The salt tolerance candidate genes family in wheat and its relationship to the phylogenetic complexity of cereals

Gaboun F., Diria G., Adenike F., Abdelwahd R., Ibriz M., Soulaymani A.

in

Porceddu E. (ed.), Damania A.B. (ed.), Qualset C.O. (ed.).
Proceedings of the International Symposium on Genetics and breeding of durum wheat

Bari : CIHEAM
Options Méditerranéennes : Série A. Séminaires Méditerranéens; n. 110

2014
pages 323-337

The salt tolerance candidate genes family in wheat and its relationship to the phylogenetic complexity of cereals

Fatima Gaboun[1], Ghizlane Diria[1], Fujah Adenike[1], Rabha Abdelwahd[1], Mohammed Ibriz[2], Abdalmajid Soulaymani[2]

[1] INRA, Research Unit of Biotechnology, Rabat, Morocco
[2] University Ibn Tofail, Faculty of Sciences, Kénitra, Morocco

Abstract. Salt is a major abiotic stress affecting crop plants worldwide. In Morocco, the problem is worsening with climate change. More than 700.000 hectares are affected by salt, therefore a large number of south Moroccan farms is affected in some way by soil and groundwater salinity. So to understand the mechanisms underlying the response of cereal species to salinity in natural systems, our study focused on reverse genetics for the characterization of genes identified as l kely to increase tolerance of wheat to salinity. Salt tolerance comes from genes that limit the rate of salt uptake by the plant from the soil and the transport of salt throughout the plant, adjust the ionic and osmotic balance of cells in roots and shoots, and regulate leaf development and the onset of senescence. Salt-tolerant candidate genes families in wheat and other cereals such as HKT, NHX, SOS, and HAK were identified and downloaded from the NCBI and wheat genomics of abiotic stress (WGAS) databases. Comparative studies and analyses, such as search for domains, multiple sequence alignments and phylogenetic tree constructions as well as primers design were carried out using bioinformatics tools. Sixty candidate genes for salt-tolerance were identified. Several protein domains have been characterized: (i) the protein family of HKT genes contains a TrKH (cation transport protein) domain, (ii) the protein family of NHX genes contains a Na+/H+ exchanger domain, (iii) *a Ktrans (K+ potassium transporter) domain for the protein family of HAK genes,* (iv) *a Na+/H+* exchanger domain and a cNMP (cyclic nucleotide binding domain) for SOS1 genes, (v) a serine/threonine protein kinase domain for SOS2 genes, and (vi) an EF-hand (calcium binding protein) domain for SOS3 genes. Multiple sequence alignments of HKT genes in wheat revealed a high frequency of glycine and serine amino acids conserved in the consensus sequence. According to the phylogenetic tree analysis, HKT genes were grouped into two subfamilies, and this division is associated with a substitution of a glycine/serine residue intended to be in first loop pores of the protein. All members of the subfamily 1 have a serine at this position, whereas members of subfamily 2 (except OsHKT1) have a glycine. The RT-PCR primers associated with the candidate genes for the studied trait were designed as markers for selection to assist the cereal breeding program.

Keywords. Bioinformatics – Salt-tolerant genes – RT-PCR primers – *Triticum .*

Famille des gènes candidats pour la tolérance à la salinité chez le blé et sa relation à la complexité phylogénétique des céréales

Résumé. *La salinisation est un stress abiotique majeur pour les plantes cultivées dans le monde entier. Au Maroc, le problème est aggravé par les effets du changement climatique. Plus de 700 000 hectares sont affectés par la salinité et donc, bon nombre d'exploitations agricoles dans le sud du Maroc sont touchées à un degré différent par la salinité du sol et des eaux souterraines. Pour comprendre les mécanismes régissant la réponse à la salinité des espèces céréalières dans les systèmes naturels, notre étude a été axée sur la génétique inverse pour la caractérisation des gènes identifiés comme susceptibles d'augmenter la tolérance du blé à la salinité. La tolérance au sel provient des gènes qui limitent la vitesse d'absorption du sel par la plante dans le sol et le transport de sel tout au long de la plante, contrôlent l'équilibre ionique et osmotique des cellules des racines et des pousses, et règlent le développement de la feuille et le début de la sénescence. Les familles de gènes candidats pour la tolérance à la salinité chez le blé et chez d'autres céréales telles que HKT, NHX, SOS, et HAK ont été identifiées et téléchargées à partir des bases de données NCBI et des données génomiques des stress abiotiques du blé (WGAS). Les études et les analyses comparatives, telles que la recherche des domaines, des alignements de séquences multiples et les constructions d'arbres phylogénétiques ainsi que la conception des amorces ont été réalisées à l'aide des outils de la bioinformatique. Soixante gènes candidats pour la tolérance à la salinité ont été identifiés. Plusieurs domaines protéiques*

ont été caractérisés : (i) la famille de protéines des gènes HKT contient un domaine TrKH (protéine de transport des cations), (ii) la famille de protéines des gènes NHX contient un domaine échangeur de Na+/H+, (iii) un domaine Ktrans (transporteur de potassium K+) pour la famille des protéines des gènes HAK, (iv) un domaine échangeur Na+/H+ et un cNMP (domaine de liaison à un nucléotide cyclique) pour les gènes SOS1, (v) un domaine sérine/thréonine protéine kinase pour les gènes SOS2, et (vi) un domaine EF-hand (protéine de liaison du calcium) pour les gènes SOS3. Des alignements de séquences multiples des gènes HKT de blé ont révélé une fréquence élevée des acides aminés de la glycine et de la sérine conservés dans la séquence consensus. Selon l'analyse de l'arbre phylogénétique, les gènes HKT ont été regroupés en deux sous-familles, et cette division est associée à une substitution d'un résidu glycine/sérine destiné à se situer dans les pores de la première boucle de la protéine. Tous les membres de la sous-famille 1 ont une serine dans cette position, alors que les membres de la sous-famille 2 (sauf OsHKT1) ont une glycine. Les amorces de RT-PCR associées aux gènes candidats pour le caractère étudié ont été conçues comme des marqueurs de sélection pour faciliter le programme d'amélioration des céréales.

