

## Full Length Research Paper

# *In silico* approach to identification of a novel gene responsive to submergence stress in rice

Hewei Du\* and Min Huang

The College of Life Science, Yangtze University, Jingzhou, Hubei 434025, P. R. China.

Received 2 April, 2015; Accepted 15 May, 2015

Submergence is one of the major constraints to rice production. Bioinformatics approach has been widely used to identify candidate genes on many biological aspects. In the present study, a novel gene involved in submergence stress in rice, *Os07g47670* was identified by *in silico* approach. The amino acid sequence of *Os07g47670* is highly homologous to hypoxia-responsive family proteins. No disordered regions are found in the *Os07g47670* protein. In the *Os07g47670* gene promoter, there are two ARE cis-regulatory elements, indicating that *Os07g47670* is associated with submergence responsiveness. The *Os07g47670* transcript levels are higher in roots of one or two-week old plants than in other tissues. Without the *Sub1A* gene, the expression level of *Os07g47670* in M202 is low under submergence, ACC treatment, and normal condition. However, in the *Sub1A* genetic background, the *Os07g47670* transcript level is strongly induced during submergence, and peaked at day 1 during submergence. The mRNA level of *Os07g47670* in M202(*Sub1A*) was also significantly increased by ACC treatment. High expression level of *Os07g47670* is correlated with the existence of the *Sub1A* gene. *Os07g47670* shares similar expression patterns with *Sub1A*, *ADH1*, *SLR1*, and *SLRL1* and are co-induced under submergence. Thus, we have documented *Os07g47670* as a novel gene associated with submergence stress response. The identification of *Os07g47670* will facilitate the understanding of the molecular mechanism of submergence tolerance in rice.

**Key words:** Rice (*Oryza sativa* L.), *in silico* approach, submergence, waterlogging.

## INTRODUCTION

Rice is one of the most important crops, as a staple food for more than half the world's population (Chen et al., 2009). However, rice is subject to various abiotic stresses, such as drought, submergence, high salinity, and low temperature, resulting in significant damage to rice.

Among these abiotic stresses, submergence is increasingly becoming a major production constraint affecting about 15-20 million hectares of rice land in South and Southeast Asia and causing a loss of up to \$ 1 billion every year (Xu et al., 2006). Rice has developed numerous strategies to

\*Corresponding author. E-mail: duhewei666@163.com. Tel: 86-716-8066182. Fax: 86-716-8066235.

**Abbreviations:** QTL, Quantitative trait locus; ARE, anaerobic response elements; ACC, 1-aminocycloprop-1-carboxylic acid; NJ, neighbor-joining; ADH, Alcohol dehydrogenase; SLR1, Slender rice-1; SLRL1, SLR1 like-1; ERF, ethylene response factors.

Author(s) agree that this article remains permanently open access under the terms of the [Creative Commons Attribution License 4.0 International License](http://creativecommons.org/licenses/by/4.0/)

cope with submergence. Much progress has been made to understand their strategies conferring submergence tolerance (Fukao and Bailey-Serres, 2008; Perata and Voesenek, 2007). *SNORKEL1* and *SNORKEL2*, two ethylene response factors (ERF) genes isolated from deepwater rice, trigger the elongation of internodes under submergence. Thus, deepwater rice grows out of the surface of water, allows gas exchange with the atmosphere, and prevents drowning (Hattori et al., 2009). However, submergence tolerance in most *Oryza sativa* cultivars is conferred by *Sub1A-1* gene. The physiological and molecular base of *Sub1A-1* has been extensively studied. The role of *Sub1A-1* is to suppress the elongation of shoot during submergence, thus limiting anaerobic catabolism and leading to the preservation of carbohydrate reserves (Xu et al., 2006; Bailey-Serres et al., 2010).

So far, *SNORKEL1*, *SNORKEL2*, and *Sub1A-1* are three major genes that have been identified in conferring submergence tolerance. The molecular model of *Sub1A-1* has been described that represents a big step forward towards understanding the regulation of submergence tolerance in rice (Hattori et al., 2009; Xu et al., 2006). However, submergence tolerance is a complex quantitative trait. The mechanism of submergence tolerance in rice is still unclear. Therefore, it is necessary to identify new genes or QTLs involved in submergence tolerance. Recently, four new QTLs were identified using mapping population derived from two moderately tolerant varieties, IR72 and Madabar. Some progeny have an even higher survival rate than the FR13A-derived tolerant control (IR40931). Four QTLs were identified on chromosomes 1, 2, 9, and 12. The QTL on chromosome 9 was *Sub1A-1* (Septiningsih et al., 2012). Thirty-two putative QTLs associated with seedling vigor in rice under submergence were detected. Two QTLs with more than 10% contribution to the total phenotypic variance were verified for involvement in shoot length determination (Manangkil et al., 2013). Microarray has been widely used to measure mRNA levels of many genes in particular cells or tissues at once (Wang et al., 2005; Shimono et al., 2003). Using oligonucleotide microarray combined with suppression subtractive hybridization, a number of submergence-responsive genes in FR13A and Goda Heenati were identified. Under submergence, two genes exhibited an opposing expression pattern between FR13A and Goda Heenati, and 324 genes were regulated by submergence only in one genotype and unchanged in their expression in the other (Xiong et al., 2012).

As databases of gene expression data continue to grow, our understanding of gene function grows as well. In the present study, we utilized available microarray data, applied bioinformatics techniques, and identified a novel gene associated with submergence response. The study of this gene will help to understand the mechanism of submergence tolerance in rice and the breeding of submergence tolerance rice varieties.

