Genome-wide identification and evolutionary analysis of neutral/alkaline invertases in *Brassica rapa*

Seung Hee Eom\(^a\), Yeonggil Rim\(^b\) and Tae Kyung Hyun\(^a\)

\(^a\)Department of Industrial Plant Science and Technology, College of Agricultural, Life and Environmental Sciences, Chungbuk National University, Cheongju, Republic of Korea; \(^b\)Department of Biochemistry, Systems & Synthetic Agrobiotech Center, Gyeongsang National University, Jinju, Republic of Korea

**ABSTRACT**

Among the several sucrose-catabolizing enzymes, neutral/alkaline invertases (NINs) play crucial roles in developmental processes as well as environmental stress responses in higher plants. Despite the fact that NINs are essential enzymes for plant life, the NIN family and their evolutionary relationships are poorly understood. Therefore, in this study, we identified 11 NINs in the *Brassica rapa* (Chinese cabbage; BraNINs) genome, and analyzed the evolutionary mechanisms of BraNIN genes. Evolution analysis suggested that the BraNIN genes were duplicated via a segmental duplication event originating 15.81–45.25 million years ago. Furthermore, two segmental duplicated pairs (BraNIN5/6 and BraNIN7/8) were subject to negative selection. Furthermore, expression analysis of BraNINs using RNA-sequencing data suggested various functions of BraNINs during responses to drought stress. Taken together, our comparative genomic analysis of NIN genes in *B. rapa* provides information that will assist future studies on sucrose metabolism in sinks and sources of higher plants.

**ARTICLE HISTORY**

Received 6 May 2019
Accepted 11 July 2019

**KEYWORDS**

Neutral/alkaline invertase; Brassica rapa; duplication; drought

**Introduction**

Sucrose plays crucial roles as a major carbohydrate transported from photosynthetic source leaves to heterotrophic sink tissues, in growth, development and responses to various stresses [1–3], and acts as an important signal molecule that regulates the activation and inactivation of genes and microRNA expression levels [4–6]. The efficient use of sucrose as a carbon and energy source depends on its cleavage catalyzed by either invertases to glucose and fructose or its conversion by sucrose synthases to UDP-glucose and fructose [3].

In higher plants, invertases are further classified into two subfamilies that display characteristic pH optima for activity [7]. Acidic invertases (optimum pH 3.5–5.5) are cell wall or vacuole invertases, whereas neutral/alkaline invertases (NINs) with optimum pH from 6.8 to 8.0 are localized to the cytosol, mitochondria, plastids and nuclei [7,8]. Cell wall and vacuole invertases have been suggested to play important roles in physiological processes, including growth, development and responses to environmental stresses [2,3,9]. Acidic invertases are transcriptionally regulated by an array of signals [10,11]; however, the activity of acidic invertases is controlled by proteinaceous inhibitors, all of which are known as small inhibitors of β-fructosidases with sizes ranging from 15 to 23 kDa [12]. This suggests that the post-transcriptional modulation of acidic invertases is required for sugar unloading to sink tissues [13]. Although accumulating evidence has supported the physiological function of acidic invertases in higher plants, very limited information is available on the function of NINs, primarily because of protein instability and low expression and enzymatic activity [8]. Since the time that genes encoding nine NINs were identified in the *Arabidopsis* genome [14], several NINs have been identified from the genomes of a range of species, including *Oryza sativa* [14], *Populus trichocarpa* [15], *Vitis vinifera* [16], *Lotus japonicas* [17], *Malus × domestica* Borkh [18], *Saccharum officinarum* [19] and *Capsicum annuum* [20]. In *S. officinarum*, transcripts of NINs were more abundant than acidic invertases in response to abiotic stresses [19]. Furthermore, the expression of CaNINV5 (*C. annuum* alkaline/neutral invertase 5) was increased...
during pepper fruit development [20], suggesting that NINs may also play important roles in physiological processes.

In this study, we conducted an in-depth in silico analysis of Chinese cabbage (Brassica rapa), a model crop for genomic study of the Brassica species, using public databases coupled with bioinformatics tools. To investigate the evolutionary relationships of B. rapa NINs (BraNINs), we evaluated their expansion and evolutionary mechanisms by their duplication. In addition, we analyzed the expression levels of BraNINs using the transcriptomes of Chinese cabbage leaves and roots under drought conditions. Taken together, our results further expand our knowledge of the physiological functions of NINs, and offer further insight into the evolutionary mechanisms of BraNINs.

