Application of the yeast one-hybrid technique to plant functional genomics studies

Yao Sun, Yao Li, Guoqing Huang, Qiong Wu and Lei Wang
Biotechnology Laboratory, Institute of Advanced Technology, Heilongjiang Academy of Sciences, Harbin, PR China

ABSTRACT
The yeast one-hybrid technique could help analyze DNA–protein interactions through the expression of reporter genes. It has become a frequently used molecular method in recent years. This mini-review briefly introduces the mechanism, the technical procedure, the major advantages and disadvantages as well as the application of the yeast one-hybrid technique to plant functional genomics studies. Prospects for future development are put forward.

INTRODUCTION

Functional genomics studies are systematic research of gene biological functions based on the information provided by genome sequences. Current research methods include expressed sequence tag (EST), serial analysis of gene expression (SAGE), RNA interference and insertion mutagenesis [1–4]. The accomplishment of the whole genome sequencing for a relevant model plant provides sufficient information for the study of gene functions. With the combination of related research techniques, plant functional genomics studies will help researchers gain a better understanding of gene functions for genetic improvement.

The yeast one-hybrid (Y1H) technique, which is a variant of the yeast two-hybrid (Y2H) technique, is applied for detection of DNA–protein interactions by expression of a reporter gene and isolation of cDNA encoding target transcription factors (TFs) [5]. The accomplishment of the whole genome sequencing for a relevant model plant provides sufficient information for the study of gene functions. With the combination of related research techniques, plant functional genomics studies will help researchers gain a better understanding of gene functions for genetic improvement.

Mechanism of the Y1H technique

The TFs involved in transcription initiation in eukaryotes are usually made up of two or more than two independent domains, including a DNA-binding domain (BD) and a transcription activation domain (AD). BD could recognize and interact with a cis-acting element, while AD could induce reporter gene expression. Although transcription cannot be activated by AD or BD alone, a fusion protein composed of BD and AD from different TFs still possesses the ability to activate the transcription process [6]. GAL4 is a typical TF applied in the Y1H technique, and its AD and BD are able to exert its function independently. When a cDNA library is constructed into a GAL4 AD vector, the expressed protein is composed of GAL4 AD and a library protein, which is equivalent to replacing GAL4 BD with the library protein; then the three key components of the Y1H technique can be listed (Figure 1). These are the bait, the prey and the reporter gene.

Once the prey is captured by the bait sequence located upstream of the reporter gene, activating the reporter gene’s transcription, the yeast cells could grow on selective medium lacking compounds or containing a toxic substance to obtain target positive clones.

Technical procedure of the Y1H technique

Let us take the screening of a cDNA encoding proteins interacting with a target sequence as an example (Figure 2): cis-element; also the bait is cloned into the plasmid vector which contains a reporter gene; then the constructed plasmid is transformed into yeast competent cells to screen positive clones. Meanwhile, synthesized cDNA library or a single TF-encoding gene is constructed into a vector expressing GAL4 AD, and the recombinant plasmid is transformed into the...
reconstructed strain containing the reporter gene and the bait sequence. Positive clones are obtained by further analysis, including sequencing, polymerase chain reaction (PCR) and other methods.

Major advantages and disadvantages of the Y1H technique

The major advantages of the Y1H technique are as follows:

(1) This method is free from the requirement for optimization of in vitro conditions compared to other biochemical methods. In addition, this method screens out protein-encoding sequences directly from the cDNA library without additional steps for protein purification and isolation. That is, it is effective, simple and reproducible.

(2) From the point of protein structure, as the interaction identification process is done in eukaryotic cells, the protein is in its native configuration, resulting in high sensitivity.

(3) From the point of vector construction, the expression of the fusion protein is regulated by strong promoters; besides, both the number of repeats of cis-element and the distance between the plant regulatory element and the minimal promoter has been optimized. Furthermore, this method enables the cloning of families of factors with conserved BDs which can interact with the same cis-acting element, and TFs of low abundances could be screened from large amount of colonies [7,8]

The problems of false positives and false negatives are the major disadvantages of this technique. False positives are usually caused by high recognition of the bait sequence by endogenous yeast TFs so that the reporter gene could be activated in the absence of prey. When generating reporter constructs, the junctions between the bait sequence and the vector sequence will create new potential binding sites, and this problem is more pronounced when screening simple bait, as the junctions represent a much larger proportion of potential sites than for larger bait [9].

