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

First submission

Guidance from your Editor

Please submit by **28 Aug 2022** for the benefit of the authors (and your token reward) .

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.

8 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.

Standout reviewing tips

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 Wang** [1], **Zaib_un Nisa** [3], **Xiaoxia Jin** [1], **Lijie Yu** [1], **Xinlei Liu** [2], **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 neighbor-joining phylogenetic tree and conserved motifs were generated by using MEME online server and MEGA 5.0 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 their 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 the 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.

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 [1, a], Yongguo Xue[2, a], Haihang Wang[1], Zaib_un Nisa[3], Xiaoxia Jin[1], Lijie Yu[1], Xinlei Liu[2], Yang Yu[4*], Chao Chen[1,*]

[1] School of Life Science and Technology, Harbin Normal University, Harbin, P.R. China
[2] Institute of Soybean Research, Heilongjiang Provincial Academy of Agricultural Sciences, Harbin, P.R. China
[3] Institute of Molecular Biology and Biotechnology, University of Lahore, Defence road campus, Lahore, Pakistan
[4] College of Life Science and Engineering, Shenyang University, Shenyang, Harbin, P.R. China

Corresponding Author:
Chao Chen; Yang Yu
Harbin Normal University, 1 South Normal University Road, Songbei Limin Economic Development Zone, Harbin 150025, P.R. China
Email address: chchao@hrbnu.edu.cn
a: These authors contributed equally to this work.

Abstract

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 neighbor-joining phylogenetic tree and conserved motifs were generated by using MEME online server and MEGA 5.0 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 their 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 the 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.

Introduction

NHL (*NDR1/HIN1-like*) family genes are known to be involved in plant resistance responses to some pathogens attack. The *HIN1* gene (harpin-induced gene), which could be rapidly activated and elicit HR (hypersensitive response) phenomenon when plants exposed to bacterial pathogens (e.g. *Pseudomonas syringae* pv. syringae), was first isolated in tobacco (Gopalan et al., 1996). Previous studies revealed that there are at least 29 *NHL* family members that show homology to *NDR1* and *HIN1* gene in Arabidopsis based on the non-redundant Genbank database and then increased this number to 45 upon the completed genome sequencing of Arabidopsis (Dormann et al., 2000;Zheng et al., 2004). The *NDR1* gene (non-race-specific disease resistance gene) was first identified to play distinct roles in response to both bacterial and fungal pathogen resistance in Arabidopsis (Century et al., 1995). Therefore, a lot of *NHL* genes were isolated and identified to play important roles in triggering plants pathogen resistance (Dagvadorj et al., 2022;Varet et al., 2003).

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

It has been documented that *NHL* family genes also play distinct roles in plant abiotic stress resistance. In pepper (*Capsicum annuum* L.), fifteen *NHL* genes were identified in a genome-wide and the responses of these genes to different abiotic stresses were characterized (Liu et al., 2020). A stress-inducing *NHL* member, *BnNHL18A*, was isolated from *Brassica napus*, and played a roles in response to different treatments including NaCl, H_2O_2, as well as ethephon and SA (Lee et al., 2006). Overexpression of pepper *NHL6* could increase the sensitive of salt and osmotic stress (Liu et al., 2020). However, the roles of *NHL* family have not explored under alkaline stress.

In previous studies, we have identified a highly adaptable saline-alkali soil tolerant wild soybean (*Glycine soja*) line (G07256). It can survive well in the saline-alkali soil (Ge et

al., 2010). By using transcriptome data, we identified some candidate genes in response to alkaline stress. In this study, the *NHL* family genes were identified in wild soybean genome and the expression patterns of *NHLs* were investigated under exposure to alkaline stress and plant hormone treatments, which might help us to understand the underlying stress responses of wild soybean involved in *NHL* family genes.

Materials and Methods

Identification of NHL family genes in wild soybean genome

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

Bioinformatics analysis of NHL family genes

The conserved motifs of all the potential *NHL* family genes were identified by MEME online server (http://meme-suite.org/) (Bailey et al., 2009). The neighbor-joining phylogenetic tree was constructed by software MEGA 5.0 (Kumar et al., 2008) software. Then, the TBtools software was used to combine the conserved motifs and phylogenetic tree. The synteny analysis was generated with Circos-0.69 (http://circos.ca/) (Krzywinski et al., 2009). To analyze the potential regulatory *cis*-acting elements in the promoters of *NHL* genes, 3000 bp upstream sequences of the above-mentioned genes were extracted based on the genome database. Then we used PLANT CARE online software (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/) to predict and analyze the regulatory cis-acting elements in promoter regions (Lescot et al., 2002). To examine the expression profiles of *NHL* family genes under alkaline stress, wild soybean transcriptome data was downloaded (DuanMu et al., 2015) and hierarchical clustering trees was generated using TM4: MeV4.9 software (Saeed et al., 2006).

Plant material, growth condition and stress treatment

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

Transcript expression analysis by qRT-PCR

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

Results

Identification of NHL genes in wild soybean

To identify the *NHL* family genes in wild soybean, the amino acid sequences of the *NHL* family genes from Arabidopsis was queried against the wild soybean genome via BLAST from NCBI. A total of 183 NHL candidate sequences were obtained based on the Hidden Markov model mentioned above in the Materials and Methods part. All candidate sequences were then subjected to Pfam and SMART database to remove the redundant

sequences or sequences with incomplete conserved domains. As a result, 59 genes were obtained as potential *NHL* family genes in wild soybean genome.

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

Phylogenetic and conserved motifs analysis of wild soybean NHL genes

To confirm the evolutionary relationship and classification of *NHL* family genes in wild soybean, a neighbor-joining phylogenetic tree was constructed with the full-length protein sequences of the predicted 59 *NHL* family genes. The result showed that the 59 wild soybean *NHL* family genes could be further classified into five groups (group 1, group 2a, group 2b, group 3a and group 3b) and the other 9 ungroup genes (Fig. 1a).

MEME motif analysis revealed that wild soybean *NHL* family proteins share six conserved motif sequences (Fig. 1b, supplementary Fig. S1). Most of wild soybean *NHL* family proteins contain conserved motif 1, motif 2 and motif 4. Interestingly, we found that some wild soybean *NHL* family proteins in the same group share a similar motif composition. For example, motif 5 is mainly present within group 1, group 2b and group 3a genes, while most genes of group 2a, group 2b group 3a and group 3b contain motif 6. Specially, the motif 3 only located in group 3b genes. The similar motif arrangement among the proteins of wild soybean *NHL* family suggested that the protein architecture was conserved within subgroups, which indicated that the proteins in the same group may have similar function in plant development and resistance responses to stress conditions. However, functions of these conserved motifs are still need to be further elucidated.

