusters of eukaryotic ortholog groups (KOG) classification.

doi:10.1371/journal.pone.0157871.g004

hundreds of species, including non-model organisms, were characterized [22], and a large number of expressed genes have been discovered, thereby accelerating our understanding of the complexity of expression, regulation and networks of gene in model and non-model organisms.

Onion is an economically important crop. By Sanger sequencing for the cDNA library, a total of 11,008 unique ESTs obtained from this crop [28], which first provided an insight into onion expressed genes. In a recent study, Rajkumar et al. *de novo* assembled and characterized the red bulb onion transcriptome, generating a total of 293,475 genes with a total length of 281 Mb [16]. Obviously, the number of genes assembled in the above study is enormous, far greater than that of other *Allium* species, such as garlic (135,360 transcripts), Chinese chive (60,031 transcripts), and Welsh onion (53,837 transcripts) [29–31]. The transcriptome characterized in this previous study was also huge (almost two times the genome size of *Arabidopsis*). Probably, there are some redundant genes in the previous onion transcriptome. In addition, as only the onion bulb was used for transcriptome analysis [16], genes that are specifically expressed in other tissues such as leaf were not identified.

In the present study, we have described the transcriptome analysis in yellow bulb onion leaves, and identified 117,189 non-redundant transcripts. However, among these transcripts, there were only 33.7% of them achieved for functional annotation. This low ratio might be caused by the large size of onion transcriptome. Our result revealed that onion had a transcriptome of about 74 MB. Because of its large size, it is difficult to assemble for the onion