Mots clés. *Bioinformatique – Gènes de tolérances à la salinité – Amorces RT-PCR – Triticum.*

I – Introduction

Abiotic stresses are a serious problem to crop production under dry land conditions in arid and semi-arid regions of the world. The area is still increasing as a result of irrigation or land-clearing (FAOSTAT, 2012). These abiotic stresses include high and low temperatures, water deficit, sodicity, alkalinity, acidity, ion deficiencies and toxicities and salinity (Javid *et al.,* 2011). The major salinity problem in Morocco is in dry lands (more than 700.000 hectares), the overuse of surface and ground water, coupled with agricultural intensification, generates soil salinity and sodicity problems (Bannari *et al.*, 2008). Whole plant tolerance to soil salinity involves numerous processes in many different tissues and cell types. For many cereals, sensitivity to salinity is due to the accumulation of sodium (Na+) to toxic concentrations in the leaves (Bryt, 2008). Recent advancements in biotechnology have led to the development of more efficient selection tools to substitute phenotype-based selection systems (Ashraf and Foolad, 2012). The mechanism of the molecular response of higher plants against water stress has been analyzed by studying a number of genes in *Arabidopsis thaliana* responding to drought, high-salinity and cold stress at the transcriptional level (Seki *et al*., 2002) (Fig.1).

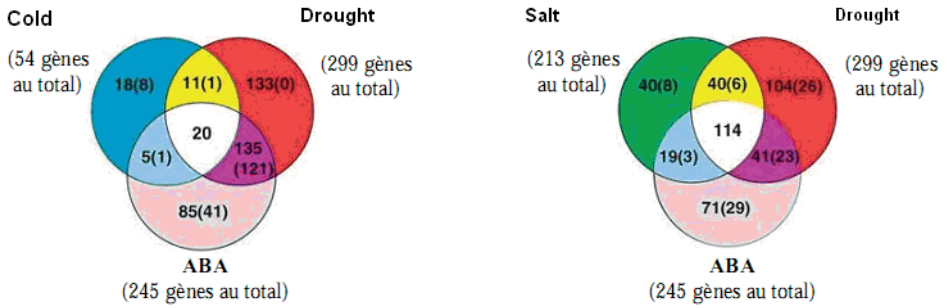

Figure 1. Venn diagrams describing the genes regulated during abiotic stress in Arabidopsis thaliana. ABA, abscisic acid. (Seki *et al*., 2002).

With the genomes of various plants having been sequenced, the total complement of potential proteins involved in Na+, K+ and Cl- transport can be surmised. Fig. 2 gives an overview of the main classes of monovalent ion transporters, often derived from large gene families (Mian *et al*. 2009).

Figure 2. Overview of gene families involved in Na+, K+ and Cl- homeostasis in rice during salt stress. Abbreviations: CCC, cation chloride co-transporter; CHX, cation/H+ exchanger; CLC, voltage gated Cl- channel; CNGC, cyclic nucleotide gated channel; GLR, glutamate like receptor; HKT, high affinity K+ transporter; KHX, K+/H+ exchanger; KIR, Shaker type K+ inward rectifier; KOR, Shaker type K+ outward rectifier; KUP/HAK, K+ uptake permease; NHX, Na+/H+ exchanger; NSCC, non-selective cation channel; TPK, two-pore K+ channel (Mian *et al*. 2009).

Improving crop plants genetically for salt tolerance represents an important part of basic plant biology (Zhu, 2000). Breeding for salt tolerance can be thought of as selecting plants that withstand salt stress most effectively (Shannon and Qualset, 1984). Whole plant tolerance to soil salinity involves numerous processes in many different tissues and cell types. For many cereals, sensitivity to salinity is due to the accumulation of sodium (Na+) to toxic concentrations in the leaves. Recent advancements in biotechnology have led to the development of more efficient selection tools to substitute phenotype-based selection systems (Ashraf and Foolad, 2012). The marker-assisted selection is a process of indirect selection in which the character in question has a high heritability, since not influenced by environmental factors. There is increased efficiency of plant breeding, reducing the number of progenies and the number of generations for the stabilization of the genotypes. The selection can be performed in early generations (Eduardo, 2011). The procedures require integrating multidisciplinarity involving researchers with backgrounds in classical plant breeding, chemistry, biochemistry, plant physiology, statistics, computer science and bioinformatics. Biotechnology and Bioinformatics carries benefits for plant researchers: it can aid in plant breeding and genetic engineering, and allow plant scientists to produce salt tolerant for the future (Mochida and Shinozaki, 2010). Hence, a detailed understanding of the basic mechanisms involved in the plant salt tolerance is an important prerequisite to improve the performance of crop plant in saline soils (Binzel and Reuveni, 1994; Eduardo, 2011).

Furthermore, promotion of comparative genomics among model and applied plants allows us to grasp the biological properties of each species and to accelerate gene discovery and functional analyses of genes (Mochida and Shinozaki, 2010). The objective of this study, for better understanding of the mechanisms that can contribute in salt tolerance of cereals, is to identify candidate genes involved in this mechanism and to develop In Silico some markers associated to those genes using bioinformatics tools.

II – Methodology

In this study, candidate genes involved in the tolerance to salt mechanism have been searched, identified and used. Twenty-three genes of the HKT family, twenty-four genes of the NHX family,

five genes of SOS family, and eight genes of the HAK family, which are homologous genes, were downloaded. Table 4 summaries the approach and tools used in this study.

1. Bioinformatics tools

Several bioinformatics tools were downloaded from the Internet and used to meet our objective in the context of our study but we opted for a few that applies to the described methodology. We, however, mostly used CLC Main Workbench because it includes several features that make certain operations effective on biological data (DNA, RNA, proteins, etc.), and accepts different file formats (GB: GenBank, GFF: Generic Feature Format, Fasta, etc.) which is limited to others. In addition, it helps to have direct access to some external databases such as NCBI (National Center for Biotechnology Information), UniProt (Universal Protein Resource), and Pfam (Protein families' database). It has a good GUI (Graphical User Interface). The list of databases and bioinformatics tools used in this study is shown in Table 1.

Table 1. List of databases and bioinformatics tools used.