Chromosomal locations and syntenic analysis

The analysis of gene duplication events could drive the potential evolution mechanisms of the wild soybean *NHL* family genes. In this study, 59 wild soybean *NHL* family genes

were found randomly distributed among 18 chromosomes, with the exception of 8 and 17 (Fig. 2). The wild soybean *NHL* family genes from the same group were mostly located in different chromosomes and located near the edges of the chromosomes, suggesting a strategy to exert their functions in the whole wild soybean genome. Further, a total of 65 pairs of *NHL* syntenic paralogs were identified in wild soybean genome. The results indicated that the wild soybean *NHL* family have been exhibited a high gene family expansion, which might play significant roles in gene functional diversity (Bowers et al., 2003).

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

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

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

To assess the potential roles of *NHL* family genes participate in the defense to alkaline stress, we generated a heat map of *NHL* family genes based on the wild soybean transcriptome data under alkaline stress. The results showed that 24 genes were differently induced under alkaline stress. Among them, 18 of *NHL* family genes were significantly up-regulated, while six genes showed down-regulation patterns (Fig. 4). To

further confirm the expression of *NHL* family genes in response to alkaline treatment, we selected 12 of the up-regulated genes to detect their expression patterns under 50 mM $NaHCO_3$ stress by using qRT-PCR analysis. As shown in Fig. 5, the expression patterns of 11 up-regulated genes were roughly consistent with the transcriptome data under alkaline stress, except that *GsNHL29* had contrary results. In addition, the expression of *GsNHL9*, *GsNHL44*, *GsNHL45* and *GsNHL47* showed higher expression levels at 3 h point than the other genes (Fig. 5e, i-k). In conclusion, the qRT-PCR analysis confirmed the results that *GsNHL* family genes possibly participate in responses to alkaline stress.

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

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

Discussion

Previously, studies have identified that *NHL* family genes are involved in plant development and pathogens attack resistance (Bao et al., 2016;Chen et al., 2021). Many *NHL* family genes have been identified in plant species, such as tomato and pepper (Dormann et al., 2000;Liu et al., 2020). However, the wild soybean *NHL* family genes have not been identified, especially the roles of *NHL* family genes in regulating alkaline

stress. Hence, this research was based on bioinformatics analysis about wild soybean *NHL* family genes in order to understand their structure and location, and mainly potential roles were investigated in response to plant hormones treatment and alkaline stress treatments.

In this study, 59 wild soybean *NHL* family genes were identified in accordance with the Arabidopsis *NHL* related genes (Table 1). On the basis of bioinformatics analysis, we found *NHL* family genes could be classified into five groups, and each groups almost share a similar motif composition, which indicated that the groups may have similar roles in plant development progress (Fig. 1a). A total of 65 pairs of wild soybean *NHL* syntenic paralogs were identified, indicating this family existed a high gene family expansion (Fig. 2). In addition, we found that conserved motif 1, 3, 4 and 5 belong to the LEA-2 domain (Fig. 1b, Supplementary Fig. S1). This result also consistent with previous study (Liu et al., 2020). Furthermore, we found that LEA-2 domain belong to the LEA_2 subgroup which are widely known as a late embryogenesis abundant proteins and play significant roles under abiotic stress responses (Jin et al., 2019) . For example, the rice LEA proteins showed accumulation during the salinity-triggered growth, while degradation in LEA proteins was observed during plant recovery from salt stress (Chourey et al., 2003). The tea plant LEA genes were significantly induced under stress conditions, such as drought, ABA, low and high temperature (Jin et al., 2019). Overexpression of *IpLEA* could show high tolerance to salt and drought stress in *Ipomoea pescaprae* by mediating water homeostasis and as a reactive oxygen species scavenger (Zheng et al., 2019). Thus, this evidence indicated the potential roles of wild soybean *NHL* family genes in response to environmental stresses.

The *cis*-acting regulatory elements play important roles in molecular switches to control various biological processes, including hormone and various stress responses (Sun et al., 2021). The *cis*-acting regulatory elements analysis showed that the promoter regions of wild soybean *NHL* family genes contains a variety of putative hormone-related and abiotic stress responsive elements (Fig. 3). Previous studies have been shown that *NHL* family genes participate in the plant hormone-mediated pathways. For example, overexpression of *AtNHL1* and *AtNHL8* in soybean could enhance plants resistance to *Heterodera glycines* by mediating the jasmonic acid and ethylene pathways, confirming

the roles of these genes in plant defense response (Maldonado et al., 2014). *NHL6* participates in the abiotic stresses-induced ABA signaling at seed germination and early seedling stages in Arabidopsis (Bao et al., 2016) . *BnNHL18A* could be significantly induced by NaCl, ethephon, methyl jasmonate or salicylic acid treatment in *Brassica napus* (Lee et al., 2006). Also, LTR is a *cis*-element responsive to low-temperature stress (Brown et al., 2001). TC-rich *cis*-element has been identified to be involved in stress mediated plant defenses responses by binding with a dehydrating responsive element gene (Mazegari et al., 2015). In conclusion, these evidence strongly suggested that *NHL* family genes may be involved in stress resistance and plant hormones responses in wild soybean.

Alkaline stress is one of the most harmful abiotic stresses, which leading to a series of regulatory mechanisms in plants, such as ion balance, osmotic adjustment, pH regulation, and ROS scavenging mechanisms. In this study, we mainly intend to explore the potential roles of wild soybean *NHL* family genes in response to alkaline stress. The transcriptome data showed that a total of 24 genes were significantly induced under alkaline stress (Fig. 4), and qRT-PCR confirmed the results that wild soybean *NHL* family genes may play positive role in response to alkaline stress (Fig. 5).

During abiotic stress responsive processes, plant hormones such as MeJA, ABA, SA, GA and Auxin also play important roles and have cross talks with these signal transduction (Ku et al., 2018). In our previous study, we also identified that plant hormones have crosstalk with plant alkaline stress resistance response. For example, the wild soybean gene *ERF71* could regulate endogenous auxin accumulation when plants treated with alkaline solution (Yu et al., 2017). The *TIFY10* gene could act as a regulator in response to alkaline stress and jasmonate signaling in wild soybean (Zhu et al., 2011). On the other hand, we found that wild soybean *NHL* family genes comprised of a variety of putative hormone-related responsive elements, and the numbers of MeJA and ABA responsive elements were significantly larger than others. Thus, we speculated that the roles of wild soybean *NHL* family genes under alkaline stress have crosstalk with MeJA or ABA signal transduction. qRT-PCR analysis showed that nine genes were up-regulated under MeJA treatment, and these genes were all up-regulated under alkaline stress (Fig. 6). This finding is consistent with the previous studies that wild soybean *TIFY10a*

