Pak. j. life soc. Sci. (2019), 17(1): 7-16 E-ISSN: 2221-7630;P-ISSN: 1727-4915

Pakistan Journal of Life and Social Sciences
www.pjlss.edu.pk

RESEARCH ARTICLE

In silico Analysis Tools Prove Basmati 2008 (Indica Rice) to be More Meaningful for Crop Improvement as Compared to Nipponbare (Japonica Rice)

Irfan Safdar Durrani[1]*, Farwa Iqbal Butt[1], Asad Jan[1], Noreen Asim[1] and S.M. Saqlan Naqvi[2]
[1]Institute of Biotechnology and Genetic Engineering, Genomics and Bioinformatics Division, The University of Agriculture Peshawar, Pakistan
[2]Pir Mehar Ali Shah University of Arid Agriculture, Murree Road, Rawalpindi, Pakistan

ARTICLE INFO
Received: Feb 22, 2019
Accepted: May 30, 2019

Keywords
Germin like proteins
Promoter Rice
Regulatory elements

***Corresponding Author:**
isdurrani@aup.edu.pk

ABSTRACT

Present study was focused on comparative *in silico* analysis of *Os*GLP12-3 promoter form Japonica rice cultivar Nipponbare and Indica rice cultivar Basmati 2008. For the purpose, the promoter region of *Os*GLP12-3 gene from both Nipponbare and Basmati 2008 was amplified by PCR using Genomic DNA. The amplicon was purified, quantified and finally sequenced. Further *in silico* studies included expression analysis, pairwise alignment, multiple sequence alignment and phylogenetic analysis. The online databases "PlantCARE" and "Plantpan3" identified many unique and important regulatory elements (cis/trans, or transcription factor binding sites) in both promoter regions. Moreover, some elements were specific to Basmati 2008 only. Basmati 2008 had greater copy number of some general transcription factor binding sites like TATA box and CAAT box as compared to Nipponbare, reflecting it to be stronger as compared to Nipponbare in terms of response to different stress conditions. The study also revealed interestingly important functions in Basmati 2008 *Os*GLP12-3 promoter region, such as defense against pathogens and other biotic and abiotic stresses. Each element was found to exist with variable copy number and positions in *Os*GLP12-3 promoter of Basmati 2008, as compared to same promoter from Nipponbare. In comparison to Nipponbare (Japonica Group), Basmati 2008 (Indica Group) *Os*GLP12-3 promoter was found to contain some unique regulatory elements that were altogether absent in Nipponbare. This is the first report which shows that *Os*GLP12-3 promoter region form Indica group of rice (Basmati 2008) contains abscisic acid response element (ABRE or G box) that is indirectly involved in defense related gene expression and LTR or H Box, which respond to cold stresses.

INTRODUCTION

Plants contain a diverse group of proteins namely germins and germin-like proteins (GLPs) that are highly resistant to proteases activity and agents like heat, sodium dodecyl sulphate and extreme pH etc. due to presence of characteristic and highly stable structure (Dunwell et al., 2008). This structure comprises six β-jellyroll monomers locked together to form a trimer of dimers, that is common both in germins and GLPs and is highly stable (Membre et al., 2000; Wei et al., 1998; Lane, 1994; Lane et al., 1993). Germins have oxalate oxidase activity while some GLPs have superoxide dismutase (SOD) or phosphodiesterase activities. The expression of germins and GLPs is stimulated by different stimuli at initiation and / or during development in different plant organs (Becerra et al., 2006: Mathieu et al., 2006; Staiger et al., 1999) or under biotic or abiotic stresses like herbivore attack, (Lou and Baldwin, 2006) and infections caused by pathogens (Schweizer et al., 1999; Park et al., 2004, Zimmermann et al., 2006; Christensen et al., 2004) or salt exposure (Hurkman and Tanaka, 1996; Nakata et al., 2002). Studies have also been carried out regarding the role of Germins and GLPs in plant defense (Banerjee and Maiti, 2010). Some events like defense-

related transcripts and increased resistance to pathogens was observed by over expression of wheat Germin in sunflowers (Hu et al., 2003). In fact, the regulation of a gene' expression of any gene is controlled by involvement of promoters and their analyses give detailed information about gene's function. As gene's regulation is governed by regulatory elements present in their promoters, that control the transcription of related gene, therefore, their identification and analysis are the field of main interest for crop biotechnologists. Developmental and environmental adaptation stages of an organism are better understood by having accurate knowledge about differential gene expression, which in large part is governed by *cis*- elements the binding sites for transcription factors. In order to understand the development of stress/disease resistant crops by using genetic modification technology, the promoter analysis is an important step as it explains functioning, regulation and response of the particular gene at given circumstances and conditions (stresses and environmental). High level of gene expression induced by biotic and abiotic stresses is directed by plant promoters that are critical to study for crop biotechnologists. Furthermore, tissue specific expression in plants can be achieved by using genetically engineered promoters. Understanding functioning of regulatory elements in promoter region is an efficient approach to have better idea about the expression of genes (Ilyas et al., 2016). Binding of transcription factors to their corresponding binding sites (cis acting elements) in the promoters are crucial for studying gene functioning at given situations. In order to develop stress/disease resistant crops by using novel promoter engineering technology, the promoter analysis is an important step as it explains functioning, regulation and response of the particular gene at given stress situation. The NCBI, Geo Profiles data of *Os*GLP12-3 gene reflects variable induction of this gene by many biotic and abiotic stimuli. Keeping in view, this study was planned to understand regulatory aspect of this gene. The main rationale of this investigation was, identification of regulatory elements in *Os*GLP12-3 gene's promoter from Basmati 2008, an Indica rice cultivar that is grown under submerged conditions, and their comparison with the same/ variable elements in the same promoter from Nipponbare, that is Japonica rice cultivar grown on dry lands.