Materials and methods

Database mining and identification of BraNIN genes

To retrieve BraNIN gene sequences, NIN protein sequences of Arabidopsis and apple were used as query sequences in the B. rapa genome, available in Phytozome v. 12.1 (https://phytozome.jgi.doe.gov; Brassica rapa FPsc v. 1.3). The Compute pl/MW tool (http://web.expasy.org/compute_pi/) was used to calculate the molecular weight (MW) and the theoretical isoelectric point (pl) of BraNIN proteins. The MitoProt web server (http://ihg.gsf.de/ihg/mitoprot.html), TargetP web server (http://www.cbs.dtu.dk/services/TargetP/) and SignalP web server (http://www.cbs.dtu.dk/services/SignalP/) were used to predict protein subcellular localizations. The exon/intron structures of BraNIN genes were drawn by comparing the coding sequences and the corresponding genomic sequences on the GSDS 2.0 (http://gsds.cbi.pku.edu.cn/).

Table 1. Neutral/alkaline invertase gene family in Brassica rapa.

| Name | Gene ID | Location | CDS (bp) | AA | Intron Nr. | pl  | kDa | Subcellular localization | SignalP | MitoProt<sup>a</sup>/ChloroP<sup>b</sup> |
|------|---------|----------|----------|----|------------|-----|-----|--------------------------|---------|-------------------------------|
| BraNIN1 | Brara.K01571.1 | Scaffold74288:56360..58901 | 1899 | 632 | 5 | 6.09 | 71.7 | Non-secretory protein | 97.56/65.0 |
| BraNIN2 | Brara.A03694.1 | A01:29634799..29637860 | 1890 | 629 | 5 | 6.15 | 71.7 | Non-secretory protein | 97.74/72.2 |
| BraNIN3 | Brara.E03276.1 | A05:26498701..26502387 | 1959 | 623 | 3 | 6.16 | 71.7 | Non-secretory protein | 97.94/63.0 |
| BraNIN4 | Brara.J01464.1 | A10:12854262..12857327 | 1833 | 610 | 5 | 5.68 | 68.8 | Non-secretory protein | 92.84/96.3 |
| BraNIN5 | Brara.E03276.1 | A05:26498701..26502387 | 1959 | 623 | 3 | 5.74 | 62.5 | Non-secretory protein | 98.45/77.0 |
| BraNIN6 | Brara.E03276.1 | A05:26498701..26502387 | 1959 | 623 | 3 | 5.74 | 62.5 | Non-secretory protein | 92.84/96.3 |
| BraNIN7 | Brara.G03094.1 | A07:24279933..24282714 | 1641 | 546 | 3 | 6.01 | 62.2 | Non-secretory protein | 42.18/27.9 |
| BraNIN8 | Brara.I03300.1 | A09:31373326..31376908 | 1623 | 540 | 3 | 6.16 | 61.6 | Non-secretory protein | 3.17/8.9 |
| BraNIN9 | Brara.C02604.1 | A03:13181651..13184810 | 1653 | 550 | 4 | 6.10 | 63.4 | Non-secretory protein | 3.17/15.2 |
| BraNIN10 | Brara.H00571.1 | A08:6930106..6932994 | 1647 | 548 | 4 | 6.76 | 62.6 | Non-secretory protein | 2.71/13.0 |
| BraNIN11 | Brara.E01857.1 | A05:15368406..15371050 | 1635 | 544 | 4 | 6.45 | 62.2 | Non-secretory protein | 3.69/27.5 |

<sup>a</sup>Probability (%) of targeting to mitochondrion.

Additional text:

Phylogenetic and gene duplication analysis of B. rapa NIN genes

The protein sequences of BraNIN genes were aligned by ClustalW and used for phylogenetic analysis using the Neighbor-Joining method in MEGA 7.0 software, with 100 replicates by default.

The Plant Genome Duplication Database (http://chibba.agtec.uga.edu/duplication/) was used to identify syntenic information among BraNIN genes. In addition, the synonymous rate (Ks), nonsynonymous rate (Ka) and evolutionary constraint (Ka/Ks) were calculated for duplicated BraNIN gene pairs using PAL2NAL (http://www.bork.embl.de/pal2nal/). The timing of the duplication event was calculated using the mean Ks values (T = Ks/2λ), assuming clock-like rates (λ) of synonymous substitution of 1.5 × 10<sup>−8</sup> substitutions/synonymous site/year [21].

