Scientia Horticulturae 216 (2017) 284–292

Contents lists available at ScienceDirect

Scientia Horticulturae

journal homepage: www.elsevier.com/locate/scihorti

Identification and characterization of expressed TIR- and non-TIR-NBS-LRR resistance gene analogous sequences from radish (*Raphanus sativus* L.) *de novo* transcriptome

Yan Wang[a,1], Zhaohui Song[a,1], Wei Zhang[a,1], Liang Xu[a], Xiaojun Su[b], Everlyne M'mbone Muleke[a], Liwang Liu[a,*]

[a] National Key Laboratory of Crop Genetics and Germplasm Enhancement, College of Horticulture, Nanjing Agricultural University, Nanjing 210095, PR China

[b] Institute of Vegetable Crops, Jiangsu Academy of Agricultural Sciences, Nanjing 210014, PR China

ARTICLE INFO

Article history:
Received 14 July 2016
Received in revised form
23 November 2016
Accepted 19 December 2016
Available online 19 January 2017

Keywords:
Radish (*Raphanus sativus* L.)
Resistance gene analogues (RGAs)
Nucleotide-binding site leucine-rich repeat (NBS-LRR)
Transcriptome sequencing

ABSTRACT

Radish is an important root vegetable crop worldwide. However, it is susceptible to a variety of diseases leading to decline in yield and quality. The plant nucleotide-binding site (NBS)-leucine-rich repeat (LRR) gene superfamily is effective against a big range of pathogens. Therefore, the isolation and identification of such *R* gene analogue (RGA) is a critical foundation for improving disease resistance in plants. So far, however, no comprehensive isolation of NBS-LRR RGAs has been reported in radish. In this study, a total of 36,613 unigenes representing 81,481 unique transcripts with an average length of 1253.22 nt were obtained from radish leaf transcriptome. In total, 38 NBS-encoding sequences were successfully identified via bioinformatics analysis. Multiple sequence alignment analysis classified these sequences into toll and interleukin-1 receptor (TIR)-NBS-LRR and non-TIR-NBS-LRR subclasses based on the presence of a TIR domain at the N-terminus. Phylogenetic tree analysis of the deduced amino acid sequences grouped these sequences into six classes. Non-synonymous to synonymous substitution ratio suggested that NBS-encoding sequences of RGAs in radish were subject to purification or negative selection. Additionally, RT-qPCR analysis revealed that several *RsRGAs* expressed differentially under three tested abiotic stimuli (SA, MeJA and ABA), suggesting that they might be involved in defense responses by activating signaling pathways. These results could provide valuable information for further identification of plant RGAs and accelerate genetic improvement of disease resistances in radish breeding programs.

1. Introduction

In plants, a myriad of potential pathogens including viruses, bacteria, fungi and nematodes, can cause extensive yield loss and quality decline in agronomic and horticultural crops. To resist pathogens, plants have activated a series of complex responses to recognize and defend pathogen infections and insect attacks, such as species-level resistance, constitutive barriers and race-specific resistance (Glazebrook, 2005; Zeng et al., 2006). The plant disease resistance genes (*R* genes) representing approximately 1% of all the genes in plant genome, could play critical roles in pathogen resistance mechanisms (Saha et al., 2013). Over the past decade, over 70 *R* genes have been isolated, and some of them have been well characterized from a wide range of plant species (Hammond-Kosack and Jones, 1997; Meyers et al., 2005; Saha et al., 2013). Within diverse *R* genes, characterization of the isolated *R* genes has uncovered their alluring attributes, and most of their proteins share a high degree of homology (Liu et al., 2007). The majority of *R* genes share striking structurally conserved domains such as nucleotide binding site (NBS), carboxyl leucine-rich repeat (LRR), serine/threonine protein kinases (PKs), leucine zipper (LZ), and trans membrane (TM) domains (Wan et al., 2013). In particular, the majority of plant disease resistance genes encoding NBS-LRR domain, and over 70 NBS-LRR coding genes have been isolated from various plant species (Moffett, 2009).

NBS-LRR class of R-genes can further be categorized into toll and interleukin-1 receptor (TIR)- and non-TIR subfamilies based

Abbreviations: ABA, abscisic acid; BLAST, basic local alignment search tool; EST, expressed sequence tag; LRR, leucine-rich repeat; MeJA, methyl jasmonate; NBS, nucleotide-binding site; RGA, resistance gene analogue; RT-qPCR, reverse transcription quantitative polymerase chain reaction; SA, salicylic acid; TIR, toll and interleukin-1 receptor.

* Corresponding author.
E-mail address: nauliulw@njau.edu.cn (L. Liu).

[1] These authors contributed equally to this work.

http://dx.doi.org/10.1016/j.scienta.2016.12.019
0304-4238/

on the presence/absence of an N-terminal domain that is parallel to the cytoplasmic domains of a TIR. In the non-TIR subfamily, a coiled-coil (CC) or leucine zipper (LZ) structure is usually found at the corresponding N-terminal domain, which corresponds to helical structures presumed to perform an integral function during its interaction with molecules involved in the signal transduction pathway (Cannon et al., 2002; Chen et al., 2007). While non-TIR sequences are present in both dicots and monocots, TIR-type genes appear only in dicots' genomes (Meyers et al., 1999). Almost all NBS-LRR gene family is involved in resistance to biotrophic pathogens (Glazebrook, 2005).

