

Genome-wide identification and characterization of *NHL* gene family in response to alkaline stress, ABA and MEJA treatments in wild soybean (*Glycine soja*) (#75299)

1

First revision

Guidance from your Editor

Please submit by **3 Nov 2022** for the benefit of the authors .



Structure and Criteria

Please read the 'Structure and Criteria' page for general guidance.



Raw data check

Review the raw data.



Image check

Check that figures and images have not been inappropriately manipulated.

Privacy reminder: If uploading an annotated PDF, remove identifiable information to remain anonymous.

Files

Download and review all files from the [materials page](#).

1 Tracked changes manuscript(s)
1 Rebuttal letter(s)
9 Figure file(s)
3 Table file(s)
8 Raw data file(s)



Structure and Criteria

Structure your review

The review form is divided into 5 sections. Please consider these when composing your review:

1. BASIC REPORTING
2. EXPERIMENTAL DESIGN
3. VALIDITY OF THE FINDINGS
4. General comments
5. Confidential notes to the editor

You can also annotate this PDF and upload it as part of your review

When ready [submit online](#).

Editorial Criteria

Use these criteria points to structure your review. The full detailed editorial criteria is on your [guidance page](#).

BASIC REPORTING

- Clear, unambiguous, professional English language used throughout.
- Intro & background to show context. Literature well referenced & relevant.
- Structure conforms to [Peerj standards](#), discipline norm, or improved for clarity.
- Figures are relevant, high quality, well labelled & described.
- Raw data supplied (see [Peerj policy](#)).

EXPERIMENTAL DESIGN

- Original primary research within [Scope of the journal](#).
- Research question well defined, relevant & meaningful. It is stated how the research fills an identified knowledge gap.
- Rigorous investigation performed to a high technical & ethical standard.
- Methods described with sufficient detail & information to replicate.

VALIDITY OF THE FINDINGS

- Impact and novelty not assessed. *Meaningful* replication encouraged where rationale & benefit to literature is clearly stated.
- All underlying data have been provided; they are robust, statistically sound, & controlled.
- Conclusions are well stated, linked to original research question & limited to supporting results.



The best reviewers use these techniques

Tip

Example

Support criticisms with evidence from the text or from other sources

Smith et al (J of Methodology, 2005, V3, pp 123) have shown that the analysis you use in Lines 241-250 is not the most appropriate for this situation. Please explain why you used this method.

Give specific suggestions on how to improve the manuscript

Your introduction needs more detail. I suggest that you improve the description at lines 57- 86 to provide more justification for your study (specifically, you should expand upon the knowledge gap being filled).

Comment on language and grammar issues

The English language should be improved to ensure that an international audience can clearly understand your text. Some examples where the language could be improved include lines 23, 77, 121, 128 - the current phrasing makes comprehension difficult. I suggest you have a colleague who is proficient in English and familiar with the subject matter review your manuscript, or contact a professional editing service.

Organize by importance of the issues, and number your points

1. Your most important issue
2. The next most important item
3. ...
4. The least important points

Please provide constructive criticism, and avoid personal opinions

I thank you for providing the raw data, however your supplemental files need more descriptive metadata identifiers to be useful to future readers. Although your results are compelling, the data analysis should be improved in the following ways: AA, BB, CC

Comment on strengths (as well as weaknesses) of the manuscript

I commend the authors for their extensive data set, compiled over many years of detailed fieldwork. In addition, the manuscript is clearly written in professional, unambiguous language. If there is a weakness, it is in the statistical analysis (as I have noted above) which should be improved upon before Acceptance.

Genome-wide identification and characterization of *NHL* gene family in response to alkaline stress, ABA and MEJA treatments in wild soybean (*Glycine soja*)

Xu Zhang^{Equal first author, 1}, Yongguo Xue^{Equal first author, 2}, Haihang Wang¹, Zaib_un Nisa³, Xiaoxia Jin¹, Lijie Yu¹, Xinlei Liu², Yang Yu^{Corresp., 4}, Chao Chen^{Corresp. 1}

¹ Harbin Normal University, Harbin, heilongjiang, China

² Heilongjiang Provincial Academy of Agricultural Sciences, Harbin, heilongjiang, China

³ University of Lahore, Lahore, Pakistan

⁴ Shenyang University, Shenyang, China

Corresponding Authors: Yang Yu, Chao Chen

Email address: yuyang_syu@163.com, chchao@hrbnu.edu.cn

Background: *NDR1/HIN1-like (NHL)* family genes are known to be involved in pathogen induced plant responses to biotic stress. Even though the *NHL* family genes have been identified and characterized in plant defense responses in some plants, the roles of these genes associated with the plant abiotic stress tolerance in wild soybean is not fully established yet, especially in response to alkaline stress.

Methods: We identified the potential *NHL* family genes by using Hidden Markov model and wild soybean genome. The maximum-likelihood phylogenetic tree and conserved motifs were generated by using MEME online server and MEGA 7.0 software, respectively. Furthermore, the syntenic analysis was generated with Circos-0.69. Then we used PLANT CARE online software to predict and analyze the regulatory *cis*-acting elements in promoter regions. Hierarchical clustering trees was generated using TM4: MeV4.9 software. Additionally, the expression levels of *NHL* family genes under alkaline stress, ABA and MEJA treatment were identified by qRT-PCR.

Results: In this study, we identified 59 potential *NHL* family genes in wild soybean. We identified that wild soybean *NHL* family genes could be mainly classified into five groups as well as exist with conserved motifs. Syntenic analysis of *NHL* family genes revealed genes location on 18 chromosomes and presence of 65 pairs of duplication genes. Moreover, *NHL* family genes consisted of a variety of putative hormone-related and abiotic stress responsive elements, where numbers of methyl jasmonate (MeJA) and abscisic acid (ABA) responsive elements were significantly larger than other elements. We confirmed the regulatory roles of *NHL* family genes in response to alkaline stress, ABA and MEJA treatment. In conclusion, we identified and provided valuable information on the wild soybean *NHL* family genes, and established a foundation to further explore the potential roles of *NHL* family genes in crosstalk with MeJA or ABA signal transduction mechanisms under alkaline stress.

1 **Genome-wide identification and characterization of**
2 ***NHL* gene family in response to alkaline stress, ABA**
3 **and MEJA treatments in wild soybean (*Glycine soja*)**
4

5 Xu Zhang^{1, a}, Yongguo Xue^{2, a}, Haihang Wang¹, Zaib_un Nisa³, Xiaoxia Jin¹, Lijie Yu¹,
6 Xinlei Liu², Yang Yu^{4*}, Chao Chen^{1,*}
7

8 ¹ School of Life Science and Technology, Harbin Normal University, Harbin, P.R. China

9 ² Institute of Soybean Research, Heilongjiang Provincial Academy of Agricultural
10 Sciences, Harbin, P.R. China

11 ³ Institute of Molecular Biology and Biotechnology, University of Lahore, Defence road
12 campus, Lahore, Pakistan

13 ⁴ College of Life Science and Engineering, Shenyang University, Shenyang, Harbin, P.R.
14 China

15

16 Corresponding Author:

17 Chao Chen; Yang Yu

18 Harbin Normal University, 1 South Normal University Road, Songbei Limin Economic
19 Development Zone, Harbin 150025, P.R. China

20 Email address: chchao@hrbnu.edu.cn

21 a: These authors contributed equally to this work.

22

23

24

25

26

27 **Abstract**

28 **Background:** *NDR1/HIN1-like (NHL)* family genes are known to be involved in pathogen
29 induced plant responses to biotic stress. Even though the *NHL* family genes have been
30 identified and characterized in plant defense responses in some plants, the roles of these
31 genes associated with the plant abiotic stress tolerance in wild soybean is not fully
32 established yet, especially in response to alkaline stress.

33 **Methods:** We identified the potential *NHL* family genes by using Hidden Markov model
34 and wild soybean genome. The maximum-likelihood phylogenetic tree and conserved
35 motifs were generated by using MEME online server and MEGA 7.0 software,
36 respectively. Furthermore, the syntenic analysis was generated with Circos-0.69. Then
37 we used PLANT CARE online software to predict and analyze the regulatory *cis*-acting
38 elements in promoter regions. Hierarchical clustering trees was generated using TM4:
39 MeV4.9 software. Additionally, the expression levels of *NHL* family genes under alkaline
40 stress, ABA and MEJA treatment were identified by qRT-PCR.

41 **Results:** In this study, we identified 59 potential *NHL* family genes in wild soybean. We
42 identified that wild soybean *NHL* family genes could be mainly classified into five groups
43 as well as exist with conserved motifs. Syntenic analysis of *NHL* family genes revealed
44 genes location on 18 chromosomes and presence of 65 pairs of duplication genes.
45 Moreover, *NHL* family genes consisted of a variety of putative hormone-related and
46 abiotic stress responsive elements, where numbers of methyl jasmonate (MeJA) and
47 abscisic acid (ABA) responsive elements were significantly larger than other elements.
48 We confirmed the regulatory roles of *NHL* family genes in response to alkaline stress,
49 ABA and MEJA treatment. In conclusion, we identified and provided valuable information
50 on the wild soybean *NHL* family genes, and established a foundation to further explore
51 the potential roles of *NHL* family genes in crosstalk with MeJA or ABA signal transduction
52 mechanisms under alkaline stress.

53

54

55

56 Introduction

57 NHL (*NDR1/HIN1-like*) family genes are previously reported to be involved in plant
58 defense response against pathogens, such as *Phytophthora infestans*, *Botrytis cinerea*,
59 *Pseudomonas syringae* (Chen et al. 2018; Chong et al. 2008; Varet et al. 2002). Previous
60 studies revealed that there are at least 29 NHL family members that show homology to
61 *NDR1* and *HIN1* gene in Arabidopsis based on the non-redundant Genbank database
62 and then increased this number to 45 upon the completed genome sequencing of
63 Arabidopsis, suggesting a potential role for Arabidopsis *NDR1/HIN1-like* family genes in
64 plant-pathogen interactions. The 45 NHL family genes were divided into four groups and
65 shared three conserved sequence motifs (Dormann et al. 2000; Zheng et al. 2004).
66 Analysis of amino acid sequence reveals that a lot of NHL family proteins contain a
67 specific conserved late embryogenesis abundant (LEA) domain and putative
68 transmembrane domain (Dormann et al. 2000; Liu et al. 2020b).

