# POJ 11(03):120-127 (2018) doi: 10.21475/poj.11.03.18.p1037

ISSN:1836-3644

 $\mathcal{P}OI$ 

# Comprehensive analysis of the NAC gene family in Elaeis guineensis

Yong Xiao<sup>1\*</sup>, Haikuo Fan<sup>1†</sup>\*, Jianwei Ma<sup>1,3†</sup>\*, Xintao Lei<sup>1</sup>, Yong Wang<sup>1</sup>, Hongxin Cao<sup>1</sup>, Lixia Zhou<sup>1</sup>, Yong Zhang<sup>3 &</sup>, Wei Xia<sup>2&</sup>

<sup>1</sup>Coconut Research Institute, Chinese Academy of Tropical Agricultural Sciences, Wenchang, P.R. China <sup>2</sup>Institute of Tropical Agriculture and Forestry, Hainan University, Haikou, P.R. China <sup>3</sup>College of Horticulture, Sichuan Agricultural University, Chengdu, P.R. China

<sup>†</sup>These authors contributed equally to this work <sup>\*</sup>Corresponding author: saizjxiawei@hainu.edu.cn

## Abstract

The NAC gene family encode transcriptional regulator that contain a conserved NAM domain near the N-terminus and participate in the regulation of plant development and response to different abiotic stresses. In this study, 129 *EgNAC* genes were identified from the genome sequence of *Elaeis guineensis* and 97 *EgNAC* located on the chromsomes with an average of 4.56 *EgNAC* genes per chromosome. About 60% of *EgNACs* contained three exons and the gene sizes varied from 541 bp to 37,294 bp. Genomic duplication analysis showed that 10 *EgNAC* genes were involved in segmental duplication events and two genes were from tandem duplication. The gene expression profiles of *EgNACs* based on transcriptome database for different oil palm tissues showed that 30 *EgNACs* with low or no expression and 24 *EgNACs* were specifically expressed in one tissue. The trancriptome comparison between the control and cold stress samples demonstrated that thirty-seven *EgNACs* were down-regulated and 82 *EgNACs* were up-regulated under cold stress. Further RT-qPCR showed that the expression for 24 out of 32 validated *EgNACs* were induced under both cold, drought and salt stresses. Our comprehensive analysis of EgNAC genes has provided clues for candidate genes involved in abiotic stress tolerance.

Keywords: Elaeis guineensis, EgNAC, transcriptome, mesocarp, abiotic stress.

# Introduction

The African oil palm (2n = 32, Elaeis guineensis) belongs to the genus Elaeis in the family Arecaceae and is an important oil crop in worldwide tropical regions. As a tropical crop, E. guineensis is highly sensitive to low temperatures; however, it has a high tolerance to salt and drought stress. The genus Elaeis contains two different species: E. guineensis (African oil palm) and E. oleifera (American oil palm). Currently, the African oil palm is commercially cultivated in tropical regions for palm oil production, especially in Indonesia and Malaysia. Because of the economic value of this species, the whole genome sequencing of E. guineensis had been completed and the sequence data have been released in 2013 (Singh et al., 2013). Applying bioinformatics to characterize gene in E. guineensis genome is the basis for the further gene function analysis. The released genome sequence and other transcriptome data could be used to analyze gene distribution, gene expansion and gene expression profiles of a specific gene family.

NAC genes encode a large family of transcription factors that contain a conserved NAM domain (Petunia NAM and Arabidopsis ATAF1, ATAF2 and CUC2) at the N-terminus (Ooka et al., 2003; Kikuchi et al., 2000) and play various roles in the plant life cycle. The mutation of one NAC gene number - CUP-SHAPED COTYLEDON - caused a defect in the separation of cotyledons (embryonic organs), the development of sepals