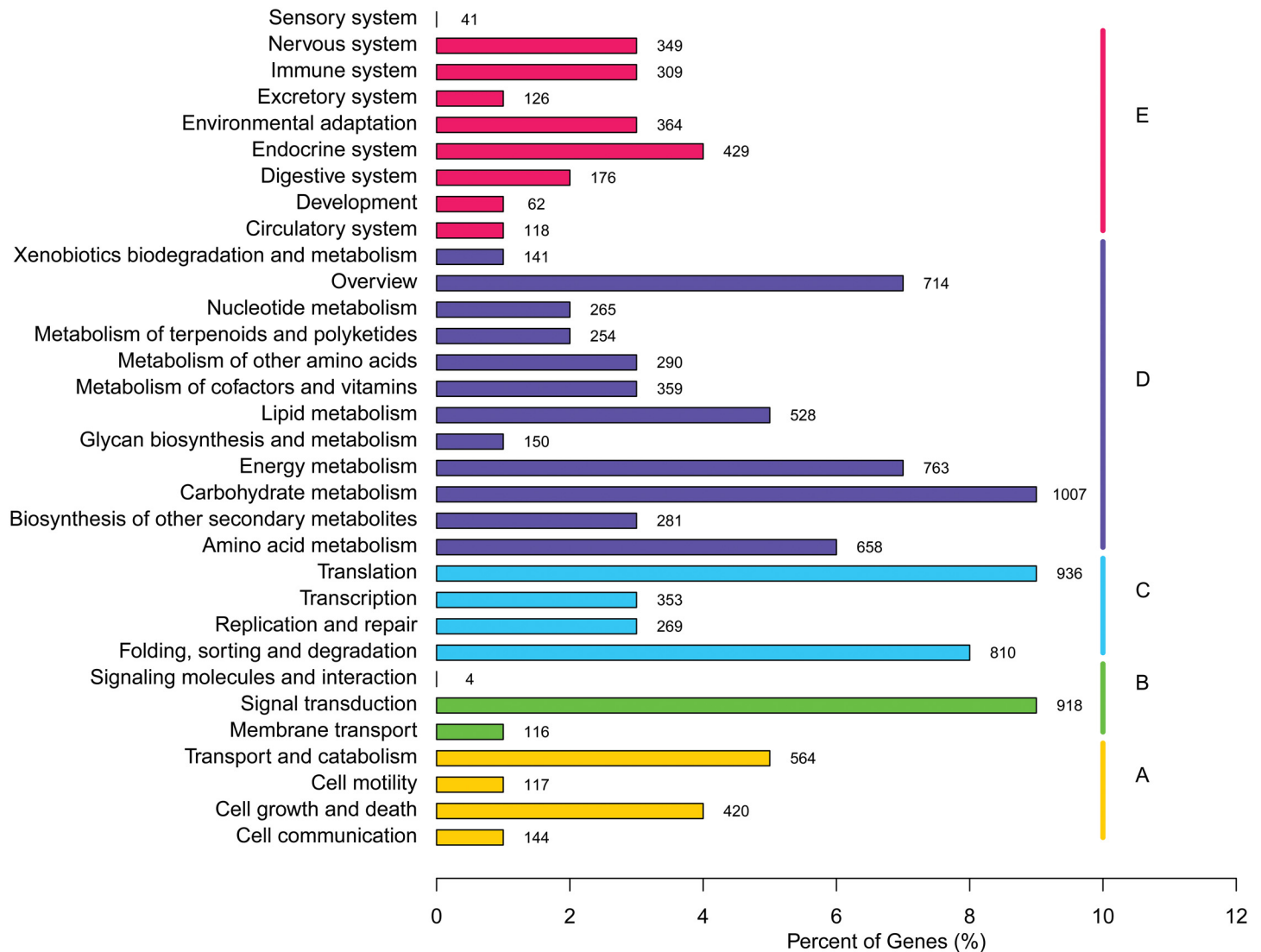


Fig 5. Classifications of transcripts according to the matching KEGG pathway; A, B, C, D, and E represent five main categories: cellular processes, environmental information processing, genetic information processing, metabolism, and organismal systems, respectively.

doi:10.1371/journal.pone.0157871.g005

transcriptome. Therefore, although that the average sequencing depth of coverage in the transcriptome was more than 170 folds in this study, many small sequence fragments still were generated, and they were difficult for annotating the function. Actually, a similar low ratio for the annotated transcripts had been observed in onion transcriptome in previous study [16]. There were substantial amount of transcripts obtained in this study, which will immensely help us to explore the major genes and accelerate our understanding in the onion growth and development.

Identification of 39 CepNAC TFs

There are 8,133 NAC genes that have been identified from 74 species in the PlantTFDB database [32]. Previous studies revealed that NAC TFs are involved in many aspects of plant growth and development, including modulation of senescence, secondary wall synthesis, stress

Table 2. The information of NAC genes of onion.