overexpression lines enhanced the alkaline stress resistance and also increased the jasmonate content of the transgenic alfalfa (Zhu et al., 2014). However, in comparison with MeJA treatment, qRT-PCR analysis showed a different expression, in which only two wild soybean *NHL* family genes were up-regulated and seven genes were down-regulated under ABA treatment (Fig. 7). For example, *GsNHL7*, *GsNHL8*, *GsNHL9*, *GsNHL12* and *GsNHL7* were up-regulated under alkaline stress and MeJA treatment, while were down-regulated under ABA treatment. In addition, our previous studies identified that overexpression of wild soybean *NAC019* or *SKP21* in *Arabidopsis* could contribute to alkaline stress tolerance, but reduced ABA sensitivity(Cao et al., 2017;Liu et al., 2015). In conclusion, these results speculated that wild soybean *NHL* family genes have crosstalk with MeJA or ABA signal transduction under alkaline stress, and some genes may display a different roles in ABA or MeJA signal transduction in response to alkaline stress.

Conclusions

In conclusion, in this study, we identified 59 potential *NHL* family genes in wild soybean. We also identified the phylogenetic analysis, conserved domain, gene duplication and *cis*-acting elements in promoters. The results suggested that wild soybean *NHL* family genes could be mainly classified into five groups as well as exist with conserved motifs, high gene family expansion, hormone-related and abiotic stress responsive elements. In addition, we confirmed that wild soybean *NHL* family genes may play important regulatory roles in response to alkaline stress, ABA and MEJA treatment. Taken together, our results established a foundation for characterization of wild soybean *NHL* family genes in response to alkaline stress, ABA and MEJA treatment. However, the potential roles of wild soybean *NHL* family genes in crosstalk with MeJA or ABA signal transduction under alkaline stress still need to explore in the future.

Abbreviations

NHL	NDR1/HIN1-like
qRT-PCR	quantitative real-time PCR
MW	molecular weight

pI	isoelectric point
ABA	abscisic acid
MeJA	methyl jasmonate
SA	salicylie acid
GA	gibberellin
MBS	drought
LTR	low temperature responsive
TC-rich	defense and stress responsive

References

Artimo P, Jonnalagedda M, Arnold K, Baratin D, Csardi G, de Castro E, et al., 2012. ExPASy: SIB bioinformatics resource portal. Nucleic Acids Research. 40:W597-W603.

Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, et al., 2009. MEME SUITE: tools for motif discovery and searching. Nucleic Acids Research. 37:W202-W8.

Bao Y, Song W-M, Pan J, Jiang C-M, Srivastava R, Li B, et al., 2016. Overexpression of the NDR1/HIN1-Like Gene NHL6 Modifies Seed Germination in Response to Abscisic Acid and Abiotic Stresses in Arabidopsis. Plos One. 11.

Bowers JE, Chapman BA, Rong JK, Paterson AH. 2003. Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. Nature. 422:433-8.

Brown APC, Dunn MA, Goddard NJ, Hughes MA. 2001. Identification of a novel low-temperature-response element in the promoter of the barley (Hordeum vulgare L) gene blt101.1. Planta. 213:770-80.

Cao L, Yu Y, Ding X, Zhu D, Yang F, Liu B, et al., 2017. The Glycine soja NAC transcription factor GsNAC019 mediates the regulation of plant alkaline tolerance and ABA sensitivity. Plant Molecular Biology. 95:253-68.

Century KS, Holub EB, Staskawicz BJ. 1995. NDR1, A LOCUS OF ARABIDOPSIS-THALIANA THAT IS REQUIRED FOR DISEASE RESISTANCE TO BOTH A BACTERIAL AND A FUNGAL PATHOGEN. Proceedings of the National Academy of Sciences of the United States of America. 92:6597-601.

Chen L, He F, Long R, Zhang F, Li M, Wang Z, et al., 2021. A global alfalfa diversity panel reveals genomic selection signatures in Chinese varieties and genomic associations with root development. Journal of Integrative Plant Biology. 63:1937-51.

Chourey K, Ramani S, Apte SK. 2003. Accumulation of LEA proteins in salt (NaCl) stressed young seedlings of rice (Oryza sativa L.) cultivar Bura Rata and their degradation during recovery from salinity stress. Journal of Plant Physiology. 160:1165-74.

Dagvadorj B, Outram MA, Williams SJ, Solomon PS. 2022. The necrotrophic effector ToxA from Parastagonospora nodorum interacts with wheat NHL proteins to facilitate Tsn1-mediated necrosis. Plant Journal. 110:407-18.

Dormann P, Gopalan S, He SY, Benning C. 2000. A gene family in Arabidopsis thaliana with sequence similarity to NDR1 and HIN1. Plant Physiology and Biochemistry. 38:789-96.

