

# Exploration of *SNORKEL1* (SK1) and *SNORKEL2* (SK2) QTLs in Deep Water Rice Germplasm Through Genotyping and *In-silico* Approach

Al Amin<sup>1,\*</sup>, KM Iftekharuddaula<sup>1</sup>, Animesh Sarker<sup>2</sup>, Md. Abdul Kader<sup>1</sup>, Ashraf Hossain Talukder<sup>2</sup>, Tamal Lata Aditya<sup>3</sup>, Md. Ansar Ali<sup>4</sup>, Md. Shahjahan Kabir<sup>5</sup>

<sup>1</sup>Plant Breeding Division, Bangladesh Rice Research Institute, Gazipur, Bangladesh

<sup>2</sup>Department of Biotechnology and Genetic Engineering, Mawlana Bhashani Science and Technology University, Tangail, Bangladesh

<sup>3</sup>Research Department, Bangladesh Rice Research Institute, Gazipur, Bangladesh

<sup>4</sup>Administration and Common Services Department, Bangladesh Rice Research Institute, Gazipur, Bangladesh

<sup>5</sup>General Office, Bangladesh Rice Research Institute, Gazipur, Bangladesh

## Email address:

alamingenetics@gmail.com (Al Amin)

\*Corresponding author

## To cite this article:

Al Amin, KM Iftekharuddaula, Animesh Sarker, Md. Abdul Kader, Ashraf Hossain Talukder, Tamal Lata Aditya, Md. Ansar Ali, Md. Shahjahan Kabir. Exploration of *SNORKEL1* (SK1) and *SNORKEL2* (SK2) QTLs in Deep Water Rice Germplasm Through Genotyping and *In-silico* Approach. *American Journal of Plant Biology*. Vol. 3, No. 4, 2018, pp. 33-40. doi: 10.11648/j.ajpb.20180304.11

**Received:** February 23, 2019; **Accepted:** March 28, 2019; **Published:** May 6, 2019

**Abstract:** The molecular screening of deep water rice landraces showed that out of 21 landraces, 15 landraces possess the *SNORKEL1* (SK1) and *SNORKEL2* (SK2) genes conferring stem elongation of paddy plant survive in increased water level of deep water ecosystem. Both genes/Quantitative trait loci (QTLs) were subjected to study through in-silico approaches. The fasta sequences, secondary structure and 3D structure of those proteins were identified and verified using bioinformatics tools. The physico-chemical properties and functions were also predicted. The in-silico study of *SNORKEL1* (SK1) and *SNORKEL2* (SK2) QTL have also revealed the capacity and way of survival of deep water rice germplasm noticeably. This study illustrates the many deep water rice varieties having *SNORKEL1* (SK1) and *SNORKEL2* (SK2) QTL and others do not possess the these QTLs though they are cultivated in deep water ecosystem. The molecular screening and in-silico study output of deep water rice landraces can be applied for establishing morphological correlation that 15 landraces are more elongating type than others and more adapted in natural deep water condition.

**Keywords:** Deep Water Rice, Molecular Screening, *SNORKEL1*, *SNORKEL2*, In-silico Study

## 1. Introduction

Deepwater rice is cultivated around river basins or delta areas in South and Southeast Asia, West Africa and South America, such as in the Mekong, Ganges, Brahmaputra, Amazon and Niger deltas, where severe floods occur in the rainy season every year, and where water levels can reach several meters. Furthermore, these deepwater floods can last for several months, depending on the amount and duration of precipitation. In these areas, few plants can survive under the long-term deepwater flood conditions, even the flash-flood-tolerant rice. It was estimated that deepwater rice is grown on

around 90,000 km<sup>2</sup> (35,000 sq mi) of land. The main areas where it is grown are in South and Southeast Asia where more than 100 million people rely its production for their livelihood [1]. In Bangladesh most of the rice grown in the low lying areas during monsoon are floating rice, generally called as deepwater rice, locally known as broadcast aman, poush dhan, etc.