| Gene ^a | Gene length | PSC ^b | PEC ^b | Predicted protein | The most similarity gene | | Group |
|-------------------|-------------|------------------|------------------|-------------------|--------------------------|---------------------|-------|
| | (bp) | (bp) | (bp) | | Gene | RS ^c (%) | |
| <i>CepNAC01</i> * | 995 | 1 | 993 | 331 | | | I |
| <i>CepNAC02</i> | 1018 | 30 | 845 | 272 | <i>ANAC036</i> | 47 | III |
| <i>CepNAC03</i> | 1349 | 125 | 1078 | 318 | <i>ENAC1</i> | 44 | III |
| <i>CepNAC04</i> | 1289 | 56 | 796 | 247 | <i>OsNAC4</i> | 80 | V |
| <i>CepNAC05</i> | 1291 | 135 | 1034 | 300 | <i>ORE1</i> | 67 | II |
| <i>CepNAC06</i> | 2081 | 71 | 1822 | 584 | <i>RIM1</i> | 61 | III |
| <i>CepNAC07</i> | 3456 | 802 | 1299 | 166 | | | III |
| <i>CepNAC08</i> | 1204 | 74 | 973 | 300 | <i>OsNAC10</i> | 56 | V |
| <i>CepNAC09</i> | 3028 | 44 | 2371 | 776 | | | I |
| <i>CepNAC10</i> | 3167 | 90 | 1097 | 336 | | | I |
| <i>CepNAC11</i> | 1122 | 93 | 989 | 299 | <i>NAC1</i> | 47 | II |
| <i>CepNAC12</i> | 2373 | 198 | 2138 | 647 | | | III |
| <i>CepNAC13</i> | 1130 | 103 | 954 | 284 | <i>NAC1</i> | 46 | II |
| <i>CepNAC14</i> | 1201 | 49 | 1026 | 326 | | | V |
| <i>CepNAC15</i> | 1689 | 34 | 1416 | 461 | | | III |
| <i>CepNAC16</i> | 1569 | 340 | 1074 | 245 | <i>OsNAC52</i> | 58 | V |
| <i>CepNAC17</i> | 2027 | 935 | 1864 | 310 | <i>PtrWND6B</i> | 51 | IV |
| <i>CepNAC18</i> * | 768 | 86 | 766 | 227 | <i>RIM1</i> | 60 | III |
| <i>CepNAC19</i> | 1417 | 250 | 978 | 243 | | | |
| <i>CepNAC20</i> | 1878 | 97 | 1338 | 414 | | | I |
| <i>CepNAC21</i> * | 675 | 2 | 673 | 224 | | | |
| <i>CepNAC22</i> | 1040 | 121 | 855 | 245 | | | II |
| <i>CepNAC23</i> | 1265 | 50 | 610 | 187 | <i>OsNAC4</i> | 82 | V |
| <i>CepNAC24</i> | 1095 | 215 | 931 | 239 | | | II |
| <i>CepNAC25</i> | 1163 | 372 | 1091 | 240 | | | II |
| <i>CepNAC26</i> | 2046 | 134 | 1774 | 547 | | | III |
| <i>CepNAC27</i> | 1217 | 245 | 1018 | 258 | | | II |
| <i>CepNAC28</i> | 1143 | 84 | 551 | 156 | <i>ENAC1</i> | 61 | III |
| <i>CepNAC29</i> | 1618 | 40 | 489 | 150 | | | V |
| <i>CepNAC30</i> | 1239 | 64 | 1215 | 384 | | | I |
| <i>CepNAC31</i> | 2037 | 451 | 1713 | 421 | <i>FSQ6</i> | 41 | III |
| <i>CepNAC32</i> | 2120 | 157 | 1953 | 599 | <i>NTL4</i> | 32 | III |
| <i>CepNAC33</i> | 1513 | 268 | 1143 | 292 | | | |
| <i>CepNAC34</i> | 1648 | 118 | 1407 | 430 | | | I |
| <i>CepNAC35</i> | 802 | 30 | 608 | 193 | <i>XND1</i> | 51 | |
| <i>CepNAC36</i> | 1172 | 298 | 948 | 217 | | | |
| <i>CepNAC37</i> | 2577 | 201 | 2180 | 660 | | | III |
| <i>CepNAC38</i> | 3414 | 296 | 3109 | 938 | | | III |
| <i>CepNAC39</i> * | 517 | 32 | 517 | 162 | <i>NAM</i> | 81 | V |

^a the asterisk indicated that the gene with partial cDNA sequence.

^b PSC and PEC represented the positions of starting and ending codon in cDNA sequence.

^c RS indicated the ratio of similarity of the protein sequence.