DuanMu H, Wang Y, Bai X, Cheng S, Deyholos MK, Wong GK-S, et al., 2015. Wild soybean roots depend on specific transcription factors and oxidation reduction related genesin response to alkaline stress. Functional & Integrative Genomics. 15:651-60.
Finn RD, Clements J, Eddy SR. 2011. HMMER web server: interactive sequence similarity searching. Nucleic Acids Research. 39:W29-W37.
Finn RD, Coggill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, et al., 2016. The Pfam protein families database: towards a more sustainable future. Nucleic Acids Research. 44:D279-D85.
Ge Y, Li Y, Zhu Y-M, Bai X, Lv D-K, Guo D, et al., 2010. Global transcriptome profiling of wild soybean (Glycine soja) roots under NaHCO3 treatment. Bmc Plant Biology. 10.
Gopalan S, Bauer DW, Alfano JR, Loniello AO, He SY, Collmer A. 1996. Expression of the Pseudomonas syringae avirulence protein AvrB in plant cells alleviates its dependence on the hypersensitive response and pathogenicity (Hrp) secretion system in eliciting genotype-specific hypersensitive cell death. Plant Cell. 8:1095-105.
Gopavajhula VR, Chaitanya KV, Khan PAA, Shaik JP, Reddy PN, Alanazi M. 2013. Modeling and analysis of soybean (Glycine max. L) Cu/Zn, Mn and Fe superoxide dismutases. Genetics and Molecular Biology. 36:225-36.
Huis R, Hawkins S, Neutelings G. 2010. Selection of reference genes for quantitative gene expression normalization in flax (Linum usitatissimum L.). Bmc Plant Biology. 10.
Jin H, Plaha P, Park JY, Hong CP, Lee IS, Yang ZH, et al., 2006. Comparative EST profiles of leaf and root of Leymus chinensis, a xerophilous grass adapted to high pH sodic soil. Plant Science. 170:1081-6.
Jin X, Cao D, Wang Z, Ma L, Tian K, Liu Y, et al., 2019. Genome-wide identification and expression analyses of the LEA protein gene family in tea plant reveal their involvement in seed development and abiotic stress responses. Scientific Reports. 9.
Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, et al., 2009. Circos: An information aesthetic for comparative genomics. Genome Research. 19:1639-45.
Ku Y-S, Sintaha M, Cheung M-Y, Lam H-M. 2018. Plant Hormone Signaling Crosstalks between Biotic and Abiotic Stress Responses. International Journal of Molecular Sciences. 19.
Kumar S, Nei M, Dudley J, Tamura K. 2008. MEGA: A biologist-centric software for evolutionary analysis of DNA and protein sequences. Briefings in Bioinformatics. 9:299-306.
Lee SB, Ham BK, Park JM, Kim YJ, Paek KH. 2006. BnNHL18A shows a localization change by stress-inducing chemical treatments. Biochemical and Biophysical Research Communications. 339:399-406.
Lescot M, Dehais P, Thijs G, Marchal K, Moreau Y, Van de Peer Y, et al., 2002. PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. Nucleic Acids Research. 30:325-7.
Liu A, Yu Y, Duan X, Sun X, Duanmu H, Zhu Y. 2015. GsSKP21, a Glycine soja S-phase kinase-associated protein, mediates the regulation of plant alkaline tolerance and ABA sensitivity. Plant Molecular Biology. 87:111-24.
Liu C, Peng H, Li X, Liu C, Lv X, Wei X, et al., 2020. Genome-wide analysis of NDR1/HIN1-like genes in pepper (Capsicum annuum L.) and functional characterization of CaNHL4 under biotic and abiotic stresses. Horticulture research. 7:93-.
Maldonado A, Youssef R, McDonald M, Brewer E, Beard H, Matthews B. 2014. Modification of the expression of two NPR1 suppressors, SNC1 and SNI1, in soybean confers partial resistance to the soybean cyst nematode, Heterodera glycines. Functional Plant Biology. 41:714-26.
Qiao Y, Lu W, Wang R, Nisa Z, Yu Y, Jin X, et al., 2020. Identification and Expression Analysis of Strigolactone Biosynthetic and Signaling Genes in Response to Salt and Alkaline Stresses in Soybean (Glycine max). DNA and Cell Biology. 39:1850-61.
Saeed AI, Hagabati NK, Braisted JC, Liang W, Sharov V, Howe EA, et al., 2006. TM4 microarray software suite. In *DNA Microarrays, Part B: Databases and Statistics.* Kimmel A, Oluver B eds. pp. 134-+.

Sazegari S, Niazi A, Ahmadi FS. 2015. A study on the regulatory network with promoter analysis for Arabidopsis DREB-genes. Bioinformation. 11:101-6.
Sun C, Zhang C, Wang X, Zhao X, Chen F, Zhang W, et al., 2021. Genome-Wide Identification and Characterization of the IGT Gene Family in Allotetraploid Rapeseed (Brassica napus L.). DNA and Cell Biology. 40:441-56.
Varet A, Hause B, Hause G, Scheel D, Lee J. 2003. The Arabidopsis NHL3 gene encodes a plasma membrane protein and its overexpression correlates with increased resistance to Pseudomonas syringae pv. tomato DC3000. Plant Physiology. 132:2023-33.
Vondrackova S, Szakova J, Drabek O, Tejnecky V, Hejcman M, Muellerova V, et al., 2015. Aluminium Uptake and Translocation in Al Hyperaccumulator Rumex obtusifolius Is Affected by Low-Molecular-Weight Organic Acids Content and Soil pH. Plos One. 10.
Wang Y, Ma H, Liu G, Xu C, Zhang D, Ban Q. 2008. Analysis of Gene Expression Profile of Limonium bicolor under NaHCO3 Stress Using cDNA Microarray. Plant Molecular Biology Reporter. 26:241-54.
Yang CW, Wang P, Li CY, Shi DC, Wang DL. 2008. Comparison of effects of salt and alkali stresses on the growth and photosynthesis of wheat. Photosynthetica. 46:107-14.
Yu Y, Duan X, Ding X, Chen C, Zhu D, Yin K, et al., 2017. A novel AP2/ERF family transcription factor from Glycine soja, GsERF71, is a DNA binding protein that positively regulates alkaline stress tolerance in Arabidopsis. Plant Molecular Biology. 94:509-30.
Zheng J, Su H, Lin R, Zhang H, Xia K, Jian S, et al., 2019. Isolation and characterization of an atypical LEA gene (IpLEA) from Ipomoea pes-caprae conferring salt/drought and oxidative stress tolerance. Scientific Reports. 9.
Zheng MS, Takahashi H, Miyazaki A, Hamamoto H, Shah J, Yamaguchi I, et al., 2004. Up-regulation of Arabidopsis thaliana NHL10 in the hypersensitive response to Cucumber mosaic virus infection and in senescing leaves is controlled by signalling pathways that differ in salicylate involvement. Planta. 218:740-50.
Zhu D, Bai X, Chen C, Chen Q, Cai H, Li Y, et al., 2011. GsTIFY10, a novel positive regulator of plant tolerance to bicarbonate stress and a repressor of jasmonate signaling. Plant Molecular Biology. 77:285-97.
Zhu D, Li R, Liu X, Sun M, Wu J, Zhang N, et al., 2014. The Positive Regulatory Roles of the TIFY10 Proteins in Plant Responses to Alkaline Stress. Plos One. 9.

