Agriculture Communications 2 (2024) 100049

Contents lists available at ScienceDirect

Agriculture Communications

journal homepage: www.sciencedirect.com/journal/agriculture-communications

Genome-wide identification and prediction of disease resistance genes in *Hirschfeldia incana*

Tingting Wu [a], Hawlader Abdullah Al-Mamun [a,b], David Edwards [a,b,c], Jacqueline Batley [a,c,*], Aria Dolatabadian [a]

[a] School of Biological Sciences, The University of Western Australia, Perth 6009, Australia
[b] Centre for Applied Bioinformatics, The University of Western Australia, Perth 6009, Australia
[c] The UWA Institute of Agriculture, The University of Western Australia, Perth 6009, Australia

HIGHLIGHTS

- Genome-wide identification found 914 RGAs in *Hirschfeldia incana*, offering a crucial genetic foundation for disease resistance in *Brassica* crops.
- Phylogenetic analysis revealed evolutionary relationships among RGAs, aiding understanding of their functional implications in disease resistance.
- We found 75 CDRHs in *H. incana*, underscoring their potential as a rich source of genetic material for breeding *Brassica* crops.
- Cis-Acting element analysis unveiled disease defence regulation in *H. incana* for manipulating gene expression to enhance crop resilience.

ARTICLE INFO

Keywords:
Hirschfeldia incana
Resistance gene analogue
Cloned disease resistance gene homolog
Blackleg
Phylogenetic analysis

ABSTRACT

Brassica species, globally cultivated as economically important vegetable and oilseed crops, face challenges from pathogens impacting their growth and productivity. Among these, blackleg, caused by the fungal pathogen *Leptosphaeria maculans*, stands out as a significant concern. Genetic resistance, primarily mediated by resistance gene analogues (RGAs), is key to sustainable blackleg control. Utilising wild relatives of *Brassica* species presents a promising avenue for enhancing resistance to blackleg in cultivated crops. In this study, we employed the newly published *Hirschfeldia incana* reference genome to identify the genome-wide RGAs in *H. incana*. A total of 914 candidate RGAs were identified; the receptor-like protein kinases (RLK) family contained the highest number with 608 (66.53%), followed by the Transmembrane coiled-coil (TM-CC) family with 167 (18.27%), nucleotide-binding site-leucine-rich repeats (NLR) family with 98 (10.72%) and receptor-like proteins (RLPs) with 41 (4.48%). We conducted duplication analysis on the 914 candidate RGAs, which revealed gene duplication occurs frequently to expand the RGAs in *H. incana* and significantly contributes to plant defence responsiveness. The phylogenetic analysis provided insights into the diversification and functional implications of the identified groups. We used the sequences of the 49 cloned R genes to identify homologs across *H. incana*. A total of 75 cloned disease-resistance gene homologs (CDRHs) were found. Cis-acting elements (CREs) were analysed in promoter sequences of 914 RGAs in *H. incana*, which confirmed their potential function in disease defence. Overall, the results suggest that the wild species *H. incana* could be a potential *R* gene source for various disease resistances, including blackleg.

1. Introduction

The Brassicaceae, a diverse family of flowering plants, encompasses 372 genera and 4060 species of great agricultural and ecological significance [1]. Among the most notable members are various *Brassica* species, including broccoli (*Brassica oleracea* var. *italica*), cabbage (*B. oleracea* var. *capitata*), turnip (*Brassica rapa*), black mustard (*Brassica nigra*), and canola (*Brassica napus*). These plants are valued for their edible parts, versatility, and adaptability in different climates and soil conditions.

Mediterranean or short pod mustard, scientifically known as *Hirschfeldia incana* (L.) Lagrèze-Fossat is an annual plant species native to the

* Corresponding author. School of Biological Sciences, The University of Western Australia, Perth 6009, Australia.
E-mail address: jacqueline.batley@uwa.edu.au (J. Batley).

https://doi.org/10.1016/j.agrcom.2024.100049
Received 11 April 2024; Received in revised form 27 May 2024; Accepted 6 June 2024

Mediterranean basin and the Middle East. However, it has now become widely distributed in warm-temperate regions across the globe, including as far as southern Australia and New Zealand, where it has established itself as an invasive weed (Buchan weed) [2]. In a recent study, the genome sequence of *H. incana* was explored to improve photosynthetic light-use efficiency, making it an excellent model for understanding this trait's genetic and physiological basis [3]. Gene flow occurs between *H. incana* and *B. napus*, as well as other commercial brassicas [4]. Therefore, one significant botanical relevance of *H. incana* is its utilisation in interspecific crossing experiments with the *Brassica* genus [5]. For instance, researchers have employed *H. incana* for these crosses, with a primary focus on quantifying the agricultural and ecological risks associated with transgenic crops in European *B. napus* production regions [6–9].

Spontaneous hybridisation of canola with *Sinapis arvensis*, *Raphanus raphanistrum* and *H. incana* has been noted during the use of male sterile oilseed rape as the recipient plant for pollen [8]. In rare instances, hybridisation has also occurred with *R. raphanistrum* and *H. incana* as the female parent [6]. Darmency and Fleury [10] examined the mating system in *H. incana* and its potential hybridisation with oilseed rape. Their study found that, throughout three years of field experiments, the average rate of spontaneous hybridisation between *H. incana* and *B. napus* was 0.6 hybrids per plant. The researchers inferred that the incorporation of *B. napus* genes into *H. incana* is not likely to be a notable phenomenon. Successful crosses with other species have been accomplished through manual pollination, as demonstrated in instances such as *Erucastrum gallicum* (Wild.) OE Schulz acting as the pollen donor [11] or by employing in vitro rescue techniques [12]. These studies have provided valuable insights into the potential impacts of transgenic crops on the environment and agricultural practices.

Plant diseases, particularly fungal infections, pose a significant challenge to *Brassica* species, affecting their growth and productivity. Blackleg, caused by the fungal pathogen *Leptosphaeria maculans*, stands out as one of the most well-known fungal diseases affecting *Brassica* crops. Blackleg can lead to devastating losses in *Brassica* production, especially in oilseed rape (*B. napus*). The fungus attacks the plant's stem, causing cankers and ultimately leading to wilting and death. Managing fungal diseases requires a combination of cultural practices, fungicide applications, and, ideally, the development of resistant crop varieties through genetic breeding programs [13]. Notably, *H. incana* has been successfully hybridised with *B. napus* in a specific study to enhance resistance to *L. maculans* [14,15]. This intergeneric hybridisation demonstrates the possibility of genetic introgression from *H. incana* into economically important crop plants, thereby offering a potential avenue for enhancing disease resistance in cultivated *Brassica* crops.

When discussing disease resistance in *Brassica* species, particularly to fungal diseases like blackleg, it is essential to highlight the role of resistance gene analogues (RGAs). RGAs are crucial in plants' defence mechanisms and host resistance against pathogens [16]. They are typically grouped into three primary categories: nucleotide-binding site--leucine-rich repeats (NLRs), receptor-like protein kinases (RLKs), and receptor-like proteins (RLPs). Among these, the NLR family stands out as the most common category of RGAs, encompassing cytoplasmic receptors that play a role in recognising specific pathogens and contributing to effector-triggered plant immunity (ETI) [17,18]. RLKs and RLPs are linked to pattern-triggered immunity (PTI), which relies on pattern recognition receptors (PRRs) to kickstart the primary line of defence by identifying pathogen elicitors [19].

RGAs have been extensively researched in resistance breeding as a valuable tool to gain deeper insights into the plant's inherent molecular defence mechanisms [20]. Given the consistent structural characteristics of RGAs in plants, analysing genome data through bioinformatics provides a robust method for comprehensive RGA prediction [21]. Numerous studies have showcased the versatility of RGAs, highlighting their utility as valuable sources of functional markers not only for marking pest resistance loci in various crops [22] and facilitating plant molecular breeding [23] but also for conducting assessments related to genetic structure and diversity [24]. Several investigations have examined the variation and evolutionary aspects of RGAs within the Brassicaceae family [25–29]. Over 30 species comprising cultivated varieties and wild counterparts have been examined within the *Brassica* genus for their RGAs content. This extensive analysis has led to the discovery of more than 30,000 RGAs [28,30–32]. In a recent study, Amas et al. [33] employed the *B. rapa* pangenome to delineate an exhaustive collection of RGAs in *B. rapa*. Their findings indicated that 309 RGAs were influenced by presence-absence variation (PAV), and 223 RGAs were absent in the reference genome. Transmembrane leucine-rich repeat (TM-LRR) RGAs exhibited higher conservation, whereas nucleotide-binding site leucine-rich repeats (NLRs) displayed greater diversity. Comparative analysis with *B. napus* revealed substantial RGA conservation (93%). Additionally, 138 candidate RGAs associated with *B. rapa* disease resistance QTL were identified, many under negative selection.