MATERIALS AND METHODS

Plant material and DNA extraction

Rice seeds (Basmati 2008) were obtained from National Agriculture Research Council (NARC) Islamabad, Pakistan. Healthy and mature seeds were de-husked manually avoiding embryos from damage. Seeds were subjected to surface sterilization by treatment with $HgCl_2$ (0.2% w/v) for five minutes and subsequently with absolute ethanol and 75 % for three and two minutes respectively. Every treatment step mediated washing with sterile distilled water for two minutes. Finally, the surface sterilized seeds were washed five times thoroughly to make sure the elimination of any residual $HgCl_2$. The sterile seed were dried with sterile filter paper and were germinated hydroponically in Murashige and Skoog medium lacking sucrose, in growth room under illuminated conditions at 25 ^oC. Young leaves from plants were harvested for DNA extraction following the methodology described by (Richards et al., 1994) with minor modifications.

Oligo design

Pair of sense and antisense oligos (Table 1) was manually designed on Chromosome 12 of *Oryza sativa* cv Nipponbare (Japonica group), to amplify the upstream regulatory sequence (Promoter) of *OsGLP12-3* gene. Oligos so designed were run on NCBI Basic local alignment tool (BLAST) database to confirm their specificity with DNA template. The properties of primers were checked using Integrated DNA Technology (IDT) Oligo analyzer available at https://eu.idtdna.com/calc/analyzer (February 27,2019).

Table 1: List of oligos used for the amplification of *Os*GLP12-3 Promoter

Oligo	Sequence
Sense Oligo	5'- GGCTTAAACATTAGAAATTT 3'
Antisense Oligo	5'- CATGTTAAGTTGATGGAACTTTG-3'

PCR amplification of *Os*GLP12-3 promoter

The genomic DNA was used as template to perform (PCR) for Indica Rice group cultivar "Basmati 2008", using 1X PCR buffer, 1.5 mM $MgSO_4$, 0.25mM each dNTPs, 0.125 mM each oligo and 1.5 Units of Hot Start prime star polymerase from Takara Bio Inc. Japan. The reaction constituents were cycled 40 times with a pre-amplification denaturation at 96^oC for 2 minutes followed by; denaturation at 96^oC for 20 seconds, annealing at 55^oC for 5 seconds and extension at 72^oC for 2 minutes.

Tris BORATE EDTA (TBE) Agarose Gel Electrophoresis

The amplified products were resolved on TBE Agarose gel (1.25%) at 110V for 40 minutes. The gel was visualized by Ethidium bromide UV trans illumination using UV tech, United Kingdom "Gel Documentation system". The "Thermo Scientific Gene Ruler" 1kb plus DNA ladder was used as marker for identification of desired amplicon.

Purification and sequencing of amplicon

The desired band of amplicon was excised using scalpel from agarose gel and further processed for

purification purpose by using DNA Gel Extraction Kit (Monarch® New England Biolabs, UK) as per manufacturer instructions. The purified samples were subjected to sequencing by acquiring services from Macrogen Incorporation, Korea.

Pairwise sequence alignment

For pairwise alignment NCBI Nucleotide BLAST tool was used. The *OsGLP12-3* promoter sequence of Indica rice cultivar Basmati 2008, obtained from Macrogen Inc, was used as query sequence, and subjected to pairwise alignment with *Os*GLP12-3 promoter sequence from Japonica rice cultivar Nipponbare that was retrieved from NCBI public database, as subject sequence. The pairwise alignment was performed to check degree of similarity between Nipponbare subject sequence that is available on public database, and newly sequenced *Os*GLP12-3 promoter region from Basmati 2008.

Multiple sequence alignment and phylogenetic analysis

Multiple sequence alignment for *Os*GLP-12-3 was performed using NCBI Protein BLAST database and Mega 7 program. The Protein sequences, that were homologous to *Os*GLP12-3 protein and other Germin Like Proteins, from both monocots and dicots, were retrieved in FASTA format from NCBI Protein database using Blast tool. The retrieved sequences were subjected to multiple sequence alignment (MSA), by importing them in to MEGA 7 program. Multiple sequence alignment was performed opting MUSCLE. The MEGA 7 program was also exploited for the construction of phylogenetic tree to determine evolutionary relationship between *Os*GLP12-3 and other Germin Like Proteins (GLPs) from other plants.

Expression analysis

For expression analysis "The Encyclopedia of Rice" database (Kawahara et al 2016), available at (http://tenor.dna.affrc.go.jp/), was used to predict / understand gene's regulatory aspect, in terms of gene expression, that is governed by cis acting elements (Transcription factor binding sites) in its promoter. The data available, with context to expression of gene, governed by its promoter, during development and in response to biotic and abiotic stresses, was retrieved from TENOR database. For data retrieval from TENOR database different parameters; including development, different abiotic stress conditions, along with time taken and time span, for the expression of *Os*GLP12-3 gene, was used.

Identification and Mapping of Regulatory Elements:

Different cis acting regulatory elements were searched in the OsGLP12-3 rice gene promoter region from Indica Rice C.V Basmati 2008 and Japonica Rice C.V Nipponbare by using online program PlantCARE database (Lescot et al., 2002) available at, http://bioinformatics.psb.ugent.be/webtools/plantcare/html/). Map of important transcription factor binding sites and corresponding transcription factors was constructed using PlantPAN 3.0 database, (Chang et al., 2008), available at *plantpan.itps.ncku.edu.tw).*

RESULTS AND DISCUSSION

PCR Amplification of OS-GLP 12-3 Promoter Region from Basmati 2008

The *Os*GLP12-3 promoter region; 2007 bp was PCR amplified using Prime Star Polymerase enzyme from Takara Bio Inc. and resolved on TBE agarose gel and finally subject to gel documentation analysis, using ethidium bromide UV transillumination as shown in Fig. 1.

