Identification of Natural Nucleotide Mutations of Salt Tolerance Candidate Genes (HvHKT1 and HvCBL4) in Barley Ecotypes by Sequencing

Adnan Al-Yassin, Raheleh Khademian, Mohsen MARDI, Seyed Mojtaba Khayam Nekouei and Babak Nakhoda

Abstract: This study was conducted to identify single nucleotide variation in salinity stress tolerance candidate genes of barley ecotypes and to search for new diversity resources. Therefore, 96 barley ecotypes, collected from different areas worldwide, were used and two salinity tolerance candidate genes (HvHKT1 and HvCBL4) were targeted through sequencing. Genomic DNA fragments of the ecotypes were amplified by eight pairs of primers of the genes (4 pairs for each gene) by PCR reaction. Total number of nucleotide variation in HvHKT1 was significantly higher than in HvCBL4. Results of comparing sequence of these two genes showed that 17.9 SNP/Kb, 1 indel per 199.3 bp for HvHKT1 and 137 SNP/Kb, 18 indel per bp for HvCBL4. SNP frequency was lower than the non-coding and non-conserved regions; also the SNP frequency in intron was more than exone. The two genes were compared and the results showed that SNP abundance in HvCBL4 was seven times more than HvHKT1, however, their SNP rate condition was reversed. Homozygous mutation in HvHKT1 gene (13.9%) was a bit higher significantly than in HvCBL4 gene (13.07). Restriction enzyme maps of these genes showed nucleotide diversity in restriction sites that are applicable for gene structure and functions change and affected significantly the changes in amino acids.

Key words: Barley · Salinity Tolerance · Single Nucleotide Diversity · HvCBL4 and HvHKT1 genes

INTRODUCTION

Barley (Hordeum vulgare L.) is grown in more than one hundred countries and it is one of the most important cereal crops worldwide. Its production ranks fourth after wheat, rice and maize [1]. Soil salinity is abiotic stress that threatens agricultural production [2]. Latest estimates showed that half of the irrigated lands worldwide, or nearly 20% of the world’s cultivated land are affected by salinity. One of the reasons of evolution of salt tolerant crops is the amount of the increased salt-affected lands globally, where the diversity of halophytes provided the base for the use of novel species in bioremediation and conservation [3]. The complexity of stress tolerance to abiotic stresses is known, but with the improve technologies of phenotyping it is easier to quantify components that contribute to abiotic stress tolerance [4]. Although, there are many physiological aspects of salt tolerance to understand, yet the interactions are complex. It requires combinations of several traits including the accumulation and compartmentation of ions for osmotic adjustment, the synthesis of compatible solutes, the ability to accumulate essential nutrients particularly K⁺ in the presence of high concentrations of the ions generating salinity (Na⁺), the ability to limit the entry of these saline ions into the transpiration stream and the ability to continue to regulate transpiration in the presence of high concentrations of Na⁺ and Cl⁻ [5].

Generally, reverse genetic techniques to improve salinity tolerance (e.g., RNAi, gene knockout, site-directed mutagenesis, transposon tagging) used transgenic materials [6]. However, new techniques such as
sequencing of interest parts of gene (genes) is a non-transgenic reverse genetic approach that has advantages of high-throughput, cost-effective and time-intensive and applicable in all plants (and animals) that can be mutagenized for discovering gene diversity.

All plants to tolerate salinity, Na\(^+\) and Cl\(^-\) uptake must be restricted while maintaining the uptake of macronutrients such as K\(^+\), NO\(_3\)\(^-\), and Ca\(^2+\). The mechanisms of Na\(^+\) and K\(^+\) transport in plants under salt stress have been extensively reviewed [7, 8]. Genes that enable crops to limit Na\(^+\) accumulation in shoot tissues represent potential sources of salinity tolerance for breeding. In barley, the locus (HvNax4) lowered shoot Na\(^+\) content between 12% and 59% (g21 DW), or not at all, depending on the growth conditions in hydroponics and a range of soil types, indicating a strong influence of environment on expression. HvNax4 was fine-mapped on the long arm of barley chromosome 1H [9]. HvCBL4, a close barley homologue of the Arabidopsis (SOS3), co-segregated with HvNax4. Therefore, HvCBL4 represents a candidate for HvNax4 that warrants. It was suggested that HvCBL4 has EF-hand motifs capable of binding Ca\(^2+\) and that Cys2 would be N-myristoylated, which are features required for SOS3 functionality. HvCBL4 co-segregated with the HvNax4 locus and is similar to the SOS3 salinity tolerance gene; decisive proof that HvCBL4 is the HvNax4 gene is currently lacking [9].