## MATERIALS AND METHODS

### Retrieval of microarray data and its analysis

The microarray data of leaves of 14-day old M202 and M202 (sub1A) seedlings that were subject to submergence for 0, 1, 6 days, respectively, were retrieved from those available at NCBI database (<http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE41103>). Microarray data were analyzed according to Jung et al. (2010).

### Database search, sequence analysis, phylogenetic tree and heatmap construction

The NCBI database (<http://www.ncbi.nlm.nih.gov>) was mined to identify genes homologous to *Os07g47670*. The amino acid sequence of the *Os07g47670* protein was used as a query sequence to search the databases using BLASTP. Multiple sequence alignment analysis was performed using ClustalX (Chenna et al., 2003; <http://bioinformatics.ubc.ca/resources/tools/clustalx>). Phylogenetic trees of the aligned protein sequences were constructed using MEGA version 5.0 (Tamura et al., 2007; <http://www.megasoftware.net>) via the neighbor-joining (NJ) method with the following parameters: Poisson correction, pairwise deletion, and bootstrap (1000 replicates; random seed).

### Bioinformatic analysis

Bioinformatic analysis of the *Os07g47670* genes, such as the nucleotide and deduced amino acid sequences, composition, physical and chemical characterization, and conserved domain sequences, was performed using the Expert Protein Analysis System (ExpASY) proteomics server of the Swiss Institute of Bioinformatics (<http://cn.expasy.org>). The solubility of the recombinant proteins when overexpressed in *Escherichia coli* was predicted using the statistical model from the University of Oklahoma (<http://biotech.ou.edu>) (Zhuang et al., 2008). The folding states of the protein were predicted using the FoldIndex program (<http://biportal.weizmann.ac.il>) (Zhuang et al., 2008). The cis-regulatory elements in the *Os07g47670* promoter were predicted using PlantCARE (<http://bioinformatics.psb.ugent.be/webtools/plantcare/html>).

### RNA isolation and expression analysis

Total RNA was extracted from leaves of seedlings at different time points under submergence. Leaf, root, shoot, stem, and panicle at different growth stages were also chosen to extract total RNA. First strand cDNA was synthesized using SuperScript-II reverse transcriptase according to the manufacturer's instructions (Invitrogen). The *actin1* gene (Os03g50890.1) was used as an endogenous control to normalize the expression data (Table 1). The qRT-PCR primers specific for the *ADH1*, *Sub1A*, *SLR1*, and *SLRL1* genes are listed in Table 1. Real-time PCR was conducted using the SYBR real-time PCR kit (Takara Japan) with IQTM SYBR® Green Supermixture according to the manufacturer's instructions (Bio-Rad USA). The reaction conditions are as follows: 94°C for 1 min; followed by 40 cycles of 95°C for 10 s; 55°C for 10 s.

### Various environmental stress and hormone treatments

All submergence, drought, and hormone treatments were replicated in at least 3 independent biological experiments. For GA and ACC treatment, 14-day old M202 and M202(sub1A) seedlings grown in

**Table 1.** Primers used in this study.

Primers	Sequence (5'-3')	Usage
Actin1 forward	ACAGGTATTGTGTTGGACTCTGG	Real-time PCR
Actin1 reverse	AGTAACCACGCTCCGTCAGG	Real-time PCR
Sub1A forward	CGGCCTCATCACAATCGGAG	Real-time PCR
Sub1A reverse	ATGTCCATGTCCATATGTCGTCG	Real-time PCR
ADH1 forward	TGTTGGAGAGGGTGTGACTG	Real-time PCR
ADH1 reverse	GCCATCACCAATCATCACAC	Real-time PCR
SLR1 forward	GATCGTCACCGTGGTAGAGC	Real-time PCR
SLR1 reverse	GAGGGAATCGAACATGGTGG	Real-time PCR
SLRL1 forward	GGCGGCGACAATAACAACAACAGT	Real-time PCR
SLRL1 reverse	TACAAACACACGCTGCTACCATCC	Real-time PCR
Os07g47670 forward	ATGGCGGAGGAGAAGAGC	Real-time PCR
Os07g47670 reverse	GCAACCTGGCGTGGATGA	Real-time PCR

germination paper were transferred into mock (0.1% DMSO), GA (10  $\mu$ M), or ACC (10  $\mu$ M) solution, respectively, and treated for 6 h. Submergence treatment was carried out as described in Fukao and Bailey-Serres's report (Fukao and Bailey-Serres, 2008). Thirty five-day old M202 and M202(sub1A) were subject to drought stress for 0, 3, 5, 7 days in greenhouse, respectively. Drought treatment was carried out according to Xiong et al. (2014).

## RESULTS

### Expression of the *Os07g47670* gene under submergence treatment

The microarray data in M202 and M202(sub1A) under submergence stress were downloaded from the NCBI database, and analyzed. The expression level of the *Os07g47670* gene under submergence is shown in Supplementary Table 1. At normal conditions, the mRNA level of *Os07g47670* in M202 is approximately 6.0 fold higher than that in M202(Sub1A). However, under submergence, the mRNA level of *Os07g47670* in M202(sub1A) was strongly induced and peaked at day 1. On the contrary, the expression level of *Os07g47670* in M202 decreased under submergence. Consequently, under submergence for 1 day, the expression level of *Os07g47670* in M202(sub1A) increased 6.5 fold, and is approximately 5.0 fold of the level in M202 (Supplementary Table 1).