The only limitation with the submergence tolerant inbred lines is short plant type which makes them unsuitable for medium to deep low land areas where water depth may reach up to 100 cm or more and persist for more than one month [2, 3]. It has been shown that few traditional deep water low

yielding rice lines are still cultivated in extremely deep water lowland areas by poor marginal farmers where none of the improved lines can grow [4]. Though Bangladesh had harboured a hand full of such rice land races, during and after green revolution, most of them were prematurely replaced by HYVs prior to any experimentation and possible utilization in breeding [5]. Experimentation and utilization of remaining few such lines which are limited to a narrow marginal rice growing regions may prove to be one of the most effective strategies for possible utilization of the different beneficiary traits available within them. However, deepwater rice varieties can survive, because of rapid stem elongation with rises in water level, sustained by internode elongation [6]. Ordinary paddy rice cultivars do not initiate internode elongation during the vegetative stage; however, deepwater rice cultivars can elongate internodes even in early growth stages and also develop aerenchyma, the snorkel-like conduits in internodes that allow gaseous exchange. In more recent times, with the development of QTL analysis, the deepwater response has been suggested to be a quantitative trait regulated by QTLs [7]. In 2009, Hattori *et al.* identified the QTL that had the strongest effect; the *SNORKEL1* (*SK1*) and *SNORKEL2* (*SK2*) are located on chromosome 12, according to positional cloning [8].

Plant molecular biology aims to study cellular processes, their genetic control, and interactions with environmental changes. Such a multi-dimensional and detailed investigation requires large-scale experiments involving entire genetic, structural, or functional components. These large scale studies are called “omics”. Major components of omics include genomics, transcriptomics, proteomics, and

metabolomics. These omics approaches are routinely used in various research disciplines of crop plants, including rice. Omics approaches have improved very rapidly during the last decade as technology advances. Subsequently, high-throughput data developed by omic experiments require extensive computational resources for storage and analysis. Thus, several online databases, analysis servers, and omics platforms have been developed. Omics is getting broader coverage and it is anticipated that several new omic fields will evolve in near future [9].

In this present study twenty one deep water landraces were experimented for their properties through genetic analysis using snorkel genes for their possible utilization in marker assisted breeding programme to develop HYV.

## 2. Materials and Methods

### 2.1. Experimental Site

The experiment was conducted at the Molecular Biology Laboratory of Plant Breeding Division, Bangladesh Rice Research Institute (BRRI), Gazipur -1701, Bangladesh.

### 2.2. Plant Materials and DNA Markers

A total of 21 landraces were tested in this experiment (Table 1). Two *SNORKEL*-region primers in the Chromosome 12 with clear amplifications were selected for molecular characterization of those 21 genotypes (Table 2). *SK1* and *SK2*, two gene-based markers specific to putative gene within *SNORKEL* QTL were used to confirm presence of *SNORKEL* QTL in the 21 landraces [10-12].

**Table 1.** List of genotypes used for molecular characterization.

SN.	Designation	Characteristics	SN.	Designation	Characteristics
1	PCR89350-B-R-3-1-2-1	Deepwater rice	12	Jole Kumari	Deepwater rice
2	Bajail 65	do	13	Gour Kajol	do
3	Horinga Digha	do	14	Saswari	do
4	Hijol Digha	do	15	Til Bajal	do
5	Laxmi Digha	do	16	Lalmoti	do
6	Lal Digha	do	17	Manikdigha	do
7	Hbj. A. VIII	do	18	Fulkari	do
8	Hbj. A. IV	do	19	Bhawalia digha	do
9	Hbj. A. II	do	20	Lal Khama	do
10	Bhawalia	do	21	Ajole digha	do
11	RD 19	do			

**Table 2.** List of primers used for molecular characterization of 21 landraces.

SN	Primers	Forward	Reverse
1	<i>Sk1</i>	ACG GTA TCC CTG AAC TAC TG	TCG TAG CGA CAG CCG TAC TG
2	<i>Sk2</i>	CAC TGG AGG CAA CGA ATG	TAA AAG GAC CAG AGG CAG C

### 2.3. Genotyping Protocol

DNA was extracted following CTAB method as described by Allen *et al.*, (2006) [13] and quantified using the Nano Drop spectrophotometer (NanoDrop 2000). For *Sk1* and *Sk2* markers, PCR was performed in 10 µl reactions containing around 25 ng/ µl of DNA template (3 µl DNA with 10X dilution factor), 1 µl 10X TB buffer (containing 200 mM

Tris-HCl pH 8.3, 500 mM KCl), 1.35 µl 25 mM MgCl<sub>2</sub>, 0.2 µl of 10 mM dNTP, 0.5 µl each of 10 µM forward and reverse primers and 0.2 µl of Taq DNA polymerase (5 U/µl) using GStorm and GenAtlas thermal cycler [14, 15]. Twelve-channel pipette was used for transferring DNA from dilution plate to PCR plate. Ten micro liter of mineral oil was added in each well to prevent evaporation and the PCR plate was wrapped with adhesive film. Then PCR plate was centrifuged

30 second at 3600 rpm with 10°C temperature for removal of bubbles within the PCR volume.