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	*GsNHL1*	*GlysoPI483463.01G116400*	1	230	25.81956	8.95
2	*GsNHL2*	*GlysoPI483463.01G198800*	1	190	21.52359	10.15
3	*GsNHL3*	*GlysoPI483463.02G162500*	2	208	23.91979	9.79
4	*GsNHL4*	*GlysoPI483463.02G162700*	2	245	27.62919	9.26
5	*GsNHL5*	*GlysoPI483463.02G199000*	2	256	28.76769	10.12
6	*GsNHL6*	*GlysoPI483463.02G230000*	2	274	30.45435	9.86
7	*GsNHL7*	*GlysoPI483463.03G161800*	3	222	24.93808	9.58
8	*GsNHL8*	*GlysoPI483463.03G162100*	3	208	24.13909	9.46
9	*GsNHL9*	*GlysoPI483463.03G162300*	3	230	26.77082	9.14
10	*GsNHL10*	*GlysoPI483463.03G162400*	3	204	23.49148	9.53
11	*GsNHL11*	*GlysoPI483463.03G162500*	3	210	23.77761	9.81
12	*GsNHL12*	*GlysoPI483463.03G213100*	3	239	26.31572	9.37
13	*GsNHL13*	*GlysoPI483463.03G218900*	3	198	21.55946	9.83
14	*GsNHL14*	*GlysoPI483463.03G162200*	3	228	26.24044	9.50
15	*GsNHL15*	*GlysoPI483463.04G093700*	4	256	28.40712	9.51
16	*GsNHL16*	*GlysoPI483463.04G185500*	4	211	23.22106	9.69
17	*GsNHL17*	*GlysoPI483463.05G146000*	5	348	40.00232	10.00
18	*GsNHL18*	*GlysoPI483463.05G185300*	5	214	22.94057	9.54
19	*GsNHL19*	*GlysoPI483463.06G095800*	6	260	28.63512	9.04
20	*GsNHL20*	*GlysoPI483463.06G124800*	6	224	25.31161	8.97
21	*GsNHL21*	*GlysoPI483463.07G008800*	7	255	27.95721	9.68
22	*GsNHL22*	*GlysoPI483463.07G045200*	7	253	28.11260	9.46
23	*GsNHL23*	*GlysoPI483463.07G092400*	7	210	24.10902	9.43
24	*GsNHL24*	*GlysoPI483463.07G092500*	7	208	24.15102	9.82
25	*GsNHL25*	*GlysoPI483463.09G134000*	9	204	23.68964	9.53
26	*GsNHL26*	*GlysoPI483463.09G151000*	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	*GsNHL57*	*GlysoPI483463.19G216600*	19	198	21.53948	9.89
58	*GsNHL58*	*GlysoPI483463.20G108700*	20	228	26.04367	9.36
59	*GsNHL59*	*GlysoPI483463.20G179000*	20	251	27.97813	9.74

2

Figure 1

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

a The phylogenetic analysis of wild soybean *NHL* family proteins. The neighbor-joining (NJ) phylogenetic tree was constructed based on 1000 replications for each branch. **b** The motif composition of *NHL* family proteins were 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.

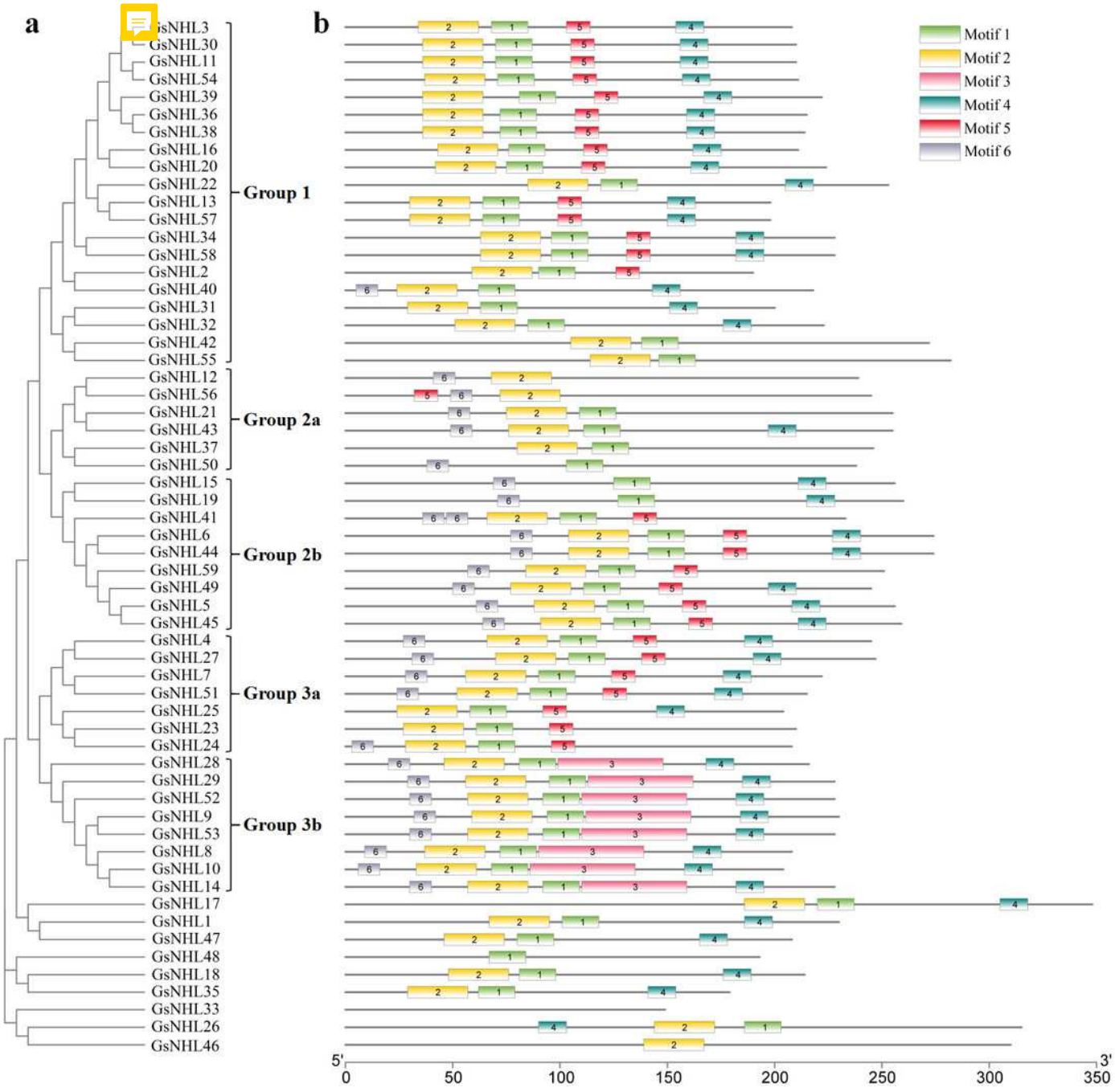
a
b
GsNHL3
GsNHL30
GsNHL11
GsNHL54
GsNHL39
GsNHL36
GsNHL38
GsNHL16
GsNHL20
GsNHL22
GsNHL13
GsNHL57
GsNHL34
GsNHL58
GsNHL2
GsNHL40
GsNHL31
GsNHL32
GsNHL42
GsNHL55
GsNHL12
GsNHL56
GsNHL21
GsNHL43
GsNHL37
GsNHL50
GsNHL15
GsNHL19
GsNHL41
GsNHL6
GsNHL44
GsNHL59
GsNHL49
GsNHL5
GsNHL45
GsNHL4
GsNHL27
GsNHL7
GsNHL51
GsNHL25
GsNHL23
GsNHL24
GsNHL28
GsNHL29
GsNHL52
GsNHL9
GsNHL53
GsNHL8
GsNHL10
GsNHL14
GsNHL17
GsNHL1
GsNHL47
GsNHL48
GsNHL18
GsNHL35
GsNHL33
GsNHL26
GsNHL46
Group 1
Group 2a
Group 2b
Group 3a
Group 3b
Motif 1
Motif 2
Motif 3
Motif 4
Motif 5
Motif 6
5'
3'
0
50
100
150
200
250
300
350

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 are identified and connected by different color lines.