69 During the last three decades, many NHL family members were isolated and identified
70 to play important roles in triggering plants defense resistance. The *HIN1* gene (harpin-
71 induced gene), which could be rapidly activated and elicit HR (hypersensitive response)
72 phenomenon when plants exposed to bacterial pathogens (e.g. *Pseudomonas syringae*
73 pv. *syringae*), was first isolated in tobacco (Gopalan et al. 1996). The *NDR1* gene (non-
74 race-specific disease resistance gene), which shows sequence similarity with tobacco
75 *HIN1*, was first identified to play distinct roles in response to both bacterial and fungal
76 pathogen resistance in Arabidopsis (Century et al. 1995; Takahashi et al. 2004).
77 Overexpression of NHL3 gene could enhanced the plant resistance to *Pseudomonas*
78 *syringae* pv. *tomato* DC3000, which was a membrane-localized protein in Arabidopsis
79 (Varet et al. 2003). In soybean (*Glycine max*), two homologs of *Arabidopsi NDR1* gene
80 named *GmNDR1a* and *GmNDR1b* were identified (Selote et al. 2014). This study showed
81 that the *NDR1* protein could interact with *RIN4* to play roles in resistance to *Pseudomonas*
82 *syringae*. The function role of *GmNDR1b*, also named *Gm-NDR1-1*, was further
83 determined that played important roles in impairing root pathogenic nematode *Heterodera*
84 *glycines* and *Meloidogyne incognita* (McNeece et al. 2017). The *StPOTHR1*, a NHL family
85 member in potato, could enhanced plants resistance to *Phytophthora infestans* through
86 effecting the MAP kinase signaling process by interacting with *NbMKK5L* (Chen et al.
87 2018). Overexpression of the pepper *CaNHL4* enhanced the expression of salicylic acid

88 (SA)- related and jasmonic acid (JA)-related genes, increased ROS production, and
89 inhibited the infection of the pathogens (Liu et al. 2020a). The interaction of ToxA with
90 NHL10 protein could induce cell death under plant pathogen stress in wheat (Dagvadorj
91 et al. 2022)

92 Saline-alkaline soils are known to have high content of sodium, bicarbonates and high
93 pH, which consequently causes growth retardation and ultimately leads to death of plants
94 growing in such soils. The total of 434 million ha of global land is affected by alkaline soils
95 (Jin et al. 2006; Wang et al. 2008). In comparison with neutral salts stress, alkaline stress
96 exerts more harmful effects on plant growth (Yang et al. 2008). Alkaline stress can inhibit
97 photosynthesis, N and sugar metabolism, as well as limits the absorption of ions, such as
98 H_2PO_4^- , Cl^- , Al^{3+} and Fe^{2+} (Vondrackova et al. 2015).

99 It has been documented that *NHL* family genes also play distinct roles in plant abiotic
100 stress resistance. In pepper (*Capsicum annuum* L.), fifteen *NHL* genes were identified in
101 a genome-wide analysis and the responses of these genes were characterized under
102 different abiotic stresses (Liu et al. 2020b). A stress-inducing *NHL* member, *BnNHL18A*,
103 was isolated from *Brassica napus*, which displayed roles in response to different
104 treatments including NaCl, H_2O_2 , as well as ethephon and SA (Lee et al. 2006). In
105 *Arabidopsis*, overexpression of *NHL6* increased the sensitivity to salt, osmotic and ABA
106 treatment, and *NHL6* could affect the seed germination and early seedling development
107 under these stresses-induced ABA signaling (Bao et al. 2016). Furthermore, studies have
108 shown that some saline or alkaline stress induced transcription factors co-expresses with
109 *NHL* network to regulate the salt or alkali response(Liu et al. 2020c). The soybean *NHLs*
110 have been identified to play important roles in regulating seed germination under chilling
111 stress and with ABA treatment (Wang et al. 2022). However, little is known about the wild
112 soybean *NHL* family genes in response to environment stresses, especially under alkaline
113 stress.

114 Soybean has been adopted as important crops in the world, particularly for protein and
115 oil production (Yu et al. 2018). Wild soybean, as the ancestor of cultivated soybean,
116 showed a better adaptation to various abiotic stress, such as salt, drought and alkaline.
117 Therefore, the wild soybean has been suggested as a valuable sources to improve the

118 agronomic traits of soybean(Wen et al. 2009). In previous studies, we have identified a
119 highly adaptable saline-alkali soil tolerant wild soybean (*Glycine soja*) line (G07256). It
120 can survive well in the saline-alkali soil (Ge et al. 2010). By using transcriptome data, we
121 identified some candidate genes in response to alkaline stress. In this study, the NHL
122 family genes in wild soybean genome were identified and their expression was
123 investigated under the influence of growth hormones in alkaline stress which may
124 enhance the stress responses.

125

126 **Materials and Methods**

127 **Identification of NHL family genes in wild soybean genome**

128 To identify all potential genes encoding NHL family genes in wild soybean genome, a
129 Hidden Markov model was first established by using the Arabidopsis and soybean amino
130 acid sequences of NHL family genes as queries (Gopavajhula et al. 2013; Wang et al.
131 2022). The HMM profile (build 2.3.2) was further used to search in wild soybean genome
132 database to get similar sequences (Finn et al. 2011). Then, the potential genes were
133 identified after removing the overlapping genes and incomplete domains genes through
134 Pfam and SMART database (Finn et al. 2016). The online software ExPASy
135 (https://web.expasy.org/compute_pi/) was used to predict the molecular weight and
136 isoelectric point values of NHL family proteins (Artimo et al. 2012).

137 **Bioinformatics analysis of NHL family genes**

138 The conserved motifs of all the potential *NHL* family genes were identified by MEME
139 online server (<http://meme-suite.org/>) (Bailey et al. 2009). The maximum-likelihood
140 phylogenetic tree was constructed by software MEGA 7.0 (Kumar et al. 2008) software.
141 Then, the TBtools software was used to combine the conserved motifs and phylogenetic
142 tree. The syntenic analysis was generated with Circos-0.69 (<http://circos.ca/>) (Krzywinski
143 et al. 2009). To analyze the potential regulatory *cis*-acting elements in the promoters of
144 *NHL* genes, 3000 bp upstream sequences of the above-mentioned genes were extracted
145 based on the genome database(Sun et al. 2014). Then we used PLANT CARE online
146 software (<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>) to predict and

147 analyze the regulatory cis-acting elements in promoter regions (Lescot et al. 2002). To
148 examine the expression profiles of *NHL* family genes under alkaline stress, wild soybean
149 transcriptome data was downloaded (DuanMu et al. 2015) and hierarchical clustering
150 trees were generated using TM4: MeV4.9 software (Saeed et al. 2006).

151 **Plant material, growth condition and stress treatment**

152 The wild soybean cultivar DN50 was grown in 1/4 Hoagland nutrient solutions in the
153 growth chamber with 22-28 °c room temperature, 70-80% relative humidity and 8 h
154 dark/16 h light. The healthy and plump seeds were rinsed with 75% ethanol for 1 min,
155 and then washed with sterile water before germination (Qiao et al. 2020). After two days,
156 the germinated seedlings were transferred in to 1/4 strength Hoagland nutrient solutions
157 to be cultured and the nutrient solutions were changed every two days. Twelve days later,
158 the young seedlings were treated with alkaline (NaHCO_3) stress or exogenous hormones
159 (ABA and MeJA), respectively. For alkaline stress, the 12-days old seedlings were
160 transferred into 1/4 Hoagland solution with 50 mM NaHCO_3 . For exogenous hormones
161 treatment, the 12-days old seedlings were transferred into 1/4 Hoagland solution with 50
162 μM ABA or 1/4 Hoagland solution with 50 μM MeJA. The roots were harvested and stored
163 in liquid nitrogen at 0 h, 1 h and 3 h after treatment for RNA extraction.

164 **Transcript expression analysis by qRT-PCR**

165 Total RNA was extracted with the OminiPlant RNA isolation kit (Kangwei), and the cDNAs
166 were synthesized using the First Stand cDNA Synthesis kit (Toyobo) for qRT-PCR. The
167 qRT-PCR was then performed with UtraSYBR Mixture (Baiouleibo) and ABI 7500
168 sequencer. The primers of wild soybean *NHL* genes and *GsGADPH* were listed in
169 supplementary Table S1 which are used in our study. Here the *GsGADPH* gene was used
170 as an internal control in wild soybean (Huis et al. 2010). The qRT-PCR data was
171 calculated with three independent biological replicates using $2^{-\Delta\Delta\text{CT}}$ method and Student's
172 *t*-test.

173 **Results**

174 **Identification of NHL genes in wild soybean**

175 To identify the *NHL* family genes in wild soybean, the amino acid sequences of the *NHL*
176 family genes from Arabidopsis and soybean were queried against the wild soybean
177 genome via BLAST from NCBI. A total of 208 *NHL* candidate sequences were obtained
178 based on the Hidden Markov model. All candidate sequences were then subjected to
179 Pfam and SMART database to remove the redundant sequences or incomplete domain
180 sequences. As a result, 59 genes were obtained as potential *NHL* family genes in wild
181 soybean genome.

182 As shown in Table 1, 59 predicted wild soybean *NHL* genes were named based on the
183 location within the reference genome from *GsNHL1* to *GsNHL59*. Then the chemical
184 properties of these proteins were determined including the protein sequence lengths,
185 molecular weights (MW), and theoretical isoelectric points (pI). The protein sequence
186 length was ranged from 149 (*GsNHL33*) to 348 (*GsNHL17*) amino acids residues. The
187 MW varied from 16.37778 (*GsNHL33*) to 40.00232 (*GsNHL17*) kDa and the pI values
188 ranged from 7.82 (*GsNHL31*) to 10.24 (*GsNHL43*).

189 **Phylogenetic and conserved motifs analysis of wild soybean *NHL* genes**

190 To confirm the evolutionary relationships of *NHL* family genes in wild soybean, a
191 maximum-likelihood phylogenetic tree was constructed with the full-length protein
192 sequences from soybean, Arabidopsis and wild soybean. The result showed that the
193 genes could be divided into six groups and the other 7 ungroup genes (supplementary
194 Fig. S1).