Compared to false positives, the occurrence of false negatives is often neglected. False negatives are usually caused by the following aspects: 1. The AD hybrid protein may be toxic to the cell, in which case transformants will not grow or may grow very slowly on selective medium. 2. The hybrid protein is not stably expressed in...
the host cell. 3. The GAL4 AD occludes the site of interaction. 4. The hybrid protein folds improperly. 5. The hybrid protein cannot be localized to the yeast nucleus.

Moreover, the large amount of positive clones and duplicate clones during cDNA library screening may infer with further analysis.

**Application of the Y1H technique to plant functional genomics studies**

The essential application of Y1H technique is to clone and isolate TFs with known cis-elements, which is still the current research focus of this technique in plant genomics studies. The Y1H technique can also be used to in the following areas.

**DNA–protein interaction detection and verification**

A typical technical procedure for interaction verification is as follows: pick up positives and re-streak to obtain genuine clones on the same selective medium; isolate plasmids from colonies; transform target plasmids into *Escherichia coli* and purify the amplified plasmid DNA to confirm the interaction by transformation of reporter and control strains with the purified plasmid. This is also applicable in case a novel DNA fragment is used as a bait sequence in the screen. The inserted sequence could also be isolated by colony PCR [10]. Several DNA–protein interactions have been verified through this method in recent studies. For example, WRKY which interacts with the W-box cis-element (5'-TTGACC/ T-3') from the promoter region of disease-resistance genes encoded by soybean root cDNA [11], ERE-binding protein EgAP2-1 from the oil palm fruit mesocarp [12] and full-length cDNA sequence encoding DREB TF *HhDREB2* isolated from *Halimodendron halodendron* [13].

**Transcription-factor–binding site identification**

In the Y1H assay, potential target genes of a TF may be screened out by cis-element analysis. Similarly, the binding site can also be identified with known TFs cDNA. Wang et al. [14] has modified an inverse Y1H assay to screen the candidate target genes of WRKY33 when studying the role of *Arabidopsis* WRKY33 in drought. In this study, GAL4 AD/WRKY33 fusion protein was used as bait to screen an *Arabidopsis* genomic library enriched with W-box sequences (5'-TTGACC/ T-3') that WRKY specifically binds to. Twenty-three putative target genes of WRKY33 were identified, including three salt stress-related genes and two defense-related genes, which is in accordance with the role WRKY33 plays in salt stress [15] and pathogen-induced defense [16,17].

Focusing on the DNA-binding activities of R2R3-MYB, a comprehensive Y1H assay was performed [18], in which a pool of 16 sequences was used as bait to screen a library of *A. thaliana* R2R3-MYB open reading frames (ORFs). Many specific patterns of interactions between subgroups and bait sequences were identified, and the variety of interactions depended on the role R2R3-MYB subgroups played. This study drew a clear picture of R2R3-MYB subgroups binding activities. It will also provide more information to help predict new putative target genes of this TF family.

**Protein biological function analysis**

Verification of the interaction is just the initial step for the Y1H assay. As the core of functional genomics study, functional analysis of target TFs has become the research hotspot for the Y1H technique. For example, Lopato et al. [7,8] screened a cDNA library prepared from wheat RNA including the factors known to be involved in the regulation of cell fate (HDZip TF) and factors which can potentially be involved in the protection of the developing grain from drought or cold stress (DREB/CFB TF). Three HDZip factors and two DREB factors encoding genes were cloned using the Y1H assay. The spatial and temporal expression patterns of isolated wheat proteins were determined by real time PCR in which cDNA derived from vegetative and floral tissues at different stages of development was required. The data demonstrate a rough correlation between the abundance of the mRNAs of the cloned TFs in the plant tissue sample and the number of independent clones. Taking HDZip as an example, the abundance of mRNAs of *TaHDZip1*, *TaHDZip2* and *TaHDZip1* in the liquid fraction of endosperm isolated 5 days after pollination (DAP) was about 18 000, 8000 and 100 copies per microgram of total RNA, while the positive clones percentage distribution was 20%, 76% and 4%, respectively. It is clear that the number of positive clones of *TaHDZip* is larger than that of TaHDZip1, which may result from similarity of the cDNA sequence with the HDZip family. Moreover, the mRNA copies of the same and different families vary a lot at different DAP, indicating the temporal dynamics of the gene exerting its functions.