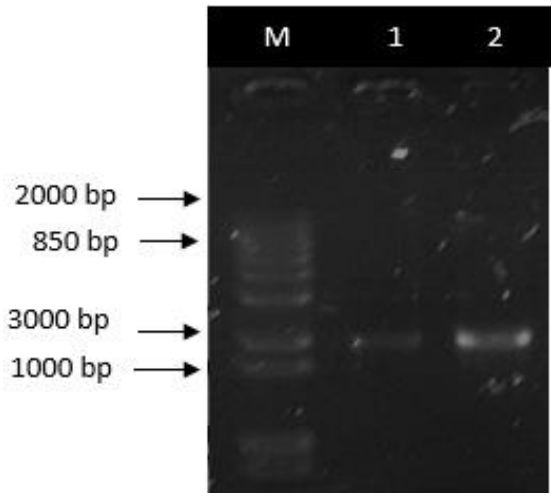

Fig. 1: PCR amplification of *Os*GLP12-3 Promoter Region from Basmati 2008 and Nipponbare

Lane M represents 1KB+ Marker; Lanes 1 and 2 represent amplified Os-GLP12-3 promoter region from Nipponbare and Basmati 2008 respectively.

Pairwise alignment

Pairwise alignment of subject promoter sequence of *Os*GLP12-3 gene from Nipponbare, with a corresponding query sequences from Indica rice cultivar Basmati 2008 was performed. Variable degrees of similarity were found between query and subject sequences in three different ranges. Range 1 spanning 300 nucleotides; 31-331 for query sequence and 57785 to 58085 for subject sequence showed 99 % similarity, while range 2 spanning 292 nucleotides for query sequence and 287 nucleotides for subject sequence i.e. from base number 31-323 for query sequence and base number 61470 to 61757 for subject sequence showed 93 % similarity respectively. Range 3 spanning 259 nucleotides for query sequence and 259 nucleotides for subject sequence i.e. 31-323 for query sequence and 61470 to 61757 for subject sequence showed 77 % similarity, respectively Figs; 2,3,4.

Range 1:
Alignment Statistics for match #1

Score	Expect	Identities	Gaps	Strand
540 bits (292)	4e-155	298/301 (99%)	0/301(0%)	Plus/Plus

```
Query  31     TGTATAGCTATAGTTGATGCTTTGCCCATGGAAATGATGCATCCATATATATAGGCCTCT  90
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  57785  TGTATAGCTATAGTTGATGCTTTGCCCATGGAAATGATGCATCCATATATATAGGCCTCT
57844

Query  91     GAGCATGAGATGATTTTGAAGTGGTCAATAACAAAAGGCTGAGCTAGCTTGTGACTGGTT
150
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  57845  GAGCATGAGATGATTTTGAAGTGGTCAATAACAAAAGGCTGAGCTAGCTTGTGACTGGTT
57904

Query  151    CATCTCCTCTAAAGTCTGAACTGATCTTGAACTCTTAATAGATTGCATCTGGGAAAGAAA
210
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  57905  CATCTCCTCTAAAGTCTGAACTGATCTTGAACTCTTAATAGATTGCATCTGGGAAAGAAA
57964

Query  211    CTATTTAATTTGATGGATTCAATCTATCTTAATTCCCAGTCAACAGCTAGGTTGGCCGAC
270
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  57965  CTATTTAATTTGATGGATTCAATCTATCTTAATTCCCAGTCAACAGCTAGGTTGGCCGAC
58024

Query  271    TAATTAATTAATTAGATTATTCCAGTGCCTGTCATGCATGATGCATGTATTAGGTGTATT
330
              |||||||||||||||||||||||||||||||||||||||||||||| ||||| | ||||
Sbjct  58025  TAATTAATTAATTAGATTATTCCAGTGCCTGTCATGCATGATGCATGCATTAGATTTATT
58084

Query  331    A  331
              |
Sbjct  58085  A  58085
```

Fig. 2: Pairwise alignment showing 99 % similarity between Nipponbare and Basmati 2008 *Os*GLP12-3 promoter sequences

Alignment Statistics for match #2

Score	Expect	Identities	Gaps	Strand
414 bits(224)	3e-117	272/294 (93%)	7/294(2%)	Plus/Plus

```
Query  31     TGTATAGCTATAGTTGATGC-TTTGCCCATGGAAATGATGCATCCATATATATAGGCCTC  89
              ||||||||| ||||  |||| |||| ||||||  |  | ||||||||||||||||||||
Sbjct  61470  TGTATAGCTGTAGTCTATGCTTTTGTCCATGG--A--A-GCATCCATATATATAGGCCTC
61524

Query  90     TGAGCATGAGATGATTTTGAAGTGGTCAATAACAAAAGGCTGAGCTAGCTTGTGACTGGT
149
              ||||||||| |||||| |||||||||||||| |||||| |||||||||| |||||| ||
Sbjct  61525  TGAGCATGAAATGATTCTGAAGTGGTCAATAAGAAAAGGATGAGCTAGCTAGTGACTTGT
61584

Query  150    TCATCTCCTCTAAAGTCTGAACTGATCTTGAACTCTTAATAGATTGCATCTGGGAAAGAA
209
              ||||||||||  |||||||||||| |||||||||||||| ||||||||||||||||||||
Sbjct  61585  TCATCTCCTCAGAAGTCTGAACTGGTCTTGAACTCTTAAT-GATTGCATCTGGGAAAGAA
61643

Query  210    ACTATTTAATTTGATGGATTCAATCTATCTTAATTCCCAGTCAACAGCTAGGTTGGCCGA
269
              |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Sbjct  61644  ACTATTTAATTTGATGGATTCAAGCTATCTTAATTCCCAGTCAACAGCTAGGTTGGCCGA
61703

Query  270    CTAATTAATTAATTAGATTATTCCAGTGCCTGTCATGCATGATGCATGTATTAG  323
              |||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Sbjct  61704  CTAATTAATTAATTAGATTATTCCAGTGCCTGTCATGCATGATGCATGCATTAG  61757
```

Fig. 3: Pairwise alignment showing 93% similarity between Nipponbare and Basmati 2008 *Os*GLP12-3 promoter sequences