To verify the results of microarray, we performed quantitative real time RT-PCR (qRT-PCR) experiments. The qRT-PCR results show that the transcript level of *Os07g47670* in M202(Sub1A) was strongly induced during submergence from day 1 to day 6 and peaked at day 1. In parallel, the transcript level of *Os07g47670* in M202 was also increased during submergence. Most importantly, the *Os07g47670* level in M202(sub1A) is significantly higher (by approximately 3-fold) than in M202 at all time points, from day 0 to day 6 (Figure 1). The

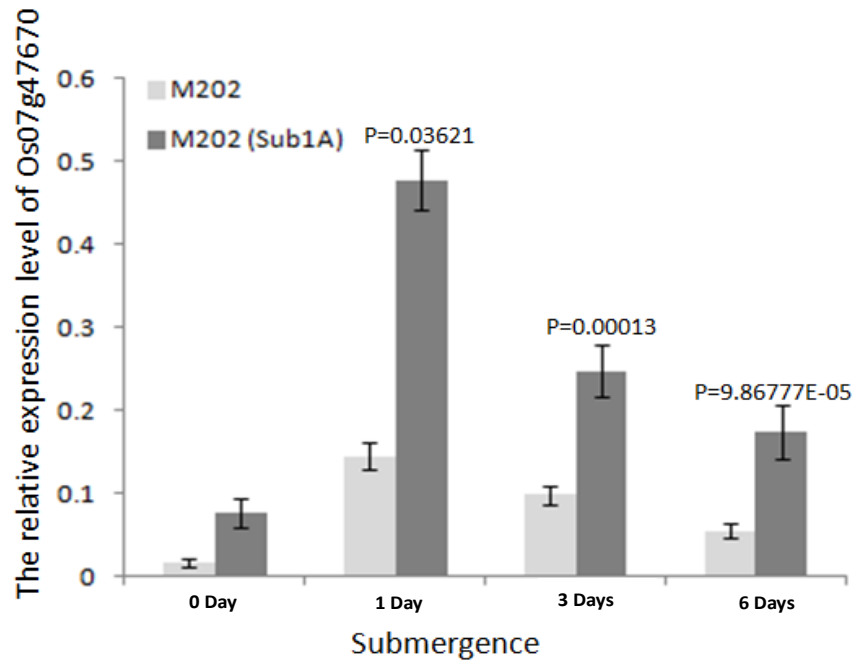
mRNA level of *Os07g47670* was induced both in M202 and M202(Sub1A) under submergence, which is not consistent with microarray results.

### Sequence alignment and biochemical property of the *Os07g47670* protein

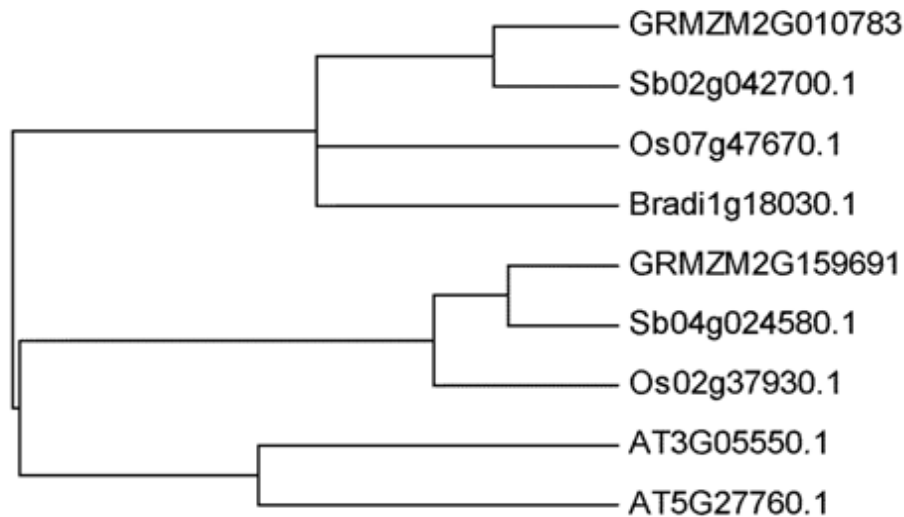
The amino acid sequence of the *Os07g47670* protein was used as a query sequence to search the databases using BLASTP. Ten homologous genes were identified. There are one homologous gene (*Os02g37930.1*) in rice, two (*GRMZM2G159691*, *GRMZM2G010783*) in maize, two (*AT3G05550.1*, *AT5G27760.1*) in *Arabidopsis*, two (*Potri.005G024500.1*, *Potri.013G015400.1*) in poplar, two (*Sb04g024580.1*, *Sb02g042700.1*) in Sorghum, and one (*Bradi1g18030.1*) in *Brachypodium*, respectively. The eleven proteins are highly homologous, and all belong to hypoxia-responsive family proteins with highly conserved amino acid sequences (Figures 2 and 3). The amino acid sequences, number of amino acids, theoretical molecular weight, theoretical pI recombinant protein solubility of *Os07g47670* are shown as Supplementary Table 2. No disordered regions are found in the *Os07g47670* protein (Supplementary Figure 1).

### Prediction of the *Os07g47670* promoter

The main cis-regulatory elements that are characteristic of a promoter sequence are revealed by PlantCARE prediction, including a Py-rich stretch element (-988 to 998 bp), conferring high transcription levels. Importantly, a number of cis-regulatory elements related to abiotic stress were found in 5' promoter of the *Nipponbare* gene, such as two anaerobic response elements (ARE) involved in anaerobic induction, three MYB binding site elements (MBS) involved in drought-inducibility, and one TC-rich repeats related to defense and stress responsiveness. One of the ARE is present on the plus-strand (-1807 to