To date, the NBS region has been identified to contain eight conserved motifs, such as Kinase P-loop, hydrophobic GLPL, Kinase-2 and Kinase-3a motifs, which are widely present in both TNL and non-TNL classes (Sood et al., 2014). While RNBS-A-TIR (LQKKLLSKLL) and RNBS-D-TIR (FLHIACFF) are characteristics of the TNL subfamily; RNBS-A-non-TIR (FDLxAWCCVSQxF) and RNBS-D-non-TIR (CFLYCELPED) are specific to the non-TNL subfamily (Pan et al., 2000; Wan et al., 2010). In addition, the last single amino acid residue of the kinase-2 motif can also be used to distinguish the two classes with 95% accuracy, which contains tryptophan (W) and aspartic acid (D) in the non-TIR and TIR-type proteins, respectively (Noir et al., 2001; Xu et al., 2005; Yuksel et al., 2005).

Currently, the resistance genes of NBS-LRR class identified in some plant species including tomato (Martin et al., 1993; Jones et al., 1994), *Arabidopsis thaliana* (Bent et al., 1994; Mindrinos et al., 1994), tobacco (Whitham et al., 1994) and rice (Zheng et al., 2001). Most of NBS-containing sequences were isolated using map-based cloning or transposon-tagging approaches. Since they share structural similarity with the NBS-LRR domain, some NBS-containing sequences have been isolated using polymerase chain reaction (PCR) amplification from plant genomes with degenerate primers (McHale et al., 2006; Sharma et al., 2009). As more and more sequences were deposited into public databases including GeneBank at the NCBI, it is possible to employ another convenient way of data-mining to find new *R* gene analogue (RGA) sequences (Wan et al., 2013). In a similar way, transcriptome databases contain many gene expression sequences, and can also be used to isolate NBS-containing sequences (Ren et al., 2014).

During the past decade, a large number of transcriptomic sequences have been obtained via the next-generation sequencing (NGS) technology in cost-effective and reliable ways (Cubry et al., 2014; Yates et al., 2014). In addition, RGAs isolated from transcriptomic databases are desirable with no interference from pseudogenes as compared to those obtained from the genome, which makes transcriptomic RGAs highly expected to be functional genes (Ren et al., 2014). In *A. thaliana*, deep sequencing is a proper way to identify novel transcripts which were not present in previous EST collections (Blanca et al., 2011). This technology provides a reliable approach for identification of RGA with conserved motifs through data mining combined with bioinformatics methods.

Radish, *Raphanus sativus* L. (2n = 2x = 18), belonging to the Brassicaceae family, is an important edible root vegetable crop worldwide, especially in East Asia (Xu et al., 2012). In the growth process of radish, a wide spectrum of pathogens including downy mildew (DM) and turnip mosaic virus (TuMV) caused huge losses of production. Currently, one of the major objectives in radish breeding is to improve the ability of the resistances against various diseases (Wang et al., 2014). Owing to no *R* genes isolated and no NBS-containing sequences reported in radish, it is necessary to characterize the genetic basis of disease resistance in radish breeding. In this study, we focused on identifying potentially expressed NBS-LRR *RsRGAs* by searching the radish transcriptome database. The structure, function and evolution of the NBS-LRR region in radish were investigated, their characteristic and phylogenetic relationships with other known NBS sequences were elucidated, and Non-synonymous (Ka) and synonymous (Ks) ratios were calculated in the evolutionary processes of NBS-LRR *RsRGAs* in radish. Furthermore, RT-qPCR analysis was performed to characterize the expression of *RsRGAs* in response to treatment with signaling molecules. The outcomes of this study would provide vital foundation to speed up the genetic improvement of disease resistances in radish breeding programs.

Table 1
Overview of transcriptome sequencing and assembly.

Items	Number
Total genes	36,613
Total isogenes	81,481
Total residuces (bp)	102,113,321
Largest isogene length (bp)	16,434
Smallest isogene length (bp)	306
Average length (bp)	1253.22

Genes and isogenes represent the number of genes and contigs, respectively.

2. Materials and methods

2.1. Plant materials

'NAU-ZQH', an advanced radish inbred line with red skin and white fleshy taproot was used in this study. The germinated seeds were grown in plastic pots and cultivated in a growth chamber at 25 °C day/18 °C night with a 14 h light/10 h dark photoperiod. Seedlings at the fourth true-leaf stage were subjected to signaling molecules of ABA (100 μM), SA (100 μM) and MeJA (100 μM), respectively. Leaves were harvested at 0, 6, 12, 24, 48 and 72 h for each treatment and were immediately frozen in liquid nitrogen, and then stored at −80 °C for further use. Equal amounts of leaf samples from three randomly selected individuals with similar size under normal condition were pooled for the cDNA library construction.

2.2. Prediction of NBS-LRR encoding domains from radish transcriptome

The transcriptomic library construction and sequencing with an Illumina HiSeq™ 2500 were performed at the Beijing Genomics Institute (BGI, Shenzhen, China). The Pfam NBS (NB-ARC) family PF00931 domain was adopted as a query, and a local tBLASTn algorithm was used to obtain potential NBS-containing genes in the transcriptome database by BioEdit program, with an E-value cut off of 10^{-4} (Ren et al., 2014). The predicted NBS sequences were used to confirm the identity of NBS domain containing sequences by running BLASTp against the NCBI protein non-redundant database (Zhu et al., 2013). Gene identification and annotation were performed using FGENESH program (http://www.softberry.com). The domain architecture of the predicted NBS-containing proteins was verified with the Pfam ver 30.0 program (http://pfam.janelia.org).

2.3. Sequence alignment and phylogenetic analysis

To remove the partial and redundant sequences, the selected sequences from the present survey were subjected to DNA Star program and only the region from the P-loop motif to the GLPL motif was used for the phylogenetic analysis. Functional complements were confirmed in the NCBI GenBank using BLAST approach. The DNAMAN software was used to assemble DNA sequences and translate to the predicted amino acid sequences. The ClustalW program with default parameters was used for sequence assembly and multiple sequence alignments of amino acid. A phylogenetic tree was constructed using MEGA 6.0 program by the neighbor-joining