After initial denaturation for 2 minutes at 95°C, each cycle comprises 30 second denaturation at 95°C, 30 second annealing at 55°C, and 25 second extension at 72°C with a final extension for 5 minutes at 72°C at the end of 32 cycles. The PCR products were mixed with bromophenol blue gel loading dye and were analyzed by electrophoresis on 6% polyacrylamide gel using mini vertical polyacrylamide gels for high throughput manual genotyping (CBS Scientific Co. Inc., CA, USA). The gels were stained in 0.5 mg/ml ethidium bromide and photos were taken using Molecular Imager gel documentation unit (AlphaImager FC).

A brief prognostic analysis has been performed by using several bioinformatics tools and techniques in order to understand the physico-chemical properties, secondary and 3D structures of conserved protein regions, function prediction, and backbone confirmation for generated 3D model.

#### 2.4. Sequence Retrieval

Twenty submergence tolerant landraces and twenty two deep water rice landraces were considered for study. The CDS of genes with the database IDs Q0P1A8 (Sub1a) and C6L7X8 (sk1) were retrieved from Uniport. The fasta formatted sequences were further used for study given below.

FASTA sequence of Sk1

>tr|C6L7X8|C6L7X8\_ORYRU Ethylene response factor  
OS=Oryza rufipogon GN=Snorkel 1 PE=4  
SV=1MCGGCLIPDELVGKPARRTRAAAAGGDSGDGWK

```

      10      20      30      40      50      60      70
      |      |      |      |      |      |      |
MCGGCLIPDELVGKPARRTRAAAAGGDSGDGWKHGRRLCPAAAPCNCKPRRRAGAADDDDVGRRRRRTTTR
ccccccccchccccchhhhhhhhhhhcccccccccccccccccccccccccccccccccccccccccccc
RAASEVRFHGIHMRSYGRWSAEIRDSSYRGHRLWIGTYATAEAAARAYDAEARRIHGAKANTNFPPPPND
chhhhhhheccccchchcccehhhhhccccceeeeeeccchhhhhhhhhhhhhhhhhhhhhhhhhhhhhhh
VDSGAPPPPPWDLEAHMRFLGEVELDDGGAEP PPPPSYGIPELLHMEPELASATQSVHGDDEPWGLDKYM
ccccccccccccchhhhhhccceccccccccccccccccccccchhhccccccccccccccccccccchhhh
RFLSEVELDDGGAPLPPPPSQHGGVAAAGSPQYGCYDYLLMMCN
hheeeeeccccccccccccccccccccccccccccccccchhhheeeeecc

```

Sequence length : 256

HNN :

Alpha helix	(Hh)	:	60 is	23.44%
3 <sub>10</sub> helix	(Gg)	:	0 is	0.00%
Pi helix	(Ii)	:	0 is	0.00%
Beta bridge	(Bb)	:	0 is	0.00%
Extended strand	(Ee)	:	19 is	7.42%
Beta turn	(Tt)	:	0 is	0.00%
Bend region	(Ss)	:	0 is	0.00%
Random coil	(Cc)	:	177 is	69.14%
Ambiguous states (?)		:	0 is	0.00%
Other states		:	0 is	0.00%

Secondary structure of Sk2

HGRRLCPAAAPCNCKPRRRAGAADDDDVGRRRRRTTTR  
TRAASEVRFHGIHMRSYGRWSAEIRDSSYRGHRLWIG  
TYATAEAAARAYDAEARRIHGAKANTNFPPPPNDVDS  
GAPPPPPWDLEAHMRFLGEVELDDGGAEP PPPPSYGI  
PELLHMEPELASATQSVHGDDEPWGLDKYMRFLSEVEL  
DDGGAPLPPPPSQHGGVAAAGSPQYGCYDYLLMM  
CN

>tr|C6L7X5|C6L7X5\_ORYSI Ethylene response factor  
OS=Oryza sativa subsp. indica OX=39946 GN=Snorkel 2  
PE=2 SV=1

MCGENDNNGAAAGSSRRLPAVGAMRGPCIEEKLKT  
VVVLSDDDDDDYEEEFRRYCENTTL

PAKGDKGRRRPAASKKQHRHRFHGIHRRKSGRWS  
AEIRDNMIKGSRSWVGTFDTAEAAA

WAYDAVARRLYGPNARTNFPLPPPPPPVAPLLPAPV  
ANKKMNSKSKKPAPKMVVAPAG

GETAAAAGEMAPVLLGNALEATNGWEFEPYSCMGL  
VVCSAVYNYADEPEPADDELQLLHL

MHGGAMADFAADGCLWSF

#### 2.5. Protein Structure Prediction

##### 2.5.1. Secondary Structure Prediction

The secondary structures were generally predicted to find out the percentage of helices, sheets and turns. Secondary structures of the Sk1 were computed using the program Hierarchical Neural Network (HNN: <http://www.expasy.org/tools/>), which showed information about beta, helix and turns.