Bioinformatics tools and databases	Functions and approach
NCBI (http://www.ncbi.nlm.nih.gov/)	Protein and nucleotide sequence download.
AMIGO (http://amigo.geneontology.org/cgi-bin/amigo/go.cgi)	Identification of genes' functions.
Pfam (http://pfam.janelia.org/)	Search for protein domains.
CLC MAIN WORKBENCH (http://www.clcbio.com)	Protein sequence analyses.
GeneFisher2 (http://bibiserv.techfak.uni-bielefeld.de/genefisher2/submission.html)	Primers design.
Oligo Calc: Oligo nucleotide Properties Calculator (http://www.basic.northwestern.edu/biotools/oligocalc.html)	Validation of the primers designed.

2. Collection of candidate genes

The availability of research platforms, such as the web tools of the National Center for Biotechnology Information (NCBI) has transformed the time-consuming task of identifying candidate genes from genetic studies to an interactive process where data from a variety of sources are obtained to select likely genes for follow-up (Sadasivam *et al.,* 2009; Sanchez *et al.,* 2011). Salt-tolerant candidate genes were known from scientific journals. Basic Local Alignment Search Tool (BLAST) is a sequence similarity search program that can be used via a web interface or as a stand-alone tool to compare a user's query to a database of sequences. Several variants of BLAST compare all combinations of nucleotide or protein queries with nucleotide or protein databases (Johnson *et al*., 2008). The nucleotide and protein sequences of salt-tolerant candidate genes were collected and saved in a Fasta file format which was later used to perform multiple sequence alignments.

The genes identified were downloaded and are detailed in the results and discussion section, with the names of candidate genes, their protein and nucleotide sequence accessions, descriptions, dates of modification and lengths in amino acids (aa) and base pairs (bp).

3. Search for candidate genes' functions

AmiGO is a web application that allows users to query, browse and visualize ontologies and related gene product annotation data (Carbon *et al*., 2009). AmiGO can be used online at the Gene Ontology (GO) website to access the data provided by the GO Consortium; it can also be downloaded and installed to browse local ontologies and annotations. AmiGO is free open source software developed and maintained by the GO Consortium. Every page in AmiGO offers a simple search box through which users can query the GO database for GO terms or gene products. AmiGO returns search results ordered by how closely the result matches the original

query; results can also be sorted by other parameters, such as accession in term searches, or gene symbol when querying for gene products..

4. Search for protein domains

Pfam is a database of protein families, where families are sets of protein regions that share a significant degree of sequence similarity, thereby suggesting homology. Similarity is detected using the HMMER3 (http://hmmer.janelia.org/) suite of programs (Finn *et al.,* 2010). This database helped to further confirm candidate gene.

5. Multiple Sequence Alignment (MSA)

CLC Main Workbench is developed for Windows, MacOSX and Linux. The software for either platform can be obtained from http://www.clcbio.com. CLC Main Workbench 5.7.1 software was installed on a Microsoft windows operating system. CLC Main Workbench was chosen amongst other bioinformatics tools such as Geneious, Bioedit, ClustalW etc., because it offers more advantages especially in the graphical presentation of sequences from multiple sequence alignments.

6. Construction of phylogenetic trees

Phylogenetic relationship was determined based on multiple sequence alignments by using the CLC Main Workbench 5.7.1.

7. Primers design

GeneFisher2 is a recent reimplementation of the original GeneFisher application which recreates the overall functionality of its predecessor while enhancing usability and user experience (Hagemeier, 2006). This new version accesses basically the same underlying tools, now turned into components, via a web application, and conducts the whole application from a Web GUI by means of AJAX (Asynchronous Java-Script and XML) technology. GeneFisher accepts single or multiple DNA and protein sequences as input. As primers are calculated for a single DNA sequence, multiple input sequences are aligned using alignment programs such as ClustalW or DCA. From the alignment, a consensus sequence is derived and used as input for the primer calculation step. GeneFisher selects PCR primers with certain criteria such as: melting temperature Tm, GC content, primer length, 3' clamp GC content and degeneration, hairpin loop structure detection, primer-primer dimers detection, primer degeneration, amplified region length, and primer uniqueness (Hagemeier, 2006). In this study, Primers were designed from the coding portion of the nucleotide sequences (mRNA) downloaded from the NCBI database.

The choice of good primer is then validated by the Oligonucleotide Properties Calculator application software which is used to calculate all the parameters as well as check for complementarily of the primers designed (Chavali *et al*., 2006). Generally, in this study, the primer length must be within the range of 18-30 nucleotides, and its composition in GC must be between 40-60%. The primers should neither form self-dimers nor cross dimers or hairpin structures.

III – Results

1. Salt-tolerant candidate genes

In this study, salt-tolerant candidate genes were searched, identified and downloaded (twenty-three HKT genes, twenty-four NHX genes, five SOS genes, eight HAK genes) (Annex 1).

2. Overview of candidate genes' functions

HKT and HAK gene families are transporters of potassium (K+) while NHX gene families are exchangers of Na+/H+. However, SOS gene families are antiporters of Na+/ H+.

3. Search for protein domains

The search for protein domain using pfam has shown that HKT genes belong to a family of proteins known as TrkH: Cation transporting proteins, which consists of various cation transport proteins (Trk) and the subunit of ATP synthase, sodium or ATPase translocation. These proteins are involved in the active absorption of sodium by using ATP in the process of absorption. Trk/ HKT transporters are reminiscent of K+ channels in that they possess in a single polypeptide chain and four domains resembling P-loops. These P-loop-like domains are weakly conserved to K+ channel P-loops (Platten *et al*. 2006).

However for NHX genes, the protein families are exchangers of sodium (Na+) and hydrogen (H+) ions. This family is also called antiporters of Na+/H+, which act as transporters that play a major role in maintaining the pH of actively metabolizing cells. The molecular mechanisms of antiporters are not clear (Rodriguez-Rosales *et al*., 2009).

As for SOS1 genes, they present a protein family of exchangers of sodium and hydrogen ions, as well as cyclic nucleotide-binding domains.

Regarding the SOS3 gene, it encodes an EF-hand type calcium-binding protein with similarities to animal neuronal calcium sensors and the yeast calcineurin B subunit. As regards SOS2 gene, it encodes a serine/threonine type of protein kinase. The SOS2 gene physically interacts with SOS3, and is activated by the latter. The SOS3-SOS2 kinase complex represents a regulatory pathway that specifically controls the homeostasis of Na+ and K+ and tolerance of salt in plants. The product of this pathway is the upregulation of SOS1 expressed under NaCl stress (Shi *et al*., 2000). In addition, HAK genes belong to a protein family of potassium carriers (Ktrans: K+ potassium transport) in different species.

