Mol Breeding (2019) 39: 131
https://doi.org/10.1007/s11032-019-1038-8

Fine mapping of a clubroot resistance gene from turnip using SNP markers identified from bulked segregant RNA-Seq

Z. Huang · G. Peng · B. D. Gossen · F. Yu

Received: 18 February 2019 / Accepted: 14 August 2019 / Published online: 29 August 2019

Abstract Clubroot caused by *Plasmodiophora brassicae* poses a serious threat to canola production around the world, but sources of clubroot resistance in canola are limited. In this study, turnip cultivar "Purple Top White Globe," which is highly resistant to pathotype 3 (Williams' system), was used as the source of resistance. Genetic studies showed that the resistance in this cultivar was controlled by a major gene, *Rcr5*. Bulked segregant RNA-Seq was used to map the gene. A total of 124.6 M raw reads were generated from the resistant (R) and susceptible (S) pooled samples, and 78 K polymorphic DNA variants between the pooled R and the pooled S samples were identified. The percentage of polymorphic variants (PPV) on A03 was much higher than that on the other chromosomes, which indicated that *Rcr5* was located on A03. *Rcr5* was further mapped into the 23–31 Mb region of A03 through analysis of PPV in the chromosome. A segregated population consisting of 824 plants was genotyped based on 15 SNP markers in the region using Kompetitive Allele Specific PCR. *Rcr5* was finely mapped between SNP_A03_83 and SNP_ A03_100, with 0.2 and 0.3 cM from *Rcr5*, respectively. The markers were located between 23,339,019 and 23,465,030 bp on A03. Eighteen genes were annotated in the *Rcr5* target region, seven genes were expressed, and DNA variants were identified in the genes. The SNP markers developed in this study could be used for marker-assisted selection for resistance to clubroot when the gene is transferred to canola.

Keywords Clubroot · *Plasmodiophora brassicae* · RNA-Seq · SNPs · Fine mapping · KASP · *Brassica napus* · *Brassica rapa*

Introduction

Clubroot, caused by *Plasmodiophora brassicae* (Wor.), is an important disease on brassica crops worldwide, especially in temperate climate regions. It can infect most of the Cruciferae, such as cabbage and Chinese cabbage (Dixon 2006; Piao et al. 2009). Clubroot is an increasing problem on canola (*Brassica napus* L.) on the Canadian Prairies, where the number of fields with confirmed clubroot infestations has steadily increased since it was first identified in 2003 (Gossen et al. 2015). The disease has also been reported in Saskatchewan and Manitoba (Cao et al. 2009; Dokkenbouchard et al. 2010). Clubroot is also a serious problem for rapeseed production in China (Hu and Peng 2016).

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s11032-019-1038-8) contains supplementary material, which is available to authorized users.

Z. Huang · G. Peng · B. D. Gossen · F. Yu (✉)
Agriculture and Agri-Food Canada, Saskatoon Research and Development Centre, 107 Science Place, Saskatoon, Saskatchewan S7N OX2, Canada
e-mail: Fengqun.Yu@Canada.ca

Z. Huang
State Key Laboratory of Crop Stress Biology for Arid Areas/ College of Agronomy, Northwest A&F University, Yangling 712100 Shaanxi, China

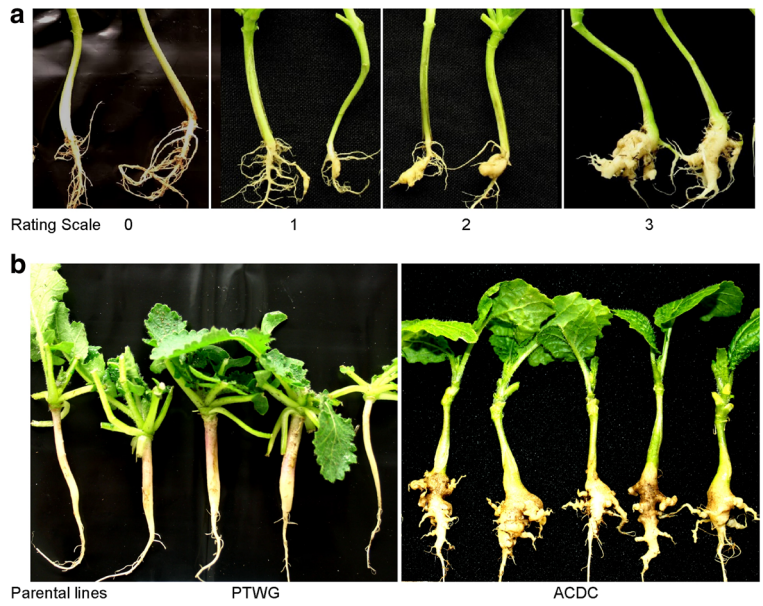

Fig. 1 Phenotypes of plants in response to inoculation with pathotype 3 of *Plasmodiophora brassicae*: (**A**) range of clubroot symptoms in the BC1 population at 5 weeks after inoculation (0 = no clubs, 3 = large clubs) and (**B**) phenotypes of resistant cultivar PTWG and susceptible line ACDC

Identification of clubroot resistance (CR) genes is a very important step toward breeding for Brassica crops resistant to clubroot. To date, > 10 genes or QTLs were mapped into chromosomes of the A-genome in *B. rapa*. *Crr2* and *PbBa1.1* were located on A01 (Chen et al. 2013; Suwabe et al. 2003); *CRc* and *Rcr8* on A02 (Sakamoto et al. 2008; Yu et al. 2017); *Crr3*, *CRa*, *CRb*, CRb^{kato}, *CRd*, *CRk*, *PbBa3.1*, *PbBa3.2*, *PbBa3.3*, *Rcr1*, *Rcr2*, and *Rcr4* on A03 (Chen et al. 2013; Chu et al. 2014; Hirai et al. 2004; Huang et al. 2017; Matsumoto et al. 1998; Pang et al. 2018; Piao et al. 2004; Sakamoto et al. 2008; Yu et al. 2016, 2017); *Crr4* on A06 (Suwabe et al. 2003); and *Crr1*, *PbBa8.1*, and *Rcr9* on A08 (Chen et al. 2013; Suwabe et al. 2006; Yu et al. 2017). Three genes, *Crr1*, *CRa*, and CRb^{kato}, have been cloned. These genes encode toll-interleukin-1 receptor, nucleotide binding site, leucine-rich repeat (TIR-NBS-LRR (TNL)) proteins (Hatakeyama et al. 2013, 2017; Ueno et al. 2012).

