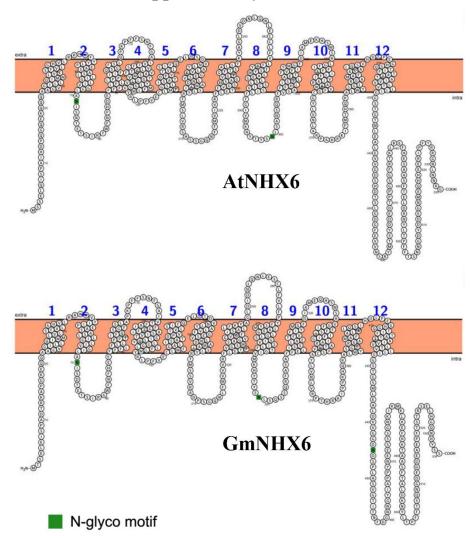
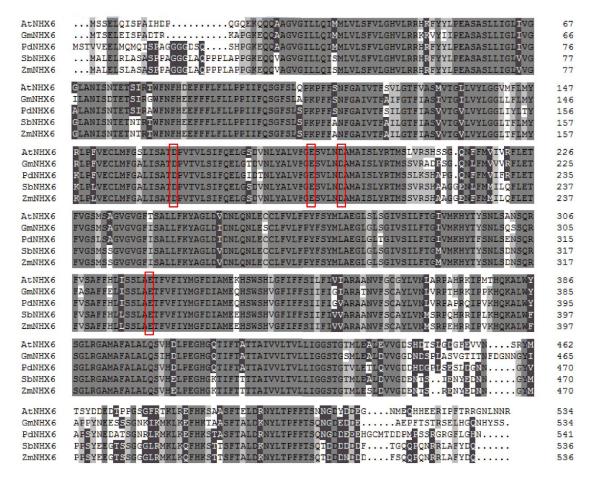


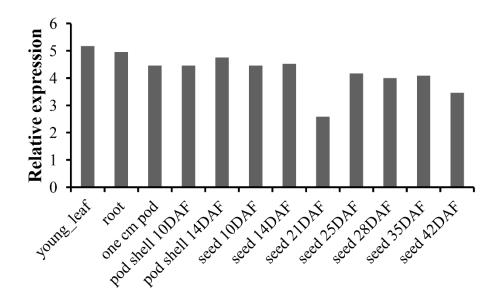
Supplementary Material



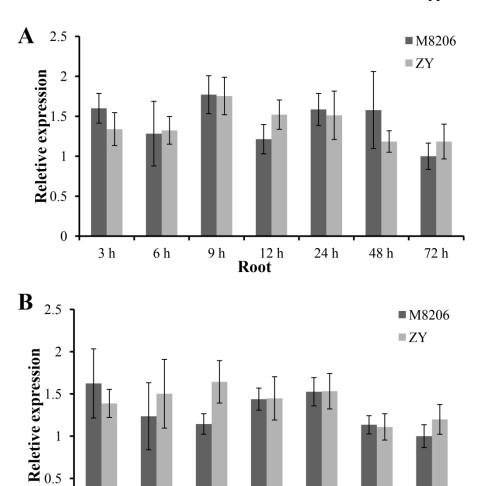
Supplementary Figure 1. The putative topology of AtNHX6 and GmNHX6 using Protter database. The topology of GmNHX6 of soybean shows 12 transmembrane domains, similar with that of AtNHX6 of *Arabidopsis*.



Supplementary Figure 2. Multiple sequence alignment of the deduced amino acid sequence of GmNHX6 and four NHX6 sequences from different plant species. Red boxes are the conserved acidic residues of NHX6. Four other NHX6 sequences include AtNHX6 (NP_178079.2) from *Arabidopsis thaliana*, PdNHX6 (XP_008811204.1) from *Phoenix dactylifera*, SbNHX6 (XP_021308686.1) from *Sorghum bicolor*, ZmNHX6 (ACG45634.1) from *Zea mays*.



Supplementary Figure 3. Tissue expression patterns of GmNHX6 in soybean. The raw reads from RNA-seq data of $Glycine\ max$ (www.soybase.org) were transformed to log_2 (FPKM + 1).