4. Multiple Sequence Alignments

Multiple sequence alignments (MSA) were carried out using protein sequences of wheat and related species such as barley (*Hordeum vulgare*), maize (*Zea mays*), rice (*Oryza sativa*), poplar (*Populus trichocarpa*), as well as selaginella (*Selaginella moellendorffii*), tomato (*Solanum lycopersicum* L.) and the model species thale cress (*Arabidopsis thaliana*), for a comparative study of these candidate genes across species.

Conserved regions, significant similitude were also observed in NHX, HAK and SOS candidate genes, respectively. MSA of HKT genes revealed conservation region which can be explained by the similarity in amino acid residues of their protein sequences. There is a high frequency of glycine and serine amino acids conserved in the consensus sequence, which is specific to the HKTs.

5. Phylogenetic tree construction

Phylogenetic trees of publicly available full-length HKT coding sequences or HKT amino acid sequences show that the gene family splits into two major branches (Fig. 2). To establish the relationship between the different salt tolerance genes in cereals , only one representative of each of the gene groups (preferentially the homologous with cut-of of 95%) was included in this analysis. In this study, we used UPGMA (Unweighted Pair Group Method with Arithmetic Mean) as our analysis method. HKT genes are homologs because they have a common ancestor. The

ancestor also known as root on this tree is *Selaginella moellendorffii* (SmHKT2). According to the phylogenetic tree, these genes were grouped into two subfamilies (subfamily 1 and subfamily 2), and this division is associated with a substitution of a glycine/serine intended to be in the first loop pores of the protein. All members of the subfamily 1 have a serine at this position, whereas members of subfamily 2 (except OsHKT1) have a glycine (Fig. 3). Functional analyses of TaHKT1, AtHKT1, and rice suggest that this particular amino acid could play a central role in determining the Na+ selection by the carrier.

The analysis of the four family of genes from cereals (Fig.4) reveals a diversification clustering in the same family of gene between species.

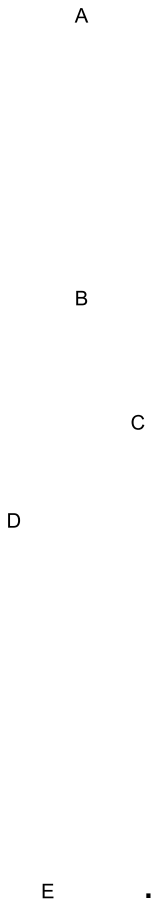

Figure 3. Phylogenetic relationships between salt tolerance genes: In the left (A) HKT genes, (B) HAK genes, (C) SOS genes and (D) NHX genes. Protein domains for HKT, HAK, NHX and SOS genes, done by CLC Main Workbench 5.7.1 using the UPGMA method with a bootstrap of 1000 (E).

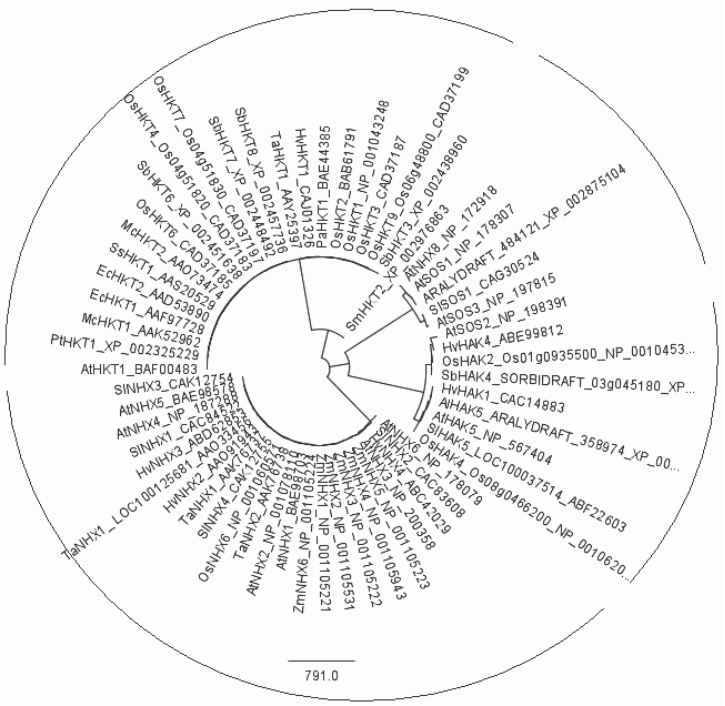

Figure 4. Phylogenetic relationships between salt tolerance genes (HKT, NHX, HAK, and SOS).

6. Primers design

The primers, designed by GeneFisher2 and validated by OligoCalc, are listed in Annex 2.

IV – Conclusion and perspectives

The search for salt-tolerant candidate genes among species via bioinformatics tools identified and characterized twenty-three HKT genes, twenty-four NHX genes, five SOS genes and eight HAK genes. Analyses such as search for genes' functions, search for protein domains, multiple sequence alignments, phylogenetic tree construction, carried out on these genes, showed that HKT and HAK genes play a role in the transport of potassium, NHX genes as exchangers of Na+/H+, and the SOS1 genes as antiporters of Na+/H+. Protein domains revealed that HKT genes belong to a protein family of TrKH involved in the active absorption of sodium by using ATP in the process of absorption while NHX family of genes are exchangers Na+/H+ that play a role in maintaining the pH of actively metabolizing cells.

References

Ashraf M., Foolad M.R., 2012. Crop breeding for salt tolerance in the era of molecular markers and marker-assisted selection. *Plant Breeding*, 132, pp. 10-20.

Bannari A., Guedon A.M., El-Harti A., Cherkaoui F.Z., El-Ghmari A., Saquaque A., 2008. Slight and moderate saline and sodic soils characterization in irrigated agricultural land using multispectral remote sensing. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 34, Part XXX, pp. 1-6.

Binzel M.L., Reuveni M., 1994. Cellular mechanisms of salt tolerance in plant Cells. *Hort. Rev*., 16, pp. 33–69.

Bryt C.S., 2008. Genes for sodium exclusion in wheat. *PhD. thesis, University of Adelaide*. Available from http://hdl.handle.net/2440/56208.