RGAs are crucial to crop disease resistance, ultimately contributing to increased agricultural productivity and sustainability. Understanding the diversity and function of R genes and RGAs in *Brassica* species, including *H. incana*, is crucial for developing disease-resistant crop varieties through breeding and biotechnological approaches. Moreover, the role of *H. incana* as a potential genetic resource in crop improvement emphasises the need for further research and exploration in this field to harness its genetic potential for the benefit of agriculture worldwide.

This study aims to investigate and characterise RGAs within *H. incana*, a significant member of the Brassicaceae family known for its agricultural and genetic importance. RGAs play a crucial role in plant defence mechanisms and disease resistance. Through advanced bioinformatics techniques, this research seeks to uncover and understand the diversity and functional aspects of RGAs in *H. incana*. By enhancing our knowledge of these genes, the study aims to improve disease resistance in *Brassica* crops, facilitating breeding and biotechnological strategies that contribute to sustainable and productive agriculture.

2. Materials and methods

2.1. *H. incana* reference genomes

The *H. incana* genome used in this study is *H. incana* isolate: *HINC_-NIJ6* Genome sequence scaffold version [3]. The genome sequence file, annotation file, protein sequence and coding sequence (CDS) utilised in this study were obtained from the website (https://www.bioinformatics.nl/hirschfeldia/, accessed on 30 September 2023).

2.2. Genome-wide mining of resistance gene analogues

The RGAugury pipeline [34] was used to perform in silico prediction of RGAs and their subfamilies (CN, CNL, NL, RN, RNL, TN, TNL, TX, OTHER-NLR, LRR-RLK, LysM-RLK, OTHER-RLK, LRR-RLP, LysM-RLP, RPW8 and TM-CC) in the genomes of *H. incana*.

2.3. Gene duplication analysis of RGAs in H. incana

Gene duplication analysis was conducted in Geneious Prime®. All the RGA proteins were subjected to BLASTp analysis against each other. The criteria employed in this study adhered to the same standards outlined in the previous study [35], requiring a minimum of 70% identity and coverage in the alignment during BLASTp comparison and an e-value smaller than e-45.

2.4. Multiple alignment and phylogenetic analysis

Multiple sequences of NLR and LRR-TM proteins were aligned separately using MUSCLE Alignment, available in Geneious Prime® with 1000 iterations. UPGMA tree build method with 1000 bootstraps and Jukes-Cantor genetic distance model were used to establish the

phylogenetic tree with aligned protein sequence in Geneious Prime® for each family. FigTree.v1.4.4 software and online software iTOL [36] (http://itol.embl.de/, accessed on 13 October 2023) were used to annotate and glorify the Phylogenetic tree.

2.5. Mining the protein sequences of the cloned disease-resistance genes

The compilation of 49 cloned R genes discovered in *Brassica crop* species and *Arabidopsis thaliana*, providing resistance to fungal and bacterial diseases affecting *Brassicaceae* species, was sourced from a prior investigation [37]. The protein sequences of the 49 cloned R genes were obtained from the UniProtKB (https://www.uniprot.org/uniprot/, verified and accessed on 1 October 2023) or NCBI (https://www.ncbi.nlm.nih.gov/, verified and accessed on 1 October 2023) website.

2.6. Identification of cloned disease-resistance gene homologs in H. incana

The RGAs from the *H. incana* genome and the 49 cloned *R* genes were aligned using BLASTp in Geneious Prime®. The BLASTp results removed a candidate gene greater than e−45 in E-value from further analysis. Then, the similarity lower than 60% in the BLASTp results was removed as the homology search was conducted in a wild species. A more detailed categorisation of RGAs was conducted based on whether they possessed a resistance domain similar to their homologous cloned R gene or if it differed.

2.7. Analysis of RGAs promoter cis-acting elements

TBtools was employed to extract DNA sequences of 2000 base pairs upstream of the identified RGA genes. The PlantCARE database [38] was used to identify cis-acting elements and select those involved in stress and defence responsiveness (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/(accessed on 11 October 2023). The results were displayed using TBtools.

3. Results

3.1. Genome-wide identification of RGAs in H. incana

A total of 914 candidate RGAs were identified in this study, accounting for 2.83% of the predicted 32,313 coding genes in the *H. incana* reference genome. The sequence length of these RGA proteins ranged from 121 to 5068 amino acid residues. The 914 RGAs consisted of 3 CN, 4 CNL, 11 NL, 7 RN, 6 RNL, 9 TN, 28 TNL, 26 TX, 4 OTHER-NLR, 248 LRR-RLK, 6 LysM-RLK, 354 OTHER-RLK, 2 LysM-RLP, 39 LRR-RLP, 161 TM-CC and 6 RPW8 (Fig. 1). We grouped CN, CNL, NL, RN, RNL, TN, TNL, TX and OTHER-NLR as NLR family; LRR-RLK, LysM-RLK and OTHER-RLK as RLK family; LysM-RLP and LRR-RLP as RLP family; TM-CC and RPW8 as TM-CC family. Among the 914 candidate genes, the RLK family contained the highest number with 608 (66.53%), followed by the TM-CC family with 167 (18.27%), NLR family with 98 (10.72%) and RLP with 41 (4.48%).

3.2. Analysis of duplications of RGAs in H. incana

In this study, duplicated genes were recognised using BLASTp by comparing all the predicted RGA proteins against each other. Duplicated genes included in the same RGA sub-family were defined as intra-sub-family duplication. In contrast, the duplicated genes occurring in the different sub-families were recorded as inter-sub-family duplication. The duplication that occurs between different RGA families was considered as inter-family duplication.

A total of 445 duplicated RGAs in *H. incana* were found (48.69% of the RGAs in *H. incana*). Among them, 435 (97.75%) were intra-sub-family duplication genes, 8 (1.80%) were inter-sub-family duplication genes, and only 2 (0.45%) inter-family duplicated genes were identified in the *H. incana* genome. This suggests that intra-sub-family duplication plays a significant role in RGA duplication in *H. incana* (Table 1).

A total of 28 NLRs (28.57% of the identified NLRs), consisting of 1 NL, 3 RN, 3 RNL, 1 TN, 8 TNL and 12 TX, were found to be involved in