Supplementary Figure 4. Relative expression of GmNHX6 in two soybean varieties under normal condition (0 mM NaHCO₃). The relative expression of GmNHX6 in soybean root tips (0 ~ 2 mm) was quantified by qRT-PCR, using GmUKN1 as the reference gene and M8206 under 0 mM NaHCO₃ at 72 h (with lowest expression level of GmNHX6) as the control (relative expression = 1). Data represents the mean \pm standard deviation of three replicates. There was no difference between two soybean varieties by the two-tailed Student's t tests.

12 h

Leaf

24 h

48 h

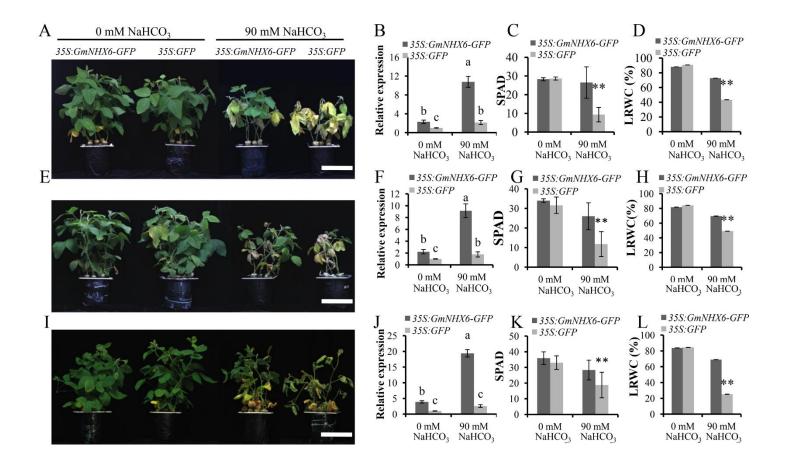
72 h

9 h

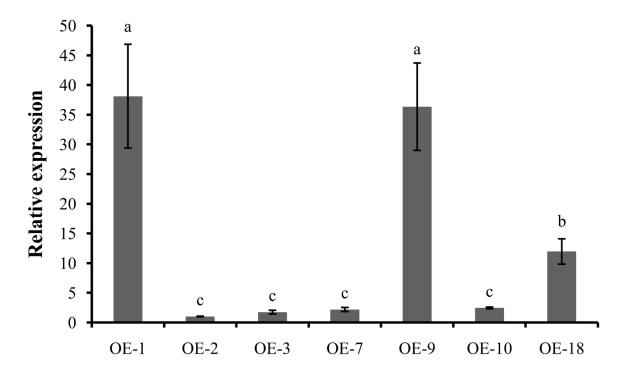
0

3 h

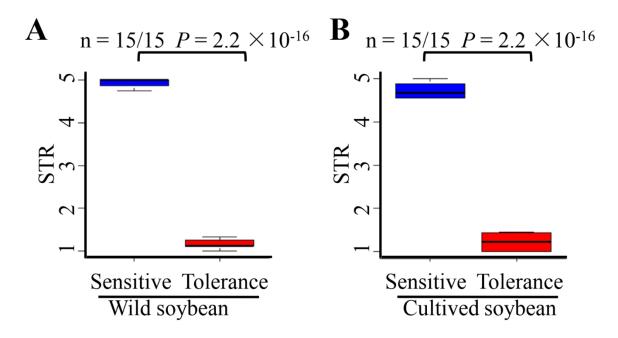
6 h



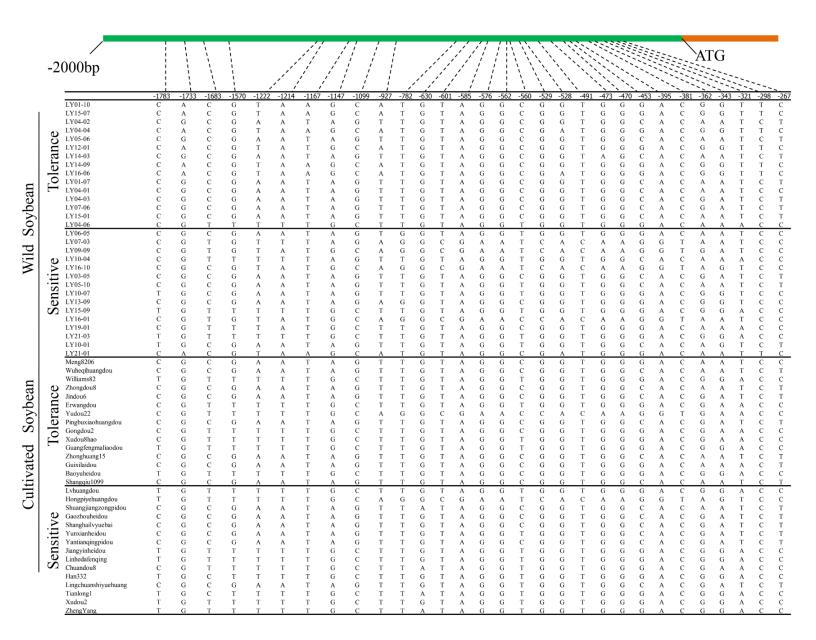
Supplementary Figure 5. Three biological replications for salt tolerance analyses of transgenic soybean composite plants. The phenotypes ($\bf A$, $\bf E$, $\bf I$), relative expression ($\bf B$, $\bf F$, $\bf J$), SPAD values for chlorophyll content ($\bf C$, $\bf G$, $\bf K$) and leaf relative water content ($\bf D$, $\bf H$, $\bf L$) of transgenic soybean composite plants under control (0 mM NaHCO₃) or alkali stress (90 mM NaHCO₃) were recorded 7 d after