Secondary structure of Sk1:



Sequence length : 258

HNN :

Alpha helix	(Hh) :	69 is	26.74%
3 <sub>10</sub> helix	(Gg) :	0 is	0.00%
Pi helix	(Ii) :	0 is	0.00%
Beta bridge	(Bb) :	0 is	0.00%
Extended strand	(Ee) :	27 is	10.47%
Beta turn	(Tt) :	0 is	0.00%
Bend region	(Ss) :	0 is	0.00%
Random coil	(Cc) :	162 is	62.79%
Ambiguous states (?)	:	0 is	0.00%
Other states	:	0 is	0.00%

2.5.2. Physico-chemical Analysis

The physico-chemical properties generally represents the molecular weight, theoretical pI, amino acid composition, atomic composition, extinction coefficient, estimated half-life, instability index, aliphatic index and grand average of hydropathicity19 (GRAVY) was computed by using

ProtParam tool (<http://web.expasy.org/protparam/>). The sequences were submitted in protparam from where both physical and chemical properties were found successfully [16, 17].

Table 3. Physico-chemical properties of proteins.

properties	Sk1	Sk2
Number of amino acids	256	258
Molecular weight	27775.0	27948.6
Total number of negatively charged residues	35	33
Total number of positively charged residues	31	32
Total number of atoms	3803	3864
Aliphatic index	56.99	62.95
Grand average of hydropathicity	-0.693	-0.506
Theoretical PI	6.14	6.60

2.5.3. 3D Modeling

As the NMR crystallographic structures of the models are not available at PDB, so the models were generated by using model building servers. The tertiary structures of proteins were predicted by using the homology modeling concept in which the template is generated and the target is compared with the template to generate more appropriate model. Modeller tool was used to generate the 3D structures of proteins. The final models were visualized by using RasMol visualization tool.

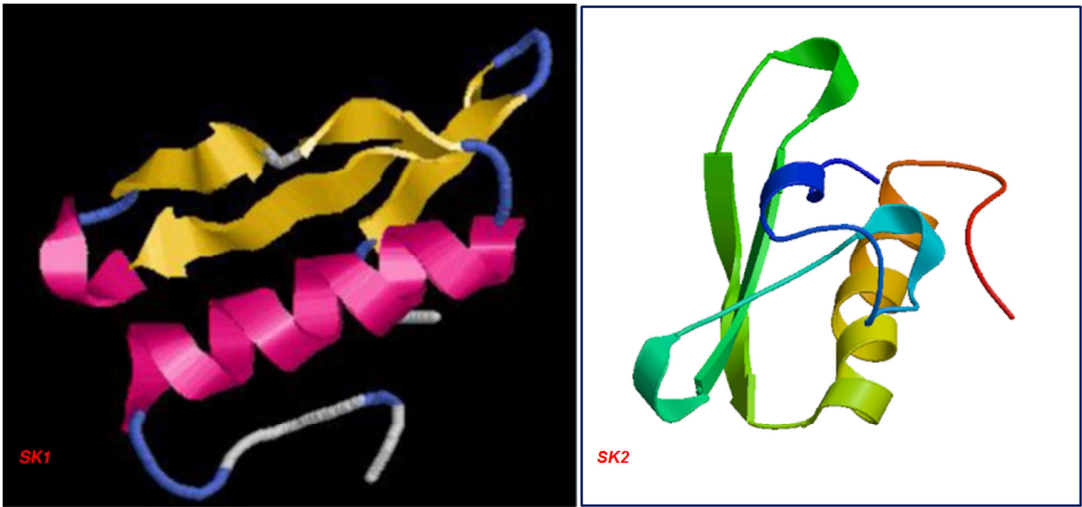


Figure 1. 3D structure of SK1 and SK2 protein.



### 2.5.4. Model Validation and Optimization

The models were then optimized by using RAMPAGE server (<http://mordred.bioc.cam.ac.uk/~rapper/rampage.php>).