Qiu [10] reported that salinity tolerance was significantly associated with K\(^+\)/Na\(^+\) ratio. Association analysis revealed that HvHKT1 mainly controls Na\(^+\) transporting under salinity stress, which was validated by further analysis of gene expression. It was reported that there is a wider genetic variation for the populations habiting in stressed environments [11]. To counter the two components of salinity stress (osmotic stress and ionic stress), glycophyte such as barley has evolved three strategies to adapt [12]. First strategy, the accumulation or synthesis of compatible solutes is often regarded as a basic strategy [13, 14]. The second strategy, is the antioxidant protection, Witzel [15] emphasized that the proteins involved in reactive oxygen species (ROS) detoxification is the key to salinity tolerance of barley. The third strategy is the regulation of ion homeostasis.

Recently, more attention has been paid to study the mechanics of homeostasis maintenance and the interaction between sodium and potassium ions [16-19]. Subsequently, the genes controlling ion homeostasis are concentrated in the relevant studies, especially the genes encoding high-affinity K\(^+\) transporters (HKT) which were considered as candidates for salinity tolerance improvement [20, 21]. In barley, the expression of HKT1 regulated by K\(^+\) was studied [22]. Moreover, Haro [23] cloned the HvHKT1 (AM000056), which mediates Na\(^+\) unipart in roots. Consequently, the HKT families involving in Na\(^+/\) K\(^+\) homeostasis were expected important in salinity tolerance.

Mutations and subsequent polymorphisms are important for understanding selection, adaptation and other evolutionary processes [24, 25]. Single nucleotide polymorphism (SNP) profiles can be used to infer the intensity of selection by comparing the frequency and distribution of SNPs of two or more sequence regions, such as genic vs. intergenic regions, exonic vs. intronic vs. untranslated regions, autosomes vs. sex chromosomes and particular gene groups with specific functions [26, 27]. To properly compare SNP frequency among groups (e.g. among functional gene groups, among chromosomes, among populations), large samples of similar size are desirable [28].

**MATERIALS AND METHODS**

Different 96 barley genotypes, collected from salt affected areas worldwide were evaluated for salinity tolerance. Three seeds of each accession were planted in treys containing peat moss, for two weeks at 14°C in growth chamber. Two salinity stress tolerance genes were selected (HvCBL4 and HvHKT1) to study single nucleotide variation in these barley ecotypes, according to the sequences of HM175878.1 (Hordeum vulgare subsp. vulgare cultivar Clipper calcineurin B-like protein 4 (CBL4) gene, complete cds) and AM000056.1 (Hordeum vulgare mRNA for high affinity sodium transporter). PCR primers for amplification of CDS fragments of these genes were designed with Primer 3.0 software according to the sequences of HM175878.1 (Hordeum vulgare subsp. vulgare cultivar Clipper calcineurin B-like protein 4 (CBL4) gene, complete cds) and AM000056.1 (Hordeum vulgare mRNA for high affinity sodium transporter) from NCBI database as showed in Table 1 for HvCBL4 and HvHKT1 genes, respectively. Leaf samples were collected after two weeks of planting and kept on the freeze drier for 3 days.

**DNA Extraction:** For DNA extraction, 0.3-0.5 gr of plant fresh leaf of each accession was grinded with grinder machine. Then, 1 ml CTAB solution (CTAB, hot and do
vortex) was added and incubated in water bath at 65°C for 1 hour. Then, the samples were kept on ice for 5 min and then 900 µl of chloroform /isoamylalcohol (24/1) were added and shacked for 30 min on a shaker. Afterward, the samples were centrifuged for 20 min. at 12, 000 rpm, aqueous phase transferred to new tubes and 6µl RNase (20 mg/ml) added and incubated for 30 min. at 7°C. Finally, 1 ml of isopropanol (-20°C) was added and incubated at 4°C for 30 min. or in -20°C freezer for 10 min.

**PCR Amplification:** PCR amplification reaction was done in 20 µl volume consist of 2 µl genomic DNA (15 ng/ µl), 1 µl primer mix containing forward and reverse primers in equal amounts (stock solution: 10 µM for each forward and reverse primer) and 17 µl of pre-prepared Master Mix containing 11.8 µl autoclaved distilled water, 2 µl 10x PCR Buffer (B55, Fermentase Inc.), 1.5 mM MgCl2 (25 mM, R0971, Fermentase Inc.), 2 µl of 2 mM dNTPs (R0192, Fermentase Inc.). All amplifications were performed on ABI Biosystem thermal cycler (Applied Biosystems, Inc.) under the following conditions: 5 min at 94°C, followed by Touch-down PCR steps, 30 s at 94°C, 30 s at 68°C and 1 min at 72°C for 10 cycles and then 30 s at 94°C, 30 s at 58°C and 1 min at 72°C for 35 cycles and 15°C for final stage.