treatment. Bar, 12 cm. Soybean variety of "TianLong1" was used. Data represents the mean \pm standard deviation of technical repeats and each repeat has at least five independent plants for each genotype (n = 5). Same letters above bars indicate no significant differences according to Duncan's multiple range test at 0.05 level. **P < 0.01 by two-tailed Student's t tests.



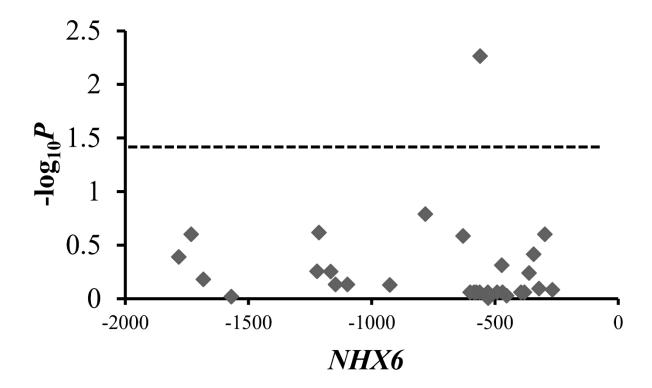
Supplementary Figure 6. The relative expression level of GmNHX6 in transgenic Arabidopsis lines. The relative expression of GmNHX6 in transgenic Arabidopsis lines was quantified by qRT-PCR, using AtACTIN7 as the reference gene and OE-2 line (with lowest expression level of GmNHX6) as the control (relative expression = 1). Data represents the mean \pm standard deviation of three replicates.



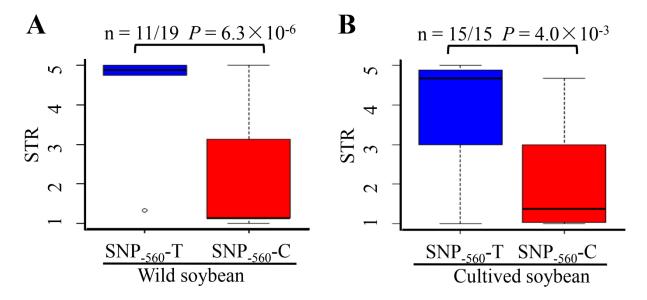
Supplementary Figure 7. Boxplots of alkaline salt tolerance in 30 wild soybean accessions and 30 cultivated soybean accessions. STR: Sodic (alkaline salt) Tolerance Rating.



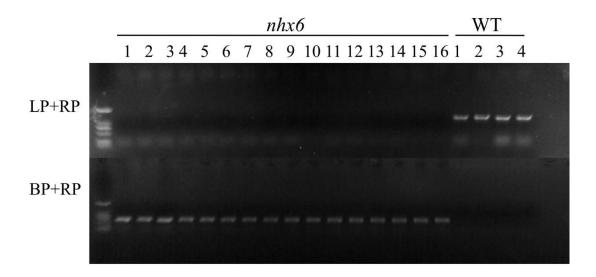
Supplementary Figure 8. Schematic graph shows the positions and polymorphisms of SNPs in the 2-kb promoter region of soybean *NHX6* gene.