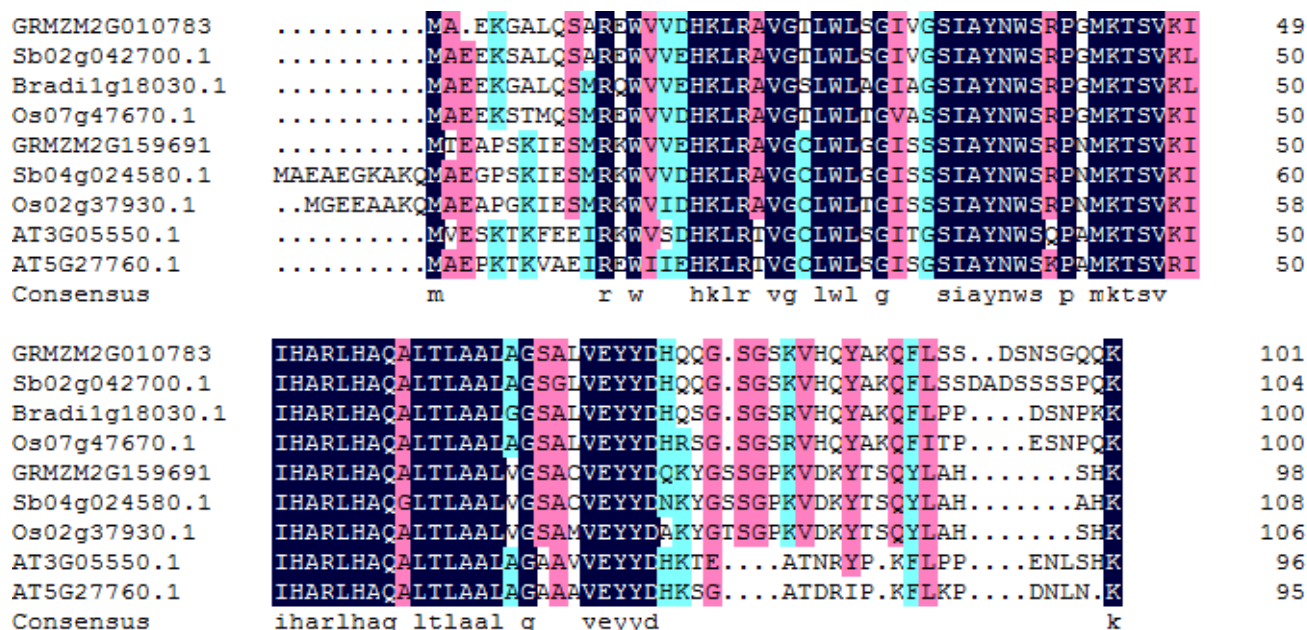
**Figure 1.** Expression level of the *Os07g47670* gene under submergence in M202 and M202(Sub1A). Fourteen-day old M202 and M202(Sub1A) seedlings were subject to submergence treatment for 0, 1, 3, and 6 days. The leaves were used to extract total RNA samples, which were used in qRT-PCR experiments. The expression levels of *Os07g47670* were calculated using  $2^{-\Delta\Delta Ct}$  values. Each bar represents the mean $\pm$ SD of 3 independent biological replicates.



**Figure 2.** Phylogenetic tree of *Os07g47670* and its homologous proteins of *Arabidopsis*, rice, maize, poplar, Sorghum, and Brachypodium. Full-length amino acid sequences were used in the multiple sequence alignment analysis using ClustalX. The phylogenetic tree of the aligned protein sequences was constructed using MEGA.

-1813 bp), and the other on negative-strand (-1845 to 1851 bp). One TC-rich repeats is present at -1389 to

-1399 bp at negative strand. For the three MBS elements, two are present on the plus-strand (-1445 to -1451 and



**Figure 3.** Sequence alignment of Os07g47670 and its homologous proteins. Multiple sequence alignment analysis was performed using ClustalX. Black color indicates amino acid residues conserved in every protein; dots represent gaps in amino acid sequences.

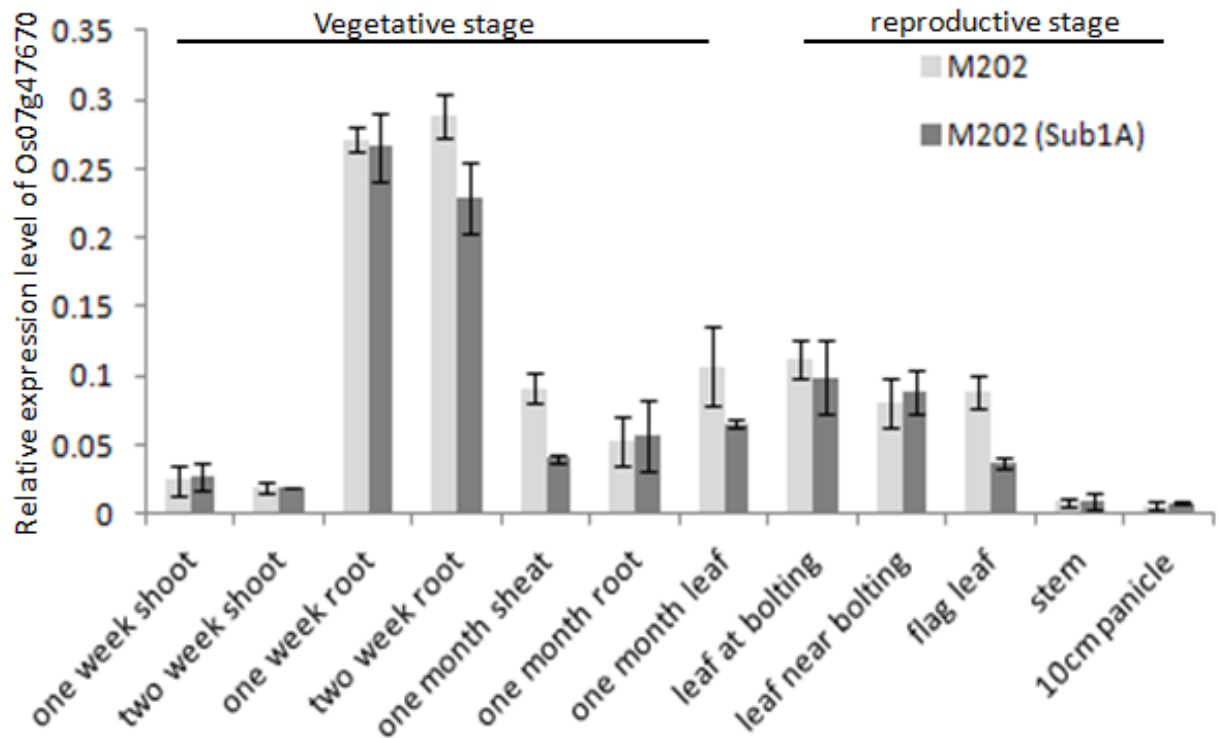
**Table 2.** Putative cis-acting elements in the *Os07g47670* gene promoter.