and stamens, and the formation of the shoot apical meristem (Aida et al., 1997). In Arabidopsis, studies showed that NAC gene participates in shoot apical meristem development (Jung et al., 2005). NAC genes also associated with auxin signal transduction and lateral root development (He et al., 2005). Moreover, the NAC transcription factors control biochemical and molecular pathways in response to different stress conditions. A NAC gene in Arabidopsis participated in the plant's defense response to necrotrophic fungal and bacterial pathogens (Wang et al., 2009). Overexpression of OsNAC genes could increase the resistance to drought stress in rice, and in Vitis vinifera, the expression of some VvNAC genes are induced under biotic and abiotic stresses (Mark et al., 2012). Since that characterization NAC genes in African oil palm may provide candidate genes involved in its stress responsive processes. genome-wide Recently, gene identification characterization of NAC genes in many different species have been conducted, including V. vinifera (74 NAC genes) (Wang et al., 2012), Arabidopsis thaliana (117) (Ooka et al., 2003), Populus trichocarpa (120) (Hu et al., 2010), Glycine max (152) (Le et al., 2011), and Oryza sativa (151) (Nuruzzaman et al., 2010). The assembled genome of E. guineensis was about 1.53 Gb of sequences and 34,802 genes were annotated on the basis of the transcriptome data from 30 tissue samples. In this study, we identified all NAC genes from E. guineensis genome and characterized gene structures, gene duplication and phylogenic relationship for all *EgNAC* genes identified. The expression profiles of these *EgNAC* genes were analyzed based on transcriptomes for different tissues, along with RT-qPCR results for *EgNAC* genes under cold, salt and drought stresses.

### Results

## Identification of NAC transcription factors in the assembled E.guineensis genome

One hundred twenty-nine genes in the assembled *E.* guineensis genome were identified using the amino acid sequences of *AtNAC* genes were used as query sequences with a cutoff E-value  $< 10^{-5}$  via BLAST analysis (Supplementary Table 1). One NAM domain was identified in each of the 129 peptide sequences via protein alignment between the 129 peptides and the Nr (non-redundant) protein database. The 129 putative *NAC* transcription factors were named *EgNAC*01 to *EgNAC*129 on the basis of their gene order in the assembled *E.* guineensis genome.

Putative biological functions of these *EgNAC* genes were predicted by BLAST against the Non-redundant protein sequences database at a threshold cutoff of E-value  $<10^{-5}$ . The BLAST results showed that all identified *EgNAC* genes belongs to the *NAC* family and fifty-seven of these putative *EgNAC* genes were othologs to *NAC* genes in database related to environmental stress, and 6 were predicted to be associated with plant growth or development.

# Chromosome position and gene structure of the putative EgNAC genes

The chromosome positions of these putative *EgNAC* genes were shown in Fig. 1. Ninety-seven putative *EgNAC* genes were mapped *in silico* on 16 chromosomes, with every chromosome containing at least one *EgNAC* gene and an average of 4.56 *EgNAC* genes per chromosome. The chromosomes for the rest 33 putative *EgNAC* genes were undetermined and located on unlinked scaffolds (Supplementary Table 1). Chromosomes 2 (12 *EgNAC*) and 3 (9 *EgNAC*) contained the top two number of *EgNAC* genes, whereas chromosomes 13 and 14 contained only three *EgNAC* genes.

The gene structures of the 129 *EgNACs* were investigated by aligning their coding sequences with the genomic sequences (Fig. 2). Almost all of the *EgNACs* contained at least two exons except for *EgNAC055* and *EgNAC087*. *EgNAC115* contained fifteen exons, which was the largest number of exons among the *EgNACs*. Nearly 60% of *EgNACs* only contained three exons and the gene sizes varied from 541 bp (*EgNAC105*) to 37,294 bp (*EgNAC014*).

The gene expansion of *EgNACs* via genomic duplication was analyzed in this study, which was done by investigating BLAST results for all protein-coding genes in oil palm genome with MCScanX software. The analysis demonstrated that 10 *EgNACs* were located in genomic regions having segmental duplication events, and two *EgNACs* were produced by tandem duplication. Two synteny blocks were detected in chromosomes 4 (*EgNAC028, EgNAC029, EgNAC030, EgNAC031,* and *EgNAC032*) and 11 (*EgNAC069, EgNAC070, EgNAC071, EgNAC072,* and *EgNAC073*).

### Phylogenetic relationship of the 129 EgNACs

The amino acid sequence of the conserved NAM domain was extracted and used to build the phylogenetic tree of the 129 *EgNAC* genes, which showed that the *EgNAC* genes could be divided into four large clusters - cluster I to cluster IV (Fig. 3). The largest cluster was cluster I with 45 *EgNAC* genes, whereas the smallest was cluster III having 23 members.