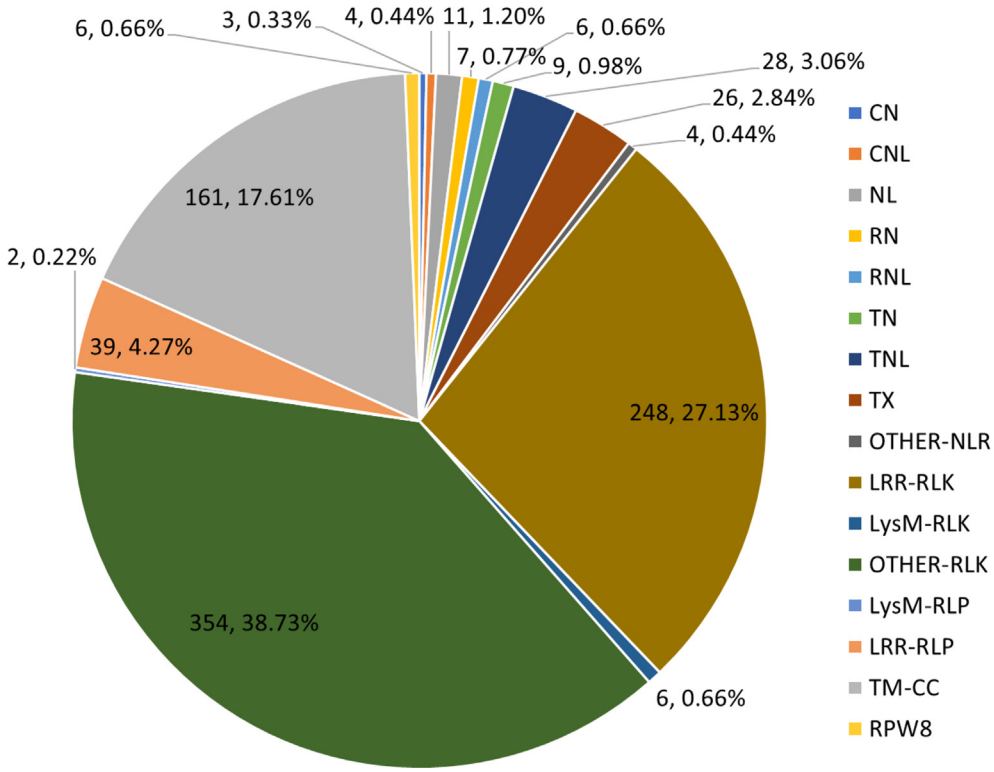

Fig. 1. The number of RGAs and sub-families detected in the *H. incana* reference genome.

Table 1
The duplication events are defined in the *H. incana* RGA.

Duplication Type	NLRs									RLKs			RLPs		TM-CC		Grand Total
	CN	CNL	NL	RN	RNL	TN	TNL	TX	OTHER-NLR	LRR-RLK	LysM-RLK	OTHER-RLK	LysM-RLP	LRR-RLP	RPW8	TM-CC	
Intra sub-family	0	0	0	3	2	0	7	11	0	129	2	206	1	11	2	61	435
Inter sub-family	0	0	1	0	1	1	1	1	0	0	0	3	0	0	0	0	8
Inter family	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	2
Total duplicated RGAs in each sub-family	0	0	1	3	3	1	8	12	0	129	2	209	1	13	2	61	445
Total RGAs in each sub-family	3	4	11	7	6	9	28	26	4	248	6	354	2	39	6	161	914
% duplicated RGAs in each sub-family	0	0	9.09	42.9	50	11.1	28.6	46.2	0	52	33.3	59	50	33.3	33.3	37.9	48.7

duplication events. Within the predicted RLKs, 55.92% were defined as duplications, with 340 genes including 129 LRR-RLKs, 2 LysM-RLKs and 209 OTHER-RLKs. A total of 14 RLPs (34.15% of the identified RLPs) had duplicated genes, consisting of 1 LysM-RLP and 13 LRR-RLPs. There were 37.72% of the identified TM-CC (61 TM-CC and 2 RPW8) genes associated with duplication events. Therefore, RLKs have the highest percentage of duplication, followed by TM-CC, RLPs and NLRs.

3.3. Phylogenetic analysis of RGAs in H. incana

The evolutionary relationship between RGAs in *H. incana* was investigated by establishing a rooted phylogenetic tree of NLRs and TM-LRRs [16], including the RLK, RLP, and TM-CC family. A total of 98 NLRs were divided into two major groups, which were further classified into six subgroups. The size of Group I and Group II significantly varied, with

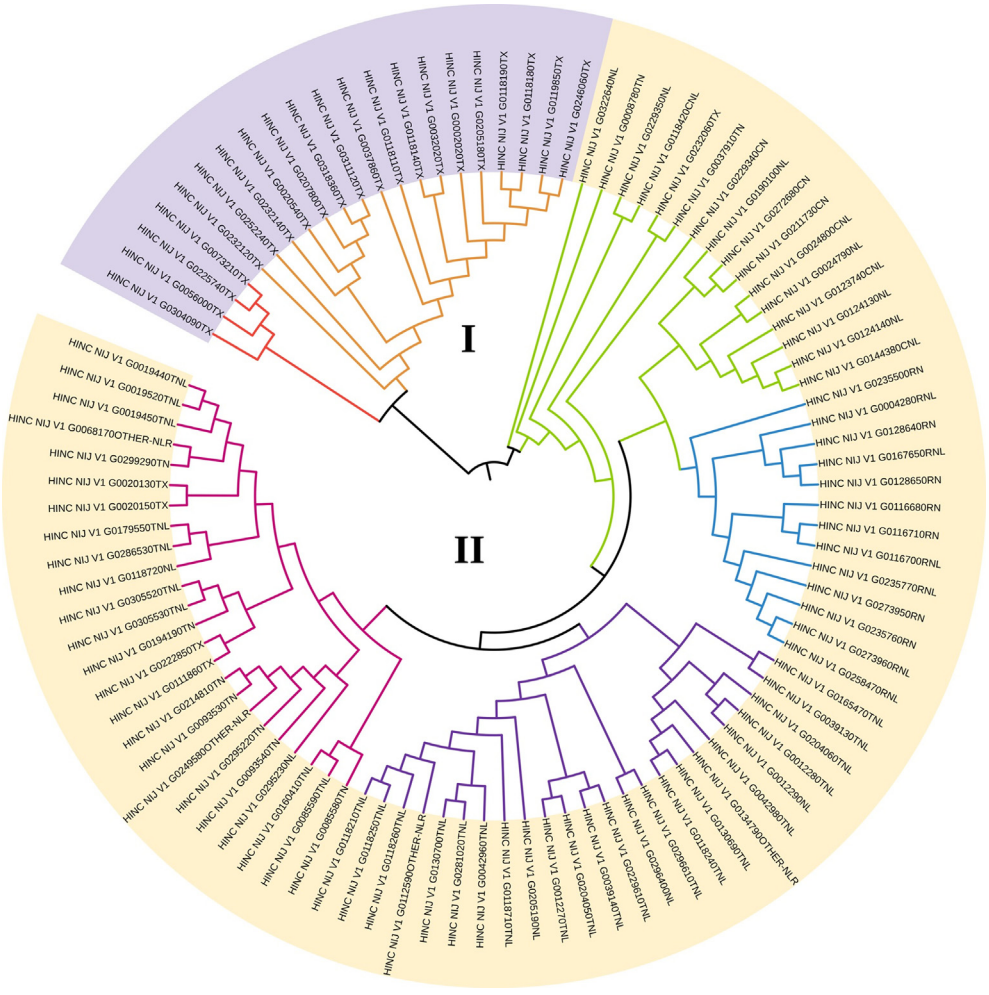

Fig. 2. Phylogenetic analysis of NLRs proteins of *H. incana*. The circular tree was generated based on the UPGMA method using Geneious Prime® 2022.2.2 software. Different colours divided these genes into two main groups and six sub-groups.

21 genes in Group I and 77 genes in Group II. The genes that made up Group I were all TX. The genes that made up Group II were NL, TN, CNL, CN, RN, RNL, TNL, OTHER-NLR, and TX (Fig. 2).

A total of 816 TM-LRRs were classified into three primary groups through phylogenetic analysis. The size of the groups was 182 in Group I, 200 in Group II and 434 in Group III. Within Group I, 180 out of the 182 genes were all from the RLK family, including 24 LRR-RLK genes, 150 OTHER-RLK genes and 6 closely linked LysM-RLK genes, suggesting the intimate evolutionary relationship and the function among LysM-RLK genes. The two exception genes are from the TM-CC family. In Group II, 198 out of 200 genes belonged to RLKs, consisting of 21 LRR-RLK genes and 177 OTHER-RLK genes. The other two genes came from the LRR-RLP subfamily. The largest Group III was further classified into IIIa, IIIb and IIIc sub-groups in the Phylogenetic tree. LRR-RLK genes were the main group members in both IIIa and IIIb. However, there were a significant number of LRR-RLP genes located in the IIIb sub-group. Noticeably, the IIIc group mainly consisted of TM-CC. The 2 LysM-RLPs were also located in the IIIc group. Therefore, RLKs were distributed in all three primary groups, while most of the RLP and TM-CC genes were restrained in Group III (Fig. 3).