^#

doi:10.1371/journal.pone.0157871.t002

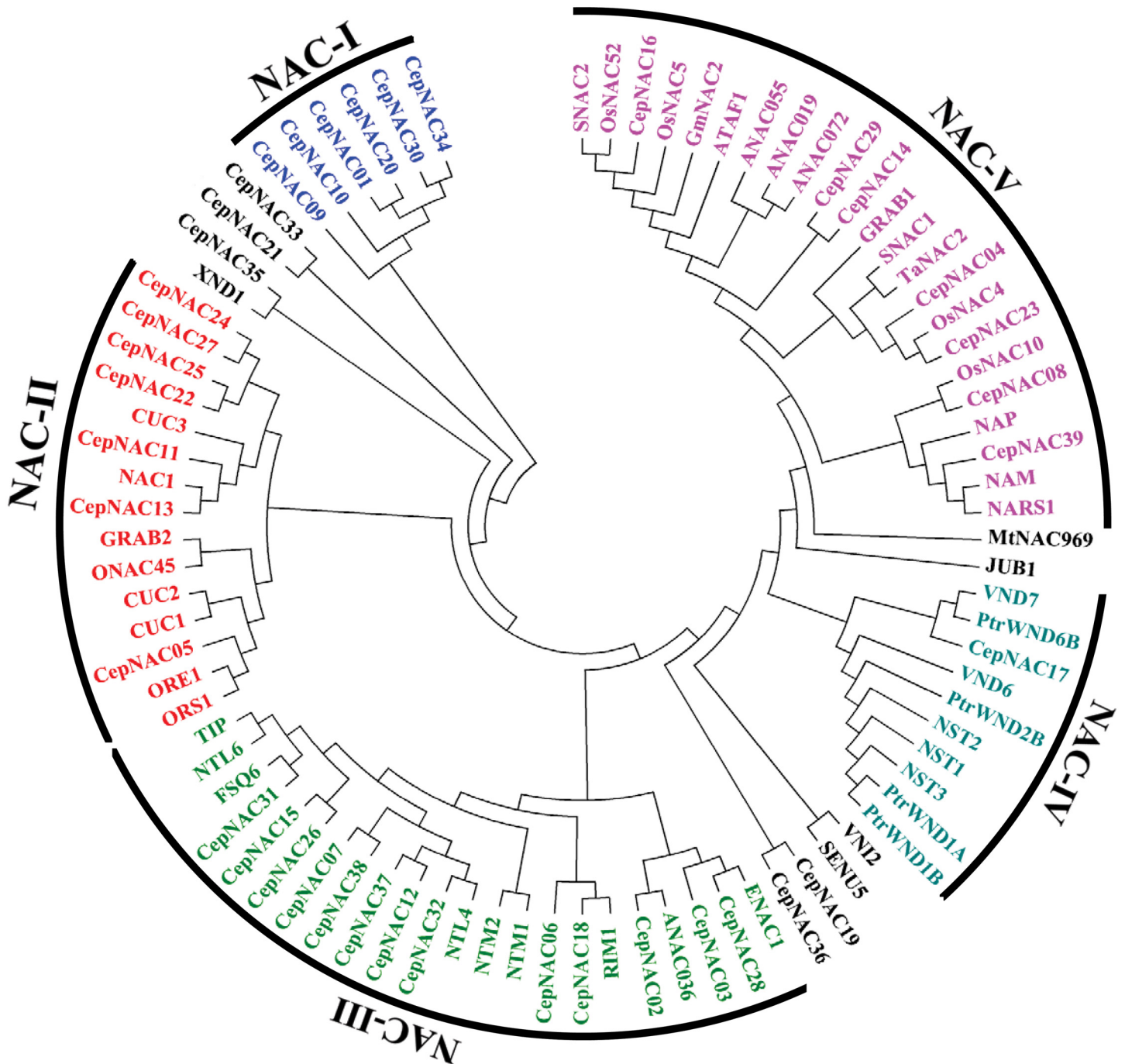


Fig 6. Phylogenetic tree of 39 CepNAC proteins and 47 NAC proteins published. Neighbor-joining method was used for constructing this unrooted tree by using the MEGA 4 program. Bootstrap values from 1000 replicates are indicated at each node, and are displayed in [S1 Fig](#).

doi:10.1371/journal.pone.0157871.g006

tolerance, and embryogenesis [11, 14, 15]. However, no study to date has identified NAC TFs family members in onion. In the present study, a total of 39 CepNAC TFs were identified for the first time. Sequence analysis revealed marked homology of 17 CepNAC TFs to NAC proteins of known function reported in other species. Among those, 7 showed more than 60% sequence similarity with the NAC proteins previously reported. The high homology in protein

sequence will provide further insights into the molecular mechanisms in which CepNAC TFs are involved.