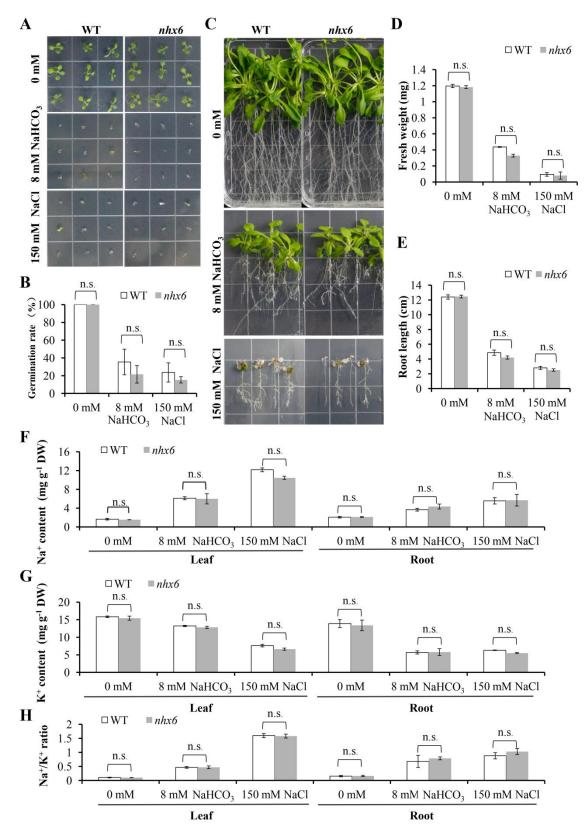
Supplementary Figure 9. Association of sodic tolerance rating (STR) with the SNPs in soybean *NHX6* promoter region. The x-axis represents the physical position of each SNP in base pairs (bp) upstream of the start codon of *GmNHX6* gene. The black dotted line represents the significance threshold of significant associations by mixed model, which is $-\log_{10}(1/30) = 1.48$.



Supplementary Figure 10. Boxplots of sodic tolerance rating (STR) for the two alleles of SNP. $_{560}$ in wild (A) and cultivated (B) soybean accessions.



Supplementary Figure 11. Verification of homozygous *nhx6* **mutant lines by the method of three primers.** LB: Left border primer of the T-DNA insertion, TCAAACAGGATTTTCGCCTGCT, RP: Right genomic primer, TTCCGTCAACAACATTAGCC, LP: Left genomic primer, TCCTCGGAAAGGAAGAAGA.



Supplementary Figure 12. Phenotypes of *Arabidopsis nhx6* mutant and wild type in response to NaHCO₃ and NaCl treatment. (A) Seed germination assay of different *Arabidopsis* lines on 1/2 MS

medium or media supplied with 8 mM NaHCO₃ or 150 mM NaCl, respectively. WT: wild type. (**B**) Seed germination rates of different *Arabidopsis* lines. (**C**) Phenotypes of *Arabidopsis* lines on 1/2 MS medium or media with 8 mM NaHCO₃ or 150 mM NaCl, respectively. (**D**, **E**) Fresh weight and root length of *Arabidopsis* lines. (**F**) Na⁺ content, (**G**) K ⁺ content, (**H**) Na⁺/K ⁺ ratio in the leaves and roots of each line receiving different treatments for 10 days, respectively. Data represents the mean \pm standard deviation of three biological replications and each repeat contained four independent plants per line for each treatment (n = 12). n.s.: no significant difference by two-tailed Student's t tests.