Moreover, we analyzed *EgNAC* gene expanded from duplication and their phylogenetic relationship. *EgNAC* pairs produced by tandem/proximal duplication had close phylogenetic relationship and tended to be clustered in one group. Further comparison of gene phylogeny and gene structure and size showed that cluster III had the largest number of average exons among the four clusters, with an average of 6.26 exons per *EgNAC*. The average exon numbers were similar among cluster I (with an average of 4.29 exons per *EgNAC*), cluster II (4.2), and cluster IV (3.8), and a similar average peptide length were also observed among the four clusters. Cluster I and cluster two contained *EgNAC* genes with 1 to 13 introns, while cluster III contained *EgNAC* genes with 2 to 15 introns.

# *Expression profiles of the 129 EgNACs in different E. guineensis tissues*

To investigate the expression profiles for the 129 EgNACs in different E. guineensis tissues, twelve raw transcriptome datasets were downloaded from the SRA database. The tissues of these transcriptomes were mesocarp (five different developmental stages), leaf, fruit, female flower, floret, shoot, and root and the RPKM value was calculated (Fig. 4). Thirty EgNACs (EgNAC08, 10, 24, 26, 31, 36, 46, 48, 49, 59, 66, 67, 72, 80, 93, 97, 98, 101, 102, 103, 105, 107, 109, 110, 112, 117, 122, 124, 126, and 129) were shown low or no expression in the 12 analyzed tissues. Twenty-four EgNACs were specifically expressed in one tissue, including sixteen were expressed only in the five different developmental mesocarp stages (EqNAC04, 07, 09, 12, 23, 27, 41, 45, 50, 51, 55, 60, 65, 74, 78, and 100), three were detected in the root (EgNAC03, 11, and 19), two was expressed in the floret (EgNAC92 and 114), one was expressed in the shoot, one in the female flower and one in the kernel.

### Expression profiles of EgNACs under cold stress

The expression fold changes of the 129 *EgNACs* between control and cold stress were examined based on two transcriptomes - one for control sample and one for mixed samples from different time points after cold treatment. Thirty-seven *EgNACs* were down-regulated and 82 *EgNACs* were up-regulated under cold stress. Among the up-regulated *EgNACs*, 35 were up-regulated at least two - folded (*EgNAC03*, 08, 09, 11, 12, 13, 16, 20, 23, 26, 27, 32, 34, 35, 38, 42, 43, 47, 48, 51, 61, 63, 65, 73, 81, 83, 84, 88, 89, 97, 99, 105, 112, 125, and 126).

# Expression profiles for 32 EgNACs under cold, drought and salt stresses

RT-qPCR was used to examine the *EgNAC* gene expression levels of under various environmental stresses and 12 *EgNAC*s



**Fig 1.** Chromosome location of 129 EgNAC transcription factors in the *Elaeis guineensis* genome. Vertical bars represent *E. guineensis* chromosomes. The chromosome number is marked at the top of each chromosome. Colored lines represent different clusters generated by phylogenetic analysis: the red line represents cluster I, the purple line represents cluster II, the blue line represents cluster III, and the green line represents cluster IV.



Fig 2. Gene structure of 129 EgNAC genes identified in this study. Exons are indicated by colored boxes, and introns are indicated by black lines.



Fig 3. Phylogenetic tree of 129 EgNAC genes, based on multiple alignments between conserved amino acid sequences. A total of four clusters are marked: red (cluster I), purple (cluster II), blue (cluster III), and green (cluster IV).



Fig 4. Expression pattern of EgNAC genes in different tissues, including the mesocarp (five different developmental stages), leaf, fruit, female flower, floret, shoot, and root.



Fig 5. Quantitative real-time PCR validation of EqNAC gene expression under salt, drought and cold treatments.