Range 3:
Alignment Statistics for match #3

Score	Expect	Identities	Gaps	Strand
137 bits(74)	**6e-34**	**213/275(77%)**	**30/275(10%)**	**Plus/Plus**

```
Query  51     TTTGCCCATGGAAATGATGCATCCATATATATAGGCCTCTGAGCATGAGATGATTTTGAA
110
              |||| ||| ||||| ||||||||||||||||||||||||||||| ||||| ||||| ||||
Sbjct  65278  TTTGGCCA-GGAAAGGATGCATCCATATATATAGGCCTCTGAGAATGAGGTGATTCTGAA
65336

Query  111    GTGGTCAATAACAAAAGGCTGA---G-CT-AGCTTGTGACTGGTTCATCTCCTCTAAAGT
165
              |||||||||||| ||||     |   | || | ||   |||| |||||||| |||||||||
Sbjct  65337  GTGGTCAATAAGAAAAATAAAAAAAGGCTGAACTAACGACTTGTTCATCTGCTCTAAAGT
65396

Query  166    CTGAACTGATCTTGAACTCTTAATAGA-TTGCATCTGGGAAAGAAACTATTTAATTTGAT
224
              |||||| | ||||||  |   |   || |||||||||||| |||||  ||||||||||||
Sbjct  65397  CTGAACCGGTCTTGA--T---A---GATTTGCATCTGGGCAAGAACTTATTTAATTTGAT
65448

Query  225    GGATTCAATCTATC--TT-AA-TT--C--CC-AGTCAACAGCTAGGTTGGCCGACTAATT
275
              ||||| | |  ||   || || ||  |  || || ||||| |  || || |||| || ||
Sbjct  65449  GGATT-AGTGGATGGATTCAAGTTGTCTTCCCAG-CAACA-C--GG-TGACCGATTAGTT
65502

Query  276    AATTAATTAGATTATTCCAGTGCCTGTCATGCATG  310
              ||| ||||||||||||  ||||||||||||||||
Sbjct  65503  AATCAATTAGATTATTTTAGTGCCTGTCATGCATG  65537
```

Fig. 4: Pairwise alignment showing 77 % similarity between Nipponbare and Basmati 2008 *Os*GLP12-3 promoter sequences

Pairwise alignment of subject promoter sequence of *Os*GLP12-3 gene from Nipponbare, with a corresponding query sequences from Indica rice cultivar Basmati 2008 was performed. Variable degrees of similarity were found between query and subject sequences in three different ranges. Range 1 spanning 300 nucleotides; 31-331 for query sequence and 57785 to 58085 for subject sequence showed 99 % similarity, while range 2 spanning 292 nucleotides for query sequence and 287 nucleotides for subject sequence i.e. from base number 31-323 for query sequence and base number 61470 to 61757 for subject sequence showed 93 % similarity respectively. Range 3 spanning 259 nucleotides for query sequence and 259 nucleotides for subject sequence i.e. 31-323 for query sequence and 61470 to 61757 for subject sequence showed 77% similarity respectively. Figs; 2,3,4.

The two promoter sequences didn't align in full stretch of 1272 base pairs. This is because indica rice cultivars and japonica rice cultivars, though have a common ancestor belong to two different subspecies (Du et al., 2017).

Multiple sequence alignment and phylogenetic analysis

Multiple Sequence alignment and phylogenetic analysis was done for finding evolutionary relationship between *Os*GLP12-3 protein with other Germin like proteins from other plants by constructing phylogenetic tree. Multiple sequence alignment and phylogenetic analysis resulted in the formation of two main groups. Group 1 contained sequences from monocot plants i.e. *Sorghum bicolor*, *Setaria italica*, *Zea mays*, *Oryza sativa* Japonica group and *Panicum hallii* (Hall's panicgrass) while Group 2 contained sequences from *Dichanthelium oligosanthes*, *Triticum urartu*, *Hordeum vulgare* subsp. *vulgare*, *Aegilops tauschii* sub. *tauschii* and *Triticum aestivum* which also belong to monocot group of plants. Group 1 was further divided into two subgroups which in turn were further divided into two sub-groups. Group two was also divided into two sub-groups. But unlike subgroup 1, subgroup 2 was found to have a cluster of a branch of GLP 12-2 belonging to *Dichanthelium oligosanthes* and a cluster of further subgroups containing sequences of GLPs from other plants. However, *Os*GLP12-3 was clustered in Subgroup 1B which contained another Germin Like Protein from *Panicum Hallii* (Hall's panicgrass: PHAL F01217) among the selected proteins as shown in Fig. 5.

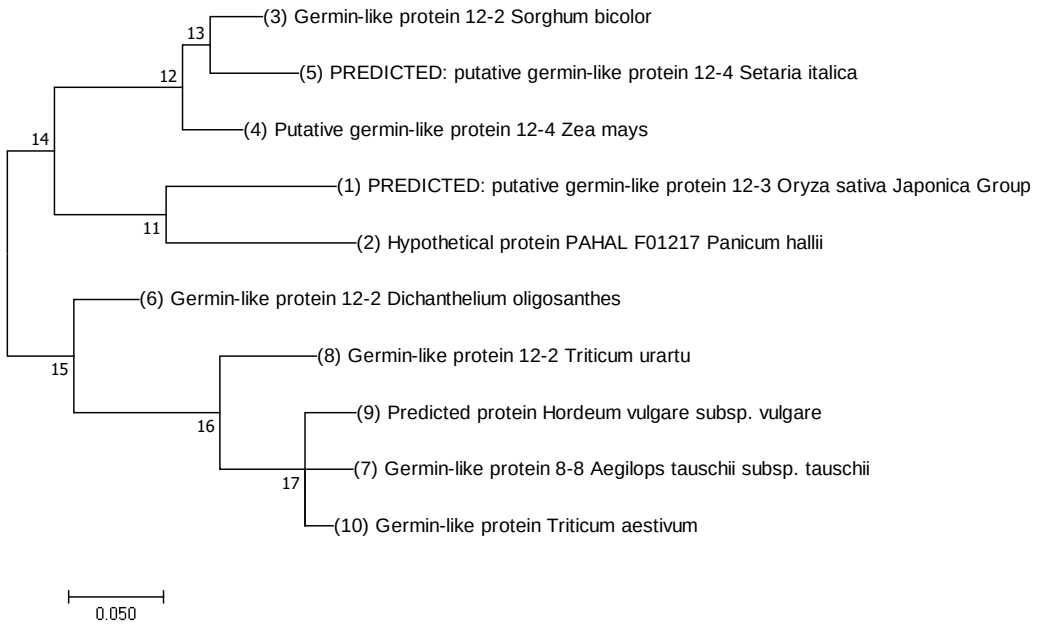

Fig. 5: Phylogenetic tree showing evolutionary relationship between *Os*GLP12-3 and other germin like proteins from different plants.