To further confirm whether *CesA8*, whose disruption has been reported to enhance the tolerance to drought stress [19], is a direct target gene of WRKY33 in *Arabidopsis*, the expression of *CesA8* in wild-type and *wrky33-1* (mutation type) plants was analyzed [14]. Semi-quantitative duplex RT-PCR assay showed that the *CesA8* mRNA
level was greatly increased in *wrky33*-1 plants compared to the wild type. The down-regulation indicates that CesA8 is a direct target gene of WRKY33. WRKY33 may play the role of a negative transcription regulator of CesA8 in drought resistance.

**Verification of the activating function of TFs**

The activating function of TFs is verified by transforming TFs-encoding cDNA and a GAL4 BD vector into yeast cells. The occurrence of positive clones suggests that the TFs may activate the expression of target genes. After analyzing the data from *Arabidopsis* ATH1 gene chip with drought treatment, Wei et al. [20] identified a cDNA designated as QRAP2 whose expression increased over 10-fold. Transformation of the plasmid expressing GAL4 BD and 112 amino acids in the N-terminal of QRAP2 could activate the reporter gene regulated by the GAL4 cis-regulating element, while no activation occurred for the C-terminal 113–292 amino acids transformation. The encoding sequence analysis showed that the glutamine residue content in the N-terminal region (residues 67–107) is 43.9%, including eight consecutive glutamine residues. Previous research [21] has shown that oligopeptides with more than six consecutive glutamine residues have a strong transcription activation ability, which may explain the transcription activation ability of QRAP2.

To study the transcriptional activity of TaMYB4, a novel TF belonging to the MYB family cloned from wheat Suwan11, Al-Attala et al. [22] transformed the full-length ORF pGBK7:TaMYB41–243 plasmid, the C-terminal deletion pGBK7:TaMYB41–112 plasmid and the control plasmid into yeast cells. The cells containing TaMYB41–112 or the negative control plasmid pGBK7 were unable to grow in the selective medium, whereas the cells containing pGBK7:TaMYB41–243 grew well, and the colonies turned blue when the same medium was supplemented with X-α-GAL. This suggests that *His3* and *LacZ* gene transcription are activated by the TaMYB4 fusion protein despite that the C-terminal deletion of TaMYB4 cDNA fails to do so.

A hybrid transcription factor (HTF) approach can be applied to identify different TFs that play a role in plant growth and development [23]. In this approach, TF-encoding sequences are fused with modules of opposite functions, that is, transcriptional activation module VP64 or a repression module EAR4. Zhang et al. [24] performed this approach with TF OsABF1, and identified a pair of HTFs designated ABF1V and ABF1E. Respective overexpression of ABF1V and ABF1E in rice resulted in demonstrating late or early flowering phenotypes. The transcriptional activity of OsABF1, ABF1V, or ABF1E was tested by a Y1H assay. The result showed that the transcriptional activity was increased by approximately 50% in ABF1V, whereas ABF1E resulted in repression of transcriptional activity by about 70%, suggesting that opposite phenotypes might be caused by the transcriptional activity differences. A similar activation assay was performed by Amalraj et al. [25] to validate the binding specificity of wheat TaRAP2.1L. In this case, TaRAP2.1L and TaRAP2.1Lmut were fused with the GAL4 AD and DRE, CRT and GCC-box were used as bait.

**Gene action mechanism study**

The action mechanism of a gene of interest and relevant hypotheses may be verified by the Y1H technique as well. For example, OsRMC, which encodes an apoplast protein described as a negative regulator of salt stress response [26], may play the part of a tuner in salt stress responses. To study the OsRMC transcriptional regulation in rice seedlings under high salinity stress, Serra et al. [27] screened two TFs, OsEREBP1 and OsEREBP2, belonging to the AP2/ERF family. As has been proved by electrophoretic mobility shift assay (EMSA) experiments in a previous study [28], these two TFs can bind to the OsRMC promoter containing the GCC-like motif S1 or the canonical GCC box. Transient expression assays were performed in *Arabidopsis* protoplasts which were transfected with the reporter plasmid pLUCm35GUS-pRMC alone, or co-transfected with the effector plasmids 35S::OsEREBP1 or 35S::OsEREBP2 to assess the two TFs’ transcriptional activities. The results showed that GUS gene expression is repressed during co-transfection of the effector plasmid and reporter plasmid, suggesting that both TFs may act as repressors during OsRMC gene expression. However, it was also reported that OsEREBP1 could serve as a transcriptional activator when it interacts with GCC box sequences [28], and its transcriptional activity may depend on the sequences around the binding sites or on the binding of other TFs; besides, additional DNA-binding motifs in the promoter sequence may also affect TF function [29].