Phylogenetic analysis was performed for 39 CepNAC TFs and 47 functionally characterized NAC proteins reported in previous studies. Among these 47 NAC proteins of known function, 10 were found to be involved in secondary wall synthesis. The results of the phylogenetic analysis revealed that 9 out of 10 NAC genes involved in secondary wall synthesis and 1 CepNAC TF (CepNAC17) were assigned to the NAC-IV subfamily, suggesting that CepNAC17 TFs share a conserved sequence with NAC TFs responsible for secondary wall synthesis. Moreover, CepNAC35 showed 51% homology with XND1, a regulator of secondary wall synthesis in *Arabidopsis* [33]. Most likely, CepNAC17 and CepNAC35 have a potential role in secondary wall synthesis in onion. In addition, 16 NAC proteins of known function fell into the NAC-V subfamily. Interestingly, all these 16 NAC TFs were found to be involved in the regulation of stress tolerance (S1 Table), suggesting that this NAC group might be a stress-responsive NAC subfamily. In light of our results, this stress-responsive NAC subfamily would contain 7 CepNAC members.

In conclusion, 39 CepNAC genes were identified for the first time in this study. Analysis of sequence conservation revealed that 2 CepNAC TFs had high homology with NAC TFs of other species that functioned as secondary wall synthesis regulators, and phylogenetic analysis found that 7 CepNAC TFs were contained in the stress-responsive NAC subfamily. The identification of the CepNAC TFs reported in this study will provide a basis for further characterization and validation of their function in the future.

Supporting Information

S1 Fig. Rectangular phylogenetic tree of NAC proteins with bootstrap values.
(TIF)

S1 File. cDNA sequences of 39 CepNAC identified genes and their predicted protein products.
(DOCX)

S1 Table. Details of the 47 functionally characterized NAC genes previously reported in other species.
(DOC)

Acknowledgments

We acknowledge the Novogene Bioinformatics Institute for the assistance with original data processing and related bioinformatics analysis.

Author Contributions

Conceived and designed the experiments: ST TL. Performed the experiments: ST XZ. Analyzed the data: SZ. Contributed reagents/materials/analysis tools: QD. Wrote the paper: TL.

References

1. Aida M, Ishida T, Fukaki H, Fujisawa H, Tasaka M. Genes involved in organ separation in *Arabidopsis*: an analysis of the cup-shaped cotyledon mutant. *Plant Cell*. 1997; 9:841–857. PMID: [9212461](#)
2. Souer E, van Houwelingen A, Kloos D, Mol J, Koes R. The no apical meristem gene of *Petunia* is required for pattern formation in embryos and flowers and is expressed at meristem and primordia boundaries. *Cell*. 1996; 85:159–170. PMID: [8612269](#)