(EgNAC09, 16, 34, 35, 38, 61, 65, 81, 83, 84, 89, and 97) that shown as induced by cold treatment from transcriptome data were selected for the expression levels detection at 0, 4, 24 and 48 hours after cold, drought and salt treatments (Fig. 5). Cross-talking were detected between different stress treatment for EgNACs, since the expression of about 75% (24 out of 32) EqNACs were induced under both cold, drought and salt stresses, including EqNAC09, 16, 34, 35, 38, 61, 65, 81, 83, 84, 89, and 97. After drought treatment, almost all of the EgNACs were up-regulated except for EgNAC44, and 75% (24 out of 32) EgNAC genes showed higher expression level at 48 hours after drought stress compared with the other three time points. Similar expression pattern were found in drought stress: half of these EgNACs showed higher expression levels at 48 hours after drought treatment, as compared with the other three time points. Under cold stress, the expression of EgNAC006 and EgNAC053 were repressed, while eight EgNACs (EgNAC009, 033, 035, 038, 039, 079, 084, and 086), 13 EgNACs (EqNAC007, 034, 037, 050, 061, 068, 069, 075, 081, 083, 090, 096, and 097) and 8 EqNACs (EqNAC001, 015, 016, 057, 058, 065, 077, and 089) had the highest expression levels at 48 hours, 24 hours and 4 hours, respectively.

## Discussion

The oil palm (*E. guineensis*, 2n = 32) is an important tropical oil crop, which is highly sensitive to low temperatures and highly tolerant to salt and drought stresses. In this study, 129 genes in the assembled *E. guineensis* genome were identified as *NAC* family members, which is more than in *A. thaliana* (117) (Ooka

et al., 2003), *P. trichocarpa* (120) (Hu et al., 2010), and *V. vinifera* (74) (Wang et al., 2012), but fewer than in *G. max* (152) (Le et al., 2011) and *O. sativa* (151) (Nuruzzaman et al., 2010). Although 42.86% of the assembled sequences are assigned on chromosomes according to genetic linkage map, 75.19% *EgNACs* identified in this study were mapped evenly *in silico* on 16 chromosomes. The amino acid lengths of the *NAC* gene products varied from 150 aa to 1,276 aa, with an average of 395 aa per *EgNAC*, which is equal to the average length of *NACs* in *Populus* (342 aa) (Hu et al., 2010).

On the basis of the conserved N-terminal NAC domain, the 129 *EgNACs* were divided into four large clusters. Previous research has shown that the *NAC* transcription factors with the same functions tend to fall into the same subgroup (Fang et al., 2008; Pinheiro et al., 2009; Shen et al., 2009). In this study, 129 *EgNAC* protein sequences were used as queries to BLAST against known proteins to predict possible functions. The alignment results showed that fifty-seven of these putative *EgNAC* genes were othologs to *NAC* genes to environmental stress and 6 were related with plant growth or development. However, each of the four large clusters contained *EgNAC* genes associated separately with abiotic and biotic factors and growth development.

A previous study has shown that gene structure diversity might be associated with the evolution of a multigene family, and has provided additional information for phylogenetic analysis (Wang et al., 2015). Some studies have shown that *NAC* genes with the same exon and intron distribution tend to cluster together in *Populus* (Hu et al., 2010). However, there is no obvious trend showed that the *NAC* gene with similar gene structure tend to have more closer phylogenetic groups in *V. vinifera* (Wang et al., 2012). In this study, the phylogenetic tree for *EgNAC* genes also showed that similar variations of gene length, extron number in different clusters and and there no relationship between gene structure diversity and phylogenetic clustering in *E. guineensis*.

As more genome sequences for different plant available, the hypothesis that majority flowering plants underwent whole genome duplication at certain ancient time points during the evolutionary process was confirmed (Van de Peer et al., 2009). Previous studies have revealed that the genome of E. guineensis underwent at least one round of genome-wide duplications followed by segmental duplication and tandem duplication (Singh et al., 2013). In this study, analysis revealed two synteny blocks in chromosome 4 and chromosome 11 in the E. quineensis genome. Five pairs of duplicate genes were located in the two synteny blocks reported by Singh et al. (2013), while most EgNAC genes were encoded in only one copy. This suggests that the duplication events has provided the elements for gene expansion and gene subfunction. Eventually, both duplicate copies might become compromised by mutation accumulation, thus destroying the original function of the ancestral gene. In this study,.