Genetic resistance is a highly effective approach to manage this disease, but CR germplasm is rare in *Brassica napus* (canola). Fortunately, sources of resistance are present in vegetable crops of *B. rapa* and *B. oleracea*, which are closely related to *B. napus* (Liu et al. 2018; Peng et al. 2014) and can be used for breeding of CR *B. napus* cultivars. Identifying resistance through phenotyping can be influenced by temperature, light, pH, and other factors. Use of molecular markers linked to CR genes for selection provides greater sensitivity and consistency (Liu et al. 2018). Kompetitive Allele Specific PCR (KASP), a single nucleotide polymorphism (SNP) genotyping technology, is fast, efficient, and accurate for molecular breeding of crops (Jatayev et al. 2017).

Bulked segregant RNA-sequencing (BSR-Seq) based on next-generation sequencing is a very robust approach for quantifying gene expression and identifying abundant SNP sites, which has been used for transcriptome analysis and gene mapping in crops (Chu et al. 2014; Huang et al. 2017; Liu et al. 2012; Yu et al. 2016). It can be used to determine the position of target genes by identifying large numbers of SNPs, which can be used to develop molecular markers and finely map the causal genes (Dakouri et al. 2018; Liu et al. 2012; Yu et al. 2016). For example, a previous study identified a high percentage of polymorphic variants (PPV) on chromosome A03 of *B. rapa* adjacent to the clubroot resistance gene *Rcr1* (Yu et al. 2016). This observation was used to map a major clubroot resistance gene in *B. oleracea* (Dakouri et al. 2018) and a major clubroot resistance gene in *B. nigra* (Chang et al. 2019).

A turnip cultivar, "Purple Top White Globe" (PTWG), was recently identified to be resistant to pathotype 3 of *P. brassicae* (Williams' system) in Canada (Peng et al. 2014). The current study describes the identification and mapping of a clubroot resistance gene from PTWG using BSR-Seq. SNP markers tightly linked to the gene were developed and the gene was finely mapped.

Table 1 Short reads from the resistant (R) and susceptible (S) bulks, number of polymorphic variants (SNP and InDel) on each chromosome uniquely identified from R sample bulks of the BC_1 population, when compared against the *Brassica rapa* reference genome v1.5 at http://brassicadb.org/brad

Chromosome	Size (Mbp)	Number of sequences ($\times 10^6$)		Accumulated length of sequences (bases $\times 10^6$)		SNP ($\times 10^3$)	InDel (10^3)	Total ($\times 10^3$)	SNP (%)	InDel (%)
		R	S	R	S					
A01	26.8	10.4	9.3	675.6	602.6	5.0	2.7	7.7	65	35
A02	26.9	9.1	8.1	592.3	528.5	4.5	2.4	6.9	65	35
A03	31.7	15.3	13.6	992.0	884.8	10.9	4	14.9	73	27
A04	19.3	8.1	7.2	529.1	466.7	2.9	1.7	4.6	63	37
A05	25.3	9.7	8.6	632.9	562.1	4.8	2.6	7.4	65	35
A06	25.2	11.8	10.3	766.9	672.1	4.8	2.8	7.6	63	37
A07	25..9	10.9	9.7	706.5	627.7	3.8	2.5	6.3	60	40
A08	208	9.5	8.5	618.2	549.9	3.9	2.1	6.0	65	35
A09	38.9	15.7	14.0	1022.9	909.2	6.9	3.9	10.8	64	36
A10	16.4	7.6	6.7	496.9	437.2	4.1	2.6	6.7	61	39
Total	25.7	108.2	96.0	7033.4	6240.7	51.9	26.9	78.8	66	34

Materials and methods

Materials

The clubroot-resistant (R) turnip cultivar PTWG (seed purchased online at https://www.westcoastseeds.com/shop/vegetable-seeds/turnip-seeds/purple-top-white-globe/) was crossed with a highly susceptible (S) doubled-haploid, canola-quality line of *B. rapa*, ACDC (developed by Dr. Kevin Falk, Saskatoon Research and Development Centre, Agriculture and Agri-Food Canada), to produce F_1. One F_1 plant (male) was crossed with the susceptible (S) line (female) to produce a BC_1 population.

Inoculation and resistance test

The parents, F_1 plants, and BC_1 populations were tested for resistance to pathotype 3 of *P. brassicae*, the most prevalent on canola in the Canadian Prairie region (Strelkov et al. 2006, 2007). A field collection of the clubroot pathogen was used to inoculate plants for studies on the inheritance and genetic mapping. Preparation of inoculums, inoculation, and clubroot severity rating followed the methods described by Chu et al. (2014). Briefly, each plant was inoculated by pouring 5 ml of *P. brassicae* resting spore suspension (1×10^7 spores/ml) into each seeded well. The inoculated seedlings were maintained in a greenhouse at about 20/18 °C day/night temperature and 16-h photoperiod. At 5 weeks after inoculation, plants were uprooted and clubroot symptoms on roots were rated using a standard 0 to 3 scale (Fig. 1A) where 0 = no clubbing; 1 = small clubbing only; 2 = moderate clubbing; and 3 = severe clubbing (Strelkov et al. 2006). A rating of 0 was considered R and ratings of 1–3 were S. The highly susceptible *B. rapa* breeding line ACDC was included as a susceptible control to ensure that the inoculation conditions were conducive for infection. Segregation of R and S phenotypes in the BC_1 population was analyzed using chi-square (χ^2) tests for goodness of fit (Sokal and Rohlf 1995).