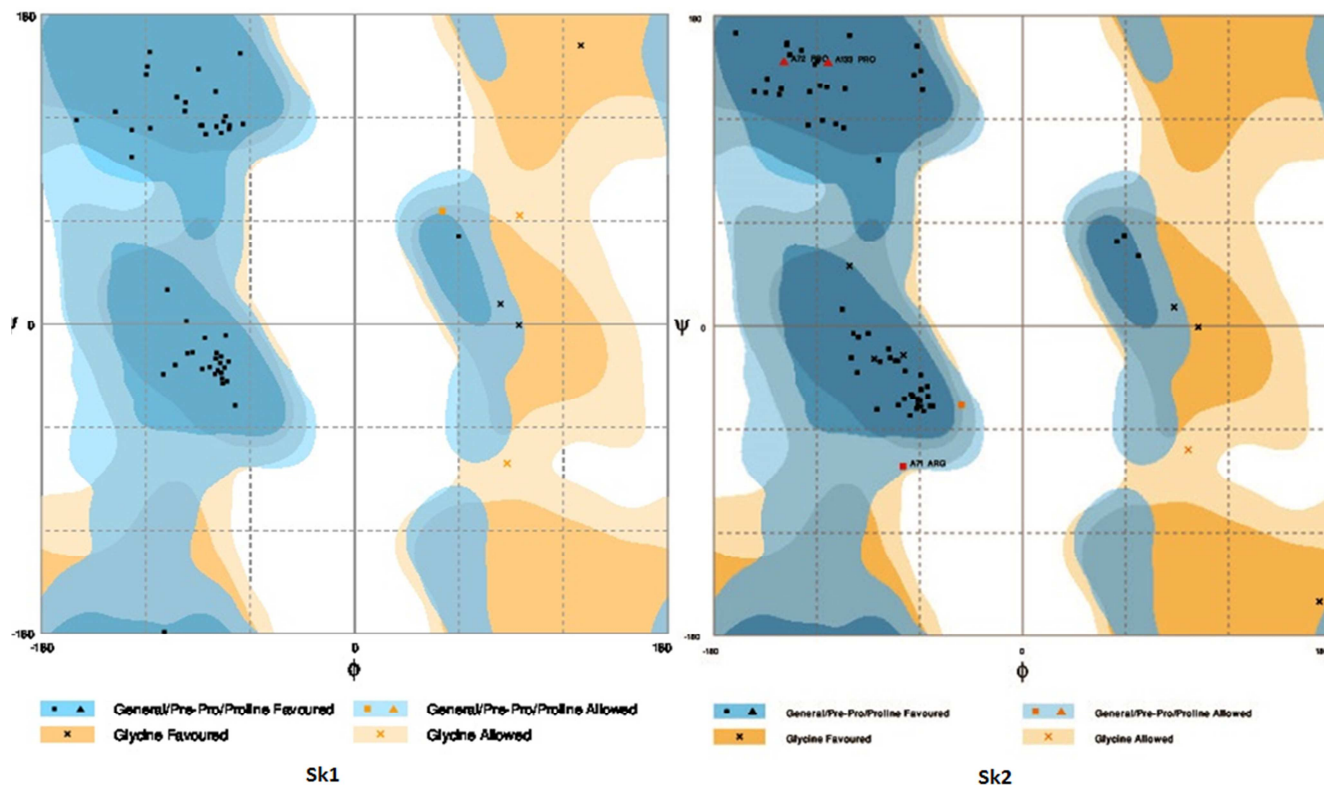


Figure 2. Ramachandran plot of 3D structure of SK1 and SK2 protein.

#### Evaluation of residues in Sk1

```
Residue [ 87 :GLY] ( 94.99, 63.16) in Allowed region
Residue [ 127 :GLY] ( 87.81, -80.67) in Allowed region
Residue [ 129 :LYS] ( 50.51, 65.64) in Allowed region
Number of residues in favoured region (~98.0% expected) : 53 ( 94.6%)
Number of residues in allowed region (~2.0% expected) : 3 ( 5.4%)
Number of residues in outlier region : 0 ( 0.0%)
```

#### Evaluation of residues in SK2

```
Residue [ 87 :GLY] ( 94.99, 63.16) in Allowed region
Residue [ 127 :GLY] ( 87.81, -80.67) in Allowed region
Residue [ 129 :LYS] ( 50.51, 65.64) in Allowed region
Number of residues in favoured region (~98.0% expected) : 53 ( 94.6%)
Number of residues in allowed region (~2.0% expected) : 3 ( 5.4%)
Number of residues in outlier region : 0 ( 0.0%)
```

### 2.5.5. Function Prediction

To comprehend the functional relationship between these proteins, function of predicted protein was retrieved by SVMProt [19, 20].

Table 4. Predicted functions of Sk1 and Sk2 genes by SVMprot.