Carbon S., Ireland A., Mungall C.J., Shu S., Marshall B., Lewis S., 2009. AmiGO: online access to ontology and annotation data. *Bioinformatics*, 25, pp. 288–289.

Chavali S., Mahajan A., Tabassum R., Maiti S., Bharadwaj D., 2005. Oligonucleotide properties determination and primer designing: a critical examination of predictions. *Bioinformatics*, 21, pp. 3918-3925.

Eduardo A.G., 2011. Molecular markers: assisted selection in soybeans, soybean - genetics and novel techniques for yield enhancement, Krezhova D. (ed), ISBN: 978-953-307-721-5, DOI: 10.5772/18934.

FAOSTAT, 2012. Faostat.fao.org.

Finn R.D., Mistry J., Tate J., Coggill P., Heger A., Pollington J.E., Gavin O.L., Gunasekaran P., Ceric G., Forslund K., 2010. The Pfam protein families database. *Nucleic Acids Res,.* 38D, pp. 211–222.

Hagemeier D., 2006. GeneFisher2 - an AJAX based implementation of GeneFisher. In *Bachelor's thesis*. University Bielefeld, Faculty of Technology.

Javid M., Nicolas M., Ford R., 2011. Current knowledge in physiological and genetic mechanisms underpinning tolerances to alkaline and saline subsoil constraints of broad acre cropping in dryland regions. In: *Abiotic Stress in Plants - Mechanisms and Adaptations. Agricultural and Biological Sciences.* Shanker A., Venkateswarlu B. (eds), pp. 193-214.

Johnson M., Zaretskaya I., Raytselis Y., Merezhuk Y., McGinnis S., Madden T., 2008. NCBI BLAST: a better web interface. *Nucleic Acids Res.,* 36, pp. W5–W9.

Mochida K., Shinozaki K., 2010. Genomics and bioinformatics resources for crop improvement. *Plant, Cell Physiol.,* 51, pp. 497–523.

Mian A.A., 2010. Improving salt stress resistance in cereals. *PhD thesis, University of York*. http://etheses.whiterose.ac.uk/id/eprint/1191.

Rodriguez-Rosales M.P., Gálvez F.J., Huertas R., Aranda M.N., Baghour M., Cagnac O., Venema K., 2009. Plant NHX cation/proton antiporters. *Plant Signal Behav.,* 4, pp. 265-276.

Sadasivam R.S., Sundar G., Vaughan L.K., Tanik M.M., Arnett D.K. 2009. Genetic region characterization (Gene RECQuest) - software to assist in identification and selection of candidate genes from genomic regions. *BMC Res. Notes,* 2, pp. 201.

Sanchez D.H., Pieckenstain F.L., Szymanski J., Erban A., Bromke M., Hannah M.A., Kraemer U., Kopka J., Udvardi M.K.,2011. Comparative functional genomics of salt stress in related model and cultivated plants identifies and overcomes limitations to translational genomics. *PLoS One,* 6:e17094

Seki M., Narusaka M., Ishida J., Nanjo T., Fujita M., Oono Y., Kamiya A., Nakajima M., Enju A., Sakurai T., Satou M., Akiyama K., Taji T., Yamaguchi-Shinozaki K., Carninci P., Kawai J., Hayashizaki Y., Shinozaki K., 2002. Monitoring the expression profiles of ca 7000 Arabidopsis genes under drought, cold and high-salinity stresses using a full-length cDNA microarray. *Plant J.,* 31, pp. 279–292.

Shannon M.C., Qualset C.O., 1984. Benefits and limitations in breeding salt-tolerant crops. *Calif. Agric.* 38(10), pp. 33-34.

Shi H., Ishitani M., Kim C., Zhu J.k., 2000. The *Arabidopsis thaliana* salt tolerance gene SOS1 encodes a putative Na+/H+ antiporter. *Proc. Nat. Acad. Sci. USA*, 97, pp. 6896-6901.

Zhu J.K., 2000. Genetic analysis of plant salt tolerance using Arabidopsis. *Plant Physiol.*, 124, pp. 941–948.

ANNEX 1. List of salt-tolerant candidate genes: HKT, NHX, SOS, and HAK genes with their protein sequence accessions, descriptions, modification dates, lengths in amino acids, nucleotide sequence accessions, and lengths in base pairs, respectively.

Gene	Protein accessions	Description	Date	Length aa	Nucleotide Accession	Length (bp)
HKT						
AtHKT1	BAF00483	Sodium transporter [*Arabidopsis thaliana*].	27/07/2006	506	AK228564	1695
PtHKT1	XP_002325229	Sodium transporter hkt1-like protein [*Populus trichocarpa*].	04/12/2009	506	XM_002325193	1772
TaHKT1	AAY25397	High-affinity potassium uptake transporter [*Triticum aestivum*].	07 /05/2005	533	DQ009003	1981
McHKT1	AAK52962	Putative potassium sodium transporter [*Mesembryanthemum crystallinum*].	14 /05/2001	505	AF367366	1904
OsHKT1	NP_001043248	Os01g0532600 [*Oryza sativa Japonica* Group].	08/06/2010	509	NM_001049783	1587
EcHKT1	AAF97728	Potassium-sodium symporter HKT1 [*Eucalyptus camaldulensis*].	13/11/2000	550	AF176035	1992
HvHKT1	CAJ01326	High-affinity sodium transporter [*Hordeum vulgare* subsp. *vulgare*].	17/11/2005	531	AM000056	1925
PaHKT1	BAE44385	High-affinity potassium transporter [*Phragmites australis*].	11/03/2008	530	AB234304	1828
SsHKT1	AAS20529	Suaeda salsa HKT1 (hkt1) mRNA, complete cds.	12/03/2008	550	AY530754	2033
SmHKT2	XP_002976863	Sodium transporter [*Selaginella moellendorffii*].	13/08/2010	745	XM_002976817	2238
McHKT2	AA073474	Mesembryanthemum crystallinum high affinity potassium transporter 2 (HKT2) mRNA, complete cds.	01/04/2003	543	AY231175	2230
EcHKT2	AAD53890	Potassium-sodium symporter HKT2 [*Eucalyptus camaldulensis*].	13/11/2004	549	AF176036	1915
OsHKT2	BAB61791	Potassium-sodium symporter [*Oryza sativa* Indica Group].	14/02/2008	530	AB061313	1781
OsHKT3	CAD37187	Putative sodium transporter [*Oryza sativa* Japonica Group].	02/07/2003	509	AJ491820	1587
SbHKT3	XP_002438960	Hypothetical protein SORBIDRAFT_10g029000 [*Sorghum bicolor*].	13/07/2009	545	XM_002438915	1638
OsHKT4 (Os04g51820)	CAD37183	Putative sodium transporter [*Oryza sativa* Japonica Group]	02/07/2003	552	AJ491816	1669
OsHKT5	----------	Oryza sativa Japonica Group hkt5 pseudogene, exons 1-3.	14/11/2006	----	AJ506745	3203
OsHKT6 (Os02g07830)	CAD37185	Putative sodium transporter [*Oryza sativa* Japonica Group]	02/07/2003	531	AJ491818	1679
SbHKT6	XP_002451638	Hypothetical protein SORBIDRAFT_04g005010 [*Sorghum bicolor*].	13/07/2009	532	XM_002451593	1899
OsHKT7 (Os04g51830)	CAD37197	Putative sodium transporter [*Oryza sativa* Japonica Group].	14/11/2006	500	AJ491853	5400
SbHKT7	XP_002448492	Hypothetical protein SORBIDRAFT_06g027900 [*Sorghum bicolor*].	13/07/2009	563	XM_002448447	1692
SbHKT8	XP_002457736	Hypothetical protein SORBIDRAFT_03g012590 [*Sorghum bicolor*].	13/07/2009	498	XM_002457691	1497
OsHKT9 (Os06g48800)	CAD37199	Putative sodium transporter [*Oryza sativa* Japonica Group].	02/07/2003	509	AJ491855	1557