195 Based on conserved motif sequences, the 59 wild soybean *NHL* family genes could be
196 further classified into five groups (group 1a, group 1b, group 2, group 3a and group 3b)
197 and the other 11 ungroup genes (Fig. 1a). MEME motif analysis also revealed that wild
198 soybean *NHL* family proteins shared ten conserved motif sequences (Fig. 1b,
199 supplementary Fig. S2). Most of wild soybean *NHL* family proteins contain conserved
200 motif 1 and motif 2. Interestingly, we found that some wild soybean *NHL* family proteins
201 in the same group shared a similar motif composition. For example, motif 5 is mainly
202 present within group 1b, group 2 and group 3b genes, while most genes of group 1a and
203 group 1b contain motif 7 and 9. The motif 3 only located in group 1a genes. The similar
204 motif arrangement among the proteins of wild soybean *NHL* family suggested that the

205 protein architecture was conserved within subgroups, which indicated that the proteins in
206 the same group may have similar function in plant development and resistance responses
207 under stress conditions. However, functions of these conserved motifs are still need to be
208 further explored.

209 **Chromosomal locations and syntenic analysis**

210 The analysis of gene duplication events could drive the potential evolution mechanisms
211 of the wild soybean *NHL* family genes. In this study, 59 wild soybean *NHL* family genes
212 were found randomly distributed among 18 chromosomes, with the exception of 8 and 17
213 (Fig. 2). Gene duplication plays significant roles in plant organismal evolution and
214 functional diversity (Bowers et al. 2003). Further, a total of 65 pairs of *NHL* syntenic
215 paralogs were identified in wild soybean genome. These results indicated that the wild
216 soybean *NHL* family have been exhibited a high gene family expansion.

217 **Identification of cis-acting elements of NHL gene promoters in wild soybean**

218 To explore the potential roles of wild soybean *NHL* family genes in response to abiotic
219 stress, the promoter sequences 3 kb upstream regions of the ATG were predicted using
220 information within the PlantCARE online tool. The results showed that the wild soybean
221 *NHL* family genes displayed a variety of putative hormone-related and abiotic stress
222 responsive elements (Fig. 3, supplementary Table S2). For example, the plant hormone-
223 related responsive elements include Methyl jasmonate (MeJA), abscisic acid (ABA,
224 ABRE), gibberellin (GA), salicylic acid (SA) and Auxin responsive elements. Interestingly,
225 we found that the numbers of MeJA and ABA responsive elements were significantly
226 larger than the other plant hormone responsive elements, indicating the potential roles of
227 *NHL* family genes in MeJA and ABA signaling pathways. We also identified some
228 response elements including MBS, LTR and TC-rich, which respond to drought, low
229 temperatures and general stress, respectively. Collectively, these results strongly
230 suggested that the roles of wild soybean *NHL* family genes are likely associated with plant
231 abiotic stresses and hormone stimuli.

232 **Expression analysis of NHL genes in response to alkaline treatment in wild** 233 **soybean**

234 To assess the potential roles of *NHL* family genes participate in the defense responses
235 towards alkaline stress, we generated a heat map of *NHL* family genes based on the wild
236 soybean transcriptome data under alkaline stress. The results showed that 24 genes were
237 differently induced under alkaline stress. Among them, 18 of *NHL* family genes were
238 significantly up-regulated, while six genes showed down-regulation patterns (Fig. 4). To
239 further confirm the expression of *NHL* family genes in response to alkaline treatment, we
240 selected 12 of the up-regulated genes to detect their expression patterns under 50 mM
241 NaHCO₃ stress by using qRT-PCR analysis. As shown in Fig. 5, the expression patterns
242 of 11 up-regulated genes were roughly consistent with the transcriptome data under
243 alkaline stress, except that *GsNHL29* had contrary results. In addition, the expression of
244 *GsNHL9*, *GsNHL44*, *GsNHL45* and *GsNHL47* showed higher expression levels at 3 h
245 point than the other genes (Fig. 5e, i-k). In conclusion, the qRT-PCR analysis confirmed
246 the results that *GsNHL* family genes possibly participate in responses to alkaline stress.

247 **Effects of different phytohormone treatments including ABA and MeJA on the** 248 **expression of NHL genes**

249 The plant hormones play regulatory roles in plant responses to various stresses. In this
250 study, we found that the numbers of MeJA and ABA responsive elements in wild soybean
251 *NHL* family genes were significantly larger than the other plant hormone responsive
252 elements. Here, to explore if the *NHL* family genes could participate in ABA and MeJA
253 signaling pathways, we analyzed the transcript expression levels of the 12 *NHL* family
254 genes mentioned above under ABA and MeJA treatments by using qRT-PCR analysis.
255 As shown in Fig. 6, nine genes were up-regulated under MeJA treatment and *GsNHL29*
256 was down-regulated. In addition, *GsNHL6* and *GsNHL11* had contrary expression pattern
257 at 1 h and 3 h (Fig. 6b, f). Under ABA treatment, only *GsNHL44* and *GsNHL51* were up-
258 regulated and seven genes were down regulated (Fig. 7). *GsNHL4* had contrary
259 expression pattern at 1 h and 3 h (Fig. 7a). *GsNHL6* and *GsNHL45* showed no significant
260 expression changes (Fig. 7b, g). Collectively, these results indicated that *NHL* family
261 genes participate in ABA and MeJA signaling pathways and play different roles in
262 response to ABA or MeJA signaling pathway.

263 **Discussion**

264 Previously, studies have identified that *NHL* family genes are involved in plant
265 development and pathogens attack resistance (Bao et al. 2016; Chen et al. 2021). Many
266 *NHL* family genes have been identified in plant species, such as tomato, pepper and
267 soybean (Dormann et al. 2000; Liu et al. 2020b; Wang et al. 2022). However, the wild
268 soybean *NHL* family genes have not been identified, especially the roles of *NHL* family
269 genes in regulating alkaline stress. Hence, this research was based on bioinformatics
270 analysis about wild soybean *NHL* family genes in order to understand their structure and
271 location, and mainly potential roles were investigated in response to plant hormones and
272 alkaline stress treatments.

273 In this study, 59 wild soybean *NHL* family genes were identified in accordance with the
274 soybean and *Arabidopsis* *NHL* related genes (Table 1). We found that *NHL* family
275 proteins varied markedly in protein sequence length and molecular weight, indicating the
276 divergent evolution in wild soybean *NHL* family genes. However, the high pI value showed
277 *NHL* families are alkaline proteins.

278 Previous studies revealed that *NHL* protein family could be classified into six groups by
279 investigating the relationship of soybean, *Arabidopsis* and rice (Wang et al. 2022). This
280 result was consistent with our findings that wild soybean *NHL* family genes be divided
281 into six groups (supplementary Fig. 1). On the basis of conserved motif analysis clustered,
282 we found most of wild soybean *NHL* family genes could be classified into five groups,
283 which was also consistent with the results of phylogenetic tree analysis of *NHL* family
284 genes in soybean (Wang et al. 2022). In addition, each group almost shared a similar
285 motif composition, which indicated that the groups may have similar roles in plant
286 development progress (Fig. 1).

287 Further, the wild soybean *NHL* family genes from the same group were mostly located
288 in different chromosomes and located near the edges of the chromosomes, suggesting a
289 strategy to exert their functions in the whole wild soybean genome. The pairs of *NHL*
290 syntenic paralogs also indicated that the wild soybean *NHL* family have been exhibited a
291 high gene family expansion, which might play significant roles in gene functional diversity
292 (Fig. 2).

293 In addition, we found that conserved motif 1, 3, 4, 5 and 10 belong to the LEA-2 domain
294 (Fig. 1b, Supplementary Fig. S2). This result also consistent with previous study (Liu et
295 al. 2020b). Furthermore, we found that LEA-2 domain belongs to the LEA_2 subgroup
296 which are widely known as a late embryogenesis abundant proteins and play significant
297 roles under abiotic stress responses (Jin et al. 2019). For example, the rice LEA proteins
298 showed accumulation during the salinity-triggered growth, while degradation in LEA
299 proteins was observed during plant recovery from salt stress (Chourey et al. 2003). The
300 tea plant LEA genes were significantly induced under stress conditions, such as drought,
301 ABA, low and high temperature (Jin et al. 2019). Overexpression of *IpLEA* could show
302 high tolerance to salt and drought stress in *Ipomoea pescaprae* by mediating water
303 homeostasis and as a reactive oxygen species scavenger (Zheng et al. 2019). Thus, this
304 evidence indicated the potential roles of wild soybean *NHL* family genes in response to
305 environmental stresses.

306 The *cis*-acting regulatory elements play important roles as molecular switches to control
307 various biological processes, including hormonal and various stress responses (Sun et
308 al. 2021). The *cis*-acting regulatory elements analysis showed that the promoter regions
309 of wild soybean *NHL* family genes contain a variety of putative hormone-related and
310 abiotic stress responsive elements (Fig. 3). Previous studies have been shown that *NHL*
311 family genes participate in the plant hormone-mediated pathways. For example,
312 overexpression of *AtNHL1* and *AtNHL8* in soybean could enhance plants resistance to
313 *Heterodera glycines* by mediating the jasmonic acid and ethylene pathways, confirming
314 the roles of these genes in plant defense response (Maldonado et al. 2014). *NHL6*
315 participates in the abiotic stresses-induced ABA signaling at seed germination and early
316 seedling stages in *Arabidopsis* (Bao et al. 2016). *BnNHL18A* could be significantly
317 induced by NaCl, ethephon, methyl jasmonate or salicylic acid treatment in *Brassica*
318 *napus* (Lee et al. 2006). Also, LTR is a *cis*-element responsive to low-temperature stress
319 (Brown et al. 2001). TC-rich *cis*-element has been identified to be involved in stress
320 mediated plant defense responses (Sazegari et al. 2015). In conclusion, these evidence
321 strongly suggested that *NHL* family genes may be involved in stress resistance and plant
322 hormones responses in wild soybean.

323 Alkaline stress is one of the most harmful abiotic stresses, which leading to a series of
324 regulatory mechanisms in plants, such as ion balance, osmotic adjustment, pH regulation,
325 and ROS scavenging mechanisms. Previously, we identified a highly adaptable saline-
326 alkali soil tolerant wild soybean line which can survive well in the saline-alkali soil. Then,
327 we explored the differentially expressed genes of wild soybean seedlings treated with 50
328 mM NaHCO₃ by RNA sequencing (DuanMu et al. 2015). In this study, we mainly intend
329 to explore the potential roles of wild soybean *NHL* family genes in response to alkaline
330 stress. According to the transcriptome data, a total of 24 genes were significantly induced
331 under alkaline stress (Fig. 4), and qRT-PCR confirmed the results that wild soybean *NHL*
332 family genes may play positive role in response to alkaline stress (Fig. 5).