Protein Family Name	R-Value	P-Value (%)
Sk1		
Zinc-binding	1.7	78.4
All DNA-binding	1.5	73.8
Copper-binding	1.0	58.6
DNA repair	1.0	58.6
EC2.1 Transferases - Transferring	1.0	58.6

Protein Family Name	R-Value	P-Value (%)
One-Carbon Groups		
Photosynthesis	1.0	58.6
Metal-binding	1.0	58.6
Sk2		
Zinc-binding	1.9	85.4
All DNA-binding	1.3	68.5
Magnesium binding	1.0	58.6
TC3.A.5 Type II (general) secretory pathway (IISP) family	1.0	58.6
DNA repair	1.0	58.6
Calcium binding	1.0	58.6
Photosynthesis	1.0	58.6
Metal-binding	1.0	58.6

### 3. Results and Discussion

#### 3.1. Molecular Screening of Local Germplasm to Identify *Sk1* QTL

This experiment was carried out to identify the SNORKEL

genes in the deepwater landraces of Bangladesh. The molecular screening of *Sk1* and *Sk2* showed that out of 21 landraces, 15 landraces possess the *Sk1* and *Sk2* genes. *Sk1* and *Sk2* genes containing landraces are indicated as present (Table 5).

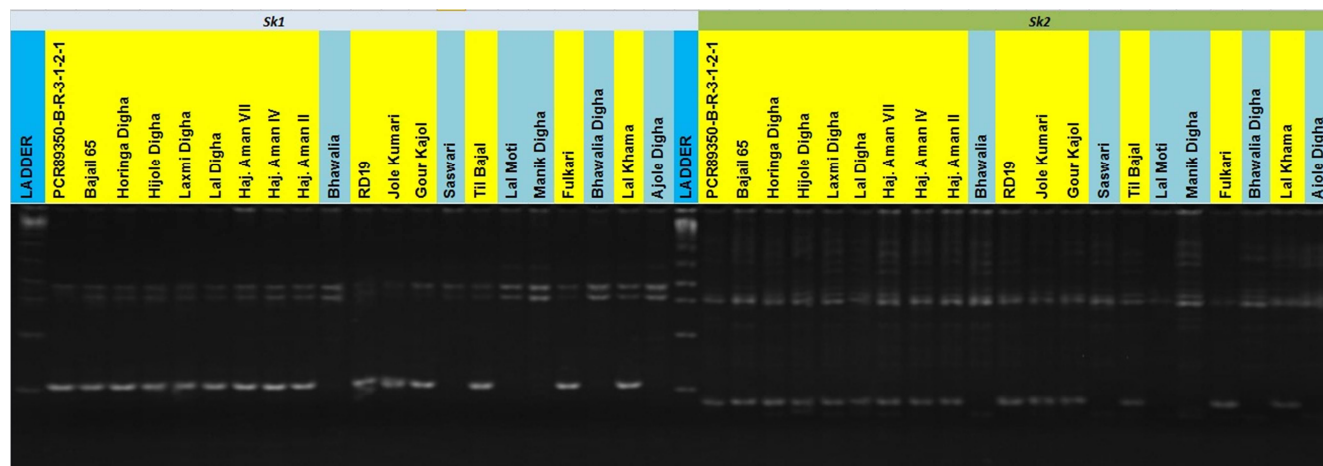


Figure 3. Partial view of the gel picture of deepwater landraces using *Sk1* and *Sk2* primer (Yellowish colored genotypes possess the *Sk1* and *Sk2* genes).

Table 5. The *Sk1* and *Sk2* genes containing landraces are indicated bold.

SN	Genotypes	<i>Sk1</i> gene status	<i>Sk2</i> gene status
1	PCR89350-B-R-3-1-2-1	Present	Present
2	Bajail 65	Present	Present
3	Horinga Digha	Present	Present
4	Hijol Digha	Present	Present
5	Laxmi Digha	Present	Present
6	Lal Digha	Present	Present
7	Hbj. A. VIII	Present	Present
8	Hbj. A. IV	Present	Present
9	Hbj. A. II	Present	Present
10	Bhawalia	Absent	Absent
11	RD 19	Present	Present
12	Jole Kumari	Present	Present
13	Gour Kajol	Present	Present
14	Saswari	Absent	Absent
15	Til Bajal	Present	Present
16	Lalmoti	Absent	Absent
17	Manikdigha	Absent	Absent
18	Fulkari	Present	Present
19	Bhawalia digha	Absent	Absent
20	Lal Khama	Present	Present
21	Ajole digha	Absent	Absent

The molecular screening of deep water rice landraces using *Sk1* and *Sk2* primer showed that out of 21 landraces, 15 landraces possess the *Sk1* and *Sk2* genes. *Sk1* and *Sk2* genes containing landraces are namely PCR89350-B-R-3-1-2-1, Bajail 65, Horinga Digha, Hijole Digha, Laxmi Digha, Lal Digha, Haj. Aman VII, Haj. Aman IV, Haj. Aman II, RD19, Jole Kumari, Gour Kajol, Til Bajal, Fulkari and Lal Khama.