dOCT

<u>uoc i</u>
-2000(M) TITCACTCTGCACTGTAAGCCGGTGAAGTGCAGTAATTGTGGTAAGCTTTAGTTTTTATCTCTCGGATCTTATACATAAGTTCCTGCATTTACCAAGGGA
$-2000(ZY)\ TTTCACTCTGCACTGTAAGCCGGTGAAGTGCAGTAATTGTGGTAAGCTTTAGTTTTTATCTCTCGGATCTTATACATAAGTTCCTGCATTTACCAAGGGA$
-1900(M) ATGTGCAGATAAAACGCATACAACGCAACACAAAAGGTTCAAAACGTTGAAGAAGGAAAATAAACAGTATAAACTACTTCACCCACGCACATACCTCGACC
-1900(ZY) ATGTGCAGATAAAACGCATACAACGCAACACAAAGGTTCAAAACGTTGAAGAAGGAAAATAAACAGTATAAACTACTTCACCCACGCACATACCTCGACG
ABRE TGACG-motif ARE
-1800(M) GCACGTGTGGCAGCAGCGAGATCCTCCTTCGCCTTAGTTCCTGAGTTGCTCTGGAATCAGAGAATCGGAGCCATCAGAGAGCGAAATTGACGGAACAAAA
$-1800(ZY)$ $G_{ m CACGTG}^{ m TGGCAGCAGCGAGGATCCTCCTTCGCCTTAGTTCCTGAGTTGCTCTGGAATCAGAGAATCGGAGCCATCAGAGAGCGAAAT _ GACGGAACA_AAA$
-1700(M) CCAAATCAATGAAATCACTGGACGCTAATTGAAGAGATATTAACACCTGCTGGCGAACGGCGGCGAGAATTCCGGCATCGACGGTGCGAATCTCGTTCTG
-1700(ZY) CCAAATCAATGAAATCATGGACGCTAATTGAAGAGATATTAACACCTGCTGGCGAACGGCGGCGAGAATTCCGGCATCGACGGTGCGAATCTCGTTCTG
MYB
-1600(M) GATTTTCTGCATCAGCGGCTCCACGCCGGAGAGAGACGCCTCTGCGTAACAACAGGCAAATTTGAGAGAAGTTAAGGTTTCGATTAGAAACGCTTGCGAT
-1600(ZY) GATTTTCTGCATCAGCGGCTCCACGCCGGATAGAGACGCCTCTGCGTAACAACACGCAAATTTGAGAGAAGTTAAGGTTTCGATTAGAAACGCTTGCGAT
MYB recognition site
$-1500 (M) TGCGAGTGAGATCGGGAATTGTGGCGTA \\ \underline{CGTTG}GGGAACATCTGGTGATGTACTCCAGAGCGCTTGACTTGTCCATTGGTGAGAATGAACGCAATCAC \\ \\ -1500 (M) TGCGAGTGAGATCGGGAATTGTGGCGTA \\ \underline{CGTTG}GGGAACATCTGGTTGATGTACTCCAGAGCGCTTGACTTGGTCCATTGGTGAGAATGAACGCAATCAC \\ \\ -1500 (M) TGCGAGTGAGATCGGGAATTGTGGCGTA \\ \underline{CGTTG}GGGAACATCTGGTTGATGTACTCCAGAGCGCTTGACTTGGTCCATTGGTGAGAATGAACGCAATCAC \\ \\ -1500 (M) TGCGAGTGAGATCGGGAATGAACGCAATCAC \\ \underline{CGTTG}GGGAACATCTGGTTGATGTACTCCAGAGCGCTTGACTTGGTCCATTGGTGAGAATGAACGCAATCAC \\ \\ -1500 (M) \underline{CGTTG}GGGAACTAC \\ \underline{CGTTG}GGGAACATCAC \\ \underline{CGTTG}GGGAACAC \\ \underline{CGTTG}GGGAACATCAC \\ CG$
$\hbox{-}1500(ZY)\ TGCGAGTGAGATCGGGAATTGTGGCGTA \hbox{-} CGTTC \hbox{-}GGGAACATCTGGTTGATTGACTCCAGAGCGCTTGACTTGTCCATTGGTGAGAATGAACGCAATCAC$
$-1400 (M) \\ A A A A G C A T G A A C C T G A T G T G T G T T A A T G T T C A G A T C T T A A C T T G C G T G T C G C A C T T T A T G C T C T C A A T T G C G G A T C A A T T G C G C A T T G C G C A T T A T G C C T C A T T G C G C A T T G C G C G A T C A A T T G C C T C A T T G C C G C A T T A T G C C T C A T T G C C C A T T G C C C C C C C C C C C C C C C C C$
-1400(ZY) AAAAGCATGAACCGTGATGGTGATTGTGATTAATGTTCAGATCTCTAAACTTGCGTGTTCGGCACCATTTATGCTCTTCAATTTGTGGCGGATCCAATTG
TGACG-motif
-1300(M) CAGATGTGTGTGCTGTGACGCTATTGTTAGTGCTCCTCCTTTTCGTCGTCTTCAACCTCTGGACACTTCTGTAATATTTTATATATA
* *