Ko et al. [30] analyzed the carrot ABA-regulated gene *DcECP31* using an ABA-responsive element containing AGCT in the *DcECP31* promoter as bait and obtained a clone including the carboxyl-terminal half of bZIP, which contains the basic and leucine zipper domains and binds to G-boxes containing the sequence AGCT. Simultaneously, it is reported that carrot plants that abnormally express *C-ABI3* possess the ability to accumulate mRNA of the *ECP31* gene in mature leaves when treated with abscisic acid (ABA) [31]. As C-ABI3 BD does not directly bind to the *DcECP31* promoter, these results support the hypothesis that C-ABI3 interacts with the G-box via protein–protein contacts with G-box–specific DNA-binding
proteins [32,33]. Then a model for regulation of *DcECP31* gene expression was proposed; that is, C-ABI3 functions as a co-activator that interacts with motifs by protein–protein contacts with bZIP factors and controls the expression of the *DcECP31* gene [30].

**Prospects**

Advances in molecular biology technology have broadened the application of Y1H. Taking advantage of Gateway cloning technology, Y1H has become an appropriate approach to conduct TF screening instead of cDNA library screening on a genomic scale [34], providing a low-cost and fast approach to investigating gene regulatory networks [35]. Even so, performance of this technique still largely depends on abundance of relevant TFs and conditions of template mRNA, which has an impact on the quality of cDNA library [5,36]. Therefore, technical improvements are still needed. When studying the genome-binding sites and functions of specific TF, a TF-centered Y1H protocol could be an optimal choice. In this case, a *cis*-element library is screened against TF bait, so that it could efficiently help eliminate false positives caused by endogenous TFs’ binding on the bait sequence [37,38]. It is crucial that the basal expression levels of a reporter gene (like toxic tolerance) are initially tested with the absence of prey to identify the inhibiting effect at different concentrations of toxic substances. A minimal inhibitory concentration is determined and applied for library screening in order to suppress basal expression of the background. To eliminate the false positives, there is another approach called reverse one-hybrid system. E.g. to identify *URA3* gene expression, 5-FOA(5-Fluoroorotic Acid) is added to medium, which may be catalyzed by the *URA3* enzyme into a toxic compound. As a result, *URA3* expression is toxic to cells grown on medium containing 5-FOA. This kind of selection is based on two conditions: (1) the reporter gene expression causes toxicity under specific conditions; and (2) construction of alleles of that gene whose expression is tightly related to a TF [39].

**Conclusions**

It is clear that plant functions are not decided by a single gene, TF, or enzyme. That is, applying Y1H in plant genomics studies should take into consideration the overall plant regulation process to target the key elements in each step, which is a comprehensive task. As the genome sequencing projects of more plant species are being completed, the yeast one-hybrid technique, in combination with the yeast two-hybrid technique and other modified methods such as the reverse one-hybrid system, Gateway cloning and other biochemical assays, will provide an innovative approach in plant functional genomics studies and will help shed light on more gene functions in plants.

**Acknowledgments**

This work was supported by the State Key Laboratory of Tree Genetics and Breeding at the Northeast Forestry University under grant number K2013102 and the National Natural Science Foundation of China under grant number 31400576.

**Disclosure statement**

No potential conflict of interest was reported by the authors.

**Funding**

This work was supported by the State Key Laboratory of Tree Genetics and Breeding at the Northeast Forestry University [grant number K2013102] and the National Natural Science Foundation of China [grant number 31400576].

**References**

[1] Kalidhasan N, Joshi D, Bhatt TK, et al. Identification of key genes involved in root development of tomato using expressed sequence tag analysis. Physiol Mol Biol Plants. 2015;21(4):491–503.