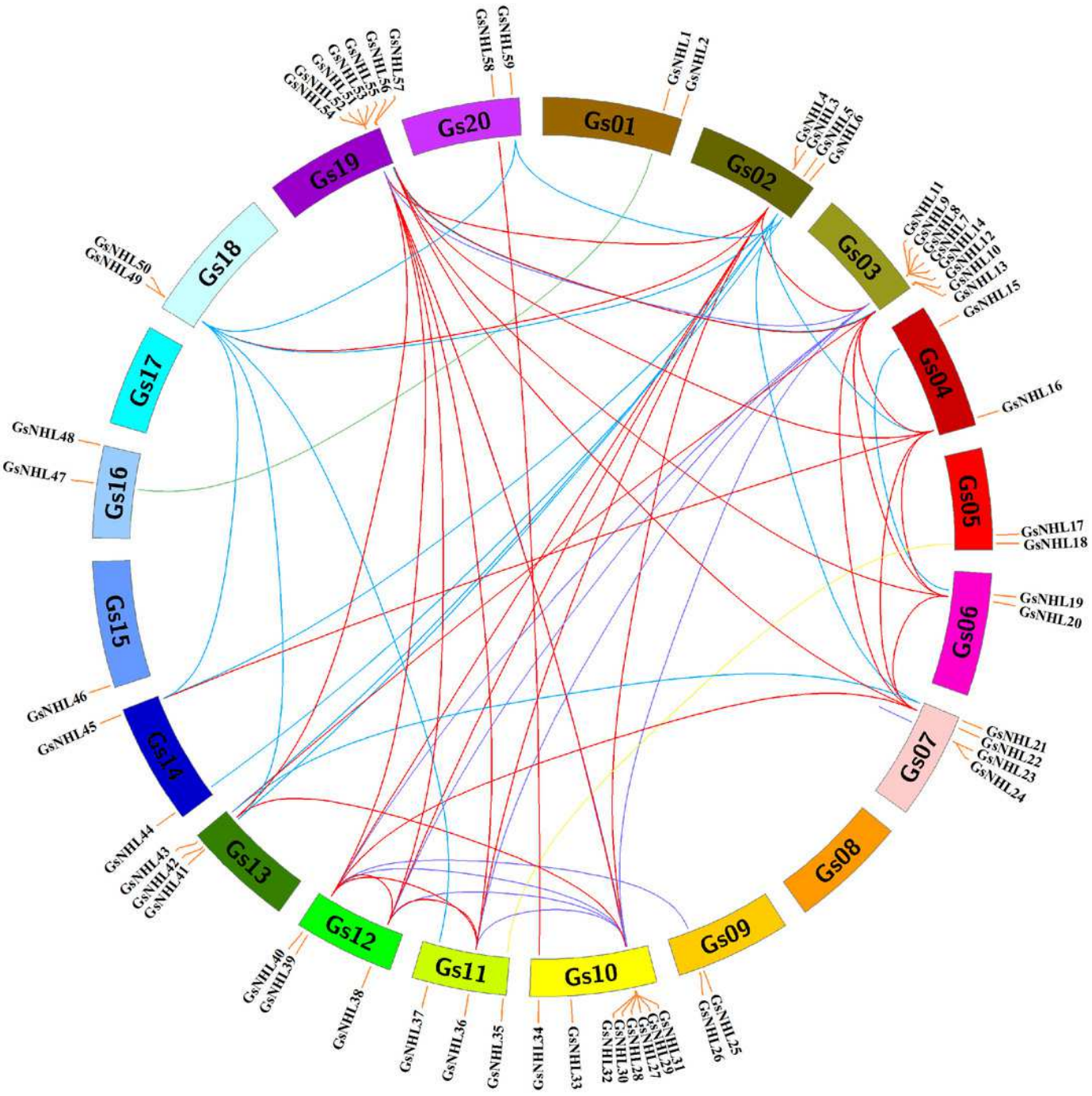
Gs01
Gs02
Gs03
Gs04
Gs05
Gs06
Gs07
Gs08
Gs09
Gs10
Gs11
Gs12
Gs13
Gs14
Gs15
Gs16
Gs17
Gs18
Gs19
Gs20
GsNHL1
GsNHL2
GsNHL3
GsNHL4
GsNHL5
GsNHL6
GsNHL7
GsNHL8
GsNHL9
GsNHL10
GsNHL11
GsNHL12
GsNHL13
GsNHL14
GsNHL15
GsNHL16
GsNHL17
GsNHL18
GsNHL19
GsNHL20
GsNHL21
GsNHL22
GsNHL23
GsNHL24
GsNHL25
GsNHL26
GsNHL27
GsNHL28
GsNHL29
GsNHL30
GsNHL31
GsNHL32
GsNHL33
GsNHL34
GsNHL35
GsNHL36
GsNHL37
GsNHL38
GsNHL39
GsNHL40
GsNHL41
GsNHL42
GsNHL43
GsNHL44
GsNHL45
GsNHL46
GsNHL47
GsNHL48
GsNHL49
GsNHL50
GsNHL51
GsNHL52
GsNHL53
GsNHL54
GsNHL55
GsNHL56
GsNHL57
GsNHL58
GsNHL59

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.