RNA-Seq and sequence alignment

Components of the BC_1 population were assessed using RNA-Seq. At 15 days post-inoculation, leaf tissues from 30 plants with a rating of 0 were combined to form a R bulk, and leaves from 30 plants with a rating of 3 were combined to form a S bulk. Together, one R and one S bulk formed a single biological replicate. Three biological replicates were assessed using RNA-Seq with an RNeasy Plant Mini Kit (Qiagen, Toronto, ON). On-column deoxyribonuclease (DNase) digestion using a

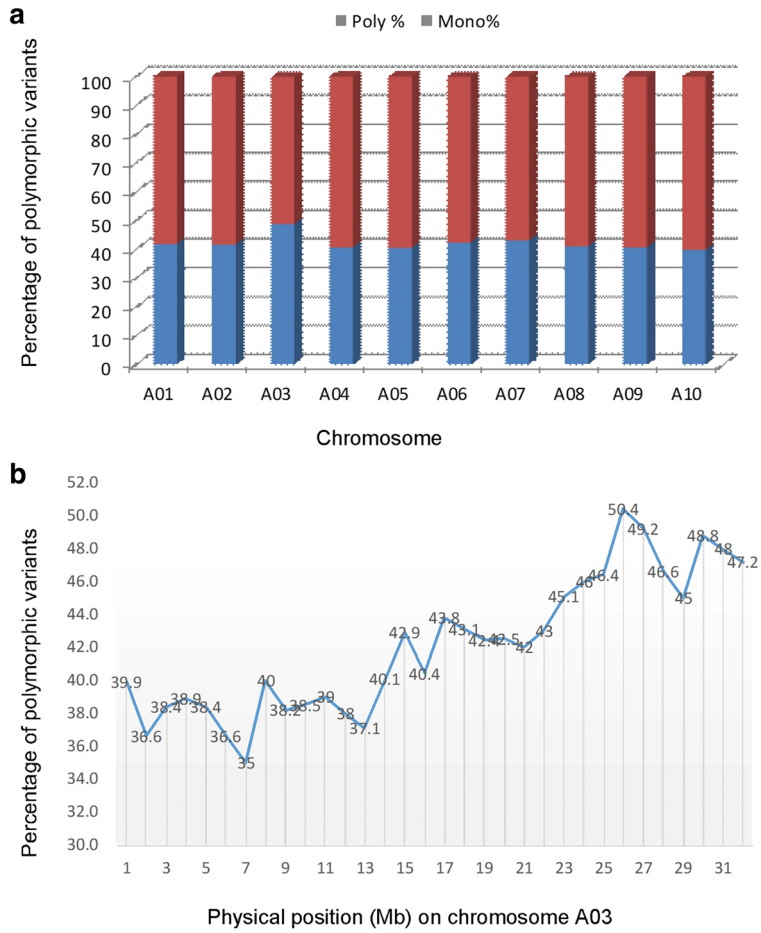

Fig. 2 Distribution of DNA variants (%): (**A**) monomorphic and polymorphic variants on each chromosome and (**B**) polymorphic variants on chromosome A03

Qiagen RNase-Free DNase set (Qiagen) was used for total RNA extraction, following the manufacturer's instructions.

The RNA concentration was checked using a NanoDrop 2000c Spectrophotometer (Thermo Scientific, Waltham, MA), and RNA quality was assessed with an Agilent Bioanalyzer 2100 (Agilent Technologies; Mississauga, ON) to ensure that the RNA integrity number (RIN) was > 8 for each sample. cDNA libraries were prepared following the TruSeq RNA Sample Preparation v2 Guide (Illumina, San Diego, CA). A NanoDrop ND-2000c spectrophotometer was used to check cDNA concentrations and purity. Quality control and qPCR analysis were used to validate the cDNA libraries. The Experion DNA 1K Analysis Kit (Bio-Rad Laboratories, Inc.) was used to confirm the size and purity of the cDNA libraries, based on a band at approximately 260 bps. The KAPA Library Quantification Kit v4.11 was used to perform qPCR of cDNA libraries.

RNA-Seq was carried out on samples from each inoculated R and S bulk using the Illumina MiSeq platform at the University of Saskatchewan (Saskatoon, SK, Canada). Short reads from a pool of the three R bulks and a pool of the three S bulks (pooled sample assembly) were assembled into the reference genome (*B. rapa*, V1.5; http://brassicadb.org/brad) as described by Yu et al. (2016) using SeqMan NGen 13 (DNASTAR, Madison, WI) software. Standard assembling and filtering parameters were used. Discovery of variants (SNPs and InDels) in comparison with the DNA sequences in the *B. rapa* "Chiifu" was performed using SeqMan Pro 13 (DNASTAR, Madison, WI) with Q call ≥ 15 and depth ≥ 5.