3. Olsen A, Ernst H, Leggio L, Skriver K. NAC transcription factors: structurally distinct, functionally diverse. *Trends Plant Sci.* 2005; 10:79–87. PMID: [15708345](#)
4. Ernst H, Olsen A, Larsen S, Leggio L. Structure of the conserved domain of ANAC, a member of the NAC family of transcription factors. *EMBO Rep.* 2004; 5:297–303. PMID: [15083810](#)
5. Hu H, Dai M, Yao J, Xiao B, Li X, Zhang Q, et al. Overexpressing a NAM, ATAF, and CUC (NAC) transcription factor enhances drought resistance and salt tolerance in rice. *Proc Natl Acad Sci USA.* 2006; 103:12987–12992. PMID: [16924117](#)
6. Kim S, Kim S, Park C. A membrane-associated NAC transcription factor regulates salt-responsive flowering via FLOWERING LOCUS T in Arabidopsis. *Planta* 2007; 226:647–654. PMID: [17410378](#)
7. Tran L, Nakashima K, Sakuma Y, Simpson S, Fujita Y, Maruyama K, et al. Isolation and functional analysis of Arabidopsis stress inducible NAC transcription factors that bind to a drought responsive cis-element in the early responsive to dehydration stress 1 promoter. *Plant Cell.* 2004; 16:2481–2498. PMID: [15319476](#)
8. Hu R, Qi G, Kong Y, Kong D, Gao Q, Zhou G. Comprehensive Analysis of NAC Domain Transcription Factor Gene Family in *Populus trichocarpa*. *BMC Plant Biol.* 2010; 10:145. doi: [10.1186/1471-2229-10-145](#) PMID: [20630103](#)
9. Pinheiro G, Marques C, Costa M, Reis P, Alves M, Carvalho C, et al. Complete inventory of soybean NAC transcription factors: Sequence conservation and expression analysis uncover their distinct roles in stress response. *Gene.* 2009; 444:10–23. doi: [10.1016/j.gene.2009.05.012](#) PMID: [19497355](#)
10. Nuruzzaman M, Manimekalai R, Sharoni A, Satoh K, Kondoh H, Ooka H, et al. Genome-wide analysis of NAC transcription factor family in rice. *Gene.* 2010; 465:30–44. doi: [10.1016/j.gene.2010.06.008](#) PMID: [20600702](#)
11. Liu T, Zhu S, Tang Q, Tang S. Identification of 32 full-length NAC transcription factors in ramie (*Boehmeria nivea* L. Gaud) and characterization of the expression pattern of these genes. *Mol Genet Genomics.* 2014; 289:675–684. doi: [10.1007/s00438-014-0842-4](#) PMID: [24691727](#)
12. Puranik S, Sahu P, Srivastava P, Prasad M. NAC proteins: regulation and role in stress tolerance. *Trends Plant Sci.* 2012; 17:369–381. doi: [10.1016/j.tplants.2012.02.004](#) PMID: [22445067](#)
13. Han X, Feng Z, Xing D, Yang Q, Wang R, Qi L, et al. Two NAC transcription factors from *Caragana intermedia* altered salt tolerance of the transgenic Arabidopsis. *BMC Plant Biol.* 2015; 15:208. doi: [10.1186/s12870-015-0591-5](#) PMID: [26297025](#)
14. Xu B, Ohtani M, Yamaguchi M, Toyooka K, Wakazaki M, Sato M, et al. Contribution of NAC transcription factors to plant adaptation to land. *Science.* 2014; 343:1505–1508. doi: [10.1126/science.1248417](#) PMID: [24652936](#)
15. Nakashima K, Takasaki H, unya Mizoi J, Shinozaki K, Yamaguchi-Shinozaki K. NAC transcription factors in plant abiotic stress responses. *Biochimica et Biophysica Acta.* 2012; 1819:97–103. doi: [10.1016/j.bbagr.2011.10.005](#) PMID: [22037288](#)
16. Rajkumar H, Ramagoni R, Anchoju V, Vankudavath R, Syed A. *De novo* transcriptome analysis of *Allium cepa* L. (Onion) bulb to identify allergens and epitopes. *PLoS ONE.* 2015; 10:e0135387. doi: [10.1371/journal.pone.0135387](#) PMID: [26284934](#)
17. Arumuganathan K, Earle ED. Nuclear DNA content of some important plant species. *Plant Mol Biol Rep.* 1991; 9: 208–218.
18. McCallum J, Samantha B, Masayoshi S, Yanbo D, Sjaak van H, Pither-joyce M, et al. AlliumMap-A comparative genomics resource for cultivated Allium vegetables. *BMC Genomics.* 2012; 13:168. doi: [10.1186/1471-2164-13-168](#) PMID: [22559261](#)
19. Liu T, Zhu S, Tang Q, Chen P, Yu Y, Tang S. *De novo* assembly and characterization of transcriptome using Illumina paired-end sequencing and identification of *CesA* gene in ramie (*Boehmeria nivea* L. Gaud). *BMC Genomics.* 2013; 14:125. doi: [10.1186/1471-2164-14-125](#) PMID: [23442184](#)
20. Liu T, Tang S, Zhu S, Tang Q, Zheng X. Transcriptome comparison reveals the patterns of selection in domesticated and wild ramie (*Boehmeria nivea* L. Gaud). *Plant Mol Biol.* 2014; 86:85–92. doi: [10.1007/s11103-014-0214-9](#) PMID: [24934879](#)
21. Lou Q, Liu Y, Qi Y, Jiao S, Tian F, Jiang L, et al. Transcriptome sequencing and metabolite analysis reveals the role of delphinidin metabolism in flower colour in grape hyacinth. *J Exp Bot.* 2014; 65:3157–3164. doi: [10.1093/jxb/eru168](#) PMID: [24790110](#)
22. Mutz K, Heilkenbrinker A, Lonne M, Walter J, Stahl F. Transcriptome analysis using next-generation sequencing. *Curr Opin Biotechnol.* 2013; 24:22–30. doi: [10.1016/j.copbio.2012.09.004](#) PMID: [23020966](#)
23. Koenig D, Jiménez-Gómez J, Kimura S, Fulopa D, Chitwooda D, Headlanda L, et al. Comparative transcriptomics reveals patterns of selection in domesticated and wild tomato. *Proc Natl Acad Sci USA.* 2013; 110:E2655–E2662. doi: [10.1073/pnas.1309606110](#) PMID: [23803858](#)