-1300(ZY) CAGATGTGTGTGTGACTGTGACGCTATTGTTAGTGCTCCTCCTTTTCGTCGTCTCTAACCTCTGGACACTTCTGTTATATTTTTTATATATA
ABRE ERE
-1200(M) CTATTTTATATATGGATGAAAACTAAAAATATGTTTTTTACTGAAAAAATCACGTAAGTATGAATTTCATTCTTGTAAAAACTTTTTA
-1200(ZY) CTATTTTATATATGGATGAAAACTAAAATATGTTTTTTACTGAAAAATCACGTGAGTATGAATTTCATTCTTGTAAAACTTTTTA <mark>TTTAAAAT</mark> ATATCTA
W box
-1100(M) GGTTCGATTGATTCTCATGTCAATGTATGATTCTATTTCTTATGTGTCACTCCCTTGTTGACCAAATTCTTGACTACTCTCTTTTTCTCATAATCAAAAC
-1100(ZY) GCTTCGATTGATTCTCATGTCAATGTATGATTCTATTTCTTATGTGTCACTCCCTTGTTGACCAAATTCTTGACTACTCTCTTTTTCTCATAATCAAAAC
-1000(M) TTCCCAAACTTGAAAAAAAGAAAGGCCTCAAATTTGCAAATTGATTTAATGCTCCTTCATCATAAATTTGACATATAGTAACAAAAAGGAAAAAGGCCTTTG
-1000(ZY)TTCCCAAACTTGAAAAAAAGAAAGGCCTCAAATTTGCAAATTTGATTTAATGCTCCTTCATCATAAATTTGACATATAGTAACAAAAAGAAAAAGAAAAAGGCTTTG
Box 4 ERE
-900(M) GGGTTAGTTTTATAGTTGACTTTTTATTACGCTAAAATTATATATTTTATTATTATTATTTAT
-900(ZY) GGGTTAGTTTATAGTTGACTTTTTATTACGCTAAAATTAATATTTAAAATTTTCTTATTATGTATTATTATTTAT
MYC
-800(M) TAAAAACATGTTTCATATTAATTTTTTTTTTTTTTATCCTTTACGTATTTATGACAATATGTTAAGGTTGAAAATTAATAAAGGTGAACATGTCTTCCTATAC
-800(ZY) TAAAAACATGTTTCATATTAATTTTTT-CTTTATTCCTTTACGTATTTATGACAATATGTTAAGTTGAAAATTAATAAAGGTGAACATGTGTTCCTATAC
MYB AT1-motif WUN-motif
-700(M) ATATATTACATAAAGGAAGACCTTCATCTCTCATCATCTAAAAATTAGCTTCAACCATATTAAAATGTTTGGAAAAAAAA
*
-700(ZY) ATATATTACATAAAGGAAGACCTTCATCTCTCATCATCTAAAAATTAGCTTCAACCATATTAAAATGTTTGAAAAAAAA
TGTCAT element Skn-1 motif TCATTT element TTGTCA element
-600(M) TICTITTITATAATAACATTIATIGTATTGACAATGUIGICALTIATIAACAFITATITATCAAATTATATGGATATAACTTIGATCAATTATATACATTG
-600(ZY) TICTITITIATAATAACATTIATIGACAATGTTGTTATTTATTTAACATTTATTTATCAAATTATATGGATATAACTTTGATCAATTATTACATTG
Box 4 GA-motif TC-rich repeats TATA
-500(M) TIGATAAACTTTICAATTAGATAATATGTAGATTAAATTAA
** -500(ZY) TTGATAAACTTTTCAATTAGATAATATGTAGATTTAAATTAATGATTGAT
Box 4 ABRE Box 4 Box 4
400(M) AAGAAATATAATTCAAATTCATTTAATATCTTATTCACATAGATTAAAAATCAATC
-400(ZY) AAGAAATATAATTCAAATTCATT <mark>TAATAT</mark> CTTATT <mark>CACGT</mark> AGATTAAAATCAATCTAGTATAAATGTTTTTAAATATAAAATATATAGAATATTTTGATT
AT-rich element
-300(M) ATCTATTTGTAGTTTAATTTTAATAAAATTTGATATAAACGATCTTAATAATTTATAAATATAATTTAAGGTGGATTA <mark>ATAAAAATCAAATTTT</mark> ATAATA
-300(ZY) ATCTATTTGTAGTTTAATTTTAATAAAAATTTGACATAAATGATCTTAATAATTTATAAATATAATTTAAGGCGGATTAATAAAAAAAA
Box 4 Box 4 ARE ATCT-motif
-200(M) TTAATATATTTAATATTCCGGAGTCCAAGAGGTGAGAAGAACAGAGAACAGAATTGAAAAACTCACAAACCAGATAGAGAGAATCTAATCAGAACAGAG
-200(ZY) TTAATATATTTAATATTCCGGAGTCCAAGAGGTGAGAAGAACAGAGAACAGAATTGAAAAACTCACAAACCAGATAGAGAGAATCTAATCAGAACAGAG
GCN4_motif
-100(M) TTCTGTTCCTTTTCTTGATTCCCGAAGAACTCAACGATCCAATCGCAACCGGAAGCAATCGATCAGTGAGTCAATTGCTTCCGGCAAGGAGAACAATA