3.4. Identification of CDRHs across H. incana

We used the sequences of the 49 cloned R genes to identify homologs across *H. incana* [37]. A total of 75 CDRHs, consisting of 1 CN, 1 CNL, 2 NL, 3 RN, 5 RNL, 9 TNL, 1 OTHER-NLR, 12 LRR-RLK, 11 OTHER-RLK, 3 LRR-RLP, 1 RPW8 (Table 2) and 26 Non-RGA (genes without RGA-related domain) (Table 3) were identified. Among the 49 RGA CDRHs, the RLK family had the highest number at 23, followed by NLRs at 22, RLPs at 3, and TM-CC at 1.

A total of 27 out of the 49 cloned *R* genes were detected to have homologous genes in *H. incana*. These genes are involved in resistance to downy mildew, powdery mildew, bacterial leaf spot, Sclerotinia stem rot, white rust, Fusarium wilt, blackleg and black rot disease (Tables 2 and 3).

Eleven out of the 27 cloned *R* genes had more than one homolog in *H. incana*. For instance, 8 CDRHs were observed for the cloned *R* gene *At_BAK1*, and 7 CDRHs were obtained for the cloned *R* gene *At_RFO1*. Interestingly, *Bna_MPK9* was found to have 14 CDRHs across *H. incana*, but none were predicted to be an RGA (Tables 2 and 3).

We found that 4 CDRHs in *H. incana; HINC_NIJ_V1_G0126980*, *HINC_NIJ_V1_G0258470*, *HINC_NIJ_V1_G0273950*, *HINC_NIJ_V1_G0273960* are homologous to more than one of the cloned *R* genes (Table 2). For example, *HINC_NIJ_V1_G0126980* is homologous to both *At_RFO1* and *Bna_Rlm9/4/7*. *HINC_NIJ_V1_G0258470*, *HINC_NIJ_V1_G0273950* and *HINC_NIJ_V1_G0273960* are homologous to both *At_NGR1a* and *At_NGR1b*.

3.5. Cis-acting element analysis of RGAs promoters in H. incana

Cis-acting elements (CREs) were analysed in promoter sequences of 914 RGAs in *H. incana* to analyse their potential regulation and function. Within the 2000 bp upstream of the transcription start site of each RGA, 36601 CREs were detected, an average of 40 CREs for each RGA. The results showed that 119 types of CREs were distributed in the 914 RGA genes. The analysis revealed the presence of eight distinctive elements.

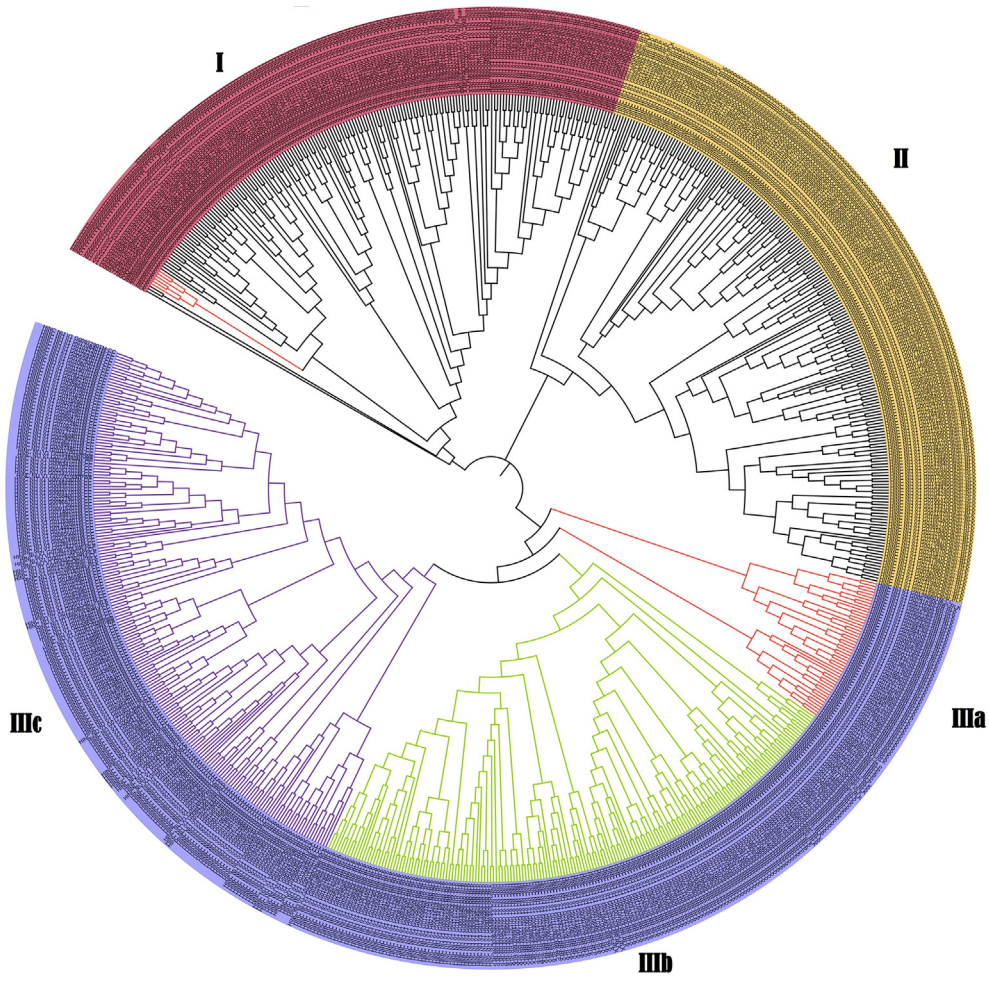

Fig. 3. Phylogenetic analysis of TM-LRRs proteins of *H. incana*. The circular tree was generated based on the FastTree method using Geneious Prime® software. These genes were divided into three main and sub-groups by different colours. In Group I, branches coloured with red were LysM-RLK genes.

Table 2
RGA CDRHs in *H. incana* and their similarity with cloned resistance genes.