SNP discovery and mapping of the causal gene

SNP discovery and mapping of the causal gene were performed using uniquely aligned short reads from the

Table 2 Details of SNP markers linked to *Rcr5*

SNP markers	Location on A03 (bp)	Primers	Sequences (5′————3′)
SNP-A03_19	24,393,825	SNP-A03_19FAM	AAAGCTCCAAACATCGTTCCCTTCA
		SNP-A03_19HEX	AAAGCTCCAAACATCGTTCCCTTCT
		SNP-A03_19Re	GTGAAGTGGAACCCCGTTGCGAA
SNP-A03_32	24,394,435	SNP-A03_32FAM	AGCTCCTCAAAGTCTTCCACT
		SNP-A03_32HEX	CTAGCTCCTCAAAGTCTTCCACG
		SNP-A03_32Re	AACAGAGATTAGAGAGAAAGTAGATGTGAT
SNP_A03_37	25,283,656	SNP-A03_37FAM	CAACATGGGTGATGACAAATCCATC
		SNP-A03_37HEX	ACAACATGGGTGATGACAAATCCATT
		SNP-A03_37Re	CGATCGGTGTCAACTTCAGCAATGAT
SNP_A03_45	27,578,239	SNP-A03_45FAM	AGAACTCTCTAGAAGCTCCTGCT
		SNP-A03_45HEX	GAACTCTCTAGAAGCTCCTGCC
		SNP-A03_45Re	AGAATGGAGCCATCTATTGCTATCGAA
SNP-A03_54	23,713,799	SNP_A03_54FAM	CTTCGAGACGAAAACTGGTGTGGATG
		SNP_A03_54HEX	CTTCGAGACGAAAACTGGTGTGGATC
		SNP_A03_54Re	GTTTAGTCATCGCCCGGAACTGGTATCT
SNP-A03_62	23,847,546	SNP_A03_62FAM	GTTCTGGGCAGGCACTGTTATCTCT
		SNP_A03_62HEX	GTTCTGGGCAGGCACTGTTATCTCC
		SNP_A03_62Re	CAGAAAGCAGAGGAACCTGAAGCAGA
SNP_A03_83	23,465,030	SNP_A03_83FAM	GGATGGAGGTCTATTTACTGTATGCCAT
		SNP_A03_83HEX	GGATGGAGGTCTATTTACTGTATGCCAC
		SNP_A03_83Re	CAGATAGATTGATAGGAGCTGACCCT
SNP_A03_84	23,465,907	SNP_A03_84FAM	GGTTCTAACGAATGCACCATTCATTCTA
		SNP_A03_84HEX	GGTTCTAACGAATGCACCATTCATTCTC
		SNP_A03_84Re	GACGGCTTTACTGTTGTTGACGTAGTGAT
SNP_A03_89	23,209,983	SNP_A03_89FAM	GCGAGGCGAGTACGAAGCTCTCTTACT
		SNP_A03_89HEX	TGCGAGGCGAGTACGAAGCTCTCTTACA
		SNP_A03_89Re	GGTGGATGGAGTTAACATTAACCACCTT
SNP_A03_90	23,213,295	SNP_A03_90FAM	CAGGCGTGTGGGCAAGTTTCTACGTT
		SNP_A03_90HEX	CAGGCGTGTGGGCAAGTTTCTACGTA
		SNP_A03_90Re	CAGAGATCAAACCAGTATCTGTAGATGA
SNP_A03_92	23,001,148	SNP_A03_92FAM	GTGCTAAAGATGGTCCTGCTTTGTCTCC
		SNP_A03_92HEX	GTGCTAAAGATGGTCCTGCTTTGTCTCA
		SNP_A03_92Re	GCTTTTGACCAACGGTGCAAGCGAAT
SNP_A03_99	23,338,743	SNP_A03_99FAM	CCACCGATCTTGGGATCAATCACATTCAA
		SNP_A03_99HEX	CCACCGATCTTGGGATCAATCACATTCAG
		SNP_A03_99Re	CCATTTGCAGCTATAGTAGGCCAAGATGA
SNP_A03_100	23,339,019	SNP_A03_100FAM	TGGCCTTATTTGGATTCCTGCATTTG
		SNP_A03_100HEX	TGGCCTTATTTGGATTCCTGCATTTA
		SNP_A03_100Re	GTTTACAGGGAGAATCTGTGGAAGAGCGT
SNP_A03_101	23,271,493	SNP_A03_101FAM	CTTCGATGGCTGGAGCTTGCTTTGGG
		SNP_A03_101HEX	CTTCGATGGCTGGAGCTTGCTTTGGT
		SNP_A03_101Re	CTCCCATGGACACAGATGCTCTGTACCGAT
SNP_A03_102	23,272,470	SNP_A03_102FAM	CTTGAGCCTGCCCTTCTTCACATCGGCG
		SNP_A03_102HEX	CTTGAGCCTGCCCTTCTTCACATCGGCT
		SNP_A03_102Re	GCATGGGAACGCAAATGTAACCTGCTAAGGT

pooled R and S bulks, as described by Yu et al. (2016). Each potential DNA variant was examined for polymorphism, and putative SNPs were identified using the Lasergene Genomics Suite13 software (DNASTAR). The DNA variants were classified into two types, monomorphic (mono) and polymorphic (poly). The location of the causal gene was determined according to PPV (Dakouri et al. 2018).

SNP genotyping and linkage analysis

Genomic DNA was extracted from young leaves of each BC_1 plant (including the BC_1 plants used for RNA-Seq) with the CTAB (cetyltrimethylammonium bromide) method (Doyle 1990). The final DNA concentration was adjusted to 50 ng/μl for each sample. SNPs identified in the target region were confirmed using KASP (http://www.lgcgroup.com/), following the manufacturer's instruction. PCR reactions were performed using a StepOne Plus Real-Time PCR System (Applied Biosystem, Mississauga, ON). A genetic linkage map was developed using JoinMap 4.1 software (Ooijen 2006).

Annotation of the genes in the target region

The gene description and gene ontology (GO) of each gene in the target region were assessed using a Blast2GO search (Conesa et al. 2005), which uses BLASTX algorithms to search for matches against the non-redundant protein database from the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov). BLAST hits were mapped to the functional information stored in the GO database to retrieve GO terms associated with the hits.