Site name	Position	Strand	Sequence	Function
5'UTR py-rich stretch	-988~-998	+	TTTCTTCTCT	conferring high transcription level
ARE	-1807~-1813	+	TGGTTT	anaerobic induction
ARE	-1845~-1851	-	TGGTTT	anaerobic induction
TC-repeats	-1389~-1399	-	ATTTTCTTCA	Cis-acting element involved in defense and stress responsiveness
MBS	-1630~-1636	+	TAACTG	MYB binding site involved in drought-inducibility responsiveness
MBS	-1445~-1451	+	TAACTG	MYB binding site involved in drought-inducibility responsiveness
MBS	-1219~-1225	-	CGGTCA	MYB binding site
ABRE	-1249~-1259	-	TACGTG	involved in the abscisic acid responsiveness
TCA-element	-1960~-1969	+	CCATCTTTTT	involved in salicylic acid responsiveness
TCA-element	-997~-1006	-	GAGAAGAATA	involved in salicylic acid responsiveness
AuxRR-core	-1186~-1193	+		

-1630 to -1636 bp) and one on the negative strand (-1219 to -1225 bp). In addition, there are four cis-regulatory elements involved in hormone responsiveness, including one ABRE involved in abscisic acid responsiveness, one AuxRR-core involved in auxin responsiveness, and two TCA-elements related to salicylic acid responsiveness. The ABRE is at -1249 to -1259 bp on the negative-strand. For the two TCA-elements, one is located at -997 to -1006 bp on the negative strand and the other at -1960 to -1969 bp on the plus strand. The AuxRR-core is located at -1186 to -1193 bp on the plus strand (Table 2).

### Expression of the *Os07g47670* gene in different tissues

The mRNA levels of the *Os07g47670* gene in different tissues were determined using qRT-PCR. At the vegetative stages of one and two-week old plants, the mRNA level of *Os07g47670* in shoots is low both in M202 and M202(Sub1A), but is high in roots. The transcript levels of *Os07g47670* in sheath, root, and leaf tissues in one-month old plants are higher than in shoots, but are lower than in roots, of one- or two-week old plants. At the



**Figure 4.** Transcript levels of the *Os07g47670* gene in various tissues at different stages. The expression levels of *Os07g47670* were calculated using  $2^{-\Delta\Delta Ct}$  values. Each bar represents the mean $\pm$ SD of 3 independent biological replicates.

reproductive stage, the expression levels of *Os07g47670* are high in leaves, including leaves at bolting, leaves near bolting, and the flag leaf, but are low in stems and panicles. The highest mRNA levels of *Os07g47670* are expressed in roots of one- or two-week old plants (Figure 4).

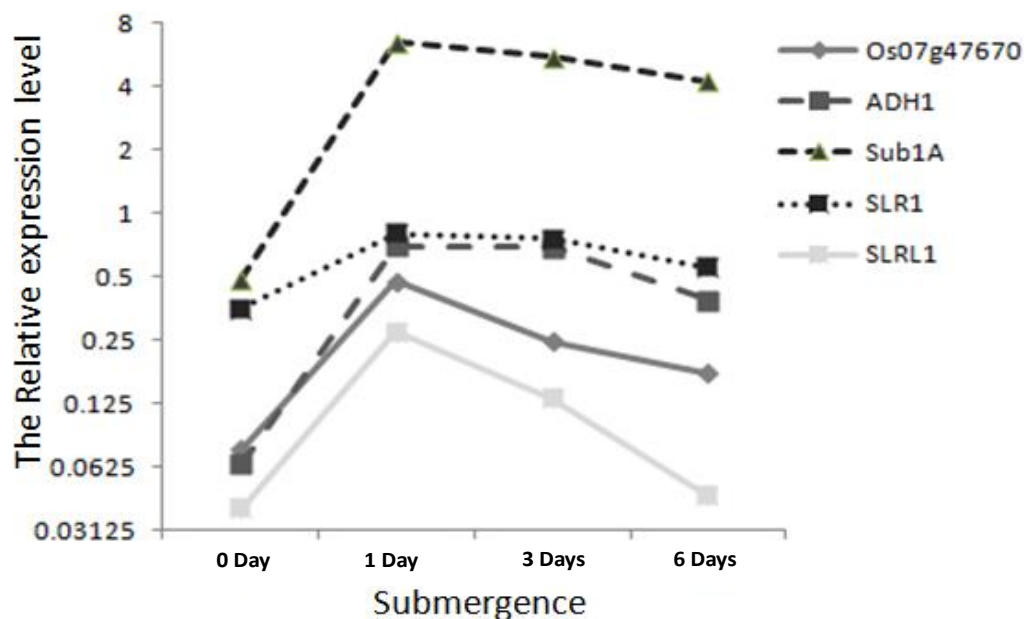
#### ***Os07g47670*, *ADH1*, *SLR1*, *SLRL1*, and *Sub1A* are co-expressed under submergence**