333 During abiotic stress responsive processes, plant hormones such as MeJA, ABA, SA,
334 GA and Auxin also play important roles and have cross talks in signal transduction
335 pathways (Ku et al. 2018). In our previous study, we also identified that plant hormones
336 have crosstalk with plant alkaline stress resistance response. For example, the wild
337 soybean gene *ERF71* could regulate endogenous auxin accumulation when plants
338 treated with alkaline solution (Yu et al. 2017). The *TIFY10* gene could act as a regulator
339 in response to alkaline stress and jasmonate signaling in wild soybean (Zhu et al. 2011).
340 On the other hand, we found that wild soybean *NHL* family genes comprised of a variety
341 of putative hormone-related responsive elements, and the numbers of MeJA and ABA
342 responsive elements were significantly larger than others. Thus, we concluded that the
343 wild soybean *NHL* family genes have crosstalk with MeJA or ABA signal transduction
344 under alkaline stress. qRT-PCR analysis showed that nine genes were up-regulated
345 under MeJA treatment, and these genes were all up-regulated under alkaline stress (Fig.
346 6). This finding is consistent with the previous studies that wild soybean *TIFY10a*
347 overexpression lines enhanced the alkaline stress resistance and also increased the
348 jasmonate content of the transgenic alfalfa (Zhu et al. 2014). However, in comparison
349 with MeJA treatment, qRT-PCR analysis showed a different expression, in which only two
350 wild soybean *NHL* family genes were up-regulated and seven genes were down-
351 regulated under ABA treatment (Fig. 7). For example, *GsNHL7*, *GsNHL8*, *GsNHL9*,
352 *GsNHL12* and *GsNHL7* were up-regulated under alkaline stress and MeJA treatment,
353 while were down-regulated under ABA treatment. In addition, our previous studies

354 identified that overexpression of wild soybean *NAC019* or *SKP21* in *Arabidopsis* could
355 contribute to alkaline stress tolerance, but reduced ABA sensitivity(Cao et al. 2017; Liu et
356 al. 2015). In conclusion, these results speculated that wild soybean *NHL* family genes
357 have crosstalk with MeJA or ABA signal transduction under alkaline stress, and some
358 genes may display different roles in ABA or MeJA signal transduction in response to
359 alkaline stress.

360 **Conclusions**

361 In conclusion, in this study, we identified 59 potential *NHL* family genes in wild soybean.
362 We identified the phylogenetic relationship, conserved domains, gene duplication events
363 and *cis*-acting elements in promoter regions. We also confirmed that wild soybean *NHL*
364 family genes may play important regulatory roles in response to alkaline stress, ABA and
365 MEJA treatment. Taken together, our results established a foundation for characterization
366 of wild soybean *NHL* family genes in response to alkaline stress, ABA and MEJA
367 treatment. However, the function analysis of up-regulated genes under ABA, MEJA or
368 alkaline stress, such as *GsNHL4*, *GsNHL44* and *GsNHL51*, is of great significance in
369 the future. On the other hand, more jobs are need to do for exploring the potential roles
370 of *NHL* family genes, especially the roles in crosstalk with MeJA or ABA signal
371 transduction pathways under alkaline stress in wild soybean.

372 **Abbreviations**

NHL	NDR1/HIN1-like
qRT-PCR	quantitative real-time PCR
MW	molecular weight
pI	isoelectric point
ABA	abscisic acid
MeJA	methyl jasmonate
SA	salicylic acid
GA	gibberellin

MBS	drought
LTR	low temperature responsive
TC-rich	defense and stress responsive

373

374

375 References

- 376 Artimo P, Jonnalagedda M, Arnold K, Baratin D, Csardi G, de Castro E, Duvaud S, Flegel V, Fortier A,
377 Gasteiger E, Grosdidier A, Hernandez C, Ioannidis V, Kuznetsov D, Liechti R, Moretti S, Mostaguir
378 K, Redaschi N, Rossier G, Xenarios I, and Stockinger H. 2012. ExPASy: SIB bioinformatics resource
379 portal. *Nucleic Acids Research* 40:W597-W603. 10.1093/nar/gks400
- 380 Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, and Noble WS. 2009. MEME
381 SUITE: tools for motif discovery and searching. *Nucleic Acids Research* 37:W202-W208.
382 10.1093/nar/gkp335
- 383 Bao Y, Song W-M, Pan J, Jiang C-M, Srivastava R, Li B, Zhu L-Y, Su H-Y, Gao X-S, Liu H, Yu X, Yang L, Cheng
384 X-H, and Zhang H-X. 2016. Overexpression of the NDR1/HIN1-Like Gene NHL6 Modifies Seed
385 Germination in Response to Abscisic Acid and Abiotic Stresses in Arabidopsis. *Plos One* 11.
386 10.1371/journal.pone.0148572
- 387 Bowers JE, Chapman BA, Rong JK, and Paterson AH. 2003. Unravelling angiosperm genome evolution by
388 phylogenetic analysis of chromosomal duplication events. *Nature* 422:433-438.
389 10.1038/nature01521
- 390 Brown APC, Dunn MA, Goddard NJ, and Hughes MA. 2001. Identification of a novel low-temperature-
391 response element in the promoter of the barley (*Hordeum vulgare* L) gene blt101.1. *Planta*
392 213:770-780. 10.1007/s004250100549
- 393 Cao L, Yu Y, Ding X, Zhu D, Yang F, Liu B, Sun X, Duan X, Yin K, and Zhu Y. 2017. The Glycine soja NAC
394 transcription factor GsNAC019 mediates the regulation of plant alkaline tolerance and ABA
395 sensitivity. *Plant Molecular Biology* 95:253-268. 10.1007/s11103-017-0643-3
- 396 Century KS, Holub EB, and Staskawicz BJ. 1995. NDR1, A LOCUS OF ARABIDOPSIS-THALIANA THAT IS
397 REQUIRED FOR DISEASE RESISTANCE TO BOTH A BACTERIAL AND A FUNGAL PATHOGEN.
398 *Proceedings of the National Academy of Sciences of the United States of America* 92:6597-6601.
399 10.1073/pnas.92.14.6597
- 400 Chen L, He F, Long R, Zhang F, Li M, Wang Z, Kang J, and Yang Q. 2021. A global alfalfa diversity panel
401 reveals genomic selection signatures in Chinese varieties and genomic associations with root
402 development. *Journal of Integrative Plant Biology* 63:1937-1951. 10.1111/jipb.13172
- 403 Chen Q, Tian Z, Jiang R, Zheng X, Xie C, and Liu J. 2018. StPOTHR1, a NDR1/HIN1-like gene in *Solanum*
404 *tuberosum*, enhances resistance against *Phytophthora infestans*. *Biochemical and Biophysical*
405 *Research Communications* 496:1155-1161. 10.1016/j.bbrc.2018.01.162
- 406 Chong J, Le Henanff G, Bertsch C, and Walter B. 2008. Identification, expression analysis and
407 characterization of defense and signaling genes in *Vitis vinifera*. *Plant Physiology and Biochemistry*
408 46:469-481. 10.1016/j.plaphy.2007.09.010

- 409 Chourey K, Ramani S, and Apte SK. 2003. Accumulation of LEA proteins in salt (NaCl) stressed young
410 seedlings of rice (*Oryza sativa* L.) cultivar Bura Rata and their degradation during recovery from
411 salinity stress. *Journal of Plant Physiology* 160:1165-1174. 10.1078/0176-1617-00909
- 412 Dagvadorj B, Outram MA, Williams SJ, and Solomon PS. 2022. The necrotrophic effector ToxA from
413 *Parastagonospora nodorum* interacts with wheat NHL proteins to facilitate Tsn1-mediated
414 necrosis. *Plant Journal* 110:407-418. 10.1111/tpj.15677
- 415 Dormann P, Gopalan S, He SY, and Benning C. 2000. A gene family in *Arabidopsis thaliana* with sequence
416 similarity to NDR1 and HIN1. *Plant Physiology and Biochemistry* 38:789-796. 10.1016/s0981-
417 9428(00)01186-4
- 418 DuanMu H, Wang Y, Bai X, Cheng S, Deyholos MK, Wong GK-S, Li D, Zhu D, Li R, Yu Y, Cao L, Chen C, and
419 Zhu Y. 2015. Wild soybean roots depend on specific transcription factors and oxidation reduction
420 related genes in response to alkaline stress. *Functional & Integrative Genomics* 15:651-660.
421 10.1007/s10142-015-0439-y
- 422 Finn RD, Clements J, and Eddy SR. 2011. HMMER web server: interactive sequence similarity searching.
423 *Nucleic Acids Research* 39:W29-W37. 10.1093/nar/gkr367
- 424 Finn RD, Coggill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, Potter SC, Punta M, Qureshi M, Sangrador-
425 Vegas A, Salazar GA, Tate J, and Bateman A. 2016. The Pfam protein families database: towards a
426 more sustainable future. *Nucleic Acids Research* 44:D279-D285. 10.1093/nar/gkv1344
- 427 Ge Y, Li Y, Zhu Y-M, Bai X, Lv D-K, Guo D, Ji W, and Cai H. 2010. Global transcriptome profiling of wild
428 soybean (*Glycine soja*) roots under NaHCO₃ treatment. *Bmc Plant Biology* 10. 10.1186/1471-
429 2229-10-153
- 430 Gopalan S, Bauer DW, Alfano JR, Loniello AO, He SY, and Collmer A. 1996. Expression of the *Pseudomonas*
431 *syringae* avirulence protein AvrB in plant cells alleviates its dependence on the hypersensitive
432 response and pathogenicity (Hrp) secretion system in eliciting genotype-specific hypersensitive
433 cell death. *Plant Cell* 8:1095-1105.
- 434 Gopavajhula VR, Chaitanya KV, Khan PAA, Shaik JP, Reddy PN, and Alanazi M. 2013. Modeling and analysis
435 of soybean (*Glycine max.* L) Cu/Zn, Mn and Fe superoxide dismutases. *Genetics and Molecular*
436 *Biology* 36:225-236. 10.1590/s1415-47572013005000023
- 437 Huis R, Hawkins S, and Neutelings G. 2010. Selection of reference genes for quantitative gene expression
438 normalization in flax (*Linum usitatissimum* L.). *Bmc Plant Biology* 10. 10.1186/1471-2229-10-71
- 439 Jin H, Plaha P, Park JY, Hong CP, Lee IS, Yang ZH, Jiang GB, Kwak SS, Liu SK, Lee JS, Kim YA, and Lim YP.
440 2006. Comparative EST profiles of leaf and root of *Leymus chinensis*, a xerophilous grass adapted
441 to high pH sodic soil. *Plant Science* 170:1081-1086. 10.1016/j.plantsci.2006.01.002
- 442 Jin X, Cao D, Wang Z, Ma L, Tian K, Liu Y, Gong Z, Zhu X, Jiang C, and Li Y. 2019. Genome-wide identification
443 and expression analyses of the LEA protein gene family in tea plant reveal their involvement in
444 seed development and abiotic stress responses. *Scientific Reports* 9. 10.1038/s41598-019-50645-
445 8
- 446 Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, and Marra MA. 2009. Circos:
447 An information aesthetic for comparative genomics. *Genome Research* 19:1639-1645.
448 10.1101/gr.092759.109
- 449 Ku Y-S, Sintaha M, Cheung M-Y, and Lam H-M. 2018. Plant Hormone Signaling Crosstalks between Biotic
450 and Abiotic Stress Responses. *International Journal of Molecular Sciences* 19.
451 10.3390/ijms19103206
- 452 Kumar S, Nei M, Dudley J, and Tamura K. 2008. MEGA: A biologist-centric software for evolutionary
453 analysis of DNA and protein sequences. *Briefings in Bioinformatics* 9:299-306.
454 10.1093/bib/bbn017