Cloned Gene (RGA Type)	Pathogen	CDRHs in *H. incana* (RGA Type)	Identity (%)
At_ADR1 (RNL)	*Hyaloperonospora arabidopsidis, Erysiphe cichoracearum and Pseudomonas syringae*	*HINC_NIJ_V1_G0004280*(RNL)	60.40%
At_BAK1 (LRR-RLK)	*P. syringae and Sclerotinia sclerotiorum*	*HINC_NIJ_V1_G0033150*(LRR-RLK)	94.90%
		HINC_NIJ_V1_G0065850(LRR-RLK)	72.80%
		HINC_NIJ_V1_G0093190(LRR-RLK)	78.90%
		HINC_NIJ_V1_G0128710(LRR-RLK)	94.20%
		HINC_NIJ_V1_G0167590(LRR-RLK)	94.30%
		HINC_NIJ_V1_G0211710(LRR-RLK)	78.00%
		HINC_NIJ_V1_G0215530(LRR-RLK)	79.00%
		HINC_NIJ_V1_G0282420(LRR-RLK)	78.20%
At_NGR1a (RNL)	*Albugo candida, H. arabidopsidis, and P. syringae*	*HINC_NIJ_V1_G0235760*(RN)	62.70%
		HINC_NIJ_V1_G0258470(RNL)	76.50%
		HINC_NIJ_V1_G0273950(RN)	65.90%
		HINC_NIJ_V1_G0273960(RNL)	73.20%
At_NGR1b (RNL)	*A. candida, H. arabidopsidis, and P. syringae*	*HINC_NIJ_V1_G0258470*(RNL)	68.60%
		HINC_NIJ_V1_G0273950(RN)	61.30%
		HINC_NIJ_V1_G0273960(RNL)	64.10%
At_FLS2 (LRR-RLK)	*P. syringae*	*HINC_NIJ_V1_G0118660*(LRR-RLK)	78.80%
At_PBS1 (OTHER-RLK)	*P. syringae*	*HINC_NIJ_V1_G0077100*(OTHER-RLK)	65.50%
At_RFO1 (OTHER-RLK)	*Fusarium oxysporum matthioliF*	*HINC_NIJ_V1_G0080510*(OTHER-RLK)	61.60%
		HINC_NIJ_V1_G0111700(OTHER-RLK)	63.70%
		HINC_NIJ_V1_G0111710(OTHER-RLK)	63.00%
		HINC_NIJ_V1_G0111720(OTHER-RLK)	62.00%
		HINC_NIJ_V1_G0126980(OTHER-RLK)	60.10%
		HINC_NIJ_V1_G0296330(OTHER-RLK)	84.60%
		HINC_NIJ_V1_G0297180(OTHER-RLK)	85.00%
At_RFO2 (LRR-RLP)	*F. oxysporum matthioli*	*HINC_NIJ_V1_G0081250*(LRR-RLP)	67.70%
		HINC_NIJ_V1_G0215150(LRR-RLK)	67.30%
		HINC_NIJ_V1_G0294990(LRR-RLK)	65.10%
		HINC_NIJ_V1_G0162980(LRR-RLP)	69.60%
At_RFO3 (OTHER-RLK)	*F. oxysporum matthioli*	*HINC_NIJ_V1_G0026000*(OTHER-RLK)	74.60%
At_RLM1b (TNL)	*Leptosphaeria maculans*	*HINC_NIJ_V1_G0019440*(TNL)	65.90%
At_RLP1 (LRR-RLP)	*Xanthomonas* spp.	*HINC_NIJ_V1_G0140450*(LRR-RLP)	65.60%
At_RPM1 (NL)	*P. syringae*	*HINC_NIJ_V1_G0190100*(NL)	80.50%
At_RPP2a (TNL)	*H. arabidopsidis*	*HINC_NIJ_V1_G0039130*(TNL)	69.00%
		HINC_NIJ_V1_G0165470(TNL)	68.30%
		HINC_NIJ_V1_G0286530(TNL)	64.50%
At_RPP2b (TNL)	*H. arabidopsidis*	*HINC_NIJ_V1_G0229610*(TNL)	73.60%
At_RPP7 (NL)	*H. arabidopsidis*	*HINC_NIJ_V1_G0211730*(CN)	62.20%
At_RPS2 (NL)	*P. syringae*	*HINC_NIJ_V1_G0123740*(CNL)	84.50%
At_RPS4 (TNL)	*P. syringae*	*HINC_NIJ_V1_G0118210*(TNL)	70.30%
		HINC_NIJ_V1_G0118250(TNL)	75.00%
		HINC_NIJ_V1_G0118260(TNL)	62.60%
At_Rpw8.2 (RNL)	*E. cichoracearum*	*HINC_NIJ_V1_G0113230*(RPW8)	67.30%
At_SOBI (LRR-RLK)	*P. syringae and S. sclerotiorum*	*HINC_NIJ_V1_G0049770*(LRR-RLK)	80.40%
		HINC_NIJ_V1_G0062030(OTHER-RLK)	80.90%
At_WRR12 (TNL)	*A. candida*	*HINC_NIJ_V1_G0249580*(OTHER-NLR)	76.60%
At_WRR4b (TNL)	*A. candida*	*HINC_NIJ_V1_G0019520*(TNL)	76.00%
At_WRR8 (TNL)	*A. candida*	*HINC_NIJ_V1_G0118720*(NL)	63.30%
Bna_Rlm9/4/7 (OTHER-RLK)	*L. maculans*	*HINC_NIJ_V1_G0126980*(OTHER-RLK)	62.60%

The MYB cis-regulatory element, belonging to the largest plant transcription factor family, emerged as the predominant factor, occurring 224 times across the 49 RGA cis-regulatory regions. Following closely, the MYC element, an important transcription factor involved in plant defence mechanisms against diverse stresses, was identified 145 times. Additionally, the G-box, AS-1, ABRE, STRE, TC-rich repeats, and WUN-motif cis-elements exhibited varying frequencies, with G-box appearing 87 times, AS-1 77 times, ABRE 71 times, STRE 55 times, TC-rich repeats 18 times, and WUN-motif 15 times (Fig. 4). The Cis-regulatory elements in the promoter region of the RGA CDRHs in *H. incana* are illustrated in Fig. 5.

4. Discussion

4.1. Genome-wide identification of RGAs in H. incana

Genome-wide exploration of RGAs in *H. incana* has revealed a noteworthy finding the count of NLRs stands at 98, marking the lowest among previously studied Brassicaceae species. In comparison, *A. thaliana* exhibits a higher NLR gene count at 205 [39], while *B. napus* demonstrate the highest among the examined species with 621 genes [40]. The average NLR gene count for the Brassicaceae species studied falls within this spectrum. This distinctive profile points out *H. incana*'s unique position within the Brassicaceae family regarding NLR gene abundance. This phenomenon could be attributed to the larger genome found in allotetraploid Brassica species.

Furthermore, wild species exhibit a reduced count of RGAs compared to domesticated crops, a trend documented in previous research studies [41]. The long domestication history may lead to an increase in the number of RGA genes. However, for the RLK, RLP and TM-CC gene families *H. incana* showed a similar number compared with other diploid Brassicaceae species like *A. thaliana*, *B. rapa*, *B. oleracea* and *B. nigra* [36]. The similarity in the numbers across these gene families might be attributed to their involvement in diverse biological processes beyond resistance mechanisms. RLKs, RLPs and TM-CC genes often participate in various cellular and developmental processes, indicating a multifunctional role [20,42–44]. For example, in addition to perceiving biotic and abiotic stimuli, the LRR-RLK subfamily is involved in several biological

Table 3
Non-RGA CDRHs in *H. incana* and their similarity with cloned resistance genes.

Cloned Gene (RGA Type)	Pathogen	CDRHs in *H. incana*	% Identity
At_NDR1 (TM-CC)	*P. syringae*	*HINC_NIJ_V1_G0146270* (Non-RGA)	79.90%
		HINC_NIJ_V1_G0292990 (Non-RGA)	75.00%
At_PBS1 (OTHER-RLK)	*P. syringae*	*HINC_NIJ_V1_G0041080* (Non-RGA)	87.60%
		HINC_NIJ_V1_G0043380 (Non-RGA)	74.70%
		HINC_NIJ_V1_G0146320 (Non-RGA)	69.00%
		HINC_NIJ_V1_G0194640 (Non-RGA)	96.40%
		HINC_NIJ_V1_G0198000 (Non-RGA)	63.40%
		HINC_NIJ_V1_G0270210 (Non-RGA)	77.10%
		HINC_NIJ_V1_G0305960 (Non-RGA)	94.70%
At_RIN4 (TM-CC)	*P. syringae*	*HINC_NIJ_V1_G0144950* (Non-RGA)	77.50%
		HINC_NIJ_V1_G0272110 (Non-RGA)	75.50%
At_WRR9 (NL)	*A. candida*	*HINC_NIJ_V1_G0019430* (Non-RGA)	66.70%
Bna_MPK9 (OTHER-RLK)	*L. maculans*	*HINC_NIJ_V1_G0024910* (Non-RGA)	76.60%
		HINC_NIJ_V1_G0027200 (Non-RGA)	95.00%
		HINC_NIJ_V1_G0054360 (Non-RGA)	67.30%
		HINC_NIJ_V1_G0071980 (Non-RGA)	80.20%
		HINC_NIJ_V1_G0081750 (Non-RGA)	72.90%
		HINC_NIJ_V1_G0093790 (Non-RGA)	75.80%
		HINC_NIJ_V1_G0138010 (Non-RGA)	72.50%
		HINC_NIJ_V1_G0162590 (Non-RGA)	76.40%
		HINC_NIJ_V1_G0182980 (Non-RGA)	76.40%
		HINC_NIJ_V1_G0198250 (Non-RGA)	82.90%
		HINC_NIJ_V1_G0206700 (Non-RGA)	71.10%
		HINC_NIJ_V1_G0208120 (Non-RGA)	74.00%
		HINC_NIJ_V1_G0221360 (Non-RGA)	71.30%
		HINC_NIJ_V1_G0249500 (Non-RGA)	74.60%

processes related to development. The model plant *A. thaliana* has been the subject of most research, with additional studies conducted on a variety of other plants [45]. The total number of LysM-RLK (2) and LysM-RLP (1) was significantly lower than LRR-RLK (129), LRR-RLP (13) and OTHER-RLKs (209), which is consistent with other species from the Brassicaceae family.