Previous studies indicate that the NAC genes are involved in regulating plant development (Matus et al., 2010; Zhao et al., 2008; Kim et al., 2007; Ko et al., 2007, Kim et al., 2007; Guo et al., 2006; Ohnishi et al., 2005; Takada et al., 2001). In this study, twenty-four *EgNACs* were specifically expressed in one tissue, sixteen in mesocarp. *EgNAC04, EgNAC12, EgNAC65* and *EgNAC74* showed high identity with GhNACs (*Gossypium hirsutum* L.) in response to leaf senescence. Moreover, *EgNAC23, EgNAC51, EgNAC60,* and *EgNAC100* displayed high similarity to the *NAC* genes of *Theobroma cacao*, which were identified by pod color.

In addition to plant development, some members of the NAC family also participate in stress response. In this study, 82 EgNACs were up-regulated under cold stress and 35 EgNACs were up-regulated at least two fold. A previous study has reported that SbSNAC1 in Sorghum bicolor increases drought and salt stress tolerance, and the amino acid sequence of SbSNAC1 shows high similarity with that encoded by *EgNAC20*. This gene is strongly induced under cold stress on the basis of transcriptome data and qRT-PCR analysis. In tobacco, overexpression of TaNAC2a, which is highly similar to EgNAC35 and EgNAC81, improves drought tolerance. In this study, on the basis of the transcriptome data, most EgNACs were induced under cold stress, among which 42.68% were up-regulated at least two fold. These results indicated that most EgNACs respond to abiotic stress, and some might be associated with the tolerance to abiotic stress.

#### **Materials and Methods**

#### **Plant materials**

*E. guineensis* (*pisifera*, thin-shelled African oil palm) introduced from Malaysia was grown in nurseries. Thirty-six oil palm plants germinated in the same week and grown in the same nursery were selected for subsequent cold, drought and salt treatments. For the cold treatment, oil palm plants were placed in a growth chamber at 26 °C for one day. Next, the remaining oil palm plants were kept at 8 °C for 0 hour, 4 hours, 24 hours, or 48 hours. Spear leaves were sampled from the

control and cold-treated seedlings and immediately frozen in liquid nitrogen. For drought treatment, when the water content reached 23% (2.3 g water per 10 g soil), spear leaves were sampled after 0 hour, 4 hours, 24 hours and 48 hours and immediately frozen in liquid nitrogen. For salt treatment, the roots of oil palm seedlings were soaked in NaCl solution (400 mmol/L), and spear leaves were sampled at 0 hour, 4 hours, 24 hours, and 48 hours after salt treatment and immediately frozen in liquid nitrogen for RNA isolation. The MRIP method was applied to extract RNA from *E.guineens is* spear leaves (Xiao et al., 2012).

# Identification of EgNAC genes in E. guineensis and their phylogenetic relationship

The whole genome sequence of E. quineensis was downloaded from the National Center of Biotechnology Information (NCBI). AtNAC gene sequences were downloaded from the Arabidopsis Information Resource (Tair) available at http://www.arabidopsis.org. To identify EgNACs in E. guineensis, AtNAC sequences were used as queries to BLAST against E. guineensis coding sequences (CDS) at a cut-off threshold of 1e<sup>-10</sup>. Multiple protein sequence alignment was used to confirm the conserved domains of EgNAC sequences. MEGA 5.0 software was used to establish a neighbor-joining (NJ) phylogenetic tree based on the amino acid sequences of the conserved NAM domain with 1000 bootstrap replicates (Larkin et al., 2007; Tamura et al., 2011).

# Genomic structure, chromosome location and gene duplication of identified EgNAC genes

A summary of the *E. guineensis* gene annotation was downloaded from NCBI. Additionally, mRNA and CDS were downloaded from NCBI. mRNA sequences were used as queries to BLAST against the whole-genome *E. guineensis* sequences to confirm *EgNAC* gene structures such as the relative position of exons and introns. Gene structures were identified by using a gene structure display server program (GSDS) (http://gsds.cbi.pku.edu.cn/). The duplication event for all of the *MeERF*s identified in this study was analyzed by using an algorithm that scans multiple genomes or subgenomes to identify putative homologous chromosomal regions and aligns these regions by using genes as anchors (Wang et al., 2012).