Gene	Protein accessions	Description	Date	Length aa	Nucleotide Accession	Length (bp)
NHX						
AtNHX1	BAE98703	Na+/H+ exchanger [*Arabidopsis thaliana*].	27/07/06	538	AK226586	2346
ZmNHX1	NP_001105221	Na+/H+ antiporter NHX1 [*Zea mays*].	14/12/2007	540	NM_001111751	1623
TaNHX1	AAK76737	Na+/H+ antiporter [*Triticum aestivum*].	28/07/2001	546	AY040245	2017
TaNHX1_ LOC100125681	AAO33456	NHX1 [*Triticum aestivum*].	02/02/2003	204	AF472486	612
SlNHX1	CAC84522	Na+/H+ antiporter, isoform 1 [*Solanum lycopersicum*].	19/06/03	534	AJ306630	2122
AtNHX2	NP_001078110	NHX2 (SODIUM HYDROGENEXCHANGER 2); sodium ion transmembrane transporter/ sodium: hydrogen antiporter [Arabidopsis thaliana].	21/08/2009	421	NM_001084641	1486
ZmNHX2	NP_001105531	Na+/H+ antiporter NHX2 [*Zea mays*].	14/12/2007	540	NM_001112061	1623
SlNHX2	CAC83608	Na+/H+ antiporter, isoform 2 [*Solanum lycopersicum*].	19/06/03	531	AJ306631	2267
TaNHX2	AAK76738	Na+/H+ antiporter [*Triticum aestivum*].	27/02/2003	538	AY040246	2422
HvNHX2	AAO91943	Vacuolar Na+/H+ antiporter [*Hordeum vulgare*]	29/01/07	546	AY247791	1941
ZmNHX3	NP_001105222	Na+/H+ antiporter NHX3 [*Zea mays*].	14/12/2007	539	NM_001111752	1620
SlNHX3	CAK12754	(Sodium/potassium)/proton exchanger 3 [*Solanum lycopersicum*].	09/05/2006	537	AM261866	1614
HvNHX3	ABD62853	Na+/H+ antiporter [*Hordeum vulgare*].	17/04/2006	541	DQ372061	1794
AtNHX3	NP_200358	ATNHX3; sodium ion transmembrane transporter/ sodium:hydrgen antiporter [*Arabidopsis thaliana*].	21/08/2009	529	NM_124929	1861
AtNHX4	NP_187288	NHX4 (SODIUM HYDROGENEXCHANGER 4); sodium iontransmembrane transporter/ sodium:hydrogen antiporter [*Arabidopsis thaliana*].	21/08/2009	503	NM_111512	2207
HvNHX4	ABC42029	Na+,K+/H+ exchanger [*Hordeum vulgare*].	14/06/06	510	DQ314285	2117
SlNHX4	CAK12755	(Sodium/potassium)/proton exchanger 4 [*Solanum lycopersicum*].	19/04/2007	536	AM261867	1611
ZmNHX4	NP_001105943	Na+/H+ antiporter NHX4 [*Zea mays*].	14/12/2007	538	NM_001112473	1617
AtNHX5	BAE98578	Na+/H+ exchanger 5 [*Arabidopsis thaliana*].	27/07/06	521	AK226435	1786
ZmNHX5	NP_001105223	Na+/H+ antiporter NHX5 [*Zea mays*].	14/12/2007	545	NM_001111753	1638
ZmNHX6	NP_001105224	Na+/H+ antiporter NHX6 [*Zea mays*].	14/12/2007	541	NM_001111754	1626
AtNHX6	NP_178079	Sodium proton exchanger, putative (NHX6) [*Arabidopsis thaliana*].	21/08/2009	535	NM_106609	1943
OsNHX6	NP_001060571	Os07g0666900 [*Oryza sativa* Japonica Group].	08/06/10	536	NM_001067106	2227
AtNHX8	NP_172918	ATNHX8; lithium:hydrogen antiporter/ sodium ion transmembrane transporter/ sodium:hydrogen antiporter [*Arabidopsis thaliana*].	21/08/2009	756	NM_101333	2471
SOS						