- 455 Lee SB, Ham BK, Park JM, Kim YJ, and Paek KH. 2006. BnNHL18A shows a localization change by stress-
456 inducing chemical treatments. *Biochemical and Biophysical Research Communications* 339:399-
457 406. 10.1016/j.bbrc.2005.10.210
- 458 Lescot M, Dehais P, Thijs G, Marchal K, Moreau Y, Van de Peer Y, Rouze P, and Rombauts S. 2002.
459 PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico
460 analysis of promoter sequences. *Nucleic Acids Research* 30:325-327. 10.1093/nar/30.1.325
- 461 Liu A, Yu Y, Duan X, Sun X, Duanmu H, and Zhu Y. 2015. GsSKP21, a Glycine soja S-phase kinase-associated
462 protein, mediates the regulation of plant alkaline tolerance and ABA sensitivity. *Plant Molecular
463 Biology* 87:111-124. 10.1007/s11103-014-0264-z
- 464 Liu C, Peang H, Li X, Liu C, Lv X, Wei X, Zou A, Zhang J, Fan G, Ma G, Ma L, and Sun X. 2020a. Genome-wide
465 analysis of NDR1/HIN1-like genes in pepper (*Capsicum annuum* L.) and functional characterization
466 of CaNHL4 under biotic and abiotic stresses. *Horticulture research* 7:93-93. 10.1038/s41438-020-
467 0318-0
- 468 Liu C, Peng H, Li X, Liu C, Lv X, Wei X, Zou A, Zhang J, Fan G, Ma G, Ma L, and Sun X. 2020b. Genome-wide
469 analysis of NDR1/HIN1-like genes in pepper (*Capsicum annuum* L.) and functional characterization
470 of CaNHL4 under biotic and abiotic stresses. *Horticulture research* 7:93-93. 10.1038/s41438-020-
471 0318-0
- 472 Liu J, Shen F, Xiao Y, Fang H, Qiu C, Li W, Wu T, Xu X, Wang Y, Zhang X, and Han Z. 2020c. Genomics-
473 assisted prediction of salt and alkali tolerances and functional marker development in apple
474 rootstocks. *Bmc Genomics* 21. 10.1186/s12864-020-06961-9
- 475 Maldonado A, Youssef R, McDonald M, Brewer E, Beard H, and Matthews B. 2014. Modification of the
476 expression of two NPR1 suppressors, SNC1 and SNI1, in soybean confers partial resistance to the
477 soybean cyst nematode, *Heterodera glycines*. *Functional Plant Biology* 41:714-726.
478 10.1071/fp13323
- 479 McNeece BT, Pant SR, Sharma K, Niruala P, Lawrence GW, and Klink VP. 2017. A Glycine max homolog of
480 NON-RACE SPECIFIC DISEASE RESISTANCE 1 (NDR1) alters defense gene expression while
481 functioning during a resistance response to different root pathogens in different genetic
482 backgrounds. *Plant Physiology and Biochemistry* 114:60-71. 10.1016/j.plaphy.2017.02.022
- 483 Qiao Y, Lu W, Wang R, Nisa Z, Yu Y, Jin X, Yu L, and Chen C. 2020. Identification and Expression Analysis of
484 Strigolactone Biosynthetic and Signaling Genes in Response to Salt and Alkaline Stresses in
485 Soybean (*Glycine max*). *DNA and Cell Biology* 39:1850-1861. 10.1089/dna.2020.5637
- 486 Saeed AI, Hagabati NK, Braisted JC, Liang W, Sharov V, Howe EA, Li J, Thiagarajan M, White JA, and
487 Quackenbush J. 2006. TM4 microarray software suite. In: Kimmel A, and Oliver B, eds. *DNA
488 Microarrays, Part B: Databases and Statistics*, 134-+.
- 489 Sazegari S, Niazi A, and Ahmadi FS. 2015. A study on the regulatory network with promoter analysis for
490 *Arabidopsis* DREB-genes. *Bioinformatics* 11:101-106. 10.6026/97320630011101
- 491 Selote D, Shine MB, Robin GP, and Kachroo A. 2014. Soybean NDR1-like proteins bind pathogen effectors
492 and regulate resistance signaling. *New Phytologist* 202:485-498. 10.1111/nph.12654
- 493 Sun C, Zhang C, Wang X, Zhao X, Chen F, Zhang W, Hu M, Fu S, Yi B, and Zhang J. 2021. Genome-Wide
494 Identification and Characterization of the IGT Gene Family in Allotetraploid Rapeseed (*Brassica
495 napus* L.). *DNA and Cell Biology* 40:441-456. 10.1089/dna.2020.6227
- 496 Sun X, Luo X, Sun M, Chen C, Ding X, Wang X, Yang S, Yu Q, Jia B, Ji W, Cai H, and Zhu Y. 2014. A Glycine
497 Soja 14-3-3 Protein GsGF14o Participates in Stomatal and Root Hair Development and Drought
498 Tolerance in *Arabidopsis thaliana*. *Plant and Cell Physiology* 55:99-118. 10.1093/pcp/pct161
- 499 Takahashi Y, Berberich T, Yamashita K, Uehara Y, Miyazaki A, and Kusano T. 2004. Identification of tobacco
500 HIN1 and two closely related genes as spermine-responsive genes and their differential
501 expression during the Tobacco mosaic virus-induced hypersensitive response and during leaf- and
502 flower-senescence. *Plant Molecular Biology* 54:613-622. 10.1023/b:Plan.0000038276.95539.39

- 503 Varet A, Hause B, Hause G, Scheel D, and Lee J. 2003. The Arabidopsis NHL3 gene encodes a plasma
504 membrane protein and its overexpression correlates with increased resistance to *Pseudomonas*
505 *syringae* pv. tomato DC3000. *Plant Physiology* 132:2023-2033. 10.1104/pp.103.020438
- 506 Varet A, Parker J, Tornero P, Nass N, Nurnberger T, Dangl JL, Scheel D, and Lee J. 2002. NHL25 and NHL3,
507 two NDR1/HIN1-like genes in Arabidopsis thaliana with potential role(s) in plant defense.
508 *Molecular Plant-Microbe Interactions* 15:608-616. 10.1094/mpmi.2002.15.6.608
- 509 Vondrackova S, Szakova J, Drabek O, Tejnecky V, Hejcman M, Muellerova V, and Tlustos P. 2015.
510 Aluminium Uptake and Translocation in Al Hyperaccumulator *Rumex obtusifolius* Is Affected by
511 Low-Molecular-Weight Organic Acids Content and Soil pH. *Plos One* 10.
512 10.1371/journal.pone.0123351
- 513 Wang J, Wu R, Shangguan T, Chen G, Zheng Y, Tao X, Li S, Wang Y, and Xu S. 2022. NDR1/HIN1-like genes
514 may regulate Glycine max seed germination under chilling stress through the ABA pathway. *Plant*
515 *Growth Regulation*. 10.1007/s10725-022-00894-x
- 516 Wang Y, Ma H, Liu G, Xu C, Zhang D, and Ban Q. 2008. Analysis of Gene Expression Profile of *Limonium*
517 *bicolor* under NaHCO₃ Stress Using cDNA Microarray. *Plant Molecular Biology Reporter* 26:241-
518 254. 10.1007/s11105-008-0037-4
- 519 Wen Z, Ding Y, Zhao T, and Gai J. 2009. Genetic diversity and peculiarity of annual wild soybean (*G. soja*
520 Sieb. et Zucc.) from various eco-regions in China. *Theoretical and Applied Genetics* 119:371-381.
521 10.1007/s00122-009-1045-y
- 522 Yang CW, Wang P, Li CY, Shi DC, and Wang DL. 2008. Comparison of effects of salt and alkali stresses on
523 the growth and photosynthesis of wheat. *Photosynthetica* 46:107-114. 10.1007/s11099-008-
524 0018-8
- 525 Yu M, Liu Z, Jiang S, Xu N, Chen Q, Qi Z, and Lv W. 2018. QTL mapping and candidate gene mining for
526 soybean seed weight per plant. *Biotechnology & Biotechnological Equipment* 32:908-914.
527 10.1080/13102818.2018.1438851
- 528 Yu Y, Duan X, Ding X, Chen C, Zhu D, Yin K, Cao L, Song X, Zhu P, Li Q, Nisa ZU, Yu J, Du J, Song Y, Li H, Liu
529 B, and Zhu Y. 2017. A novel AP2/ERF family transcription factor from Glycine soja, GsERF71, is a
530 DNA binding protein that positively regulates alkaline stress tolerance in Arabidopsis. *Plant*
531 *Molecular Biology* 94:509-530. 10.1007/s11103-017-0623-7
- 532 Zheng J, Su H, Lin R, Zhang H, Xia K, Jian S, and Zhang M. 2019. Isolation and characterization of an atypical
533 LEA gene (IpLEA) from *Ipomoea pes-caprae* conferring salt/drought and oxidative stress tolerance.
534 *Scientific Reports* 9. 10.1038/s41598-019-50813-w
- 535 Zheng MS, Takahashi H, Miyazaki A, Hamamoto H, Shah J, Yamaguchi I, and Kusano T. 2004. Up-regulation
536 of Arabidopsis thaliana NHL10 in the hypersensitive response to Cucumber mosaic virus infection
537 and in senescing leaves is controlled by signalling pathways that differ in salicylate involvement.
538 *Planta* 218:740-750. 10.1007/s00425-003-1169-2
- 539 Zhu D, Bai X, Chen C, Chen Q, Cai H, Li Y, Ji W, Zhai H, Lv D, Luo X, and Zhu Y. 2011. GsTIFY10, a novel
540 positive regulator of plant tolerance to bicarbonate stress and a repressor of jasmonate signaling.
541 *Plant Molecular Biology* 77:285-297. 10.1007/s11103-011-9810-0
- 542 Zhu D, Li R, Liu X, Sun M, Wu J, Zhang N, and Zhu Y. 2014. The Positive Regulatory Roles of the TIFY10
543 Proteins in Plant Responses to Alkaline Stress. *Plos One* 9. 10.1371/journal.pone.0111984

544

Table 1 (on next page)

Protein information of *NHL* family genes in wild soybean.