*Sub1A* is an important gene for submergence tolerance in rice and it confers submergence tolerance by augmenting accumulation of the GA signaling repressors *SLR1* and *SLRL1*. *ADH1* is a well-known marker gene for submergence tolerance. Therefore, we carried out real-time RT-PCR to assess possible co-expression of these genes and *Os07g47670* under submergence. Under submergence, the expression levels of *Os07g47670*, *ADH1*, *Sub1A*, *SLR1*, and *SLRL1* were all induced, reaching the peak at day 1 of submergence, and then gradually decreased from day 3 to day 6 under submergence. The *Os07g47670* gene exhibits a similar expression pattern with the *Sub1A*, *ADH1*, *SLR1*, and *SLRL1* genes during submergence. Therefore, the *Os07g47670* gene is co-expressed with *Sub1A*, *ADH1*,

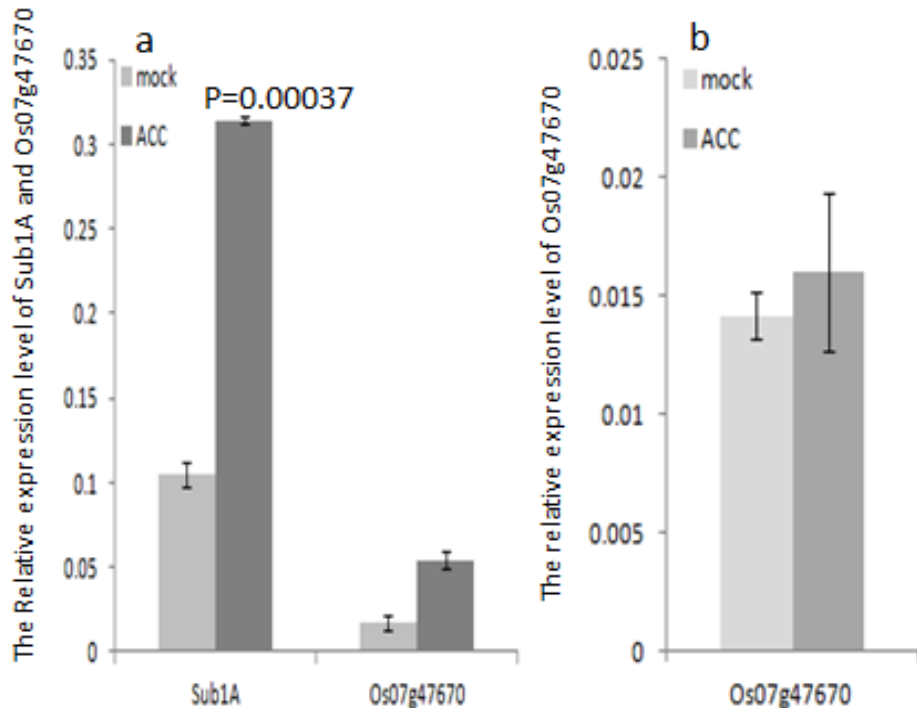
*SLR1*, and *SLRL1* during submergence stress (Figure 5).

#### ***Os07g47670* and *Sub1A* are co-regulated by GA and ACC**

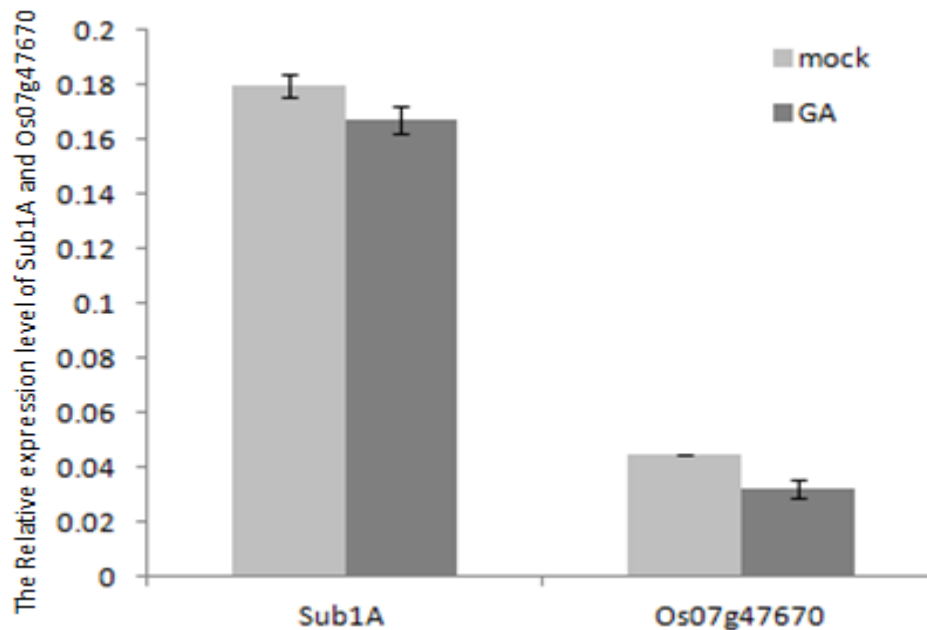
*Sub1A* encodes an ethylene-responsive transcription factor (ERF) and is strongly induced by ACC, the precursor of ethylene. *Sub1A* confers submergence tolerance by modulating the GA signal pathway in rice. Therefore, we performed real-time PCR to assess the expression of *Os07g47670* and *Sub1A* under GA and ACC treatments, respectively. With ACC treatment, *Os07g47670* and *Sub1A* gene transcript levels in M202(Sub1A) were strongly induced. The mRNA level of *Os07g47670* in M202(Sub1A) was induced to about 3.0 fold compared to mock (without ACC treatment), indicating that ACC stimulates the expression of *Os07g47670* (Figure 6a). However, the mRNA level of *Os07g47670* in M202 did not significantly increase with ACC treatment (Figure 6b). With GA treatment, the mRNA levels of *Os07g47670* and *Sub1A* were modestly decreased in M202(Sub1A), indicating that GA suppresses the expression of *Os07g47670* and *Sub1A*, in contrast to ACC action. These results demonstrate that *Os07g47670* and *Sub1A* share the same expression patterns in