### Expression profiles of EgNACs in various tissues

Raw data for nine transcriptomes from different tissues including mesocarp (four different developmental stages), leaf, fruit, flower, root, and shoot were downloaded from the NCBI SRA (Sequence Read Archive) database. The RPKM value (reads per kb per million reads) was used to calculate the gene expression level by using the following formula (Mortazavi et al., 2008):

$$RPKM = \frac{10^{6}C}{NL/_{10^{3}}}$$

in which C is the number of reads exclusively aligned to one expressed sequence, N is the total number of reads aligned to all expressed sequences, and L is the basic number in the coding sequence of the corresponding expression sequences.

## Quantitative real-time PCR

In this study, quantitative real-time PCR was performed with a standard SYBR Premix ExTaqTM Kit (TaKaRa) protocol in 96-well optical plates (Axygen) with a final volume of 10  $\mu$ L. Real-time qPCR reactions were incubated in 0.2 mL tubes in a Mastercycler Ep Realplex4 (Eppendorf) machine as follows: 95 °C for 5 seconds, 55 °C for 15 seconds and 68 °C for 20 seconds. The procedure was completed with a melt-curve ramping from 60°C to 95°C for 20 minutes to ensure PCR specificity.

### Conclusion

NAC transcription factors had been found to be associated with pant development and tolerance to abiotic stress in various species. In the study, we identified 129 NAC transcription factors in the genome of *Elaeis guineensis*. We analyzed twelve raw transcriptome data sets downloaded from the SRA (Sequence Read Archive) database for different oil palm tissues. Sixteen *EgNACs* were specifically expressed at five different developmental mesocarp stages. To determine whether these *EgNACs* responded to abiotic stress, we performed a comparative transcriptome analysis and found that most *EgNACs* were induced under cold stress, with at least a two-fold up-regulation for one-third of the *EgNACs*.

### Acknowledgement

This work was supported by the Scientific and Technological Cooperation Projects of Hainan province (No. SQ2015GJXM0076) and the fundamental Scientific Research Funds for Chinese Academy of Tropical Agriculture Sciences (1630152016006).

#### References

- Aida M, Ishida T, Fukaki H, Fujisawa H, Tasaka M (1997) Genes involved in organ separation in Arabidopsis: an analysis of the cup-shaped cotyledon mutant. Plant Cell. 9: 841-857.
- Fang Y, You J, Xie K, Xie W, Xiong L (2008) Systematic sequence analysis and identification of tissue-specific or stress-responsive genes of NAC transcription factor family in rice. Mol Genet Genomics. 280: 547-563.
- Guo Y, Gan S (2006) AtNAP, a NAC family transcription factor, has an important role in leaf senescence. Plant J. 46, 601-612.
- He XJ, Mu RL, Cao W, Chen SY (2005) AtNAC2, a transcription factor downstream of ethylene and auxin signaling pathways, is involved in salt stress response and lateral root development. Plant J. 44: 903-916.
- Hu R, Qi G, Kong Y, Kong D, Gao Q, Zhou G (2010) Comprehensive analysis of NAC domain transcription factor gene family in *Populus trichocarpa*. BMC Plant Biol. 10: 145.
- Jung JH, Yun J, Seo YH, Park CM (2005) Characterization of an Arabidopsis gene that mediates cytokinin signaling in shoot apical meristem development. Mol Cells. 19: 342-349.
- Kikuchi K, Ueguchi-Tanaka M, Yoshida KT, Nagato Y, Matsusoka M (2000) Molecular analysis of the NAC gene family in rice. Mol Gen Genet. 262: 1047-1051.