Lu et al. (2010) work was focused on Identification and analysis of the Germin-like gene family in soybean. They suggested that the Germin and Germin-like genes of the plant species might be evolved by independent gene duplication events. This might be the reason that our phylogenetic analysis results show close relationship of *Os*GLP12-3 with GLP from Hali's grass and variability with other GLPs. In addition, *Os*GLP12-3 gene can be seen located on chromosome 12 and duplication of same gene has evolved another gene located on chromosome 8 in *Aegiolps tauchii* subsp. tauchii, another grass from Poaceae family (Fig.5).

Expression analysis of OsGLP12-3 gene in rice

In silico analysis for expression data of *Os*GLP12-3 gene was retrieved from TENOR database. Parameters for data retrieval included development and abiotic stress conditions including hormones, along with time taken for the expression of *OsGLP12-3* gene.

Expression during development

The *Os*GLP12-3 can be observed to be involved in plant development majorly roots. Highest level of expression was observed in first three hours after germination and gene was found to be continuously expressed after commencement of first expression. These results are in agreement with previous findings by (Yasmin, 2009), who observed involvement of *Os*GLP-1 gene in root elongation, when it was transgenicaly overexpressed in tobacco.

Expression in response to abiotic stress conditions and hormonal stimuli

Data retrieved, regarding response of *Os*GLP12-3 gene against abiotic stress conditions included, high salinity, Phosphorus and Cadmium both at high and low concentrations, drought (dry), flood, cold treatment, osmotic stresses and hormone reveal induction of *Os*GLP-12-3 gene by all abiotic factors including hormones i.e. abscisic acid (Lescot et al., 2002; Yasmin, 2009; Weirauch et al., 2014; Kawahara et al., 2016) and Jasmonic acid (Lane et al., 1993; Blackwood and Kadonaga, 1998; Nakata et al., 2002; Park et al., 2004; Weirauch et al., 2014; Allen and Taatjes 2015). Two types of tissues including shoots and roots were targeted for analysis.

As shown, highest level of expression in response to high Cadmium stress five days after plant's first exposure to Cd stress in roots. However, the expression commenced after just one hour after exposure to high Cd and continued at moderate levels throughout plant development. The expression of *Os*GLP-12-3 gene can also be seen in shoots (Fig. 6) in response to high cadmium stress. The gene expression can be observed to be initiated minimally twelve hours after first exposure to High Cd stress and expression continues exponentially for five days only (Fig. 6).

In case of other abiotic stresses including dry condition, flood, osmotic stress (flood induced), the expression level of *Os*GLP-12-3 can be observed at moderate levels in case of osmotic and dry stresses, moreover, lower moderate level expression can be observed for flood and low Cd stresses. It is here to emphasize that for Cd stress at higher concentrations, that the gene responds majorly in roots, and, very minimally in shoots, similar behavior of gene can be observed at low Cd concentrations. In case of flood stress much lower

moderate expression of gene is depicted in Fig 6. In all cases, it is quite obvious that the *Os*GLP12-3 gene expresses in response to various abiotic stresses, though variably, but majorly in roots.

Concerning salinity stress, *Os*GLP-12-3 gene activity was also observed in high salinity stress both in roots and shoots. Interestingly in both plant organs the expression trend can be seen similar, however, the expression level is two folds in roots as compared to shoots. Response of *Os*GLP-12-3 gene to hormones in shape of its expression, including Abscisic acid and Jasmonic acid, is also positive in both roots and shoots, but continuous expression in response to both hormones is in roots only.

Looking into *In silico* findings as stated above it becomes quite obvious that *Os*GLP12-3 gene expresses continuously and majorly in roots during development and abiotic stress conditions including high salinity, high and low Phosphorus, High cadmium, dry, flood, cold and osmotic stresses.

The induction of *Os*GLP-12-3 gene expression in response to hormones including ABA and MeJa (methyl Jasmonate) is also continuous in roots. The induction of *Os*GLP-12-3 gene expression in response to hormones showed no continuous expression of gene. This indicates that *Os*GLP12-3 gene expresses majorly in roots and minimally in shoots Fig 6. Lu et al. (2010) from their study report that indole acetic acid (IAA) stimulates expression of the GmGER genes from soybean. This phenomenon reflects that expression of Germin like protein genes is affected by hormones.

Many transcription factors binding sites / cis and trans acting elements, for both general and specific binding sites were found, from PlantCARE database, spanning 1272 bp promoter region of *Os*GLP12-3 gene from Nipponbare and Basmati 2008. Only, the elements with known functions were addressed.

Interestingly, some elements were exclusively found in Basmati 2008 *Os*GLP12-3 promoter, and altogether absent in same promoter from Nipponbare, whose whole genome has been sequenced. The unique elements include Abscisic acid response element (ABRE), G-box and low temperature response element (LTR or H-Box).

The ABRE, Abscisic acid responsive element, is the binding site of TGA1 and NPR1 transcription factors that in turn is responsible for increased the binding of TGA2 to the ABRE element. NPR1 is essential in activating systemic, inducible plant defense responses (Despres et al., 2000).