1 **Table 1 Protein information of *NHL* family genes in wild soybean.**

Number	Gene ID	Gene Name	Chr	Amino acid residues	MW (kDa)	pI
1	<i>GsNHL1</i>	<i>GlysoPI483463.01G116400</i>	1	230	25.81956	8.95
2	<i>GsNHL2</i>	<i>GlysoPI483463.01G198800</i>	1	190	21.52359	10.15
3	<i>GsNHL3</i>	<i>GlysoPI483463.02G162500</i>	2	208	23.91979	9.79
4	<i>GsNHL4</i>	<i>GlysoPI483463.02G162700</i>	2	245	27.62919	9.26
5	<i>GsNHL5</i>	<i>GlysoPI483463.02G199000</i>	2	256	28.76769	10.12
6	<i>GsNHL6</i>	<i>GlysoPI483463.02G230000</i>	2	274	30.45435	9.86
7	<i>GsNHL7</i>	<i>GlysoPI483463.03G161800</i>	3	222	24.93808	9.58
8	<i>GsNHL8</i>	<i>GlysoPI483463.03G162100</i>	3	208	24.13909	9.46
9	<i>GsNHL9</i>	<i>GlysoPI483463.03G162300</i>	3	230	26.77082	9.14
10	<i>GsNHL10</i>	<i>GlysoPI483463.03G162400</i>	3	204	23.49148	9.53
11	<i>GsNHL11</i>	<i>GlysoPI483463.03G162500</i>	3	210	23.77761	9.81
12	<i>GsNHL12</i>	<i>GlysoPI483463.03G213100</i>	3	239	26.31572	9.37
13	<i>GsNHL13</i>	<i>GlysoPI483463.03G218900</i>	3	198	21.55946	9.83
14	<i>GsNHL14</i>	<i>GlysoPI483463.03G162200</i>	3	228	26.24044	9.50
15	<i>GsNHL15</i>	<i>GlysoPI483463.04G093700</i>	4	256	28.40712	9.51
16	<i>GsNHL16</i>	<i>GlysoPI483463.04G185500</i>	4	211	23.22106	9.69
17	<i>GsNHL17</i>	<i>GlysoPI483463.05G146000</i>	5	348	40.00232	10.00
18	<i>GsNHL18</i>	<i>GlysoPI483463.05G185300</i>	5	214	22.94057	9.54
19	<i>GsNHL19</i>	<i>GlysoPI483463.06G095800</i>	6	260	28.63512	9.04
20	<i>GsNHL20</i>	<i>GlysoPI483463.06G124800</i>	6	224	25.31161	8.97
21	<i>GsNHL21</i>	<i>GlysoPI483463.07G008800</i>	7	255	27.95721	9.68
22	<i>GsNHL22</i>	<i>GlysoPI483463.07G045200</i>	7	253	28.11260	9.46
23	<i>GsNHL23</i>	<i>GlysoPI483463.07G092400</i>	7	210	24.10902	9.43
24	<i>GsNHL24</i>	<i>GlysoPI483463.07G092500</i>	7	208	24.15102	9.82
25	<i>GsNHL25</i>	<i>GlysoPI483463.09G134000</i>	9	204	23.68964	9.53
26	<i>GsNHL26</i>	<i>GlysoPI483463.09G151000</i>	9	315	34.59544	9.85

27	GsNHL27	GlysoPI483463.10G070800	10	247	27.67418	9.09
28	GsNHL28	GlysoPI483463.10G070900	10	216	24.78271	8.71
29	GsNHL29	GlysoPI483463.10G071000	10	228	26.25031	9.64
30	GsNHL30	GlysoPI483463.10G071100	10	210	24.00494	10.02
31	GsNHL31	GlysoPI483463.10G071200	10	200	22.61333	7.82
32	GsNHL32	GlysoPI483463.10G071700	10	223	24.65133	9.10
33	GsNHL33	GlysoPI483463.10G105800	10	149	16.37778	9.39
34	GsNHL34	GlysoPI483463.10G214600	10	228	26.04372	9.56
35	GsNHL35	GlysoPI483463.11G015000	11	179	19.59170	9.55
36	GsNHL36	GlysoPI483463.11G145400	11	215	23.56234	9.96
37	GsNHL37	GlysoPI483463.11G186300	11	246	27.22966	8.88
38	GsNHL38	GlysoPI483463.12G084400	12	214	23.57643	9.96
39	GsNHL39	GlysoPI483463.12G153200	12	222	24.646.62	9.34
40	GsNHL40	GlysoPI483463.12G178100	12	218	24.24817	9.81
41	GsNHL41	GlysoPI483463.13G250400	13	233	26.33392	7.93
42	GsNHL42	GlysoPI483463.13G263900	13	272	29.99915	10.01
43	GsNHL43	GlysoPI483463.13G298400	13	255	27.84143	10.24
44	GsNHL44	GlysoPI483463.14G036900	14	274	30.37919	9.65
45	GsNHL45	GlysoPI483463.14G166200	14	259	29.21918	10.15
46	GsNHL46	GlysoPI483463.15G014800	15	310	34.24409	9.95
47	GsNHL47	GlysoPI483463.16G088600	16	208	22.74147	9.23
48	GsNHL48	GlysoPI483463.16G179000	16	193	20.53805	9.30
49	GsNHL49	GlysoPI483463.18G044000	18	245	27.85761	9.97
50	GsNHL50	GlysoPI483463.18G046800	18	238	26.32375	8.06
51	GsNHL51	GlysoPI483463.19G161400	19	215	24.27526	9.74
52	GsNHL52	GlysoPI483463.19G161500	19	228	26.41171	9.34
53	GsNHL53	GlysoPI483463.19G161600	19	228	26.43265	9.34
54	GsNHL54	GlysoPI483463.19G161700	19	281	23.90965	9.81
55	GsNHL55	GlysoPI483463.19G161800	19	282	31.18464	8.22
56	GsNHL56	GlysoPI483463.19G209900	19	245	26.96660	9.77

57	<i>GsNHL57</i>	<i>GlysoPI483463.19G216600</i>	19	198	21.53948	9.89
58	<i>GsNHL58</i>	<i>GlysoPI483463.20G108700</i>	20	228	26.04367	9.36
59	<i>GsNHL59</i>	<i>GlysoPI483463.20G179000</i>	20	251	27.97813	9.74

2

Figure 1

Figure 1. The phylogenetic and conserved domain analysis of *NHL* family proteins.

a The phylogenetic analysis of wild soybean *NHL* family proteins. The maximum-likelihood (ML) phylogenetic tree was constructed based on 1000 replications for each branch. **b** The motif composition of *NHL* family proteins was identified using MEME online software, and the motif were displayed by boxes of different numbers and colors. The TBtools software was used to combine the conserved motifs and phylogenetic tree.

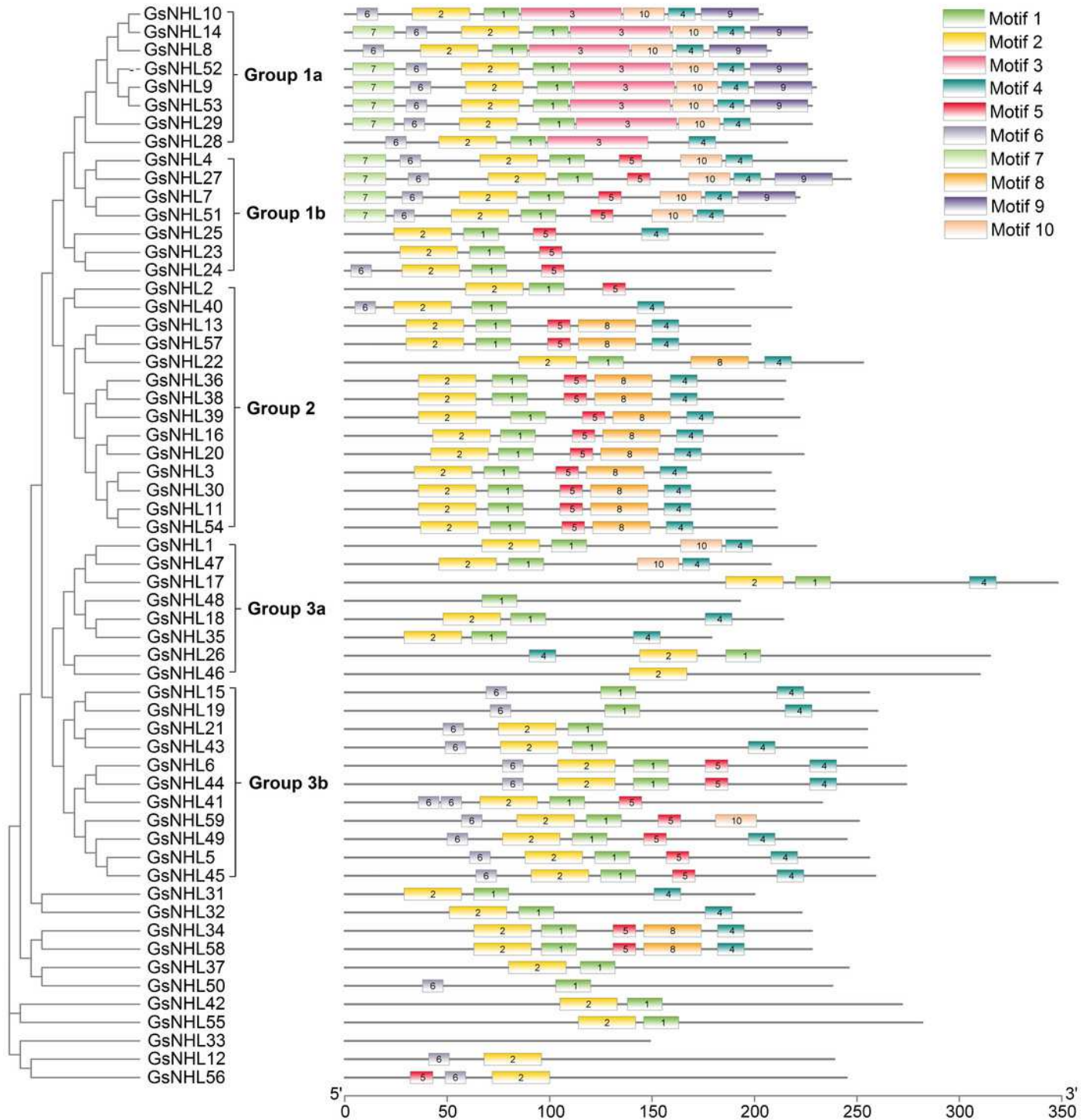


Figure 2

Chromosomal locations and syntenic analysis of *NHL* family genes in wild soybean.