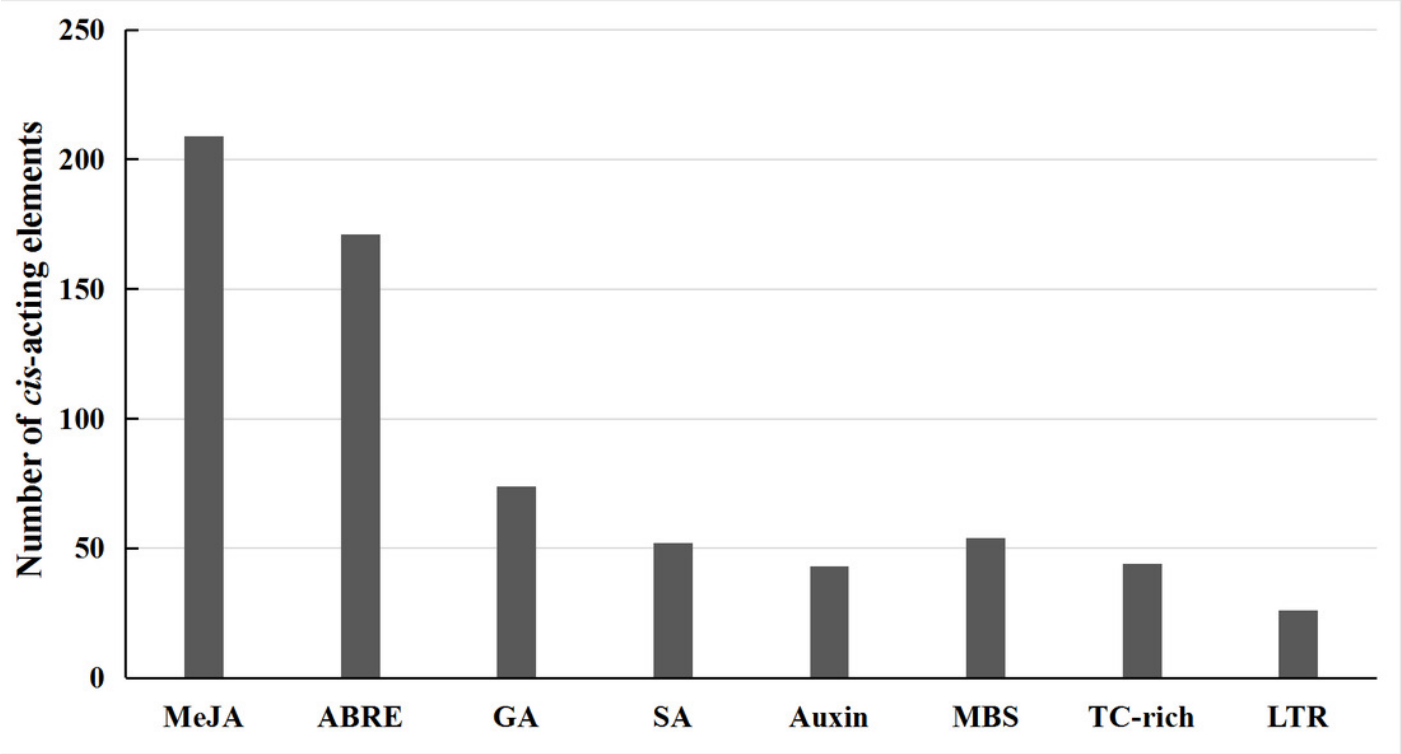

Figure 4

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 (|Log2 fold change| > 2, $P < 0.05$), respectively.

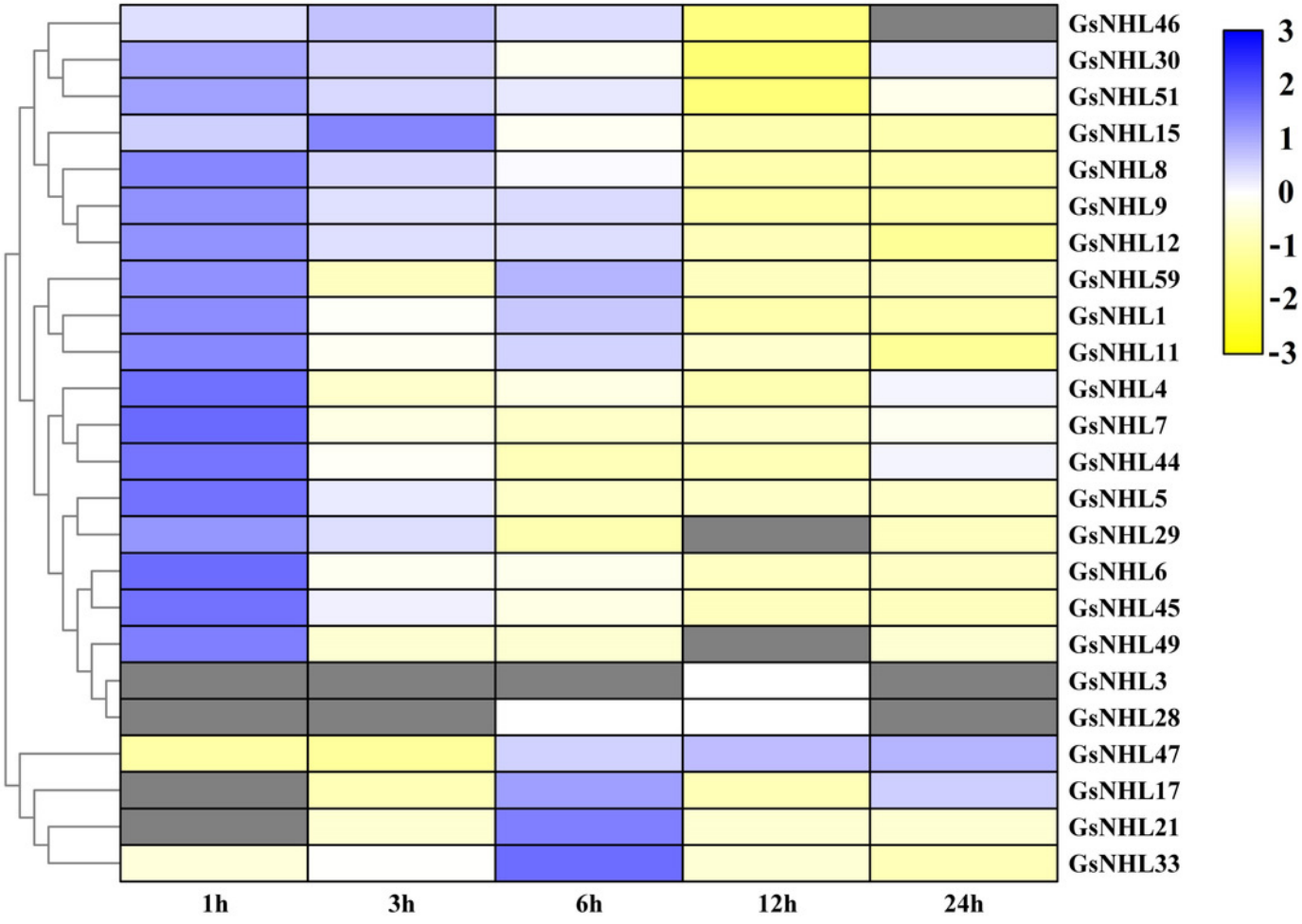

Figure 5

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

a-l The wild soybean seedlings were treated with 50 mM $NaHCO_3$ for 0, 1 and 3 h. The qRT-PCR results were analyzed using the $2^{-\Delta\Delta CT}$ method and Student's *t*-test.