4.2. RGA duplication in H. incana

Gene duplication is an acritical driving force of evolutionary innovation and gene family expansion in plants, allowing them to be resilient against selective pressure in the natural environments [46]. In this study, about 48.65% of the identified RGAs have evolved through a duplication event, suggesting that gene duplication occurs frequently to expand the RGAs in *H. incana* and significantly contributes to plant defence responsiveness. Among them, 97.75% of duplications are considered intra-sub family duplication, which plays a major role in enlarging the RGAs. It highlights the importance of subfamily-level duplication events in shaping the diversity and abundance of genes involved in plant defence mechanisms. Defence-associated LRR-RLKs have reportedly experienced several duplication events, the majority of which are enormously lineage-specific expansions caused mostly by tandem duplication [47].

Interestingly, inter-sub family duplication events contribute to a lesser extent, accounting for only 0.45% of RGA duplications in *H. incana*. This suggests that such events have a minimal impact on the expansion of defence-related genes in this plant species. The dominance of intra-sub family duplications may indicate a prioritised mechanism for maintaining and enhancing the plant's defence capabilities. Distinct patterns of NLR gene duplications and clustering are observed in species such as *B. rapa*, *B. oleracea* [31], *B. napus* [32], and *Brassica juncea* [48]. These patterns have been linked to varying pathogen selection pressures across these species.

4.3. Phylogenetic analysis

The phylogenetic analysis revealed interesting patterns within the gene families, providing insights into the diversification and functional implications of the identified groups. The grouping of 98 NLRs into groups Group I and Group II implies the complexity of the evolutionary landscape. The presence of clustered RGA genes is common across various plant genomes [49,50]. In *A. thaliana*, for instance, over 71% of NBS-LRR genes are organised into 38 clusters [51]. This clustering pattern is similarly observed in the rice genome, where NBS-LRR genes also form clusters [52]. The size difference between Group I (21 genes) and Group II (77 genes) suggests potential differences in functional roles or selective pressures acting on these two groups.

Moreover, the composition of genes within each group, particularly the presence of specific motifs such as TX, NL, TN, CNL, CN, RN, RNL, TNL, and OTHER-NLR, provides a detailed understanding of the diversity in the *H. incana* genome. The analysis of 816 TM-LRRs further expands the evolutionary perspective by classifying them into three groups. Group I, dominated by RLKs, exhibits a specific clustering of LRR-RLK, OTHER-RLK and LysM-RLK genes, indicating a shared evolutionary history and potential functional similarities among these genes. Group II and the larger Group III also display distinct compositions, with RLKs predominantly present in both groups and a noteworthy presence of LRR-RLP genes in Group II and TM-CC in Group III. The subdivision of Group III into IIIa, IIIb, and IIIc subgroups reveals differential patterns among LRR-RLK, LRR-RLP, and TM-CC within this expansive group. The observed distribution of RLKs across all three primary groups suggests a diverse array of functions within the *H. incana* genome. On the other hand, the more restricted distribution of RLPs and TM-CC in Group III suggests a potential specialisation of these genes in specific biological processes or responses. Overall, this phylogenetic analysis provides a comprehensive view of the evolutionary relationships among RGAs in *H. incana*, offering a foundation for future investigations into the functional significance of the identified gene groups.

4.4. Identification of CDRHs across H. incana

These findings highlight the extensive diversity of *R* gene homologs in *H. incana*, signifying the plant's adaptive capacity against various pathogens. Enriching RLK and NLR families among the identified CDRHs suggests their crucial role in mediating resistance responses. The presence of multiple homologs for specific *R* genes, such as *At_RFO1* and *Bna_MPK9*, suggests a complex interplay of redundancy and specificity in the defence mechanisms of *H. incana*. The similar situation has been reported in other *Brassica* species. In a 2022 study, researchers employed the same 49 cloned *R* genes to identify homologs across *Brassica species and A. thaliana.* They found 136 CDRHs in *B. juncea*, followed by 119 in *Brassica carinata,* 101 in *B. napus,* 80 in *B. rapa*, 78 in both *B. nigra* and *B. oleracea* and 76 in *A. thaliana*. This suggests that each cloned *R* gene has an average of 1.9

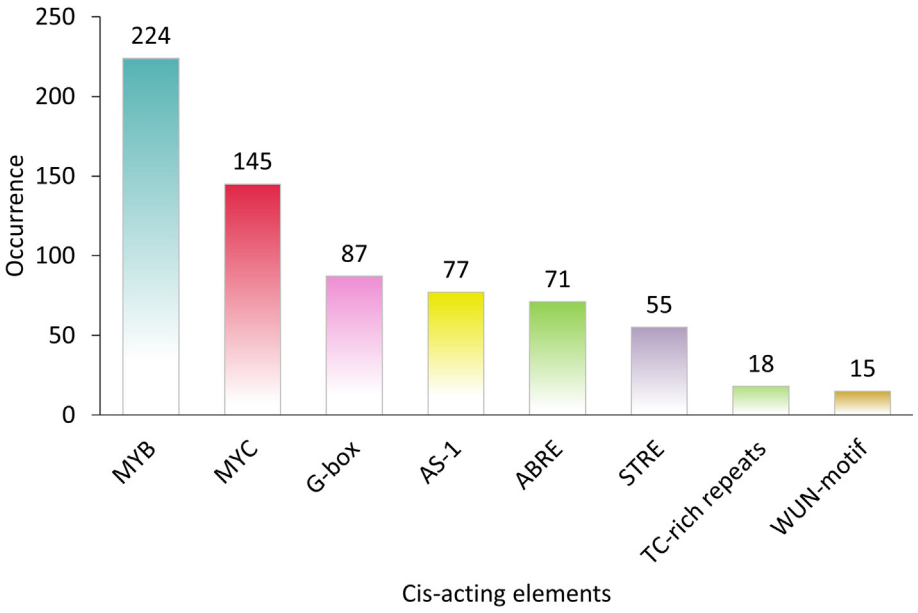

Fig. 4. The number of cis-acting elements identified among the RGAs in *H. incana*.