G Box in fact is a unique Abscisic acid responsive element found only in basmati 2008. This was initially found in Arabidopsis dehydration-responsive gene rd22, that is a dehydration responsive gene, and its promoter itself doesn't contain Abscisic acid (ABA) responsive element however, its activity is mediated by ABA. Iwasaki et al. (1995) reported that the rd22 promoter, that was indeed non-responsive to ABA,

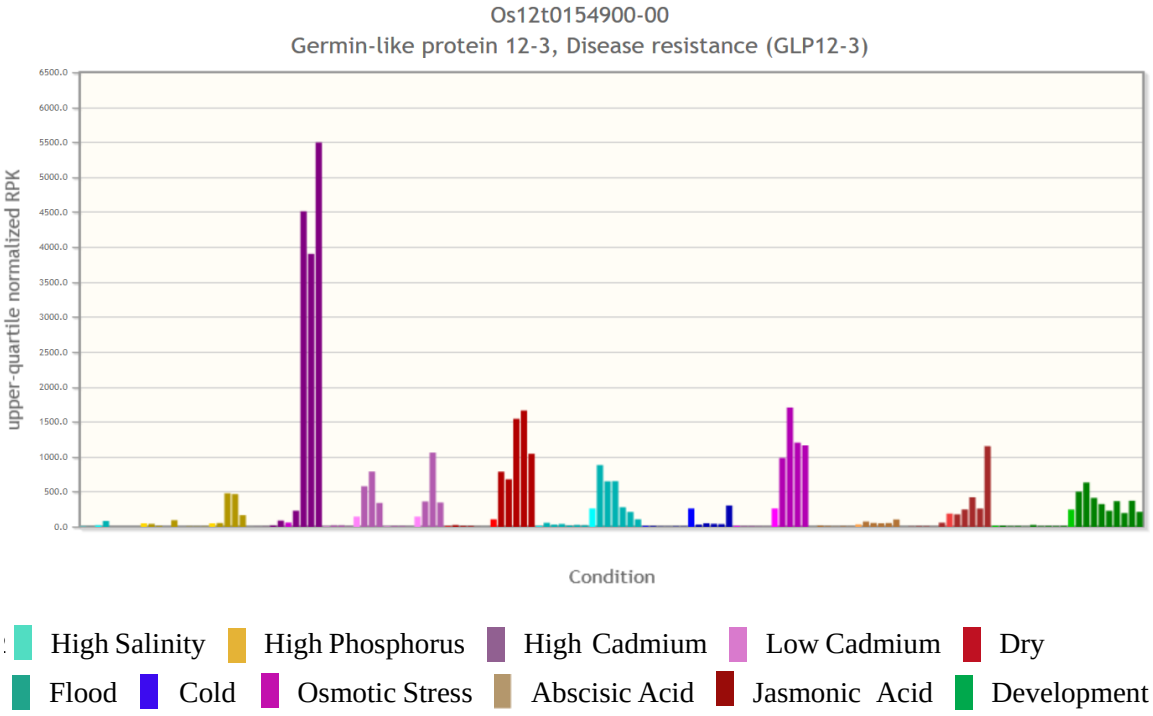

Fig. 6: Expression Profile of *OsGLP12-3* Gene during Development and Induction by Hormones and Different Stress Conditions.

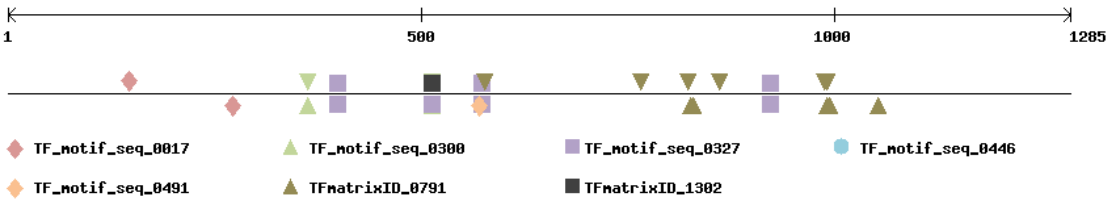

Fig. 7: Arrangement and localization of selected cis/ trans acting elements in *Os*GLP12-3 promoter region of basmati-2008. Legend: Upper Region of Line Represents Plus Strand; Lower Region of Line Represents Minus Strand

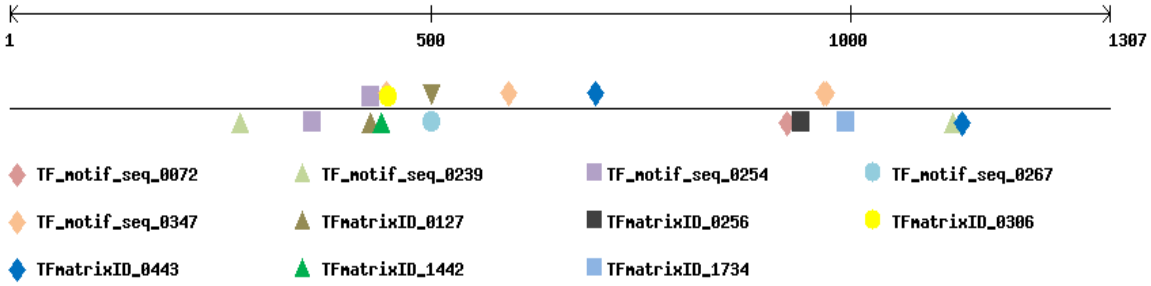

Fig. 8: Arrangement and localization of selected cis/ trans acting elements in *Os*GLP12-3 promoter region of nippponbare. Legend: Upper Region of Line Represents Plus Strand; Lower Region of Line Represents Minus Strand.

engineered with 67 bp fragment containing G- Box conferred responsiveness to dehydration and ABA on a previously non-responsive promoter for these factors. According to them the 67-bp fragment contains the sequences of the recognition sites for some transcription factors, such as MYC, MYB, and GT-1 and accumulation of rd22 mRNA requires protein synthesis raises the possibility that the expression of rd22 might be regulated by one of these trans-acting protein factors whose de novo synthesis is induced by dehydration or Abscisic acid (Lescot et al., 2002, Yasmin, 2009, Weirauch et al., 2014, Kawahara et al., 2016).

Consensus sequence of H-boxes was found in *Phaseolus vulgaris* on chrosome15 gene promoter that is essential for both light regulation and elicitor induction. Similar sequence was found in tobacco Tnt1 retrotransposon promoter (LTR). Tnt1 is induced by wounding and by abiotic stress. KAP-2 is reported to bind to the H-box and stimulate the transcription of promoters having the H-box (Loake et al., 1992, Mhiri et al., 1997, Lindsay et al., 2002).