Results

Inheritance of clubroot resistance in PTWG

At 5 weeks after inoculation, 14 plants each of the parental lines and F_1 were evaluated for resistance to pathotype 3 of *P. brassicae*. All of the PTWG plants were resistant (rating = 0), all ACDC plants were highly susceptible (rating = 3) (Fig. 1B), and all F_1 plants were

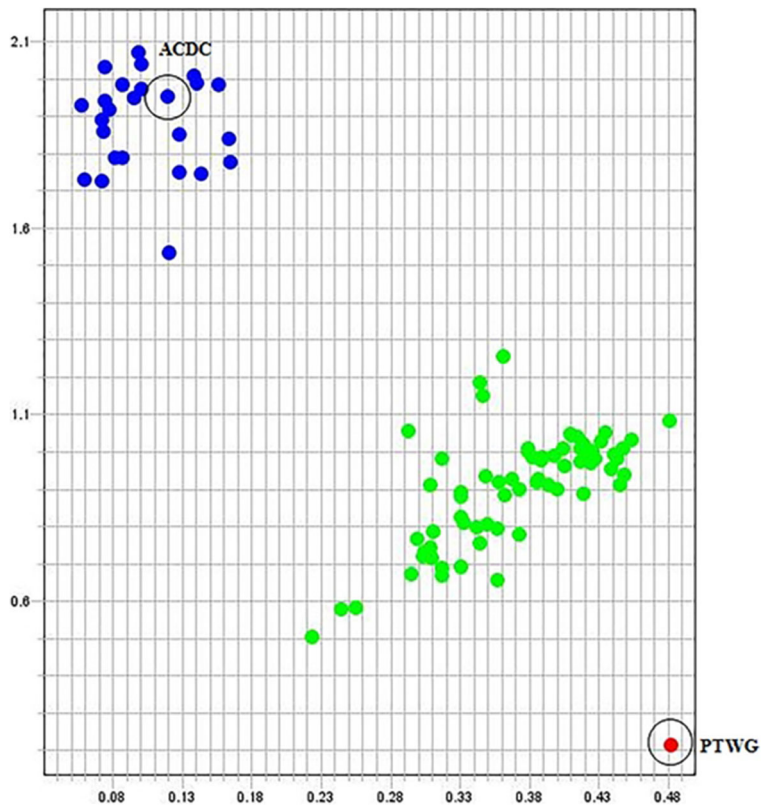

Fig. 3 Analysis of Kompetitive Allele Specific PCR (KASP) using SNP_A03_84 in the parental lines PTWG (R) and ACDC (S) and BC_1 plants of ACDC × (ACDC × PTWG). The green dots are clubroot-resistant plants and blue dots are susceptible plants. The circled dots are the parental lines

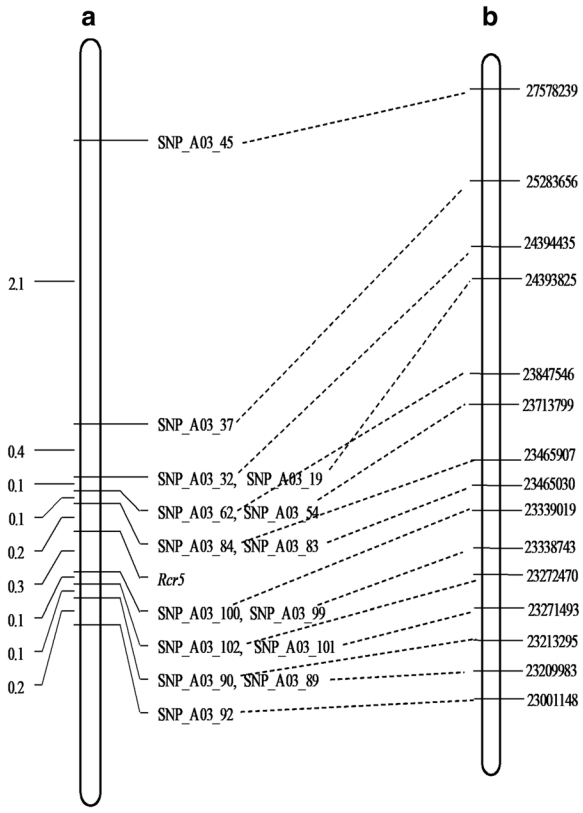

Fig. 4 Mapping of the *Rcr5* region in the BC_1 population: (**A**) linkage map of the region surrounding the *Rcr5* gene (y) and (**B**) a partial physical map of A03 chromosome of *Brassica rapa* showed the homologs of mapped marker sequences. Dotted lines indicate the relationship of the two maps

resistant (rating = 0). Separation of R and S was observed in the BC_1 population, where the numbers of plants with ratings of 0, 1, 2, and 3 were 435, 27, 71, and 291, respectively, and the number of the clubroot-R and -S plants were 435 and 389, respectively, which was consistent with an expected ratio of 1:1 ($\chi^2 = 2.56$, $P = 0.11$). These results indicated that resistance in PTWG was associated with a single dominant gene, designated *Rcr5* (*Resistance to clubroot 5*).

Assembly of RNA-Seq short reads into the reference genomes of *B. rapa*

There are a total of 108.2 M sequences, 2062.2 Mb in length, with 27-fold coverage of the reference A-genome from the three pooled R bulks, and 96.0 M sequences, 6240.7 Mb in length, with 27-fold coverage from the three pooled S bulks (Table 1). The sequence counts assembled into the genome for each chromosome were significantly correlated to chromosome length for the R bulks ($r = 0.90$, $P = 0.004$) and S bulks ($r = 0.91$, $P = 0.0003$).