Gene	Protein accessions	Description	Date	Length aa	Nucleotide Accession	Length (bp)
SlSOS1	CAG30524	Putative plasmalemma Na+/H+ antiporter [Solanum lycopersicum].	26/05/2005	1151	AJ717346	3823
AtSOS1	NP_178307	SOS1 (SALT OVERLY SENSITIVE); sodium:hydrogen antiporter	21/08/2009	1146	NM_126259	3682
AtSOS2	NP_198391	SOS2 (SALT OVERLY SENSITIVE 2); kinase/ protein kinase [*Arabidopsis thaliana*].	21/08/2009	446	NM_122932	1757
AtSOS3	NP_197815	SOS3 (SALT OVERLY SENSITIVE 3); calcium ion binding / calcium-dependent protein serine/threonine phosphatase [*Arabidopsis thaliana*].	21/08/2009	222	NM_122333	759
ARALYDRAFT_484121	XP_002875104	Hypothetical proteinARALYDRAFT_484121 [*Arabidopsis lyrata* subsp. *lyrata*].	11/06/2010	1135	XM_002875058	3537
HAK						
HvHAK1	CAC14883	Putative potassium transporter [*Hordeum vulgare*].	14 /11/2006	255	AJ297888	766
OsHAK2 (Os01g0935500)	NP_001045320	Putative HAK2 [*Oryza sativa* Japonica Group].	16/02/2008	783	NM001051855	2939
LOC100037514	ABF22603	HAK5 [*Solanum lycopersicum*].	23/01/2007	786	DQ489721	2361
HvHAK4	ABE99812	Potassium transporter HAK4 [*Hordeum vulgare*].	19/11/2009	785	DQ465924	2808
SbHAK4 (SORBIDRAFT_ 03g045180)	XP_002456904	Hypothetical protein SORBIDRAFT_03g045180 [*Sorghum bicolor*].	13/07/2009	783	XM_002456859	2765
OsHAK4 (Os08g0466200)	NP_001062000	Os08g0468200 [*Oryza sativa* Japonica Group].	08/06/2010	916	NM_001068535	2579
AtHAK5	NP_567404	HAK5 (HIGH AFFINITY K+ TRANSPORTER 5); potassium ion transmembrane transporter/ potassium: sodium symporter [*Arabidopsis thaliana*].	21/08/2009	785	NM_117416	2623
AlHAK5 (ARALYDRAFT_358974)	XP_002863132	Hypothetical protein ARALYDRAFT_358974 [*Arabidopsis lyrata* subsp. *lyrata*].	11/06/2010	645	XM_002863086	1938

Annex 2. List of primers designed from candidate genes of salt tolerance.

Gene	Nucleotide accession	Forward Primer	Reverse Primer	%GC (F /R)	℃Tm (F/R)	Length (bp) (F/R)	Product size
HKT							
AtHKT1	AK228564	CCATCACCGTCTCTTCCA	TTAGGAGCCAGATGAGA	56/47	54.9 /48.7	18 /17	500
PtHKT1	XM_002325193	ACTTTCCAGGGCTAGAGA	GAGACTGTGATAGCGAGA	50/50	52.7/51.8	18/18	500
TaHKT1	DQ009003	GTCGCTGAAACCAAGCA	TAGTGAGAAGAGCACGA	53 /47	53.8/49.7	17/17	501
McHKT1	AF367366	CATAGGAAGAGGGAGCAA	CACCACAAACTGAACCA	50/47	51.8/50.3	18/17	500
OsHKT1	NM_001049783	CGCAGTAGGTTTCAGTCA	CCAGTGGACAACCCAA	50 /56	53.3/50.5	18/16	501
EcHKT1	AF176035	CCTGATTCTCATCCCTCA	CAGGTAGGAAAGCTGTGA	50/50	51.5/52.5	18/18	500
HvHKT1	AM000056	GTACAGGGTTGAAGAGGA	GGCCATCAAACACGGAA	50/53	51.7/53.4	18/17	500
PaHKT1	AB234304	GCACCACCAAACAGAGA	TTGCTCCACTAGGATCCA	53/50	52.6/53.2	17/18	500
SsHKT1	AY530754	CCATTTTTCGGCCTCGAA	CCTAGCATCCATATCGA	50/47	54.7/47.7	18/17	500
SmHKT2	XM_002976817	CTGTGGCTTTTCCCCAA	CCACGAAGTAGATCGAGA	53 /50	52.7/51.8	17/18	500
McHKT2	AY231175	ACTACCACAACCACCACA	GAACCCGAAAACGGTCA	50/53	54.4/53.1	18/17	500
EcHKT2	AF176036	CCTCTGTCTCCTTGGCTA	GGTCGACACAGTGGTGA	56/59	53.8/54.4	18/17	499
OsHKT2	AB061313	CACCCATTCTGGATCCAA	TGTACCAGAGAACCAGCA	50/50	52.8/53.8	18/18	500
OsHKT3	AJ491820	TGCTACTCATTGGCCAGA	TTGAAGCCGAGTGGTGA	50/53	54.3/53.8	18/17	501
SbHKT3	XM_002438915	GGAGAGAAGCTCACCAA	TGAGGAGAACAAGGAGCA	53/50	51.1/53.8	17 /18	500
OsHKT4	AJ491816	TTGGGAGAAAGCTCAGCA	GTACCTGGTTTTGCGACA	50/50	54.9/54.3	18/18	501
OsHKT5	AJ506745	ACTTCTGTGTCCACAGCA	ATCATTTACCGGGGTGA	50/47	54.8/50.2	18/17	500
OsHKT6)	AJ491818	TCTGAACTCCGCATGGAA	AGCTCCACTCCAAAGAGA	50/50	54.7/53.5	18/18	500
SbHKT6	XM_002451593	CTGAACTCCACTTGGAA	GAGCTCCATTCTAGAGA	47/47	48.9/47.0	17/17	501
OsHKT7	AJ491853	CCATTTCCCAGTTCGTGA	TTCCTAAACCAGGCTGCA	50/50	53.6/54.8	18/18	501
SbHKT7	XM_002448447	GGGGAGAAGCTGTCCAA	TGCCTCCTTTCATGCTGA	59/50	53.9/54.9	17/18	500
SbHKT8)	XM_002457691	TCCACTTCACCTTCACCA	AATGGAGAGCCTGAGGAA	50/50	54.0/53.5	18/18	501
OsHKT9)	AJ491855	CAAGAGGAGCTGCCACA	GGAGGTTAGCCCTGCAA	59/59	54.8/54.4	17/17	501