The CAAT Box element sequence is found in promoters and enhancers of most genes that act as binding site of corresponding transcription factors. In case of promoters it mediates initiation of transcription, whereas enhancer is a short region of DNA that are binding site of activator proteins that increase the likelihood, that transcription of a particular gene will occur (Blackwood and Kadonaga, 1998; Pennacchio et al., 2013). These proteins (activator proteins) are usually referred to as transcription factors. Enhancers are cis-acting elements, located up to 1 Mega bp (1,000,000 bp) away from the gene, upstream or downstream from the start site (Pennacchio et al., 2013; Maston et al., 2006; Smemo et al., 2014). The enhancer was first discovered in eukaryotes in the immunoglobulin heavy chain gene in 1983 (Banerji et al., 1983; Gillies et al., 1983). This enhancer element was reported to be involved in transcriptional activation of rearranged Vh gene promoters while unrearranged Vh promoters remained inactive.

Basmati 2008 (*Indica* cultivar) was found to possess greater number of copies CAAT box, that is a transcription rate enhancer element, as compared to Nipponbare (*Japonica* cultivar)

. By this we can demonstrate that owing to greater copy number of CAAT box, *Os*GLP12-3 promoter in Basmati 2008 has got greater potential of enhancing transcription rate of related gene/genes, in response to stress conditions as compared to Nipponbare.

Two copies of another cis acting element “W-box”, WRKY63 and WRKY40 (Weirauch et al., 2014) were identified in Basmati 2008 and Nipponbare as well, which is the binding site of WRKY transcription factors. The expression of WRKY class of transcription factors is triggered in response to wound or fugal attack via gibberellin signaling and found in gene promoter in Parsley (Eulgem et al, 1999). Many groups report involvement of W Box class and corresponding WRKY transcription factors directly or indirectly in plant defense in case of wound response (Eulgem et al., 1999; Knoth et al., 2007) or in response to invasion attempt by bacterial or fungal pathogens (Banerjee et al., 2010). This reflects that *Os*GLP12-3 promoter

having W box, LTR or H box, unique G-box and multiple copies of CAAT box in Basmati 2008 (Indica cultivar) OsGLP12-3 promoter has a potential for developing crops that are resistant to pathogens, cold, and other abiotic stress conditions, provided the copy numbers of respective elements in promoter regions is increased or they may be genetically exploited some other way using recombinant DNA technology.

In conclusion, the *Os*GLP12-3 gene's promoter from indica rice cultivar Basmati 2008 has got greater potential for making disease and stress resistant crops as compared to same promoter from Nipponbare.

Authors' contribution

All authors contributed equally to this work

Acknowledgements

I cannot forget to acknowledge Dr. S. M. S. Naqvi for equipping me with knowledge and research skills.

REFERENCES

Allen BL and Taatjes DJ, 2015. The Mediator complex: a central integrator of transcription. Nature Reviews Molecular Cell Biology, 16: 155-166.

Banerjee J, N Das, P Dey and MK Maiti, 2010. Transgenically expressed rice germin-like protein1 in tobacco causes hyper-accumulation of H_2O_2 and reinforcement of the cell wall components. Biochemical and Biophysical Research Commmunications, 402: 637-643.

Banerjee J and MK Maiti, 2010. Functional role of rice germin-like protein1 in regulation of plant height and disease resistance. Biochemical and Bio-physcal Research Communtions, 394: 178-183.

Banerji J, L Olson and WJC Schaffner, 1983. A lymphocyte-specific cellular enhancer is located downstream of the joining region in immuno-globulin heavy chain genes. Cell, 33: 729-740.

Becerra C, P Puigdomenech and CM Vicient, 2006. Computational and experimental analysis identifies Arabidopsis genes specifically expressed during early seed development. BMC Genomics, 7: 38.

Blackwood EM and JT Kadonaga, 1998. Going the distance: a current view of enhancer action. Science, 281: 60-63.

Chang WC, TY Lee, HD Huang, HY Huang and RL Pan, 2008. PlantPAN: Plant promoter analysis navigator, for identifying combinatorial cis-regulatory elements with distance constraint in plant gene groups. BMC Genomics. 9: 561.

Christensen AB, HT Christensen, G Zimmermann, T Gjetting, MF Lyngkjaer, R Dudler and P Schweizer, 2004. The germin like protein GLP4 exhibits superoxide dismutase activity and is an important component of quantitative resistance in wheat and barley. Molecular Plant-Microbe Interactions, 17: 109-117.

Despres C, C DeLong, S Glaze, E Liu and PR Fobert, 2000. The Arabidopsis NPR1/NIM1 protein enhances the DNA binding activity of a sub-group of the TGA family of bZIP transcription factors. The Plant Cell, 12: 279-290.

Du H, Y Yu, Y Ma, Q Gao, Y Cao, Z Chen, B Ma, M Qi, Y Li and X Zhao, 2017. Sequencing and de novo assembly of a near complete indica rice genome. Nature Communications, 8: 15324.

Dunwell JM, JG Gibbings, T Mahmood and SS Naqvi, 2008. Germin and germin-like proteins: evolution, structure, and function, Critical Reviews in Plant Sciences, 27: 342-375.

Eulgem T, PJ Rushton, E Schmelzer, K Hahlbrock and IE Somssich, 1999. Early nuclear events in plant defence signalling: rapid gene activation by WRKY transcription factors, The European Molecular Biology Oranization Journal, 18: 4689-4699.

Gillies SD, SL Morrison, VT Oi and SA Tonegawa, 1983. A tissue-specific transcription enhancer element is located in the major intron of a rearranged immunoglobulin heavy chain gene. Cell, 33: 717-728.

Hu X, DL Bidney, N Yalpani, JP Duvick, O Crasta, O Folkerts and G Lu, 2003. Overexpression of a gene encoding hydrogen peroxide-generating oxalate oxidase evokes defense responses in sunflower. Plant Physiology, 133: 170-181.

Hurkman WJ and CK Tanaka, 1996. Germin gene expression is induced in wheat leaves by powdery mildew infection. Plant Physiology, 111: 3: 735-739.