Identification of DNA variants and mapping of *Rcr5*

A total of 78.8 K poly variants (51.9 K SNPs, about 66% of the variants, and 26.9 K InDels, about 34% of the variants) were identified in 10 chromosomes from the BSR-Seq project (Table 1). The number of poly variants was also significantly associated with the length of chromosome ($r = 0.74$; $P = 0.010$). The PPV was usually 40–43%, whereas, for A03, the PPV was higher than that of the other chromosomes, reaching to 49% (Fig. 2A), indicating that *Rcr5* was located on chromosome A03. The PPV on chromosome A03 was further analyzed. The highest PPV was located in the physical range from approximately 23–31 Mb (Fig. 2B), and no other regions in the rest of chromosomes showed such a high level of PPV (Supplemental Figure 1), indicating that *Rcr5* was located in the genomic region of the chromosome. There were 824 plants in the BC_1 population analyzed with 15 poly SNP loci (Table 2) spanning the 23–27-Mb region of chromosome A03 using the KASP method (Fig. 3). A total of 1253 unique SNPs from the R bulk were identified in the 23–27-Mb region from the BSR-Seq project (Supplemental Table 1). The 15 SNP loci (Table 2) consisted of 11 identified from the project (11 out of 1253 SNPs) and 4 SNPs (SNP-A03_19, SNP-A03_32, SNP-A03_54, and SNP-A03_62) from the previous mapping of *Rcr1* (Yu et al. 2016) and *Rcr2* (Huang et al. 2017) projects. A linkage map composing of 3.6 cM was constructed (Fig. 4A), confirming that *Rcr5* resides in the region. *Rcr5* was flanked by SNP_A03_83 & 84 and SNP_A03_99 &100, in an interval of 0.5 cM (Fig. 4A).

DNA variants and gene annotation in the target region

The two closest SNP markers, SNP_ A03_100 and SNP_ A03_83, flanking *Rcr5* were developed based on the SNP sites at 23,339,019 and 23,465,030 bp (Fig. 4B). These markers were mapped within the genes *Bra012595* and *Bra012578* on A03, respectively. The physical distance between these two markers was 126,012 bases. There are 18 genes in this region, based on the reference genome v1.5 (Table 3). A very low

Table 3 Numbers of polymorphic variants (SNPs and InDels) and the level of gene expression in RPKM (Reads Per Kilobase of transcript per Million mapped reads) among the 18 genes located in the *Rcr5* interval that could be identified

Gene	Start (base)	End (base)	No. of SNP	No. of InDel	Total no. of variant	No. of short reads in R	No. of short reads in S	R bulks (RPKM)	S bulks (RPKM)
Bra012595	23,337,781	23,339,208	9	0	9	26	19	282.4	326.9
Bra012594	23,340,092	23,343,812	0	0	0	0	0	1.3	1.2
Bra012593	23,353,318	23,353,554	0	0	0	0	0	0.0	0.0
Bra012592	23,357,893	23,358,129	0	0	0	0	0	0.0	0.0
Bra012591	23,366,900	23,368,123	12	0	12	21	9	14.4	15.8
Bra012590	23,369,671	23,372,052	0	0	0	0	0	0.2	0.2
Bra012589	23,376,061	23,376,893	0	0	0	0	0	0.1	0.1
Bra012588	23,379,457	23,381,516	2	0	2	12	12	4.1	2.2
Bra012587	23,382,122	23,387,287	0	0	0	0	0	0.0	0.1
Bra012586	23,398,926	23,401,869	17	0	17	33	18	21.1	20.6
Bra012585	23,403,421	23,408,434	0	0	0	0	0	0.1	0.1
Bra012584	23,421,093	23,421,668	0	0	0	0	0	1.0	1.0
Bra012583	23,433,359	23,434,717	0	0	0	0	0	0.5	0.6
Bra012582	23,448,511	23,449,257	0	0	0	0	0	0.1	0.0
Bra012581	23,457,614	23,459,115	2	0	2	25	24	16.5	16.3
Bra012580	23,460,027	23,460,952	4	1	5	7	3	283.2	288.4
Bra012579	23,462,221	23,463,664	0	0	0	0	0	0.0	0.0
Bra012578	23,464,101	23,467,994	9	0	9	23	18	5.8	6.0

level of gene expression was found in most of the genes except seven of them (*Bra012595*, *Bra012591*, *Bra012588*, *Bra012586*, *Bra012581*, *Bra012580*, and *Bra012578*) (Table 3). Poly variants in the coding regions were detected in all seven genes (Table 3, Supplemental Table 2), but no TNL genes were identified in the region (Supplemental Table 2).

Discussion

Genetic resistance to *P. brassicae* can provide effective management of clubroot, but sources of resistance in *B. napus* are very limited, so identification and introgression of CR from related species into *B. napus* is important. In this study, more than ten KASP markers closely linked to *Rcr5* were developed, which can be used in the detection of these loci in *B. napus* or *B. rapa* breeding programs.

A previous study identified a high proportion of PPV on chromosome A03 of *B. rapa* adjacent to the clubroot resistance gene *Rcr1* (Yu et al. 2016). Dakouri et al. (2018) successfully map *Rcr7* through identification of PPV. In this study, *Rcr5* was found to be associated with chromosome A03 of *B. rapa* since A03 carried the highest PPV in the BSR-Seq project. A high PPV was identified within the physical interval 23–31 Mb of chromosome A03 in *B. rapa*, which indicated that *Rcr5* was likely located in this region. KASP and linkage analysis of the BC_1 population based on several SNP markers further narrowed the interval to 0.126 Mb. This supported the previous report that identification of PPV could be used for genetic mapping of genes of interest (Dakouri et al. 2018). We conclude that identification of PPV in combination with KASP analysis is a powerful approach for fine mapping of causal genes that control agronomic traits of crops.