**Sequencing Data Analysis:** After amplification, all PCR products from 96 accessions (8 products for each accession) were purified and sequenced using ABI 3100 automated sequencer following the manufacturer’s instructions (Applied Biosystems, Inc.). Sequenced fragments obtained for each gene in each accession (4 fragments) were pooled by CodonCode Aligner software and assembled related contigs. The consensus sequence for any contig was formed and employed as a reference sequence to which individual sequence reads were aligned using CodonCode Aligner.

**RESULTS AND DISCUSSIONS**

The **Nucleotide Diversity of HvHKT1 and HvCBL4:** In this study, sequence data of 96 barley genotypes were obtained from separate PCR reactions. The whole length of 1,925 and 2,333bp genomic DNA sequences for HvHKT1 and HvCBL4 were assembled, respectively. The full length of two coding regions was 1,595 and 657 bp, respectively, which consisted of one exon in HvHKT1 and seven exons separated by six introns in HvCBL4. These regions were amplified by four PCR amplification of each gene. We considered natural nucleotide diversity (or nucleotide mutation) on these fragments. The polymorphism density including SNPs and Indels were found, with 17.9 SNP/kb, 1 indel per 199.3 bp for HvHKT1 and 137SNP/ kb, 18 indel per bp for HvCBL4. Total number of nucleotide variation in HvHKT1 was significantly more than that in HvCBL4 because the whole length of fragments studied in HvHKT1 was higher than that of HvCBL4. But SNP rate condition in them was reverse. This indicates that the total number of SNPs correlated positively with sequence whole length, but
different rate of SNP mutations was observed in the fragments. This could be attributed to the effect of other factors such as coding or non-coding regions on SNP rate. The lower rate of SNP in HvHKT1 was due to more coding and conserved regions in HvHKT1 than to HvCBL4 gene. The characteristics of nucleotide diversity in two genes are provided in Table 2.

According to cereals SNP database (http://autosnspdb.appliedbioinformatics.com), up to now, 454989 sequence reads of barley containing 25674 contig registered in this database and 29447 SNP reported. The SNP rate reported in this database is 240 SNP/bp. Where others reported very varied range of SNP rate in from 27-200 SNP/bp [30, 31].

Available information on the rate of appearance of new mutations and their effects at the molecular and phenotypic levels is prerequisite for understanding of mechanisms of evolution. Such information on spontaneous mutations are now available in a variety of organisms, general patterns have emerged for the scaling of mutation rate with genome size and for the likely mechanisms that drive this pattern [32]. Hyung Eo and De Woody [28] studied the influence of contig depth on the estimation of SNP frequencies and they tested the hypothesis that non-coding transcripts are less conserved than protein-coding transcripts by comparing their SNP frequencies. They found little difference in SNP frequency in the contigs, with a trend of a higher frequency of SNPs in long contigs representing non-coding transcripts relative to protein-coding transcripts. These results support the hypothesis that long non-coding transcripts are less conserved than long protein-coding transcripts. Hyung Eo and De Woody [28] found positive associations between the number of SNPs and contig length (r = 0.32; P<0.001) and between the number of SNPs and contig depth and contig depth is more efficient to SNP discovery (r = 0.46; P<0.001).

The characteristics of nucleotide diversity in two genes based on contig position are shown in Table 2. Moreover, this study determined various types of nucleotide diversity in amplified PCR product from different primers of the two genes and the results are shown in Table 3. A sample of heterozygous mutation was shown in Fig. 1 where ecotype 12 contains 3 nucleotide substitutions in amplified fragment of HvCBL4-2 primer those are located on 109, 130 and 179 nucleotides.

HvHKT1 and HvCBL4 Primers Amplified Product Analysis: The results of the genomic DNA fragments of the genotypes amplified by two primers (Table 3), revealed that HvHKT1-2 has the highest number of mutant base (1430 bases), homozygote mutation in coding region (90 mutations), the number of mutations that cause to change amino acid (1297) and indels (6) than to other three segments of HvHKT1 gene. In HvCBL4 gene, HvCBL4-1 primer had the maximum number of mutant base (1229 bases), the highest number of mutation that cause to change amino acid (1086) and homozygote mutation (193 mutations) in coding region in all genotypes. The highest number of indels (82) was observed in HvCBL4-4 amplified fragments. Although, the most mutant base in HvHKT1 gene was higher than that of HvCBL4, but homozygote mutation number in HvCBL4 was higher than that of HvHKT1 gene.