The Circos-0.69 software was used to generate the chromosomes as a circle. The pair of duplication genes were identified and connected by different color lines.

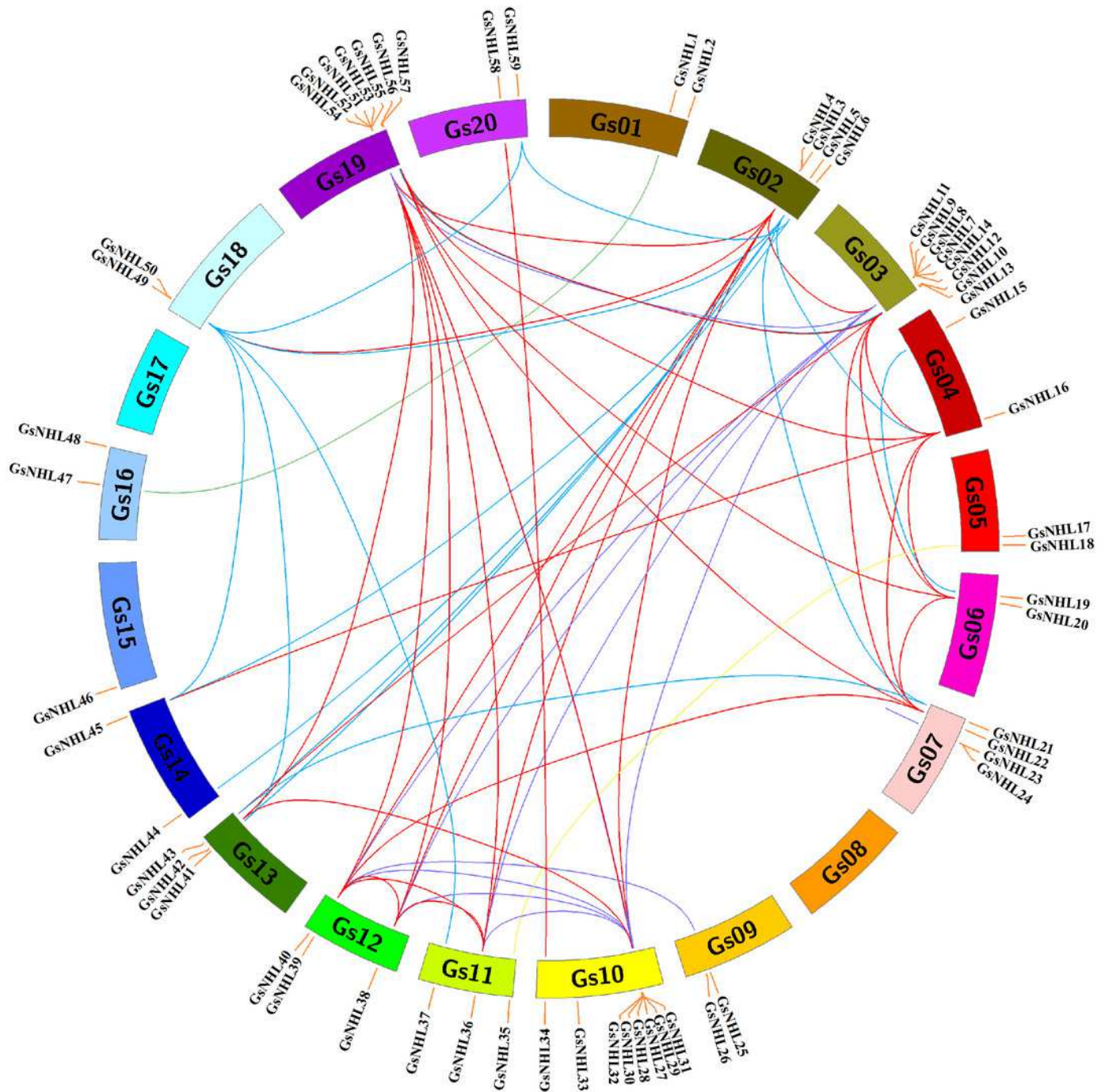


Figure 3

Analysis of *cis*-acting elements of putative *NHL* family gene promoters related to hormones and abiotic stress responses.

The potential regulatory *cis*-acting elements were analyzed in the 3000 bp upstream of translation start site by using the PLANT CARE online software.

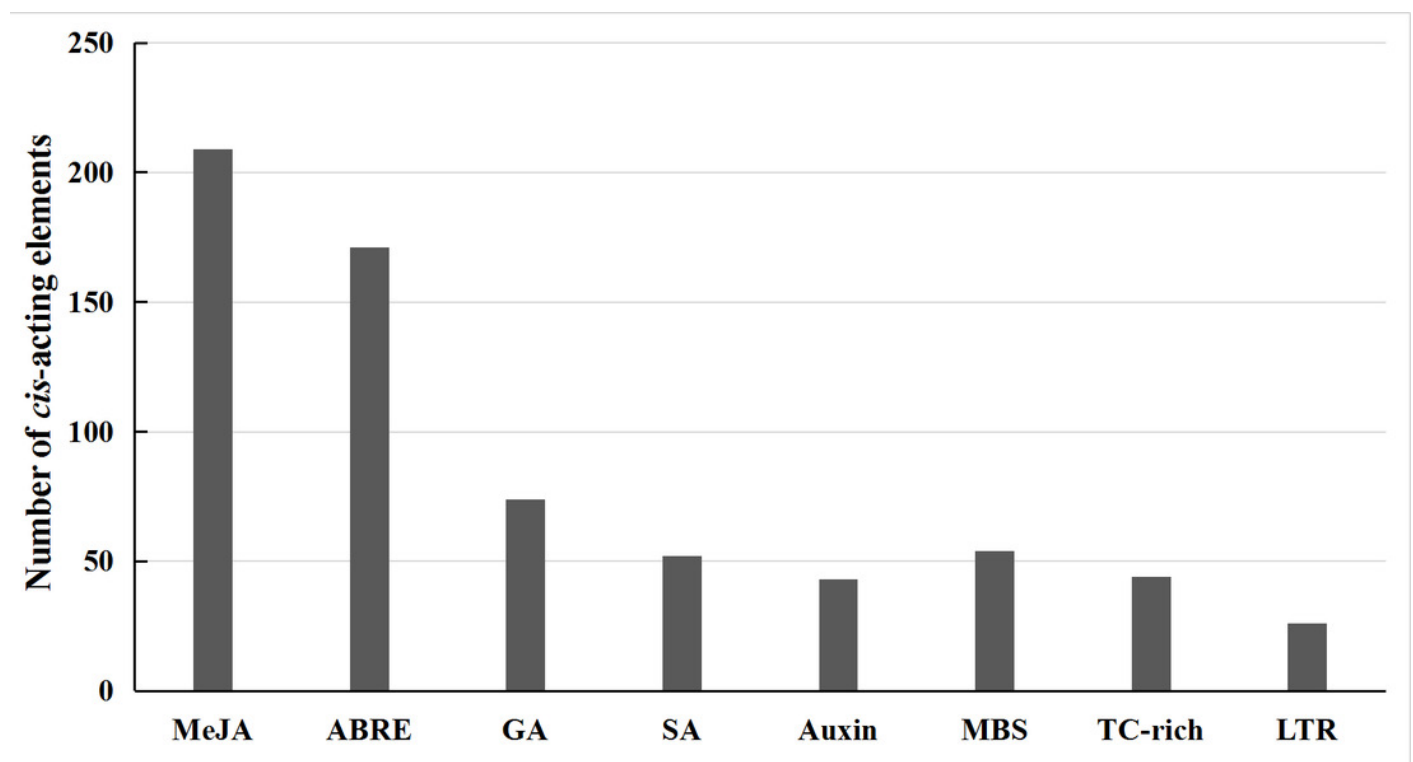


Figure 4

Expression patterns of *NHL* family genes in response to alkaline stress.

The wild soybean transcriptome data were used to detect the expression pattern of *NHL* family genes under alkaline stress. The TM4: MeV4.9 software was used to generate the heat map. The blue and yellow colors represent high or low expression levels ($|\text{Log}_2$ fold change > 2 , $P < 0.05$), respectively.