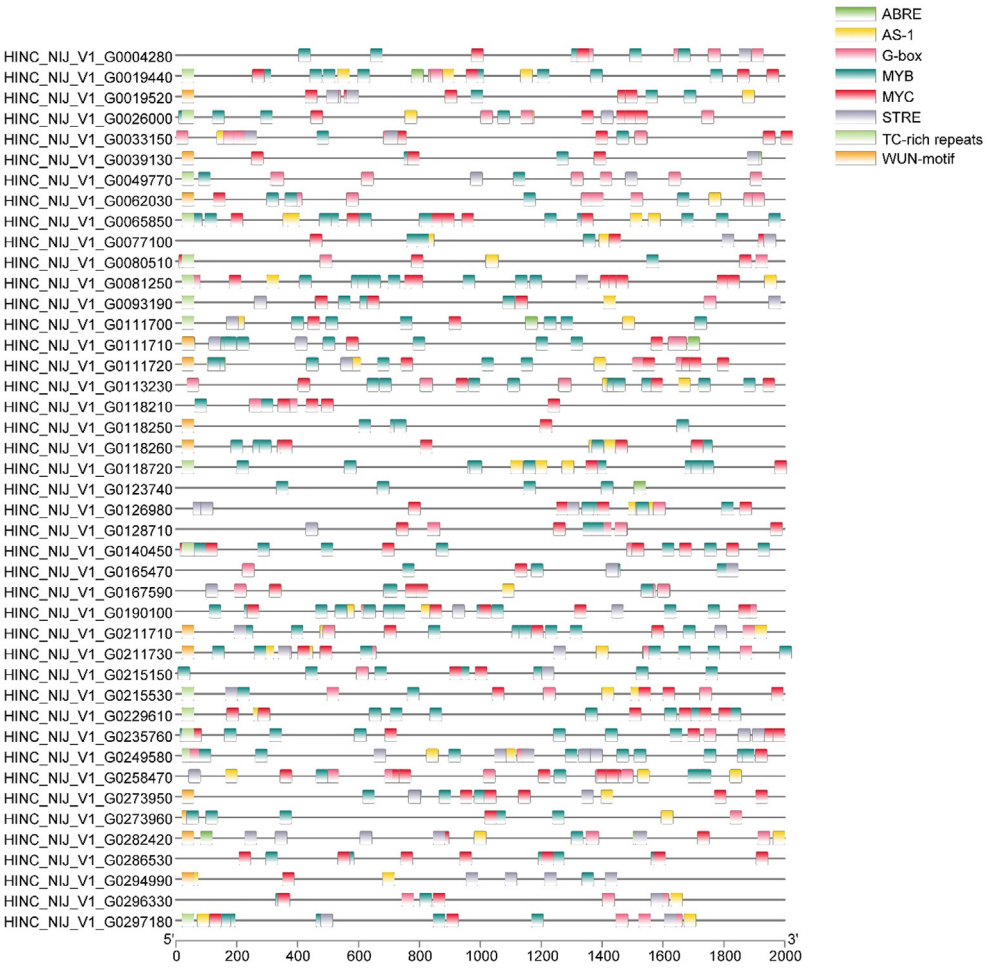

Fig. 5. Cis-regulatory elements in the promoter region of the RGA CDRHs in *H. incana*. The prediction of cis-regulatory elements for 2 kb sequence upstream of CDRHs. Some stress and defence response elements are displayed: ABRE, AS-1, G-box, MYB, MYC, STRE, TC-rich repeats, and WUN-motif.

CDRHs in these seven species [37]. Moreover, the identification of four CDRHs in *H. incana* (*HINC_NIJ_V1_G0126980*, *HINC_NIJ_V1_G0258470*, *HINC_NIJ_V1_G0273950*, *HINC_NIJ_V1_G0273960*) homologous to more than one cloned *R* gene raises interesting possibilities for crosstalk and synergistic interactions in the plant's defence network. For instance, *HINC_NIJ_V1_G0126980* is homologous to both *At_RFO1* and *Bna_Rlm9/4/7*, suggesting potential convergent evolution in pathogen recognition.

The results suggest that the wild species *H. incana* could be a potential *R* gene source for various disease resistances. Based on the CDRHs having RGA domains, it is a potential source of resistance against downy mildew, powdery mildew, bacterial leaf spot, sclerotinia stem rot, white rust, fusarium wilt, blackleg and black rot. The cloned R gene sequence with a confirmed function against pathogen to search homologs in *H. incana* can narrow down the RGA candidate gene identification and facilitate their application in Brassica crop disease resistance breeding programs.

4.5. Cis-acting element analysis of RGAs promoters in H. incana

The results investigate the functional classification of CREs based on their roles, which include hormone-responsive, light-responsive, defence-responsive, stress-responsive, and some possessing protein binding sites. The identified CREs include core promoter elements such as the TATA box and CAAT box, as well as elements associated with light response (ACE, G-box, GT1-motif, SP1, AAAC-motif, 3-AF1), auxin response (TGA, TGA-box, AuxRE, GATA-box), gibberellin response (TATC-box, P-box, GARE-motif, anaerobic induction element ARE), MeJA response (CGTCA-motif, TGACG-motif), salicylic acid response (TCA, SARE), cell cycle regulation (MAS-like), and low-temperature response (LTR). The study highlights the frequency of specific elements, such as G-box, AS-1, ABRE, STRE, TC-rich repeats, and WUN-motif, in the promoter regions. G-box, known to bind with G/HBF-1 and MYC to induce defence gene expression after pathogen attacks, was identified 87 times, underscoring its significance in defence against pathogens. AS-1, which plays a role in defence against pathogen attacks in rice genes, was also identified with high frequency (77 times) in the study. Additionally, ABRE (71 times), STRE (55 times), TC-rich repeats (18 times), and WUN-motif (15 times) were all found, indicating their involvement in stress and defence responsiveness. Notably, the presence of these elements aligns with previous studies, providing consistency and validation of the results.

5. Conclusion

In this study, we predicted a large number of RGAs in the reference genome of *H. incana*, an important member of the *Brassicaceae* family with practical and genetic significance. We investigated and characterised RGAs within *H. incana* that mainly consisted of RLK genes. The gene duplication analysis revealed that a significant portion of RGAs experienced duplication through evolution. We also perform phylogenetic analysis with the RGAs predicted in *H. incana* to gain insights into their evolution. Cloned gene homologues identification and cis-acting elements analysis of these genes further confirmed their potential role in disease defence. This research uncovers the diversity and functional aspects of RGAs in *H. incana* using advanced bioinformatics techniques. This exploration is essential for enhancing disease resistance in *Brassica* crops through breeding and biotechnological approaches, ultimately contributing to sustainable and productive agriculture.

Authors' contributions

TW: Data curation, formal analysis and writing original draft. HAAM: Formal analysis, RGAugury analysis. DE: Provided bioinformatics resources and software. JB: Conceptualisation, supervision, writing-review and editing. AD: Formal data analysis, visualisation, writing-review and editing.

Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. DE is Editorial Board Member of Agriculture Communications, but had no involvement in the peer review of this article and has no access to information regarding its whole review process.

Acknowledgments

This study is funded by the Australian Research Council projects DP200100762, DP210100296 and FL230100030.

References

[1] Tamokou JDD, Mbaveng AT, Kuete V. Chapter 8-Antimicrobial activities of African medicinal spices and vegetables. In: Kuete V, editor. Medicinal spices and vegetables from Africa. Academic Press; 2017. p. 207–37.
[2] Siemens J. Hirschfeldia. In: Kole C, editor. Wild crop relatives: genomic and breeding resources. Heidelberg: Springer Berlin; 2011. p. 171–6.
[3] Garassino F, Wijfjes RY, Boesten R, Reyes Marquez F, Becker FFM, Clapero V, et al. The genome sequence of *Hirschfeldia incana*, a new Brassicaceae model to improve photosynthetic light-use efficiency. Plant J 2022;112:1298–315.
[4] Liu Y, Wei W, Ma K, Li J, Liang Y, Darmency H. Consequences of gene flow between oilseed rape (*Brassica napus*) and its relatives. Plant Sci 2013;211:42–51.
[5] Chevre AM, Eber F, Kerlan MC, This P. Cytogenetics and molecular markers for genome analysis of crucifers. Comptes Rendus de l'Academie d'Agriculture de France 1991;77:59–67.
[6] Lefol E, Danielou V, Darmency H. Predicting hybridisation between transgenic oilseed rape and wild mustard. Field Crops Res 1996;45:153–61.
[7] Chevre AM, Eber F, Renard M. Transgenic rape and environmental hazards. Biofutur 1997;172:44–8.
[8] Chevre AM, Eber F, Baranger A, Vallée P, Pierre J, Renard M. Impact of oilseed rape genetic transformation. Cah Agric 1998;7:525–30.
[9] Chevre AM, Eber F, Jenczewski E, Darmency H, Renard M. Gene flow from oilseed rape to weedy species. Acta Agric Scand 2003;53:22–5.
[10] Darmency H, Fleury A. Mating system in *Hirschfeldia incana* and hybridisation to oilseed rape. Weed Res 2000;40:231–8.
[11] Lefol E, Séguin-Swartz G, Downey RK. Sexual hybridisation in crosses of cultivated *Brassica* species with the crucifers *Erucastrum gallicum* and *Raphanus raphanistrum*: potential for gene introgression. Euphytica 1997;95:127–39.
[12] Scheffler JA, Dale PJ. Opportunities for gene transfer from transgenic oilseed rape (*Brassica napus*) to related species. Transgenic Res 1994;3:263–78.
[13] Mazzucotelli E, Mastrangelo AM. Coping with fungal diseases in crops: new advances in genomics, breeding and management. Genes 2023;14(9):1758.
[14] Plümper B. Somatische und sexuelle Hybridisierung für den Transfer von Krankheitsresistenzen auf *Brassica napus* L. Ph.D. Thesis. Free University of Berlin; 1995.
[15] Scholze P, Hammer K. Evaluation of resistance to *Plasmodiophora brassicae*, alternaria and phoma in Brassicaceae. International Symposium Brassica 97, Xth Crucifer Genetics Workshop 1997;459:363–72.
[16] Sekhwal MK, Li P, Lam I, Wang X, Cloutier S, You FM. Disease resistance gene analogs (RGAs) in plants. Int J Mol Sci 2015;16:19248–90.
[17] Ravensdale M, Bernoux M, Ve T, Kobe B, Thrall PH, Ellis JG, et al. Intramolecular interaction influences binding of the Flax L5 and L6 resistance proteins to their AvrL567 ligands. PLoS Pathog 2012;8(11):e1003004.
[18] Le Roux C, Huet G, Jauneau A, Camborde L, Trémousaygue D, Kraut A, et al. A receptor pair with an integrated decoy converts pathogen disabling of transcription factors to immunity. Cell 2015;161:1074–88.
[19] Afzal AJ, Wood AJ, Lightfoot DA. Plant receptor-like serine threonine kinases: roles in signaling and plant defense. Mol Plant Microbe Interact 2008;21:507–17.
[20] Cortaga CQ, Latina RA, Habunal RR, Lantican DV. Identification and characterisation of genome-wide resistance gene analogs (RGAs) of durian (*Durio zibethinus* L.). J Genet Eng Biotechnol 2022;20(1):29.
[21] Lantican DV, Cortaga CQ, Manohar ANC, Sison MLJ. Resistance gene analogs of mango: insights on molecular defenses and evolutionary dynamics. Philipp J Sci 2020;149:915–34.
[22] Hasan N, Choudhary S, Naaz N, Sharma N, Laskar RA. Recent advancements in molecular marker-assisted selection and applications in plant breeding programmes. J Genet Eng Biotechnol 2021;19:1–26.
[23] Amiteye S. Basic concepts and methodologies of DNA marker systems in plant molecular breeding. Heliyon 2021;7(10):e08093.
[24] Ren J, Yu Y, Gao F, Zeng L, Lu X, Wu X, et al. Application of resistance gene analog markers to analyses of genetic structure and diversity in rice. Genome 2013;56:377–87.
[25] Fritz-Laylin LK, Krishnamurthy N, Tör M, Sjölander KV, Jones JDG. Phylogenomic analysis of the receptor-like proteins of rice and *Arabidopsis*. Plant Physiol 2005;138:611–23.