Sub1A and Snorkel genes play an important role in submergence and deep water rice tolerance evolved from the ancestor FR13A and WO120 variety of *O. sativa*. The mechanism involves in sub1 gene submergence tolerance and snorkel gene tolerance mechanism is antagonistic to each other. Sub1-genes mechanism involve storage of energy in

the stressed condition and utilize it for plant regeneration, the antagonistic is escape from the level of water by utilizing current resource. Sub1-genes are useful when period of stress is for 1-2 weeks but when it comes to deep water or stagnant water rice snorkel genes are useful. The secondary structure and model which generated by using the protein of *Sub1A* and *Sk1* gave the best structure and most of the residues lie in favored regions, so this protein is more suitable for further study.

#### 3.2. Sequence Retrieval and Secondary Structure Prediction

The amino acid sequences were retrieved from uniprot and subjected for secondary structure prediction. The secondary structure of proteins shows the percentage of helix, sheet and coil of proteins (In M&M section 3.28.1).

#### 3.3. Physico-chemical Analysis

From the physico-chemical analysis it is found that the protein of *Sk1* and *Sk2* are distinct in physico-chemical nature (Table 3 in M&M section)

#### 3.4. Tertiary Structure Prediction Analysis

The amino acid sequences of the protein were subjected for generating homology models and then visualized by using various visualization tools. The helices are denoted with pink, sheets with yellow and loops in green colors respectively. Amino acids are the building blocks of the proteins and the backbone confirmation generally refers to phi-psi bonds, i.e. the regions where the amino acids are bind together by the peptide bonds due to electrostatic force, vander-walls force. So when more amino acid residues lie in the favored regions, then it gives better result on the further analysis (Table 4).

**Table 6.** Backbone confirmation of proteins.

Gene ID	Number of residues in favored region	Number of residues in allowed region	Number of residues in outlier region
Sk1	53 (94.6%)	3 (5.4%)	0%
Sk2	53 (94.6%)	3 (5.4%)	0%

## 4. Conclusion

The molecular screening of Sk1 and Sk2 showed that out of 21 landraces, 15 landraces possess the Sk1 and Sk2 genes. Sk1 and Sk2 genes containing landraces are namely PCR89350-B-R-3-1-2-1, Bajail 65, Horinga Digha, Hijole Digha, Laxmi Digha, Lal Digha, Hbj. Aman VII, Hbj. Aman IV, Hbj. Aman II, RD19, Jole Kumari, Gour Kajol, Til Bajal, Fulkari and Lal Khama. The proteins of Sk1 and Sk2 genes were characterized successfully. Secondary structure, tertiary structure, Physico-chemical analysis, function prediction were illustrated. The model prediction will be helpful in breeding programs and it will give breeders a better look to choose the gene of introgression. Deepwater rice having Sk1 and Sk2 genes will be used as donor parent in breeding program, which will accelerate the varietal development of the deepwater rice. Introgression of more potential gene to the elite varieties lacking the snorkel genes will improve their deep and stagnant water tolerance mechanism for rainfed lowland rice and deep water rice varieties.

## Acknowledgements

The author is grateful to BRRI authorities for providing supports in this research work. This research activity was financially supported the Integrated Agricultural Productivity Project (IAPP) funded by the Global Agriculture and Food Security Program (GAFSP).