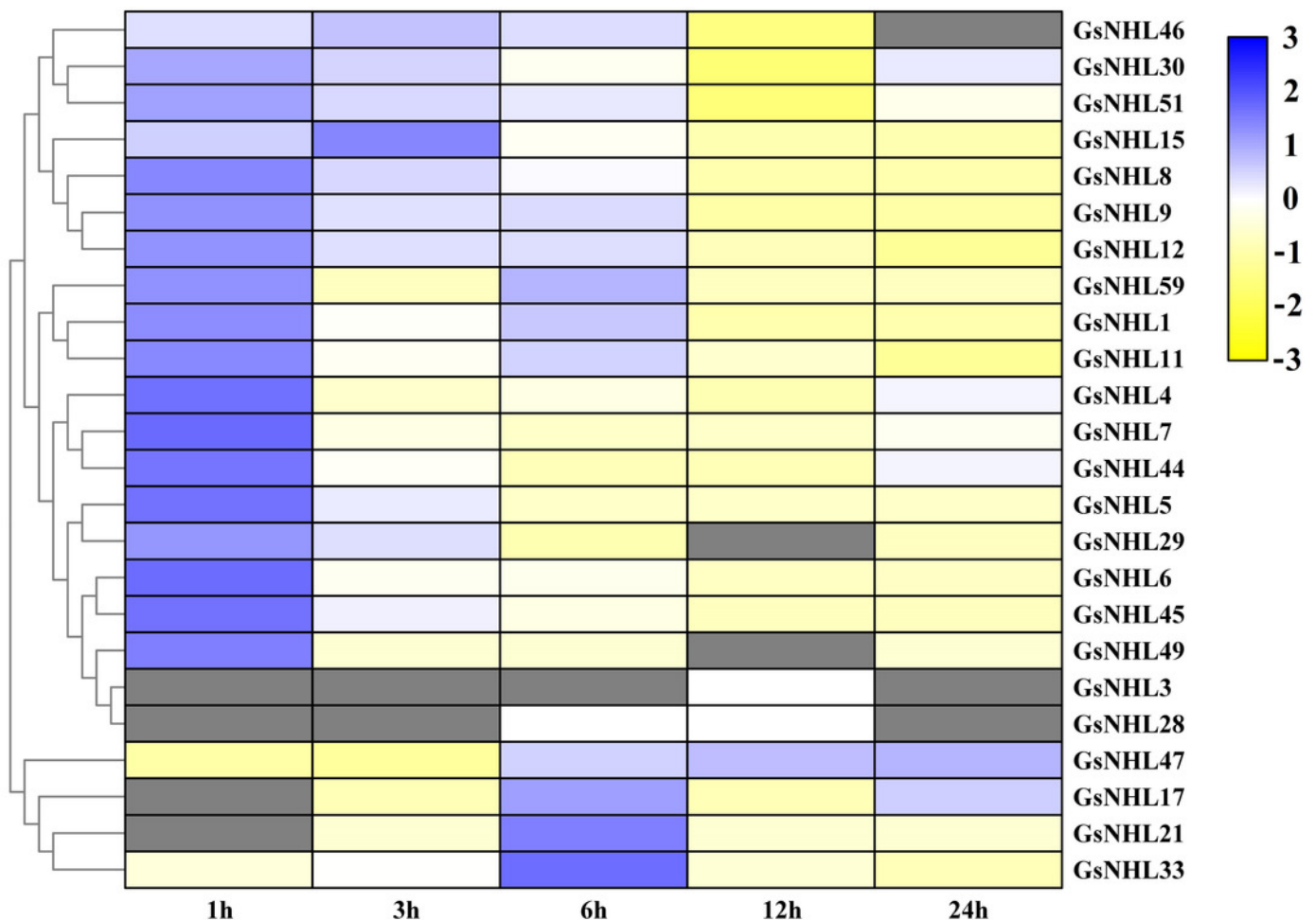


Figure 5

Expression analysis of wild soybean *NHL* family genes in response to alkaline stress.

a-l The wild soybean seedlings were treated with 50 mM NaHCO₃ for 0, 1 and 3 h. The qRT-PCR results were analyzed using the 2^{-ΔΔCT} method and Student's *t*-test.

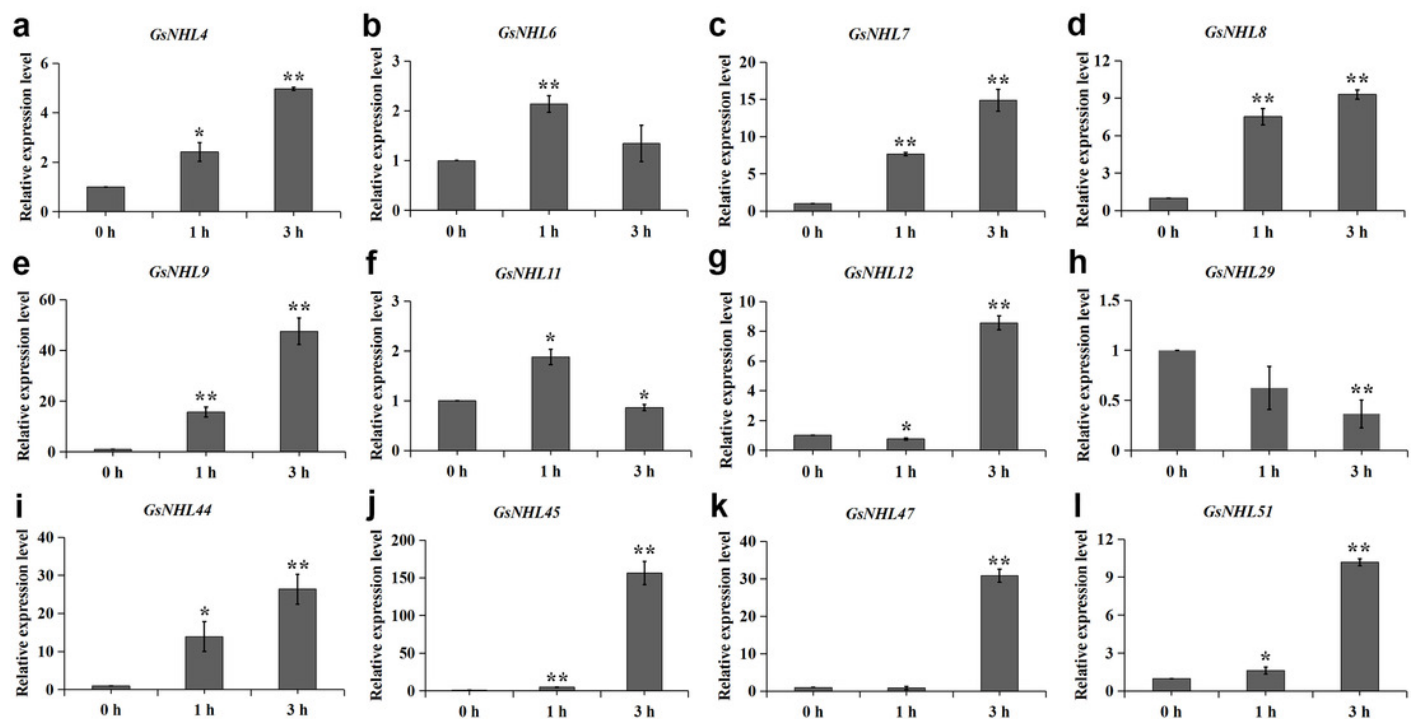


Figure 6

Expression analysis of wild soybean *NHL* family genes in response to MeJA.

The wild soybean seedlings were treated with 50 μ M MeJA for 0, 1 and 3 h. The qRT-PCR results were analyzed using the $2^{-\Delta\Delta CT}$ method and Student's *t*-test.

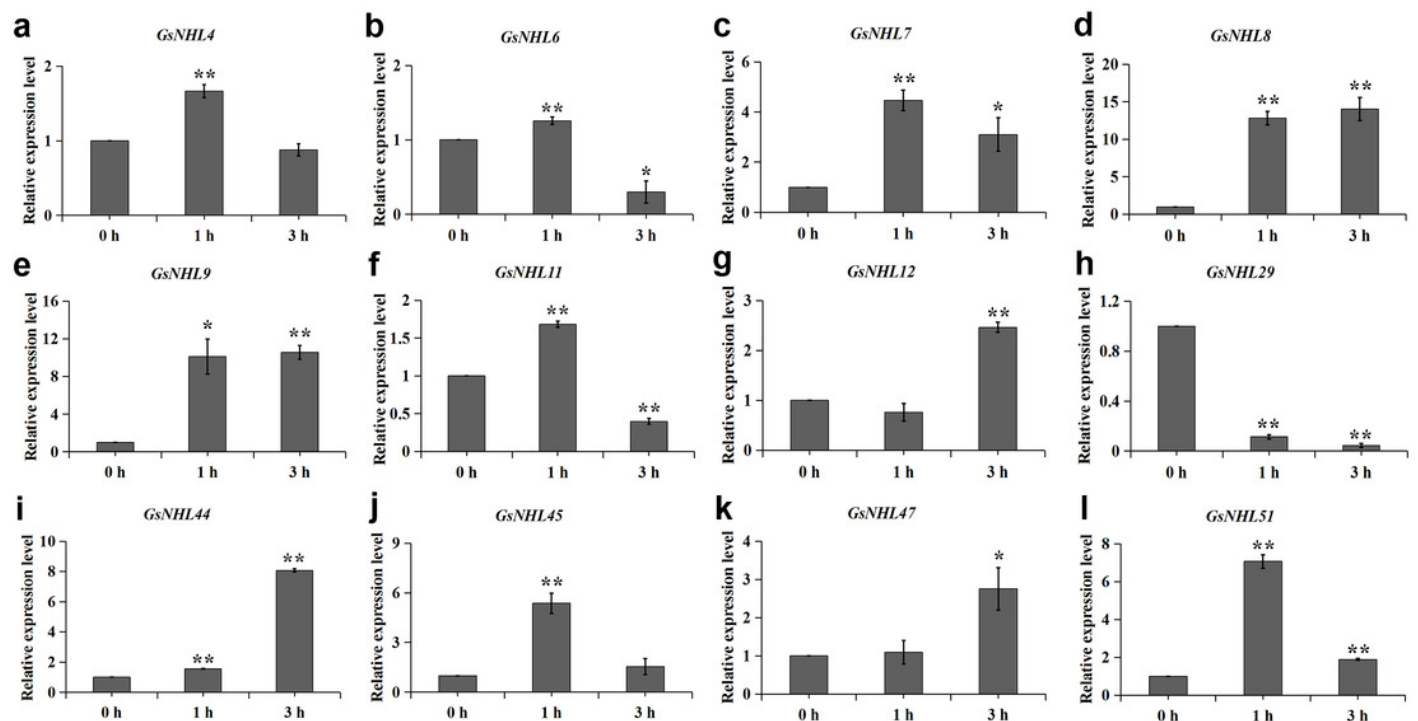


Figure 7

Expression analysis of wild soybean *NHL* family genes in response to ABA.

a-l The wild soybean seedlings were treated with 50 μ M ABA for 0, 1 and 3 h. The qRT-PCR results were analyzed using the $2^{-\Delta\Delta CT}$ method and Student's *t*-test.