Ilyas M, SMS Naqvi and T Mahmood, 2016. *In silico* analysis of transcription factor binding sites in promoters of germin-like protein genes in rice. Archives of Biological Sciences, 68: 863-876.

Iwasaki T, KY Shinozaki and KJM Shinozaki, 1995. Identification of a cis-regulatory region of a gene in *Arabidopsis thaliana* whose induction by dehydration is mediated by abscisic acid and requires protein synthesis. Molecular and General Genetics, 247: 391-398.

Kawahara Y, Y Oono, H Wakimoto, J Ogata, H Kanamori, H Sasaki, S Mori, T Matsumoto and T Itoh, 2016. TENOR: database for comprehensive mRNA-Seq experiments in rice. Plant Cell Physiology, 57: e7.

Knoth C, J Ringler, JL Dangl and T Eulgem, 2007. Arabidopsis WRKY70 is required for full RPP4-mediated disease resistance and basal defense against *Hyaloperonospora parasitica*. Molecular Plant-Microbe Interactions, 20: 120-128.

Lane BG, JM Dunwell, J Ray, M Schmitt and A Cuming, 1993. Germin, a protein marker of

early plant development, is an oxalate oxidase. Journal of Biological Chemistry, 268: 12239-12242.

Lane BG, 1994. Oxalate, germin, and the extracellular matrix of higher plants. The Federation of American Societies for Experimental Biology Journal, 8: 294-301.

Lescot M, P Dehais, G Thijs, K Marchal, Y Moreau, YV de Peer, P Rouze and S Rombauts, 2002. PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. Nucleic Acids Research, 30: 325-327.

Lindsay WP, FM McAlister, Q Zhu, XZ He, WD Laser, S Hedrick, P Doerner, C Lamb and RA Dixon, 2002. KAP-2, a protein that binds to the H-box in a bean chalcone synthase promoter, is a novel plant transcription factor with sequence identity to the large subunit of human Ku autoantigen. Plant Molecular Biology, 49: 503-514.

Loake GJ, O Faktor, CJ Lamb and RA Dixon, 1992. Combination of H-box [CCTACC (N) 7CT] and G-box (CACGTG) cis elements is necessary for feed-forward stimulation of a chalcone synthase promoter by the phenylpropanoid-pathway intermediate p-coumaric acid. Proceedings of the National Academy of Sciences, 89: 19: 9230-9234.

Lou Y and IT Baldwin, 2006. Silencing of a germin-like gene in *Nicotiana attenuata* improves performance of native herbivores. Plant Physiology, 140: 1126-1136.

Lu M, YP Han, JG Gao, XJ Wang and WB Li, 2010. Identification and analysis of the germin-like gene family in soybean. BMC Genomics, 11: 620.

Maston GA, SK Evans and MR Green, 2006. Transcriptional regulatory elements in the human genome. Annual Reviews Genomics Human Genetics, 7: 29-59.

Mathieu M, MAL Walter, AS Blervacq, H David, S Hawkins and G Neutelings, 2006. Germin-like genes are expressed during somatic embryogenesis and early development of conifers. Plant Molecular Biology, 61: 615-627.

Membre N, F Bernier, D Staiger and A Berna, 2000. *Arabidopsis thaliana* germin-like proteins: common and specific features point to a variety of functions. Planta, 211: 345-354.

Mhiri C, JB Morel, S Vernhettes, JM Casacuberta, H Lucas and MA Grandbastien, 1997. The promoter of the tobacco Tnt1 retrotransposon is induced by wounding and by abiotic stress. Plant Molecular Biology, 33: 257-266.

Nakata M, T Shiono, Y Watanabe and T Satoh, 2002. Salt stress-induced dissociation from cells of a germin-like protein with Mn-SOD activity and an increase in its mRNA in a moss, *Barbula unguiculata*. Plant Cell Physiology, 43: 1568-1574.

Park CJ, JM An, YC Shin, KJ Kim, BJ Lee and KH Paek, 2004. Molecular characterization of pepper germin-like protein as the novel PR-16 family of pathogenesis-related proteins isolated during the resistance response to viral and bacterial infection. Planta, 219: 797-806.

Pennacchio LA, W Bickmore, A Dean, MA Nobrega and G Bejerano, 2013. Enhancers: five essential questions. Nature Reviews Genetics, 14: 288.

Richards E, M Reichardt and S Rogers, 1994. Preparation of genomic DNA from plant tissue. Current Protocols in Molecular Biology, 27: 2.3.1-2.3.7.

Schweizer P, A Christoffel and R Dudler, 1999. Transient expression of members of the germin-like gene family in epidermal cells of wheat confers disease resistance. The Plant Journal, 20: 541-552.

Smemo S, JJ Tena, KH Kim, ER Gamazon, NJ Sakabe, C Gomez-Marin, I Aneas, FL Credidio, DR Sobreira and NF Wasserman, 2014. Obesity-associated variants within FTO form long-range functional connections with IRX3. Nature, 507: 371.

Staiger D, K Apel and G Trepp, 1999. The Atger3 promoter confers circadian clock-regulated transcription with peak expression at the beginning of the night. Plant Molecular Biology, 40: 873-882.

Wei Y, Z Zhang, CH Andersen, E Schmelzer, PL Gregersen, DB Collinge, V Smedegaard-Petersen and H Thordal-Christensen, 1998. An epidermis/papilla-specific oxalate oxidase-like protein in the defence response of barley attacked by the powdery mildew fungus. Plant Molecular Biology, 36: 101-112.

Weirauch, MT, A Yang, M Albu, AG Cote, A Montenegro-Montero, P Drewe, HS Najafabadi, SA Lambert, I Mann and K Cook, 2014. Determination and inference of eukaryotic transcription factor sequence specificity. Cell, 158: 1431-1443.

Yasmin T, 2009. Cloning and over-expression of a germin-like protein gene for its functional analysis. PhD Thesis Dissertation. PMAS Arid Agriculture University, Rawalpindi, Pakistan.

Zimmermann G, H Baumlein, HP Mock, A Himmelbach and P Schweizer, 2006. The multigene family encoding germin-like proteins of barley. Regulation and function in basal host resistance. Plant Physiology, 142: 181-192.