Several genes for clubroot resistance have previously been mapped into chromosome A03 of *B. rapa*, including *Crr3* (Hirai et al. 2004), *CRa* (Matsumoto et al. 2005; Ueno et al. 2012), *CRb* (Piao et al. 2004; Zhang et al. 2014), CRb^{kato} (Hatakeyama et al. 2017; Kato et al. 2013), *CRk* (Sakamoto et al. 2008), *Rcr1* (Chu et al. 2014; Yu et al. 2016), *Rcr2* (Huang et al. 2017), and *Rcr4* (Yu et al. 2017). Two cloned genes *CRa* and CRb^{kato} are

identical (Hatakeyama et al. 2017) while *Rcr1*, *Rcr2*, and *Rcr4* were co-localized with *CRa*/*CRb*^{kato} (Huang et al. 2017; Yu et al. 2016, 2017), in the genomic region of 24.35 and 24.39 Mb of the A03 chromosome (Ueno et al. 2012; Zhang et al. 2014). *CRb* was tightly linked to *CRa*/*CRb*^{kato}, located between 23.67 and 23.76 Mb, and *Crr3* and *CRk* were located between 15.2 and 15.4 Mb (Zhang et al. 2014). In the current study, two SNP markers in the *CRa*/*CRb*^{kato} region (A03_SNP_32, 24.39 Mb; A03_SNP_19, 24.39 Mb) and two SNP markers in the *CRb* region (A03_SNP_54, 23.71 Mb; A03_SNP_58, 23.72 Mb) were analyzed. These four SNP markers identified several recombinants in the BC_1 population. *Rcr5* was mapped between 23.33 and 23.46 Mb, which demonstrated that it was located at a different location from the CR genes published previously. In addition, the resistance reaction of the parental resistant line, PTWG, against some of the Canadian pathotypes differed from the lines carrying the other resistant genes (unpublished data).

Only 18 genes lie in the small interval into which *Rcr5* was mapped. Poly variants unique to the R samples were expressed in 7 out of 18 genes, but none of the 18 genes encode proteins known to be associated with TNL or with other components of plant disease resistance. Previous studies have shown that there are DNA sequence gaps in the *Rcr5* target region in the *B. rapa* reference genome (*B. rapa*, V1.5; http://brassicadb.org/brad). It is possible that *Rcr5* is located in a sequence gap or that *Rcr5* encodes a novel protein that differs from the previous cloned CR resistance genes. However, these different hypotheses can only be resolved by cloning *Rcr5*, which was beyond the scope of the current study.

Acknowledgments We thank Dr. M. Hecker for providing the MiSeq facility for sequencing, J. Wang and L. McGregor for assistance on phenotyping, Dr. K. Falk for proving line ACDC, and Dr. Q. Chen for assistance on analyzing KASP markers.

Funding information We thank Agriculture and Agri-Food Canada for providing funding for the study.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

Cao T, Manolii VP, Hwang SF, Howard RJ, Strelkov SE (2009) Virulence and spread of *Plasmodiophora brassicae* [clubroot] in Alberta, Canada. Can J Plant Pathol 31:321–329

Chang A, Lamara A, Wei Y, Hu H, Parkin IAP, Gossen BD, Peng G, Yu F (2019) Clubroot resistance gene *Rcr6* in *Brassica nigra* resides in a genomic region homologous to chromosome A08 in *B. rapa*. BMC Plant Biol. https://doi.org/10.1186/s12870-019-1844-5

Chen J, Jing J, Zhan Z, Zhang T, Zhang C, Piao Z (2013) Identification of novel QTLs for isolate-specific partial resistance to *Plasmodiophora brassicae* in *Brassica rapa*. PLoS One 8:e85307

Chu M, Song T, Falk KC, Zhang X, Liu X, Chang A, Lahlali R, Mcgregor L, Gossen BD, Peng G (2014) Fine mapping of Rcr1 and analyses of its effect on transcriptome patterns during infection by *Plasmodiophora brassicae*. BMC Genomics 15:1166

Conesa A, Götz S, García-Gómez JM, Terol J, Talón M, Robles M (2005) Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 21:3674–3676

Dakouri A, Zhang X, Peng G, Falk KC, Gossen BD, Strelkov SE, Yu F (2018) Analysis of genome-wide variants through bulked segregant RNA sequencing reveals a novel gene for resistance to *Plasmodiophora brassicae* in *Brassica oleracea*. Sci Rep 8:17657

Dixon GR (2006) Vegetable brassicas and related crucifers. CABI

Dokkenbouchard FL, Bouchard AJ, Ippolito J, Peng G, Strelkov S, Kirkham CL, Kutcher HR (2010) Detection of *Plasmodiophora brassicae* in Saskatchewan, 2008. Canadian plant disease survey= Inventaire des maladies des plantes au Canada

Doyle J (1990) Isolation of plant DNA from fresh tissue. Focus 12: 13–15

Gossen BD, Strelkov SE, Manolii VP, Cao T, Hwang SF, Peng G, McDonald MR (2015) Spread of clubroot on canola in Canada, 2003–2014. Old pathogen, new home. Can J Plant Pathol 37:403–413

Hatakeyama K, Suwabe K, Tomita RN, Kato T, Nunome T, Fukuoka H, Matsumoto S (2013) Identification and characterization of Crr1a, a gene for resistance to clubroot disease (*Plasmodiophora brassicae* Woronin) in *Brassica rapa* L. PLoS One 8:621–626

Hatakeyama K, Niwa T, Kato T, Ohara T, Kakizaki T, Matsumoto S (2017) The tandem repeated organization of NB-LRR genes in the clubroot-resistant CRb locus in *Brassica rapa* L. Mol Gen Genomics 292:397–405

Hirai M, Harada T, Kubo N, Tsukada M, Suwabe K, Matsumoto S (2004) A novel locus for clubroot resistance in *Brassica rapa* and its linkage markers. Theor Appl Genet 108:639–643