Homogygote mutation is very valuable in genetic and mutation studies [33], because, homogygote mutation convert one amino acid to another one. A lot of homogygote mutations can change amino acid, as well as, shown in Table 3, a large number (more than 70%) of total mutations in each fragment led to change amino acid. This type of mutations, especially in coding regions, causes changes in gene product and then gene function alteration. Therefore, the effect of them on gene operation and use their beneficial efficacy for gene improving is considered. The INDELS have a major effect on gene structure and function particularly if they located on coding regions of gene.

Singh et al. [33] studied BADH1 gene in local rice varieties and found 17 SNPs in intron and 3 SNPs in exone region of this gene and they realized that all exonic SNPs
Table 3: Nucleotide diversity in amplified products of *HvHKT1* and *HvCBL4* using two primers

<table>
<thead>
<tr>
<th>Gene</th>
<th>Primer</th>
<th>Total number of mutant nucleotide</th>
<th>The number of genomic sites bearing mutation</th>
<th>The number of homozygous mutation in coding region</th>
<th>The number of mutations led to change amino acid</th>
<th>Indels</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>HvHKT1</em></td>
<td>HKT1-1</td>
<td>218</td>
<td>83</td>
<td>3</td>
<td>148</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>HKT1-2</td>
<td>1430</td>
<td>110</td>
<td>90</td>
<td>1297</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>HKT1-3</td>
<td>144</td>
<td>49</td>
<td>1</td>
<td>95</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>HKT1-4</td>
<td>225</td>
<td>69</td>
<td>7</td>
<td>153</td>
<td>1</td>
</tr>
<tr>
<td><em>HvCBL4</em></td>
<td>CBL4-1</td>
<td>1229</td>
<td>173</td>
<td>193</td>
<td>1086</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>CBL4-2</td>
<td>203</td>
<td>58</td>
<td>5</td>
<td>144</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>CBL4-3</td>
<td>494</td>
<td>90</td>
<td>7</td>
<td>402</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>CBL4-4</td>
<td>274</td>
<td>134</td>
<td>3</td>
<td>71</td>
<td>82</td>
</tr>
</tbody>
</table>

Fig. 1: Three INDELS located on *HvCBL4*-2 primer amplified fragment in 12 barley ecotypes

Fig. 2: Amplified product of *HvHKT1*-1 primer with five restriction sites in different parts and two SNPs placed on two restriction sites

Fig. 3: *HvHKT1*-2 amplified DNA segment with one restriction site

Fig. 4: Two different restriction sites on *HvHKT1*-3 amplified segment both were created by EcoRI enzyme

Fig. 4: Two restriction sites on *HvHKT1*-3 amplified segment both created by EcoRI enzyme

Fig. 5: No enzyme restriction site was created on *HvHKT1*-4 amplified DNA fragment

Fig. 6: *HvCBL4*-1 amplified product restriction map in contig 1, with two restriction sites created by two different enzymes
Fig. 7: HvCBL4-1 amplified product restriction map in contig 2

Fig. 8: HvCBL4-1 amplified product restriction map in contig 3

Fig. 8: HvCBL4-2 amplified product restriction map, with two restriction sites

Fig. 10: HvCBL4-3 amplified product restriction map with five restriction enzyme sites created by four enzymes

Fig. 11: HvCBL4-4 amplified product restriction map with nine restriction sites from five enzymes

Table 4: Mutation description in barley ecotypes as located in the map

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Mutant type</th>
<th>The number of genotypes bearing this mutation</th>
<th>Restriction site (Base)</th>
<th>Mutation description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BamHI</td>
<td>Indel</td>
<td>One genotype</td>
<td>100</td>
<td>One base deletion</td>
</tr>
<tr>
<td>BSTxl</td>
<td>Homozygote mutation</td>
<td>One genotype</td>
<td>254</td>
<td>G→A</td>
</tr>
<tr>
<td>EcoRI</td>
<td>Nucleotide Substitution</td>
<td>One genotype</td>
<td>310</td>
<td>A→T: Resulted in conversion of Arginine to Serine.</td>
</tr>
<tr>
<td>BamHI</td>
<td>Nucleotide Substitution</td>
<td>nine genotype</td>
<td>387</td>
<td>- Two genotypes bearing G→C</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>- Two genotypes bearing G→A</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-Five genotypes bearing G→T</td>
</tr>
<tr>
<td>XmiI</td>
<td>Nucleotide Substitution</td>
<td>One genotype</td>
<td>209</td>
<td>Converting C→G</td>
</tr>
<tr>
<td>HincII</td>
<td>Nucleotide Substitution</td>
<td>Two genotype</td>
<td>324</td>
<td>Converting G→T</td>
</tr>
<tr>
<td>Sall</td>
<td>Nucleotide Substitution</td>
<td>Two genotype</td>
<td>517</td>
<td>Converting T→C</td>
</tr>
<tr>
<td>XmiII</td>
<td>Nucleotide Substitution</td>
<td>One genotype</td>
<td>518</td>
<td>Converting C→G</td>
</tr>
</tbody>
</table>