Supplementary Figure 13. The cis-element analysis of the 2-kb promoter region upstream of start codon ATG of GmNHX6 from soybean varieties of M8206 (M) and ZY.

Supplementary Table 1. List of 60 soybean accessions with extreme alkaline salt tolerance (sodic tolerance rating, STR).

Accession	Category	Average STR	Alkaline salt tolerance
LY01-10	Wild soybean	1.00	Tolerant
LY15-07	Wild soybean	1.00	Tolerant
LY04-02	Wild soybean	1.13	Tolerant
LY04-04	Wild soybean	1.13	Tolerant
LY05-06	Wild soybean	1.13	Tolerant
LY12-01	Wild soybean	1.13	Tolerant
LY14-03	Wild soybean	1.13	Tolerant
LY14-09	Wild soybean	1.13	Tolerant
LY16-06	Wild soybean	1.13	Tolerant
LY01-07	Wild soybean	1.13	Tolerant
LY04-01	Wild soybean	1.25	Tolerant
LY04-03	Wild soybean	1.25	Tolerant
LY07-06	Wild soybean	1.25	Tolerant
LY15-01	Wild soybean	1.25	Tolerant
LY04-06	Wild soybean	1.33	Tolerant
LY06-05	Wild soybean	4.75	Sensitive
LY07-03	Wild soybean	4.75	Sensitive
LY09-09	Wild soybean	4.75	Sensitive
LY10-04	Wild soybean	4.88	Sensitive
LY16-10	Wild soybean	4.88	Sensitive
LY03-05	Wild soybean	5.00	Sensitive
LY05-10	Wild soybean	5.00	Sensitive
LY10-07	Wild soybean	5.00	Sensitive
LY13-09	Wild soybean	5.00	Sensitive
LY15-09	Wild soybean	5.00	Sensitive
LY16-01	Wild soybean	5.00	Sensitive
LY19-01	Wild soybean	5.00	Sensitive
LY21-03	Wild soybean	5.00	Sensitive
LY10-01	Wild soybean	5.00	Sensitive
LY21-01	Wild soybean	5.00	Sensitive
Meng8206	Cultivated soybean	1.00	Tolerant
Wuheqihuangdou	Cultivated soybean	1.00	Tolerant
Williams82	Cultivated soybean	1.00	Tolerant
Zhongdou8	Cultivated soybean	1.00	Tolerant

Accession	Category	Average STR	Alkaline salt tolerance
Jindou6	Cultivated soybean	1.00	Tolerant
Erwangdou	Cultivated soybean	1.00	Tolerant
Yudou22	Cultivated soybean	1.05	Tolerant
Pingbuxiaohuangdou	Cultivated soybean	1.20	Tolerant
Gongdou2	Cultivated soybean	1.16	Tolerant
Xudou8	Cultivated soybean	1.44	Tolerant
Guangfengmaliaodou	Cultivated soybean	1.33	Tolerant
Zhuanghuang15	Cultivated soybean	1.37	Tolerant
Guixilaidou	Cultivated soybean	1.44	Tolerant
Baoyuheidou	Cultivated soybean	1.44	Tolerant
Shangqiu1099	Cultivated soybean	1.44	Tolerant
Lvhuangdou	Cultivated soybean	4.55	Sensitive
Hongpiyehuangdou	Cultivated soybean	4.56	Sensitive
Shuangjiangzongpidou	Cultivated soybean	4.56	Sensitive
Gaozhouheidou	Cultivated soybean	4.56	Sensitive
Shanghailvyuebai	Cultivated soybean	4.56	Sensitive
Yunxianheidou	Cultivated soybean	4.56	Sensitive
Yantianqingpidou	Cultivated soybean	4.67	Sensitive
Jiangyinheidou	Cultivated soybean	4.67	Sensitive
Linhedafenqing	Cultivated soybean	4.70	Sensitive
Chuandou8	Cultivated soybean	4.78	Sensitive
Han332	Cultivated soybean	4.89	Sensitive
Lingchuanshiyuehuang	Cultivated soybean	4.89	Sensitive
Tianlong1	Cultivated soybean	5.00	Sensitive
Xudou2	Cultivated soybean	5.00	Sensitive
ZhengYang	Cultivated soybean	5.00	Sensitive

Supplementary Table 2. Primers used for qRT-PCR, gene cloning, and vector construction.