Expression analysis of BraNINs

To analyze the expression patterns of BraNIN genes, we used RNA-seq data of drought treated-Chinese cabbage [22]. L1 (R1), L2 (R2) and L3 (R3) indicated leaf (root) samples obtained from Stage 1 (non-stressed condition), Stage 2 (soil water content of 20%) and Stage 3 (soil water content of 5%) plants, respectively. Transcript levels were calculated using SAMtools (http://samtools.sourceforge.net/) and relative transcript abundances were analyzed using DEseq as described by Eom et al. [22].

Results and discussion

Identification and characterization of neutral/alkaline invertases in Chinese cabbage

To identify the NIN genes, we analyzed the B. rapa genome database (Brassica rapa FPsc v1.3) using previously identified NIN sequences in Arabidopsis and...
apple. Subsequently, the redundant sequences were removed, resulting in a total of 11 putative NIN genes (Table 1). The full-length coding sequences of BraNINs ranged from 1623 bp (BraNIN8) to 1959 bp (BraNIN3) encoding 540 to 652 amino acids. These putative BraNINs have a calculated molecular weight ranging from 61.6 to 73.7 kDa and a theoretical pI ranging from 5.68 to 8.68 (Table 1). In addition, a phylogenetic tree with 11 BraNIN protein sequences was constructed using the Neighbor-Joining method, to investigate the phylogenetic relationships among NIN family members in Chinese cabbage. As shown in Figure 1A, the phylogenetic tree divided the BraNINs into two major groups that differed consistently at eight amino acid residues in the conserved motifs (C273V, C277S, Y287H, Y289H, V388L, S389Q, R460P and V471T based on the amino acid numbering of BraNIN5, Figure 1C). The α group contains BraNIN1-4, encoded by six exons with conserved location, whereas the β group BraNINs (BraNIN5-11) had a different number of exons (Figure 1B), indicating that the α and β groups arose from distinct ancestral genes [18].

Since it has been assumed that plant NINs accumulate in the cytoplasm, the subcellular localization analysis of rice and Poncirus trifoliate NINs demonstrated that plant NINs located in organelles including mitochondria, plastids and chloroplasts [23,24]. Based on subcellular location analyses, plant NINs belonging to the α group were predicted to have mitochondrial or chloroplast localizations, whereas the other NINs were predicted to be cytosolic proteins [14,23,25]. Similarly, the prediction of subcellular localization using computational analysis indicated that α group BraNINs were located in mitochondria or chloroplasts (Table 1). In Arabidopsis, sucrose hydrolysis by chloroplastic A/N-Inv (Arabidopsis NIN) is required for controlling chloroplast-cytosolic carbon partitioning [26], suggesting that BraNINs might also be involved in the controlling carbon balance between cytoplasm and chloroplasts.

Intron phase analysis of BraNIN genes revealed that their first exons flanked by intron phase 0 (Figure 1B) defined that introns positioned between two codons [27], similar to other NIN genes in higher plants including sugar cane [19] and apple [18]. In the α group, these genes were observed to present conserved intron-exon structures, whereas BraNIN genes of the β group exhibited different intron-exon structures and different number of exons. This suggests that α and β groups of BraNIN genes derived from different ancestral genes, and the conserved intron phases in genes of α group suggest stability during evolution.
Evolutionary patterns of the BraNIN family

Gene duplication resulting from unequal crossing over, retrotransposition or chromosomal duplication has been suggested as the main reason for the generation of new genes and gene family expansion [28]. In addition, it has been suggested that tandem and segmental duplications are the major sources of diversity for the evolution of large gene families in plants [29]. Chromosomal location and phylogenetic analyses of BraNIN genes and proteins (Table 1 and Figure 1A) indicated that the expansion of BraNINs is not due to tandem duplication. In poplar (P. trichocarpa), segmental duplication has been identified to play a leading role in the expansion of the NIN family [30]. Similarly, duplication analysis regarding the identification of chromosomal homologous segments within the genome showed two pairs of segmental duplicated paralogs (BraNIN5/6 and BraNIN7/8) (Figure 2), indicating that segmental duplication contributed to NIN family expansion in Chinese cabbage. Furthermore, the Ka/Ks ratio of the two identified BraNIN paralogous gene pairs was calculated to trace the types and divergence times of duplications. Ks in segmental duplication in BraNIN genes ranged from 0.4744 to 1.3575, dating the duplication event to have occurred between 15.81 million years ago (Mya) to 45.25 Mya (Table 2). Segmental duplications in Populus NIN genes occurred between 3.29 Mya (Ks = 0.0598) and 13.49 Mya (Ks = 0.2456) [30], suggesting that the segmental expansion in the Populus NIN family underwent a duplication event more recently than did the BraNIN family. In the BraNIN family, the Ka/Ks ratio of the two segmental duplication pairs was lower than 1 (Table 2), suggesting that these segmental duplication pairs were subjected to negative selection [31].