[2] Delporte M, Legrand G, Hilbert JL, et al. Selection and validation of reference genes for quantitative real-time PCR analysis of gene expression in *Cichorium intybus*. Front Plant Sci. 2015 [cited 2017 Mar 23];6(651):651. DOI:10.3389/fpls.2015.00651

[3] Hu Z, Parekh U, Maruta N, et al. Down-regulation of fusarium oxysporum endogenous genes by host-delivered RNA interference enhances disease resistance. Front Chem. 2015 [cited 2017 Mar 23];3:1. DOI:10.3389/fchem.2015.00001

[4] Hyun Y, Kim J, Cho SW, et al. Site-directed mutagenesis in *arabidopsis thaliana* using dividing tissue-targeted rgen of the crispr/cas system to generate heritable null alleles. Planta. 2015;241(1):271–284.

[5] Wang MM, Reed RR. Molecular cloning of the olfactory neuronal transcription factor Olf-1 by genetic selection in yeast. Nature. 1993;364(3643):121–126.

[6] Brent R, Ptashne M. A eukaryotic transcriptional activator bearing the dna specificity of a prokaryotic repressor. Cell. 1985;43(3):729–736.

[7] Pyyovarenko T, Lopato S. Isolation of plant transcription factors using a yeast one-hybrid system, Plant Methods. 2006;754(2):292–307.

[8] Lopato S, Bazanova N, Moran S, et al. Isolation of plant transcription factors using a modified yeast one-hybrid system, Plant Methods. 2006 [cited 2017 Aug 31];2:3. DOI:10.1186/1746-4811-2-3