Hu X, Peng C (2016) A new solution to the problem of clubroot in rapeseed: breeding a new resistant variety. Seed Technol 34: 25–26

Huang Z, Peng G, Liu X, Deora A, Falk KC, Gossen BD, Mcdonald MR, Yu F (2017) Fine mapping of a clubroot

resistance gene in Chinese cabbage using SNP markers identified from bulked segregant RNA sequencing. Front Plant Sci 8:1448

Jatayev S, Kurishbayev A, Zotova L, Khasanova G, Serikbay D, Zhubatkanov A, Botayeva M, Zhumalin A, Turbekova A, Soole K (2017) Advantages of amplifluor-like SNP markers over KASP in plant genotyping. BMC Plant Biol 17:254

Kato T, Hatakeyama K, Fukino N, Matsumoto S (2013) Fine mapping of the clubroot resistance gene CRb and development of a useful selectable marker in *Brassica rapa*. Breed Sci 63:116–124

Liu S, Yeh CT, Tang HM, Dan N, Schnable PS (2012) Gene mapping via Bulked Segregant RNA-Seq (BSR-Seq). PLoS One 7:e36406

Liu Y, Xu A, Liang F, Yao X, Wang Y, Liu X, Zhang Y, Dalelhan J, Zhang B, Qin M, Huang Z, Shaolin L (2018) Screening of clubroot-resistant varieties and transfer of clubroot resistance genes to *Brassica napus* using distant hybridization. Breed Sci 68:258–267

Matsumoto E, Yasui C, Ohi M, Tsukada M (1998) Linkage analysis of RFLP markers for clubroot resistance and pigmentation in Chinese cabbage (*Brassica rapa* ssp. *pekinensis*). Euphytica 104:79–86

Matsumoto E, Hayashida N, Sakamoto K, Ohi M (2005) Behavior of DNA markers linked to a clubroot resistance gene in segregating populations of Chinese cabbage (*Brassica rapa ssp. pekinensis*). J Japan. Soc Hort Sci 74:367–373

Ooijen JWV (2006) JoinMap® 4.0: software for the calculation of genetic linkage maps in experimental population. Kyazma BV, Wageningen.

Pang W, Fu P, Li X, Zhan Z, Yu S, Piao Z (2018) Identification and mapping of the clubroot resistance gene CRd in Chinese cabbage (*Brassica rapa* ssp. *pekinensis*). Front. Plant Sci 9:653

Peng G, Falk KC, Gugel RK, Franke C, Yu FQ, James B, Strelkov SE, Hwang SF, McGregor L (2014) Sources of resistance to *Plasmodiophora brassicae* (clubroot) pathotypes virulent on canola. Can J Plant Pathol 36:89–99

Piao ZY, Deng YQ, Choi SR, Park YJ, Lim YP (2004) SCAR and CAPS mapping of CRb , a gene conferring resistance to *Plasmodiophora brassicae* in Chinese cabbage (*Brassica rapa ssp. pekinensis*). Theor Appl Genet 108:1458–1465

Piao ZY, Ramchiary N, Lim YP (2009) Genetics of clubroot resistance in brassica species. J Plant Growth Regul 28: 252–264

Sakamoto K, Saito A, Hayashida N, Taguchi G, Matsumoto E (2008) Mapping of isolate-specific QTLs for clubroot resistance in Chinese cabbage (*Brassica rapa* L. ssp. *pekinensis*). Theor Appl Genet 117:759–767

Sokal RR, Rohlf FJ (1995) Biometry: the principles and practice of statistics in biological research, 3rd edn. WH Freeman and Company, New York

Strelkov SE, Tewari JP, Smith-Degenhardt E (2006) Characterization of Plasmodiophora brassicae populations from Alberta, Canada. Can J Plant Pathol 28:467–474

Strelkov SE, Manolii VP, Cao T, Xue S, Hwang SF (2007) Pathotype classification of *Plasmodiophora brassicae* and its occurrence in *Brassica napus* in Alberta, Canada. J Phytopathol. 155:706–712

Suwabe K, Tsukazaki H, Iketani H, Hatakeyama K, Fujimura M, Nunome T, Fukuoka H, Matsumoto S, Hirai M (2003) Identification of two loci for resistance to clubroot (*Plasmodiophora brassicae* Woronin) in *Brassica rapa* L. Theor Appl Genet 107:997–1002

Suwabe K, Tsukazaki H, Iketani H, Hatakeyama K, Kondo M, Fujimura M, Nunome T, Fukuoka H, Hirai M, Matsumoto S (2006) Simple sequence repeat-based comparative genomics between *Brassica rapa* and *Arabidopsis thaliana*: the genetic origin of clubroot resistance. Genetics 173:309–319

Ueno H, Matsumoto E, Aruga D, Kitagawa S, Matsumura H, Hayashida N (2012) Molecular characterization of the CRa gene conferring clubroot resistance in *Brassica rapa*. Plant Mol Biol 80:621–629

Yu F, Zhang X, Huang Z, Chu M, Song T, Falk KC, Deora A, Chen Q, Zhang Y, McGregor L, Gossen BD, McDonald MR, Peng G (2016) Identification of genome-wide variants and discovery of variants associated with *Brassica rapa* clubroot resistance gene Rcr1 through bulked segregant RNA sequencing. PLoS One 11:e0153218

Yu F, Zhang X, Peng G, Falk KC, Strelkov SE, Gossen BD (2017) Genotyping-by-sequencing reveals three loci for clubroot resistance to six pathotypes of *Plasmodiophora brassicae* in *Brassica rapa*. Sci Rep 7:4516

Zhang T, Zhao Z, Zhang C, Pang W, Su RC, Yong PL, Piao Z (2014) Fine genetic and physical mapping of the CRb gene conferring resistance to clubroot disease in *Brassica rapa*. Mol Breed 34:1173–1183