Expression analysis of BraNINs in response to drought stress

Drought or water deficit is the dominant factor affecting the growth and development of crops. Agricultural drought has become a major problem in global agricultural production. The activity of NIN in higher plants is affected by abiotic stresses including wounding, drought, salinity and low temperature [32]. In addition, PtrA/NINV (Poncirus trifoliata alkaline/neutral invertase)-overexpression improves tolerance to drought stress [24], suggesting that NINs play an important role when plants are subjected to abiotic stresses. To gain insight into the potential functions of BraNIN genes during responses to drought stress, we analyzed RNA-seq data of drought treated-Chinese cabbage. As shown in Figure 3, the transcript level of most BraNIN genes was down-regulated, whereas an
increased level of BraNIN1, BraNIN4 and BraNIN5 transcripts was observed in leaves and roots, when plants were treated with drought stress. Furthermore, BraNIN3 was down-regulated in the leaves but up-regulated in the roots, and BraNIN6 was up-regulated in the leaves but down-regulated in the roots under drought-stress conditions. This suggested that BraNIN1, BraNIN4 and BraNIN5 might be important BraNIN family members in terms of drought responses in Chinese cabbage, since they were up-regulated by drought stress, suggesting a divergence in the function of BraNINS in response to drought stress.

Conclusions

In this study, based on genome-wide analysis, we conducted a detailed analysis on the NIN gene family in B. rapa, and developed new insights into how these genes have evolved in B. rapa. These data will support a solid foundation for further understanding of the underlying evolutionary mechanisms in NIN genes in higher plants. An in-depth analysis of BraNIN transcription pattern under drought stress provides an important starting point for future efforts to understand the physiological function of BraNINS.

Disclosure statement

The authors declare no conflict of interest.

Funding

This work was carried out with the support of "Cooperative Research Program for Agriculture Science and Technology Development (Project No. PJ01261303)" Rural Development Administration, Republic of Korea.

References

[1] Koch K. Sucrose metabolism: regulatory mechanisms and pivotal roles in sugar sensing and plant development. Curr Opin Plant Biol. 2004;7:235–246.
[2] Tauzin AS, Giardina T. Sucrose and invertases, a part of the plant defense response to the biotic stresses. Front Plant Sci. 2014;5:293.
[3] Wan H, Wu L, Yang Y, et al. Evolution of sucrose metabolism: the dichotomy of invertases and beyond. Trends Plant Sci. 2018;23:163–177.
[4] Yang L, Xu M, Yeonjong K, et al. Sugar promotes vegetative phase change in Arabidopsis thaliana by repressing the expression of MIR156A and MIR156C. Elife. 2013;2:e00260.
[5] Huang H, Xie S, Xiao Q, et al. Sucrose and ABA regulate starch biosynthesis in maize through a novel transcription factor, ZmEREB156. Sci Rep. 2016;6:27590.
[6] Huang H, Long J, Zheng L, et al. Identification and characterization of microRNAs in maize endosperm response to exogenous sucrose using small RNA sequencing. Genomics. 2016;108:216–223.
[7] Roitsch T, Gonzalez MC. Function and regulation of plant invertases: sweet sensations. Trends Plant Sci. 2004;9:606–613.
[8] Vargas WA, Salerno GL. The Cinderella story of sucrose hydrolysis: alkaline/neutral invertases, from cyanobacteria to unforeseen roles in plant cytosol and organelles. Plant Sci. 2010;178:1–8.
[9] Ruan YL. Sucrose metabolism: gateway to diverse carbon use and sugar signaling. Annu Rev Plant Biol. 2014;65:33–67.
[10] Proels R, Roitsch T. Extracellular invertase Lin6 of tomato: a pivotal enzyme for integration of metabolic, hormonal, and stress signals is regulated by a diurnal rhythm. J Exp Bot. 2009;60:1555–1567.
[11] Hyun TK, Hoffmann A, Sinha AK, et al. Tomato mitogen activated protein kinases regulate the expression of extracellular invertase Lin6 in response to stress related stimuli. Functional Plant Biol. 2009;36:1088–1097.
[12] Rausch T, Greiner S. Plant protein inhibitors of invertases. Biochim Biophys Acta. 2004;1696:253–261.
[13] Tang X, Su T, Han M, et al. Suppression of extracellular invertase inhibitor gene expression improves seed weight in soybean (Glycine max). J Exp Bot. 2017;68:469–482.
[14] Ji X, van den Ende W, van Laere A, et al. Structure, evolution, and expression of the two invertase gene families of rice. J Mol Evol. 2005;60:615–634.
[15] Bocock PN, Morse AM, Dervinis C, et al. Evolution and diversity of invertase genes in Populus trichocarpa. Planta. 2008;227:565–576.
[16] Nonis A, Ruperti B, Pierasco A, et al. Neutral invertases in grapevine and comparative analysis with Arabidopsis, poplar and rice. Planta. 2008;229:129–142.
[17] Welham T, Pike J, Horst I, et al. A cytosolic invertase is required for normal growth and cell development in the model legume, Lotus japonicus. J Exp Bot. 2009;60:3353–3365.
[18] Hyun TK, Eom SH, Kim J-S. Genomic analysis and gene structure of the two invertase families in the domesticated apple (Malus × domestica Borkh.). Plant Omics. 2011;4:391–399.