[9] Reece-Hoyes JS, Marian WAJ. Yeast one-hybrid assays: a historical and technical perspective. Methods. 2012;57(4):441–447.
[10] Yang PC, Zhou B. Construction of yeast one-hybrid library for screening of G-box binding proteins. Lett Biotech. 2012;23:532–536.
[11] Liu QX, Guo CH, Bi YD, et al. Construction of fusion gene expression cDNA library of soybean root with yeast one-hybrid method. Soyb Sci. 2013;32:165–167.
[12] Omidvar V, Abdullah SNA, Ebrahimi M, et al. Gene expression of the oil palm transcription factor EgAP2-1 during fruit ripening and in response to ethylene and ABA treatments. Biol Plant. 2013;57(4):646–654.
[13] Lei Z, Yin CC, Li YL, et al. Cloning and analysis of anti-stress transcription factor gene. J Agric Sci Tech China. 2014;16(4):71–78.
[14] Wang X, Du B, Liu M, et al. Arabidopsis transcription factor WRKY33 is involved in drought by directly regulating the expression of CesA8. Am J Plant Sci. 2013;4(6A):21–27.
[15] Jiang YQ, Deyholos MK. Functional characterization of Arabidopsis NaCl-inducible WRKY25 and WRKY33 transcription factors in abiotic stresses. Plant Mol Biol. 2009;69(1):91–105.
[16] Zheng Z, Qamar SA, Chen Z, et al. Arabidopsis WRKY33 transcription factor is required for resistance to necrotrophic fungal pathogens. Plant J. 2006;48(5):592–605.
[17] Lippok B, Birkenbihl RP, Rivory G, et al. Expression of AtWRKY33 encoding a pathogen- or PAMP-responsive WRKY transcription factor is regulated by a composite DNA motif containing W box elements. Mol Plant Microbe. 2007;20(4):273–283.
[18] Kelemen Z, Sebastian A, Xu W, et al. Analysis of the DNA-binding activities of the Arabidopsis R2R3-MYB transcription factor family by one-hybrid experiments in yeast. Plos One. 2015 [cited 2017 Aug 31];10(10):e0141044. DOI: 10.1371/journal.pone.0141044
[19] Chen Z, Hong X, Zhang H. Disruption of the cellulose synthase gene, AtCesA8/IRX1, enhances drought and osmotic stress tolerance in Arabidopsis. Plant J. 2005;43(2):273–283.
[20] Wei G, Lei J, Gong W, et al. Clone, expression, DNA binding ability and analysis of transcription activation in vitro of gene QRAP2 in Arabidopsis. Chinese Sci Bull. 2005;50(17):1873–1878.
[21] Gerber HP, Seipel K, Georgiev O. Transcriptional activation modulated by homopolymeric glutamine and proline stretches, Science. 1994;263(5148):808–811.
[22] Al-Attala MN, Wang X, Abou-Attia MA. A novel TaMYB4 transcription factor involved in the defence response against Puccinia strifor mis f. sp. tritici and abiotic stresses. Plant Mol Biol. 2014;84(4):598–603.
[23] Zhao T, Liu J, Li HY. Using hybrid transcription factors to study gene function in rice. Sci China Life Sci. 2015;58(11):1160–1162.
[24] Zhang C, Liu J, Zhao T. A drought-inducible bZIP transcription factor OsABF1 delays reproductive timing in rice. Plant Physiol. 2016;171(1):334–343.
[25] Amalraj A, Luang S, Kumar MY. Change of function of the wheat stress-responsive transcriptional repressor TaRAP2.1L by repressor motif modification. Plant Biotechnol J. 2016;14(2):820–832.
[26] Zhang L, Tian LH, Zhao JF, et al. Identification of an apoplastic protein involved in the initial phase of salt stress response in rice root by two-dimensional electrophoresis. Plant Physiol. 2009;149(2):916–928.
[27] Serra TS, Figueiredo DD, Cordeiro AM. OsRMC, a negative regulator of salt stress response in rice, is regulated by two AP2/ERF transcription factors. Plant mol biol. 2013;82(4):439–455.
[28] Yong HC, Kim CY, Min CK, et al. BWMK1, a rice mitogen-activated protein kinase, locates in the nucleus and mediates pathogenesis-related gene expression by activation of a transcription factor. Plant Physiol. 2003;132(4):1961–1972.
[29] Flavia B, Elizabeth C, Patricia D. The Arabidopsis ABA-INSENSITIVE (ABI) 4 factor acts as a central transcription activator of the expression of its own gene, and for the induction of ABI5 and SBE2.2 genes during sugar signaling. Plant J. 2009;59(3):359–374.
[30] Ko S, Kamada H. Isolation of carrot basic leucine zipper transcription factor using yeast one-hybrid screening. Plant Mol Biol Rep. 2002;20(3):301–304.
[31] Shiota H, Satoh R, Watabe K, et al. C-ABI3, the carrot homologue of the Arabidopsis ABI3, is expressed during both zygotic and somatic embryogenesis and functions in the regulation of embryo-specific ABA-inducible genes. Plant Cell Physiol. 1998;39(11):1184–1193.
[32] Hattori T, Terada T, Hamasuna S. Regulation of the Oserm gene by abscisic acid and the transcriptional activator VP1: analysis of cis-acting promoter elements required for regulation by abscisic acid and VP1. Plant J. 1995;7(6):913–925.
[33] Hobo T, Asada M, Koywama Y. ACGT-containing abscisic acid response element (ABRE) and coupling element 3 (CE3) are functionally equivalent. Plant J. 1999;19(6):679–689.
[34] Breton G, Kay SA, Pruneda-Paz JL. Identification of Arabidopsis transcriptional regulators by yeast one-hybrid screening using a transcription factor ORFeome. Methods Mol Biol. 2016 [cited 2017 Mar 23];1398:107. DOI: 10.1007/978-1-4939-3356-3_10
[35] Pruneda-Paz JL, Ghislain B, Nagel DH. A genome-scale resource for the functional characterization of Arabidopsis transcription factors. Cell Rep. 2014;8(2):622–632.
[36] Li JJ, Herskowitz I. Isolation of ORC6, a component of the yeast origin recognition complex by a one-hybrid system. 1993;262(5141):1870–1874.
[37] Yanai K. A modified yeast one-hybrid system for genome-wide identification of transcription factor binding sites. Methods Mol Biol. 2013;977:125–136.
[38] Keiko TY, Yoshiko K, Takashi H. Identification and characterization of guttation receptor-binding sites in the human genome. J Recept Signal Transduct Res. 2010 [cited 2017 Mar 23];30(2):88–105. DOI: 10.3109/10799891003614816
[39] Vidal M, Brachmann RK, Fattaeya E. Reverse two-hybrid and one-hybrid systems to detect dissociation of protein–protein and DNA–protein interactions. Proc Natl Acad Sci USA. 1996;93(19):10315–10320.