**Figure 5.** Time course of *Os07g47670*, *SLR1*, *SLRL1*, *ADH1*, and *Sub1A* expression under submergence. Fourteen-day old seedlings were subjected to submergence and GA treatments. The mRNA levels in leaves were determined by qRT-PCR. The expression levels were calculated using  $2^{\Delta\Delta Ct}$  values. Each bar represents the mean  $\pm$  SD of 3 independent biological replicates.



**Figure 6.** Expression levels of *Os07g47670* and *Sub1A* after ACC treatment. **a.** Expression levels of *Os07g47670* and *Sub1A* in M202(Sub1A) with and without ACC treatment. **b.** Expression levels of *Os07g47670* in M202 with and without ACC treatment. 14-day old M202 and M202(Sub1A) seedlings were subject to mock and ACC treatment, respectively. The transcript levels of *Sub1A* and *Os07g47670* were determined by qRT-PCR. The expression level was calculated using its  $2^{\Delta\Delta Ct}$  value. Each bar represents the mean  $\pm$  SD of 3 independent biological replicates.



**Figure 7.** Expression level of *Os07g47670* in M202(Sub1A) after GA treatment. 14-day old M202(Sub1A) seedling were subject to mock and GA treatment, respectively. The expression level of *Os07g47670* was measured by real-time PCR and calculated using its  $2^{-\Delta\Delta Ct}$  value. Each bar represents the mean $\pm$ SD of 3 independent biological replicates.

response to GA and ACC treatments (Figure 7).

## DISCUSSION

The candidate gene approach has been at the forefront in studying many biological aspects. The use of bioinformatics tools is an affordable, fast, and efficient method for researchers to mine candidate genes responsive to various environmental stress or disease challenges. Submergence is one of the major constraints to rice production. In the present study, we attempted to identify novel genes associated with submergence tolerance in a silico approach. The microarray data were downloaded from the NCBI GEO database and analyzed for differential gene expression patterns between M202 and M202(Sub1A) under submergence stress. One novel gene, *Os07g47670*, was found to respond to submergence. The mRNA levels of the *Os07g47670* gene are low in M202 both in normal and submergence conditions, while are strongly induced in M202(Sub1A) during submergence. BLASTP and ClustalX analyses show that the *Os07g47670* protein is highly homologous to protein members that belong to a hypoxia-responsive protein family. Analysis of the *Os07g47670* promoter reveals that there are two ARE cis-regulatory elements associated with hypoxia responsiveness. This *in silico* prediction was based on best available knowledge on cis-elements. Therefore, the

prediction indicates that *Os07g47670* is a novel gene responsive to submergence stress in rice.

In the submergence-intolerance cultivar M202, the Sub1 region, covering 182 kb on chromosome 9, encodes two ERF genes, *Sub1B* and *Sub1C*. In tolerance near-isogenic line, M202(Sub1A), this locus encodes an additional ERF gene, namely *Sub1A*. Thus, *Sub1A* mediates the extinguished submergence tolerance of other cultivated rice. Although the molecular mechanism of *Sub1A* has been reported (Fukao and Bailey-Serres., 2008), genes that interact with *Sub1A* to confer submergence tolerance remain unclear. In the present study, in the absence of the *Sub1A* gene, the transcript levels of *Os07g47670* remain low in M202 both in normal and submergence conditions. On the contrast, in the presence of the *Sub1A* gene, the *Os07g47670* mRNA level is significantly increased in M202(Sub1A) during submergence. Thus, the induction of *Os07g47670* expression is specific to *Sub1A*.

*Sub1A* is strongly induced by ACC treatment (Fukao and Bailey-Serres., 2008). In M202 absent of *Sub1A*, the *Os07g47670* transcript level was not significantly increased with ACC treatment. However, in the presence of *Sub1A*, the expression level of *Os07g47670* in M202(sub1A) is significantly induced by ACC treatment. *ADH1*, *SLR1*, and *SLRL1* are submergence tolerance marker genes (Fukao and Bailey-Serres, 2008). *Os07g47670* responds to submergence stress similarly



as *ADH1*, *SLR1*, *SLRL1*, and *Sub1A*. Thus, *Os07g47670* is co-expressed with *Sub1A*, *ADH1*, *SLR1*, and *SLRL1* under submergence, indicating that they may function in the same pathway. Our results confirm that *Os07g47670* is a novel gene responsive to submergence stress. Therefore, the silico approach to identify candidate gene of submergence tolerance has demonstrated its validity.

Our results suggest that *Sub1A* regulates the expression of *Os07g47670*. *Sub1A* is an ERF transcription factor and can potentially bind to the two identified ARE cis-elements in the *Os07g47670* promoter to activate the promoter. Alternatively, *Sub1A* can activate other transcription factors that bind to the two ARE cis-elements and activate *Os07g47670* expression. It remains to be determined how *Sub1A* regulates *Os07g47670* expression.

### Conflict of interests

The author(s) did not declare any conflict of interest.