cause change in amino acid and finally significant change in gene function. Likewise, Xin et al. [15] studied mutation effect on Cardiac myosin-Binding protein C gene causes severe neonatal hypertrophic cardiomyopathy and reported that skipping of the 140-bp exone 30 led to a frame shift and premature stop codon in exone 31.

Restriction Maps: Restriction maps for eight amplified PCR products are shown in Figures (2-11) and their information are provided in Table (4). The restriction sites in different parts of amplified product ranged between nine (HvCBL4-4 primer; Figure 11) and zero (HvHKT1-4 or 2; Figure 3). As shown in Table 4, in all three cases (HincII, Sall, XmiI enzymes) one amino acid changed to another and influenced gene structure and gene function.

Contig 2 and 3, with five and two restriction sites, respectively from HvCBL4-1 amplified segment were observed (Figs. 7 and 8). These sites were created by using four and two enzymes for contig 2 and 3, respectively and these results indicate that there is no match between restriction sites and mutation places.

Restriction Map Evaluation: Restriction mapping requires the use of restriction enzymes. In molecular biology, restriction maps are used as a reference to engineer plasmids or other relatively short pieces of DNA and sometimes for longer genomic DNA [34]. Restriction mapping is a very useful technique when used for determining the orientation of an insert in a cloning vector, by mapping the position of an off-center
restriction site in the insert [35]. In organisms that lack robust DNA transformation methods, mapping with visible or selected single nucleotide polymorphism (SNP) markers to progressively finer genomic intervals is the traditional route to ascertain identity of the mutant gene [36-40]. While resequencing a genome to identify mutant alleles is being used more frequently, in some cases it is more efficient to sequence only a portion of a genome. For example, sequencing of a single chromosome, a defined genomic interval, exonic sequences, or a single locus can be more cost effective when there is evidence that a mutation resides within a specific genome feature [41, 42, 43]. We adopted this method (sequencing of coding regions of two genes) in the current study to make restriction enzyme maps for two genes and found SNPs in restriction sites. These SNPs are very important for functional genomics. Because we can study the effect of type and location of different SNPs on gene structure and operation. Also, they can be very useful for gene manipulating and engineering. In future, we will able to use these SNPs as markers to other genetic studies, too.

O’Rourke [44] reported that the use of restriction site-associated DNA (RAD) polymorphism markers for rapidly mapping mutations after chemical mutagenesis and mutant isolation. Genomic interval pull-down sequencing (GIPS) could be also used to selectively capture and sequence megabase-sized portions of a mutant genome. Therefore, the information obtained from our study (finding SNPs in restriction sites) is valuable for genetic studies because some nucleotide mutation place on important part of the gene (or genes) and have direct or indirect effect. There are different restriction enzymes with different efficiency to restrict DNA fragments, which means there are various scissors to cut different size of genomic DNA and it is the best material for gene manipulating.

**CONCLUSIONS**

This study was conducted to determine single nucleotide diversity of two salinity tolerance genes (HvHKT1 and HvCBL4) in barley ecotypes collected from different salt affected area worldwide. The results showed very important and valuable SNPs that are applicable in other genetic studies, gene function change and engineering. In some genomic area, such as coding and conserved regions of gene, SNP frequency was lower than the non-coding and non-conserved regions. Also, the frequency of SNPs in intron is more than exone. When the two genes were compared, SNP abundance in HvCBL4 was seven times more than HvHKT1. Homozygous mutation in HvHKT1 gene (13.9%) was a bit higher than that of HvCBL4 gene (13.07). This type of mutation is very noteworthy because, it is the cause of almost half of amino acid changes in genome. It can use the SNP diversity as markers to mapping, too. Restriction enzyme mapping help to map mutations and let to investigate interest sites for gene manipulating and engineering. We found very good nucleotide diversity in restriction sites of maps obtained from amplified fragments that they are applicable for gene structure and functions change.

**REFERENCES**


35. Dale, J. and M. Von Schantz, 2003. From Genes to Genomes. West Sussex: John Wiley and Sons Ltd.