Gene	Nucleotide accession	Forward Primer	Reverse Primer	%GC (F /R)	℃Tm (F/R)	Length (bp) (F/R)	Product size
NHX							
AtNHX1	AK226586	GCAGTGAGCTCAATCCTA	GTCGCATGAAGGAGTCA	50/53	52.6/52.1	18/17	501
ZmNHX1	NM_001111751	AGCCAAGATGAGACACCA	GACTTACCTGGGGTGTCA	50/56	54.2/54.3	18/18	501
TaNHX1	AY040245	ACCGTGTTCTTCTGTGGA	GTCTTGGTGGGATGAGGA	50 /56	54.5/54.4	18/18	500
TaNHX1	AF472486	AAGCAATTCCTCCGCAA	GCCACTCAGATCCAGCA	47/59	52.4/54.4	17/17	500
SlNHX1	AJ306630	GGGGTGGTAAATGATGCA	CTTCCAACCAGAACCAGA	50/50	53.5/52.6	18/18	501
AtNHX2	NM_001084641	AGAAGCATCAGAGCGAGA	TCATGGCTGCTTCTCTCA	50/50	54.1/54.3	18/18	500
ZmNHX2	NM_001112061	TGATGTTGGTCCACTCGA	GAAAGACAGGGTGGCAA	50/53	54.2/52.4	18/17	501
SlNHX2	AJ306631	CCTTGGTGGAGTTACGTA	CGACACAAATCGCTGTGA	50/50	52.2/54.8	18/18	500
TaNHX2	AY040246	ATGACCACCAAGGGGAA	GACGTATGCACTAAGCA	53/47	52.7/49.8	17/17	500
HvNHX2	AY247791	TCATTCTGCTCTGCACCA	GCACTAAGCAATCCAGCA	50/50	54.9/54.3	18/18	500
ZmNHX3	NM_001111752	TCTATCTACTGCCGCCAA	CAGGACCATAATGGCCAA	50/50	53.8/53.2	18/18	501
SlNHX3	AM261866	AAGAGTCACCACCAAGCA	AAGTGGCACGATGAAGGA	50/50	54.7/54.9	18/18	500
HvNHX3	DQ372061	TTCCAGGGTCACAACCAA	AAGGACTTTGGGCTGGA	50/53	54.6/53.1	18/17	500
AtNHX3	NM_124929	GGGGTTGAGTGCTAGAGA	CTCCGCAATAAAGGACA	56/47	54.1/49.6	18/17	501
AtNHX4	NM_111512	CCTGGTCAGTCGATTGGA	GCGGAACACTATTCTCA	56 /47	54.9/49.3	18/17	499
HvNHX4	DQ314285	GCAGAGTTTGAGCACGA	CCACAGAGGCAAGACGA	53/59	53.2/54.4	17/17	500
SlNHX4	AM261867	GGATACCGAGAAGTGGAA	TGGTGCGGTAAGTAGCA	50/53	51.9/53.5	18/17	501
ZmNHX4	NM_001112473	GTAGTGAACGATGCCACA	TCTGCCTAGCATGACCA	50/53	53.7/52.7	18/17	500
AtNHX5	AK226435	CAGGTTAAGCAGCAGCAA	CTCCAAAGACCAAAGCA	50/47	54.6/50.0	18/17	501
ZmNHX5	NM_001111753	AGGAGAACAGATGGCTCA	CGAAGACTTGGAGTGCA	50/53	53.5/52.4	18/17	500
ZmNHX6	NM_001111754	CAATATTGGAGCCCTCGA	AATGAGAGGGTCGCGAA	50/53	52.7/53.6	18/17	500
AtNHX6)	NM_106609	TAAATCCGTCGAGGCGTA	CAGGATCAGTGGCTGAGA	50/56	54.3/54.5	18/18	500
OsNHX6	NM_001067106	GAGTTTGCCAGTGACAGA	CGCCAGTAGTAGTGGACA	50/56	53.5/54.6	18/18	501
ATNHX8	NM_101333	CGTACAAATCACCGGAGA	CCAAGCTCCTTTAGCAA	50 /47	53.1/49.7	18/17	501

Gene	Nucleotide accession	Forward Primer	Reverse Primer	%GC (F /R)	℃Tm (F/R)	Length (bp) (F/R)	Product size
SOS							
SlSOS1	AJ717346	GAAAGCGAGGAAGAAGGA	TGTAAGGCTTTCCCACCA	50/50	52.9/54.4	18/18	500
AtSOS1	NM_126259	TGCTCTTGGATCTCTCGA	CGCCAGACCAATGCCTA	50/59	53.4/54.9	18/17	501
AtSOS2	NM_122932	GGTTACGATGGTTCAGCA	CAGCCTCAATGTTAGCA	50/47	53.6/49.9	18/17	501
AtSOS3	NM_122333	CCACCGGGATATGAGGA	GAGCGATGGATTCAAGGA	59/50	52.7/53.1	17/18	499
ARALYD	XM_002875058	ATGCTTGATGAGGGCAGA	TTCAAACCGGTCTGGACA	50/50	54.6/54.8	18/18	501
HAK							
HvHAK1	AJ297888	TGAGCTTGGCTTTCCAGA	AGGACCACTTGTGTCTGA	50/50	54.8/53.7	18/18	500
OsHAK2	NM_001051855	GGTGGATCTATGGACAGA	TCATGAAAGGCAGGGAGA	50/50	51.1/53.8	18/18	500
LOC1000	DQ489721	TCAGACAGGGTTTGGAGA	TTGGCCACTGTAAGCTGA	50/50	53.3/54.8	18/18	501
HvHAK4	DQ465924	GGTGGATCTATGGGCAGA	TCATGAAAGGCAGGGAGA	56/50	54.2/53.8	18/18	500
SbHAK4	XM_002456859	GCAGTTGTTGGTAGCCAA	CAAGGCTTGGACCCAGA	50/59	54.5/54.2	18/17	501
OsHAK4	NM_001068535	ATGTTGAAGCCGGACAGA	GAGTCCATCAATCGCTGA	50/50	54.9/53.1	18/18	501
AtHAK5	NM_117416	GGTCACTTCAGTGTTCGA	CTATGGAGCCAAAGACGA	50/50	53.2/52.7	18/18	500
AlHAK5	XM_002863086	GGAAGCCATGTTTGCTGA	AGAACGTAGCGATCCACA	50/50	54.5/54.5	18/18	501