### ACKNOWLEDGEMENTS

We thank Pro. Xuewei Chen (Sichuan Agricultural University) for providing M202 and M202(*Sub1A*) seeds. We are grateful to Dr. Chern Mawsheng (University of California, Davis) for his critical reading and editing of the manuscript. This work was supported by National Natural Science Foundation (31271741) and the Hubei Province Natural Science Foundation (2011CDB006 and 2012FFA051).

### REFERENCES

- Bailey-Serres J, Fukao T, Ronald PC, Ismail A, Heuer S, Mackill D (2010). Submergence tolerant rice: SUB's journey from landrace to modern cultivar. *Rice*, 3:138-147.
- Chen H, Lin YJ, Zhang QF (2009). Review and prospect of transgenic rice research. *Chinese Sci. Bull.* 54(22):4049-4068.
- Chenna R, Sugawara H, Koike T, Lopez R, Gibson TJ, Higgins DG, Thompson JD (2003). Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Res.* 31:3497-3500.
- Fukao T, Bailey-Serres J (2008). Submergence tolerance conferred by *Sub1A* is mediated by *SLR1* and *SLRL1* restriction of gibberellins responses in rice. *PNAS*, 105(43):16814-16819.
- Hattori Y, Nagai K, Furukawa S, Song XJ, Kawano R, Sakakibara H, Wu J, Matsumoto T, Yoshimura A, Kitano H, Matsuoka M, Mori H, Ashikari M (2009). The ethylene response factors *SNORKEL1* and *SNORKEL2* allow rice to adapt to deep water. *Nature*, 460:1026-1031.
- Manangkil OE, Vu HTT, Mori N, Yoshida S, Nakamura C (2013). Mapping of quantitative trait loci controlling seedling vigor in rice (*Oryza sativ L.*) under submergence. *Euphytica*, 192:63-75
- Perata P, Voeselek LACJ (2007). Submergence tolerance in rice requires *Sub1A*, an ethylene-response-factor-like gene. *Trends Plant Sci.* 12(2):43-46.
- Septiningsih EM, Sanchez DL, Singh N, Sendon PMD, Pamplona AM, Heuer S, Machill DJ (2012). Identifying novel QTLs for submergence tolerance in rice cultivars IR72 and Madabaru. *Theor. Appl. Genet.* 124:867-874.
- Tamura K, Dudley J, Nei M, Kumar S (2007). MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol. Biol. Evol.* 24:1596-1599.
- Xiong HY, Yang J, Li YS (2012). Identification of submergence-responsive genes in two indica rice genotypes carrying *SUB1A-1* but exhibiting differential tolerance. *J. Plant Biol.* 55:233-241.
- Xu KN, Xu X, Fukao T, Canlas P, Maghirang-Rodriguez R, Heuer S, Ismail AM, Bailer-Serres J, Ronald PC, Machill DJ (2006). *Sub1A* is an ethylene-response-factor-like gene that confers submergence tolerance to rice. *Nature*, 442:705-708.
- Zhuang J, Cai B, Peng RH, Zhu B, Jin XF, Xue Y, Gao F, Fu XY, Tian YS, Zhao W, Qiao YS, Zhang Z, Xiong AS, Yao QH (2008). Genome-wide analysis of the AP2/ERF gene family in *Populus trichocarpa*. *Biochem. Biophys. Res. Commun.* 371:468-474.

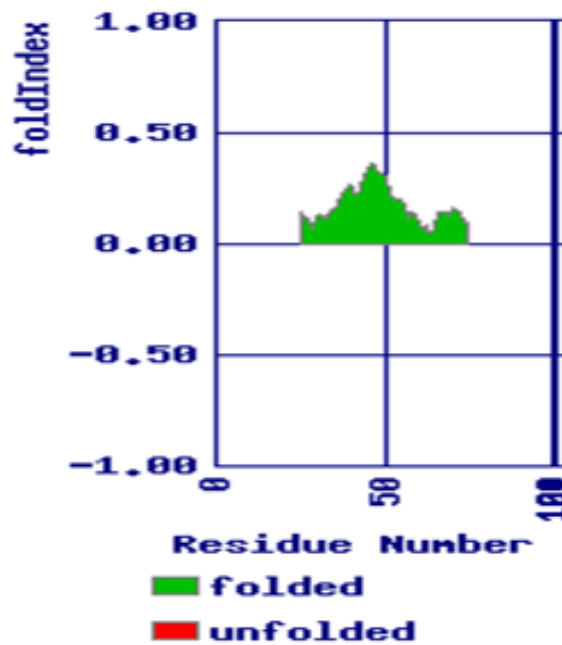
**Supplementary Table 1.** Microarray data of the Os07g47670 gene under submergence.

$\log_2$ s/m_0d	$\log_2$ s/m_1d	$\log_2$ s/m_6d	$\log_2$ 1d/0d_sub1	$\log_2$ 6d/0d_sub1	$\log_2$ 1d/0d_m202	$\log_2$ 6d/0d_m202
-2.66329	2.254901	0.706373	2.743458	1.576116	-2.17473	-1.79355

s and sub1, M202(Sub1A); m, M202; 0d, 1d, 6d, number of days subject to submergence.

**Supplementary Table 2.** The composition and physical and chemical characterization of amino acid sequences in Os07g47670.

Genbank No.	Number of amino acid	Theoretical Mw/Da	Theoretical pI	Percentage of Aliphatic amino acid /%	Percentage of Aromatics amino acid /%	Percentage of Positive amino acid /%	Percentage of Negative amino acid /%	Recombinant Protein Solubility(per cent of insolubility)/ %
Os07g47670	101	11269.85	9.63	84	8	17	8	97.4

**Supplement Figure 1.** Prediction of the folding state of Os07g47670.