[19] Wang L, Zheng Y, Ding S, et al. Molecular cloning, structure, phylogeny and expression analysis of the invertase gene family in sugarcane. BMC Plant Biol. 2017;17:109.

[20] Shen LB, Yao Y, He H, et al. Genome-wide identification, expression, and functional analysis of the alkaline/neutral invertase gene family in pepper. IJMS. 2018;19:224.

[21] Koch MA, Haubold B, Mitchell-Olds T. Comparative evolutionary analysis of chalcone synthase and alcohol dehydrogenase loci in Arabidopsis, Arabis, and related genera (Brassicaceae). Mol Biol Evol. 2000;17:1483–1498.

[22] Eom SH, Baek SA, Kim JK, et al. Transcriptome analysis in Chinese cabbage (Brassica rapa ssp. pekinensis) provides the role of glucosinolate metabolism in response to drought stress. Molecules. 2018;23:E1186.

[23] Murayama S, Handa H. Genes for alkaline/neutral invertase are located in plant mitochondria and also in plastids. Planta. 2007;225:1193–1203.

[24] Dahro B, Wang F, Peng T, et al. PtrA/NINV, an alkaline/neutral invertase gene of Poncirus trifoliata, confers enhanced tolerance to multiple abiotic stresses by modulating ROS levels and maintaining photosynthetic efficiency. BMC Plant Biol. 2016;16:76.

[25] Xiang L, Le Roy K, Bolouri-Moghaddam MR, et al. Exploring the neutral invertase-oxidative stress defence connection in Arabidopsis thaliana. J Exp Bot. 2011;62:3849–3862.

[26] Vargas WA, Pontis HG, Salerno GL. New insights on sucrose metabolism: evidence for an active A/N-Inv in chloroplasts uncovers a novel component of the intracellular carbon trafficking. Planta. 2008;227:795–807.

[27] Barvkar VT, Pardeshi VC, Kale SM, et al. Phylogenomic analysis of UDP glycosyltransferase 1 multigene family in Linum usitatissimum identified genes with varied expression patterns. BMC Genomics. 2012;13:175.

[28] Magadum S, Banerjee U, Murugan P, et al. Gene duplication as a major force in evolution. J Genet. 2013;92:155–161.

[29] Cannon SB, Mitra A, Baumgarten A, et al. The roles of segmental and tandem gene duplication in the evolution of large gene families in Arabidopsis thaliana. BMC Plant Biol. 2004;4:10.

[30] Chen Z, Gao K, Su X, et al. Genome-wide identification of the invertase gene family in populus. PLoS One. 2015;10:e0138540.

[31] Wagner A. Selection and gene duplication: a view from the genome. Genome Biol. 2002;3:1012.1–1012.3.

[32] Liu J, Han L, Huai B, et al. Down-regulation of a wheat alkaline/neutral invertase correlates with reduced host susceptibility to wheat stripe rust caused by Puccinia striiformis. J Exp Bot. 2015;66:7325–7338.