24. Grabherr M, Haas B, Yassour M, Levin J, Thompson D, Amit I, et al. Full length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnol.* 2011; 29:644–652.
25. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* 1997; 25:4876–4882. PMID: [9396791](#)
26. Tamura K, Dudley J, Nei M, Kumar S. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol Biol Evol.* 2007; 24:1596–1599. PMID: [17488738](#)
27. Metzker M. Sequencing technologies-the next generation. *Nature reviews genetics.* 2010; 11:31–46. doi: [10.1038/nrg2626](#) PMID: [19997069](#)
28. Kuhl J, Cheung F, Yuan Q, Martin W, Zewdie Y, McCallum J, et al. A unique set of 11,008 onion expressed sequence tags reveals expressed sequence and genomic differences between the monocot orders *Asparagales* and *Poales*. *Plant Cell.* 2004; 16:114–125. PMID: [14671025](#)
29. Liu T, Zeng L, Zhu S, Chen X, Tang Q, Mei S, et al. Large-scale development of expressed sequence tag-derived simple sequence repeat markers by deep transcriptome sequencing in garlic (*Allium sativum* L.). *Mol Breeding.* 2015; 35:204.
30. Zhou S, Chen L, Liu S, Wang X, Sun X. *De novo* assembly and annotation of the Chinese chive (*Allium tuberosum* Rottler ex Spr.) transcriptome using the Illumina platform. *PLoS ONE.* 2015; 10:e0133312. doi: [10.1371/journal.pone.0133312](#) PMID: [26204518](#)
31. Sun X, Yu X, Zhou S, Liu S. *De novo* assembly and characterization of the Welsh onion (*Allium fistulosum* L.) transcriptome using Illumina technology. *Mol GenetGenomics.* 2016; 291: 647–659.
32. Jin J, Zhang H, Kong L, Gao G, Luo J. PlantTFDB 3.0: a portal for the functional and evolutionary study of plant transcription factors. *Nucleic Acids Res.* 2013;: doi: [10.1093/nar/gkt1016](#)
33. Zhao C, Avci U, Grant E, Haigler C and Beers E. XND1, a member of the NAC domain family in *Arabidopsis thaliana*, negatively regulates lignocellulose synthesis and programmed cell death in xylem. *Plant J.* 2008; 53, 425–436. PMID: [18069942](#)