## References

- [1] Catling, David. (1992). Rice in Deep Water. 10.1007/978-1-349-12309-4.
- [2] Iftekharruddaula K M, Newaz M A, Salam MA, Ahmed H U, Mahbub M A A, Septiningsih E M, Collard B C Y, Sanchez D L, Pamplona A M and Mackill D J. 2011. Rapid and High-Precision Marker Assisted Backcrossing to Introgress the SUB1 QTL into BR11, the Rainfed Lowland Rice Mega Variety of Bangladesh. *Euphytica*, 178, 83-97.
- [3] Xu K, Xia X, Fukao T, Canlas P, Maghirang-Rodriguez R, Heuer S, Ismail AM, Bailey-Serres J, Ronald PC, Mackill DJ. Sub1A is an ethylene response factor-like gene that confers submergence tolerance to rice. *Nature*. 2006; 442: 705–708. doi: 10.1038/nature04920.
- [4] Bailey-Serres J, Fukao T, Ronald P, Ismail A, Heuer S, Mackill D. Submergence Tolerant Rice: SUB1's Journey from Landrace to Modern Cultivar. *Rice*. 2010; 3: 138–147. doi: 10.1007/s12284-010-9048-5.
- [5] Das KK, Panda D, Sarkar RK, Reddy JN, Ismail AM. Submergence tolerance in relation to variable floodwater conditions in rice. *Environmental and Experimental Botany*. 2009; 66: 425–434.
- [6] Catling D: Rice in Deepwater. London: Macmillan Press; 1992. This is a comprehensive and specialized book on deepwater rice, dealing with such topics as the phenotype of deepwater rice, local cultures, agroecosystem and climate.
- [7] Jackson, M. B. and Ram, P. C. (2003) Physiological and Molecular Basis of Susceptibility and Tolerance of Rice Plants to Complete Submergence. *Annals of Botany*, 91, 227-241. <http://dx.doi.org/10.1093/aob/mcf242>
- [8] 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. The ethylene response factors SNORKEL1 and SNORKEL2 allow rice to adapt to deep water. *Nature*. 2009; 460: 1026–1030. doi: 10.1038/nature08258.
- [9] Rupesh Deshmukh, Humira Sonah, Gunvant Patil, Wei Chen, Silvas Prince, Raymond Mutava, Tri Vuong, Babu Valliyodan and Henry T. Nguyen (2014) Integrating omic approaches for abiotic stress tolerance in soybean. *Frontiers in plant science*, doi: 10.3389/fpls.2014.00244.
- [10] Hattori Y, Miura K, Asano K, Yamamoto E, Mori H, Kitano H, Matsuoka M, Ashikari M: A major QTL confers rapid internode elongation in response to water rise in deepwater rice. *Breed Sci* 2007, 57: 305-314.
- [11] Hattori Y, Nagai K, Mori H, Kitano H, Matsuoka M, Ashikari M: Mapping of three QTL that regulate internode elongation in deepwater rice. *Breed Sci* 2008, 58: 39-46.
- [12] Hattori, Y., Nagai, K. and Ashikari, M. (2011) Rice Growth Adapting to Deep Water. *Current Opinion in Plant Biology*, 14, 100-105. <http://dx.doi.org/10.1016/j.pbi.2010.09.008>.
- [13] Allen G C, Flores-Vergara M A, Krasnyanski S, Kumar S, Thompson W F. 2006. A modified protocol for rapid DNA isolation from plant tissues using cetyltrimethylammonium bromide. *Nat. Protoc.*, 1 (5), pp. 2320–2325.
- [14] Chen X, Temnykh S, Xu Y, Cho Y G, and McCouch SR. 1997. Development of a microsatellite framework map providing genome-wide coverage in rice (*Oryza sativa* L.). *Theor Appl Genet*, 95, 553-567.
- [15] Neeraja C N, Rodriguez R M, Pamplona A, Heuer S, Collard B C Y, Septiningsih E M, Vergara G, Sanchez D, Xu K, Ismail A M and Mackill D J. 2007. A marker-assisted backcross approach for developing submergence-tolerant rice cultivars. *Theor Appl Genet*, 115 (6): 767-776.
- [16] Gasteiger E., Hoogland C., Gattiker A., Duvaud S., Wilkins M. R., Appel R. D., Bairoch A.; Protein Identification and Analysis Tools on the ExpASY Server.
- [17] John M. Walker (ed): The Proteomics Protocols Handbook, Humana Press (2005). pp. 571-607.
- [18] Xu K, Xia X, Fukao T, Canlas P, Maghirang-Rodriguez R, Heuer S, Ismail AM, Bailey-Serres J, Ronald P C and Mackill D J. 2006. Sub1A Is an Ethylene Response Factor-Like Gene That Confers Submergence Tolerance to Rice. *Nature*, 442, 705-708.

- [19] Y. H. Li, J. Y. Xu, L. Tao, X. F. Li, S. Li, X. Zeng, S. Y. Chen, P. Zhang, C. Qin, C. Zhang, Z. Chen, F. Zhu, Y. Z. Chen. (2016) SVM-Prot 2016: A Web-Server for Machine Learning Prediction of Protein Functional Families from Sequence Irrespective of Similarity. PLoS One. 11 (8): e0155290.
- [20] C. Z. Cai, L. Y. Han, Z. L. Ji, X. Chen, Y. Z. Chen. (2003) SVM-Prot: Web-Based Support Vector Machine Software for Functional Classification of a Protein from Its Primary Sequence. Nucleic Acids Res.. 31: 3692-3697.