Primer sets	Forward primer sequences	Reverse primer sequences	Purpose	Amplification Efficiency (Mean±SD)
q- GmNHX6	TGCTGCCTCTTTTACTGCATTAGAT	ACCATGAAGTTCCGATCTTGTAGAA	qRT-PCR	1.01 ± 0.01
GmUKN1	TGGTGCTGCCGCTATTTACTG	GGTGGAAGGAACTGCTAACAATC	qRT-PCR	0.99 ± 0.02
AtACTIN7	CCTCAGCACCTTCCAACAGATG	AACTCACCACCACGAACCAGAT	qRT-PCR	1.02±0.01
c-GmNHX6	GTGAGTCAATTGCTTCCGGC	GCACAAGTAATGCGCACCAA	Gene cloning	
c-GmNHX6-promoter	CTTCACCCACGCACATACCT	GAGGACGAAGGACAAGACCAG	Promoter cloning	
v-pBinGFP4-GmNHX6	TCCCCGGGATGAC CTCGGAGTTAGAAAT	CGCGGATCCTGATGAA TAGTGGTTCTGACC	Vector construction	
v-pCAMBIA3301-GmNHX6	tctagaggatctcgaggcgcgcc ATGACCTCGGAGTTAGAAAT	attcgagctcactagttaattaa TGATGAATAGTGGTTCTGACC	Vector construction	
v-pAN580-GmNHX6	CTAGCTAGCATGACC TCGGAGTTAGAAATTTCTC	CCCCCGGGTGATGA ATAGTGGTTCTGAC	Vector construction	
c-GmNHX6-promoter-LUC	ctatagggcgaattgggtacc CTTCACCCACGCACATACCTC	atcgataccgtcgacctcgag TATTGTTCTCCTTGCCGGAAGC	Vector construction	

Supplementary Table 3. The summary of the Ct values of the reference gene GmUKN1 for qRT-PCR across all samples in this study.

Minimum	Maximum	Mean	Standard Deviation
19.99	20.99	20.53	0.25

Note: A total of 540 Ct values from 60 samples × 3 biological replications × 3 technical repeats.

Supplementary Material

Supplementary Table 4. Regional association study of the alkaline salt tolerance.

SNP (Chr_bp)	Position ^a	Allele (major: minor)	MAF ^b (%)	P-value (MLM ^c)
9_1430060	-267	T:C	38.33%	0.7210
9_1430091	-298	C:T	11.67%	0.2249
9_1430114	-321	T:A	35.00%	0.7803
9_1430136	-343	A:G	38.33%	0.4516
9_1430155	-362	A:G	40.00%	0.7390
9_1430174	-381	C:T	10.00%	0.8274
9_1430188	-395	A:G	10.00%	0.8274
9_1430246	-453	C:G	40.00%	0.8948
9_1430263	-470	G:A	10.00%	0.8274
9_1430266	-473	G:A	11.67%	0.5628
9_1430284	-491	T:C	10.00%	0.8274
9_1430321	-528	G:A	15.00%	0.9868
9_1430322	-529	G:C	10.00%	0.8274
9_1430353	-560	C:T	43.33%	0.0055
9_1430355	-562	G:A	10.00%	0.8274
9_1430369	-576	G:A	10.00%	0.8274
9_1430378	-585	A:G	10.00%	0.8274
9_1430394	-601	T:C	10.00%	0.8274
9_1430423	-630	G:A	6.67%	0.2558
9_1430575	-782	T:G	13.33%	0.1598
9_1430720	-927	T:A	23.33%	0.8753
9_1430892	-1099	G:C	50.00%	0.6015
9_1430940	-1147	A:G	50.00%	0.6015
9_1430960	-1167	T:A	10.00%	0.5348
9_1431007	-1214	A:T	35.00%	0.2001
9_1431015	-1222	A:T	46.67%	0.5750
9_1431363	-1570	G:T	35.00%	0.9620
9_1431476	-1683	C:T	36.67%	0.6813
9_1431526	-1733	G:A	11.67%	0.2249
9_1431576	-1783	C:T	25.00%	0.4274

^aPosition: position from the start codon; ^bMAF: minor allele frequency; ^cMLM: Mixed linear model.