- Kim SY, Kim SG, Kim YS, Seo PJ, Bae M, Yoon HK (2007) Exploring membrane-associated NAC transcription factors in Arabidopsis: implications for membrane biology in genome regulation. Nucleic Acids Res. 35: 203-213.
- Kim SG, Park CM (2007) Membrane-mediated salt stress signaling in flowering time control. Plant Signal Behav. 2: 517-518.
- Ko JH, Yang SH, Park AH, Lerouxel O, Han KH (2007) ANAC012, a member of a plant-specific NAC transcription factor family, negatively regulates xylary fiber development in *Arabidopsis thaliana*. Plant J. 50: 1035-1048.
- Larkin MA, Blackshields G, Brown N (2007) Clustal W and Clustal X version 2.0. Bioinformatics. 23: 2947-2948.
- Le DT, Nishiyama R, Watanabe Y, Mochida K, Yamaguchi-Shinozaki K, Shinozaki K (2011) Genome-wide survey and expression analysis of the plant-specific NAC transcription factor family in soybean during development and dehydration stress. DNA Res. 18: 263-276.
- Mark CFR, Strasser RJ, Jeong JS, Kim YS, Kim JK (2011) The use of JIP test to evaluate drought-tolerance of transgenic rice overexpressing OsNAC10. Plant Biotechnol Rep. 5: 169-175.
- Matus JT, Poupin MJ, Canon P, Bordeu E, Alcalde JA, Arce-Johnson P (2010) Isolation of WDR and bHLH genes related to flavonoid synthesis in grapevine (*Vitis vinifera* L.). Plant Mol Biol. 72: 607-620.
- Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B (2008) Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat Methods. 2008, 5, 621-628.
- Nuruzzaman M, Manimekalai R, Sharoni AM, Satoh K, Kondoh H, Ooka H (2010) Genome-wide analysis of NAC transcription factor family in rice. Gene. 465: 30-44.
  - Ohnishi T, Sugahara S, Yamada T, Kikuchi K, Yoshiba Y, Hirano (2005) OsNAC6, a member of the NAC gene family, is induced by various stresses in rice. Genes Genet Syst. 80: 135-139.
  - Ooka H, Satoh K, Doi K, Nagata T, Otomo Y, Murakami K (2003) Comprehensive analysis of NAC family genes in Oryza sativa and Arabidopsis thaliana. DNA Res. 10: 239-247.
  - Pinheiro GL, Marques CS, Costa MD, Reis PAB, Alves MSk, Carvalho CM (2009) Complete inventory of soybean NAC transcription factors: sequence conservation and expression analysis uncover their distinct roles in stress response. Gene. 444: 10-23.
  - Shen H, Yin YB, Chen F, Dixon R (2009) A bioinformatic analysis of NAC genes for plant cell wall development in relation to lignocellulosic bioenergy production. Bioenerg Res. 2: 217-232.
  - Singh R, Ong-Abdullah M, Low ET, Manaf MA, Rosli R, Nookiah (2013) Oil palm genome sequence reveals divergence of interfertile species in Old and New worlds. Nature. 2500: 335-339.
  - Takada S, Hibara K, Ishida T, Tasaka M (2001) The CUP-SHAPED COTYLEDON1 gene of Arabidopsis regulates shoot apical meristem formation. Development. 128: 1127-1135.
  - Van de Peer Y, Maere S, Meyer A. The evolutionary significance of ancient genome duplication. Nat Rev Genet. 2009, 10, 725-732.
  - Wang N, Zheng Y, Xin H, Fang L, Li S (2012) Comprehensive analysis of NAC domain transcription factor gene family in *Vitis vinifera*. Plant Cell Rep. 10: 145.

- Wang X, Basnayake BM, Zhang H, Li G, Li W, Li W, et al. The Arabidopsis ATAF1, a NAC transcription factor, is a negative regulator of defense response against necrotrophic fungal and bacterial pathogens. *Mol* Plant Microbe Interact. 2009, 22, 1227-1238.
- Wang J, Pan C, Wang Y, Ye L, Wu J, Chen L. Genome-wide identification of MAPK, MAPKK, and MAPKKK gene families and transcriptional profiling analysis during development and stress response in cucumber. BMC Genomics. 2015, 16, 386.
- Xiao Y, Yang Y, Cao H, Fan H, Ma Z, Lei X, et al. Efficient isolation of quality RNA from tropical palms for RNA-seq analysis. Plant Omics. 2012, 5, 584-589.
- Zhao C, Avci U, Grant EH, Haigler CH, Beers EP (2008) XND1, a member of the NAC domain family in *Arabidopsis thaliana*, negatively regulates lignocellulos synthesis and programmed cell death in xylem. Plant J. 53: 425-436.

- Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol. 28: 2731-2739.
- Wang Y, Tang H, Debarry JD, Tan X, Li J, Wang X (2012) MCScan X: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. Nucleic Acids Res. 40: e49.