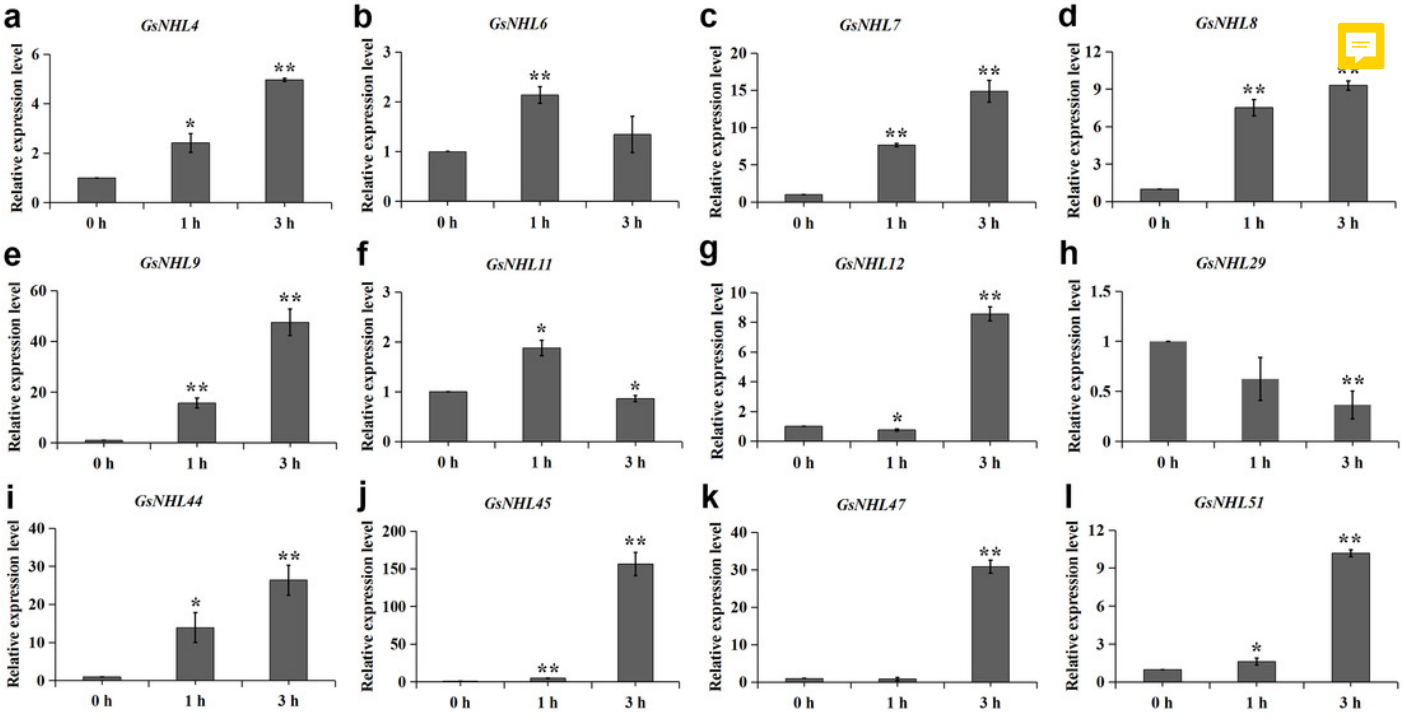

Figure 6

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

The wild soybean seedlings were treated with 50 μM 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.

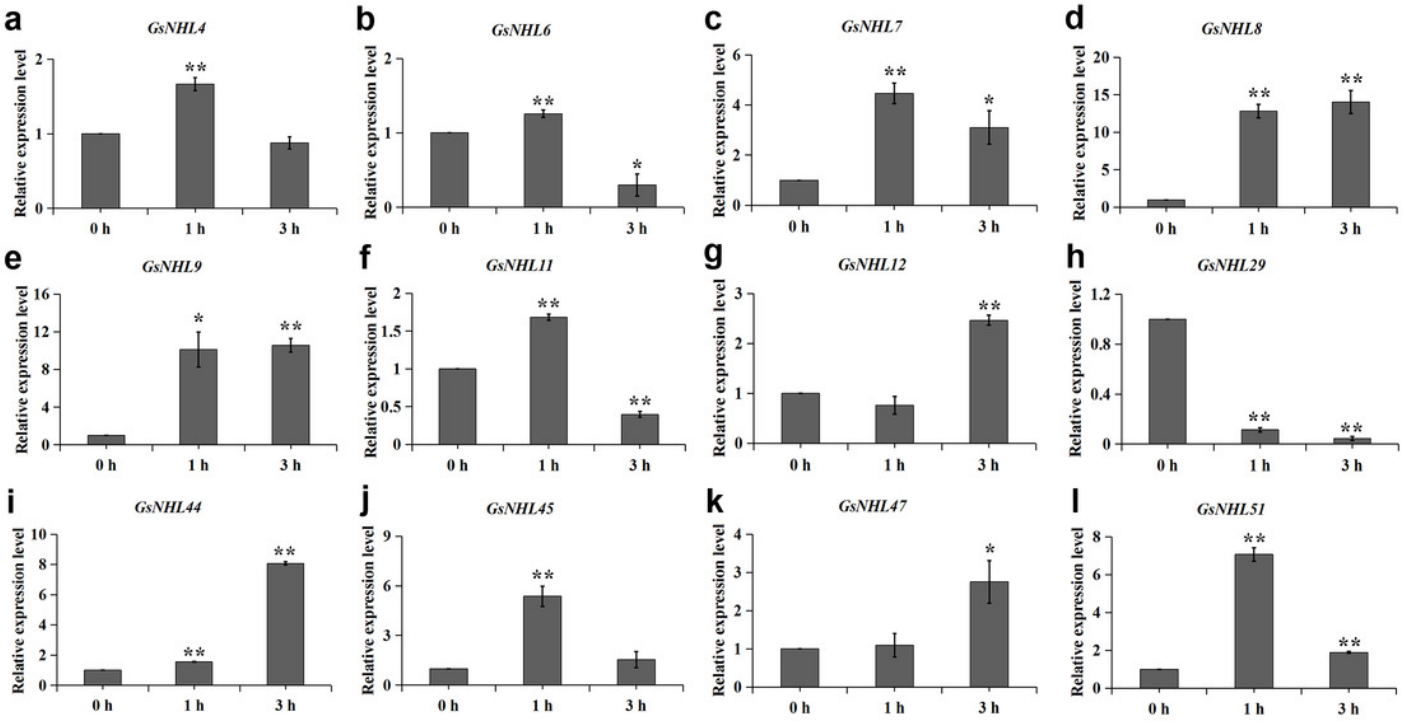

Figure 7

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

a-l The wild soybean seedlings were treated with 50 μM 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.