[26] Wang G, Ellendorff U, Kemp B, Mansfield JW, Forsyth A, Mitchell K, et al. A genome-wide functional investigation into the roles of receptor-like proteins in *Arabidopsis*. Plant Physiol 2008;147:503–17.
[27] Shao ZQ, Zhang YM, Hang YY, Xue JY, Zhou GC, Wu P, et al. Long-term evolution of nucleotide-binding site-leucine-rich repeat genes: understanding gained from and beyond the legume family. Plant Physiol 2014;166:217–34.
[28] Yu F, Zhang X, Huang Z, Chu M, Song T, Falk KC, et al. Identification of genome-wide variants and discovery of variants associated with *Brassica rapa* clubroot resistance gene *Rcr1* through bulked segregant RNA sequencing. PLoS One 2016; 11(4):e0153218.
[29] Zhang YM, Shao ZQ, Wang Q, Hang YY, Xue JY, Wang B, et al. Uncovering the dynamic evolution of nucleotide-binding site-leucine-rich repeat (NBS-LRR) genes in Brassicaceae. J Integr Plant Biol 2016;58:165–77.
[30] Hofberger JA, Zhou B, Tang H, Jones JDG, Schranz ME. A novel approach for multi-domain and multi-gene family identification provides insights into evolutionary dynamics of disease resistance genes in core eudicot plants. BMC Genomics 2014; 15:1–21.
[31] Bayer PE, Golicz AA, Tirnaz S, Chan CKK, Edwards D, Batley J. Variation in abundance of predicted resistance genes in the *Brassica oleracea* pangenome. Plant Biotechnol J 2019;17:789–800.
[32] Dolatabadian A, Bayer PE, Tirnaz S, Hurgobin B, Edwards D, Batley J. Characterisation of disease resistance genes in the *Brassica napus* pangenome reveals significant structural variation. Plant Biotechnol J 2020;18:969–82.
[33] Amas JC, Bayer PE, Tan WH, Tirnaz S, Thomas WJW, Edwards D, et al. Comparative pangenome analyses provide insights into the evolution of Brassica rapa resistance gene analogues (RGAs). Plant Biotechnol J 2023;21:2100–12.
[34] Li P, Quan X, Jia G, Xiao J, Cloutier S, You FM. RGAugury: a pipeline for genome-wide prediction of resistance gene analogs (RGAs) in plants. BMC Genomics 2016; 17:852.
[35] Yang H, Bayer PE, Tirnaz S, Edwards D, Batley J. Genome-wide identification and evolution of receptor-like kinases (RLKs) and receptor like proteins (RLPs) in *Brassica juncea*. Biol (Basel) 2020;10(1):17.
[36] Letunic I, Bork P. Interactive Tree of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. Nucleic Acids Res 2021;49:293–6.
[37] Cantila AY, Thomas WJW, Bayer PE, Edwards D, Batley J. Predicting cloned disease resistance gene homologs (CDRHs) in radish, underutilised oilseeds, and wild Brassicaceae species. Plants 2022;11:3010.
[38] Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, et al. TBtools: an integrative toolkit developed for interactive analyses of big biological data. Mol Plant 2020;13:1194–202.
[39] Lescot M, Déhais P, Thijs G, Marchal K, Moreau Y, Van de Peer Y, et al. PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. Nucleic Acids Res 2002;30:325–7.
[40] Cheng C, Krishnakumar V, Chan AP, Thibaud-Nissen F, Schobel S, Town CD. Araport11: a complete reannotation of the Arabidopsis thaliana reference genome. Plant J 2017;89:789–804.
[41] Zhang H, Mittal N, Leamy LJ, Barazani O, Song B. Back into the wild-Apply untapped genetic diversity of wild relatives for crop improvement. Evol Appl 2017;10:5–24.
[42] Smýkal P, Nelson MN, Berger JD, Von Wettberg EJB. The impact of genetic changes during crop domestication. Agronomy 2018;8:119.
[43] Sun R, Wang S, Ma D, Liu C. genome-wide analysis of LRR-RLK gene family in four *Gossypium* species and expression analysis during cotton development and stress responses. Genes 2018;9:592.
[44] Rody HVS, Bombardelli RGH, Creste S, Camargo LEA, Van Sluys MA, Monteiro-Vitorello CB. Genome survey of resistance gene analogs in sugarcane: genomic features and differential expression of the innate immune system from a smut-resistant genotype. BMC Genomics 2019;20:1–17.
[45] Zhu L, Li F, Xie T, Li Z, Tian T, An X, et al. Receptor-like kinases and their signaling cascades for plant male fertility: loyal messengers. New Phytol 2024;241:1421–34.
[46] Soltabayeva A, Dauletova N, Serik S, Sandybek M, Omondi JO, Kurmanbayeva A, et al. Receptor-like kinases (LRR-RLKs) in response of plants to biotic and abiotic stresses. Plants 2022;11(19):2660.
[47] Xu G, Guo C, Shan H, Kong H. Divergence of duplicate genes in exon-intron structure. Proc Natl Acad Sci USA 2012;109:1187–92.
[48] Shiu SH, Karlowski WM, Pan R, Tzeng YH, Mayer KFX, Li WH. Comparative analysis of the receptor-like kinase family in *Arabidopsis* and rice. Plant Cell 2004;16:1220–34.
[49] Bai J, Pennill LA, Ning J, Lee SW, Ramalingam J, Webb CA, et al. Diversity in nucleotide binding site–leucine-rich repeat genes in cereals. Genome Res 2002;12: 1871–84.
[50] Meyers BC, Kozik A, Griego A, Kuang H, Michelmore RW. Genome-wide analysis of NBS-LRR–encoding genes in *Arabidopsis*. Plant Cell 2003;15:809–34.
[51] David P, Chen NWG, Pedrosa-Harand A, Thareau V, Sévignac M, Cannon SB, et al. A nomadic subtelomeric disease resistance gene cluster in common bean. Plant Physiol 2009;151:1048–65.
[52] Guo YL, Fitz J, Schneeberger K, Ossowski S, Cao J, Weigel D. Genome-wide comparison of nucleotide-binding site-leucine-rich repeat-encoding genes in *Arabidopsis*. Plant Physiol 2011;157:757–69.