Plant Science xxx (2015) xxx–xxx

Contents lists available at ScienceDirect

Plant Science

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

Genomic-assisted phylogenetic analysis and marker development for next generation soybean cyst nematode resistance breeding

Suhas Kadam[a,*], Tri D. Vuong[a], Dan Qiu[a], Clinton G. Meinhardt[a], Li Song[a], Rupesh Deshmukh[a], Gunvant Patil[a], Jinrong Wan[a], Babu Valliyodan[a], Andrew M. Scaboo[a], J.Grover Shannon[b], Henry T. Nguyen[a,*]

[a] Division of Plant Sciences and National Center for Soybean Biotechnology (NCSB), University of Missouri, Columbia, MO 65211, USA
[b] Division of Plant Sciences and NCSB, University of Missouri, Portageville, MO 63873, USA

ARTICLE INFO

Article history:
Received 2 June 2015
Received in revised form 20 August 2015
Accepted 22 August 2015
Available online xxx

Keywords:
Soybean
KASPar assay
Copy Number Variation
Digital-PCR assay
Phylogenetic diversity
SCN resistance genes/QTL
SNP genotyping

ABSTRACT

Soybean cyst nematode (SCN, *Heterodera glycines* Ichinohe) is a serious soybean pest. The use of resistant cultivars is an effective approach for preventing yield loss. In this study, 19,652 publicly available soybean accessions that were previously genotyped with the SoySNP50K iSelect BeadChip were used to evaluate the phylogenetic diversity of SCN resistance genes *Rhg1* and *Rhg4* in an attempt to identify novel sources of resistance. The sequence information of soybean lines was utilized to develop KASPar (KBioscience Competitive Allele-Specific PCR) assays from single nucleotide polymorphisms (SNPs) of *Rhg1*, *Rhg4*, and other novel quantitative trait loci (QTL). These markers were used to genotype a diverse set of 95 soybean germplasm lines and three recombinant inbred line (RIL) populations. SNP markers from the *Rhg1* gene were able to differentiate copy number variation (CNV), such as resistant-high copy (PI 88788-type), low copy (Peking-type), and susceptible-single copy (Williams 82) numbers. Similarly, markers for the *Rhg4* gene were able to detect Peking-type (resistance) genotypes. The phylogenetic information of SCN resistance loci from a large set of soybean accessions and the gene/QTL specific markers that were developed in this study will accelerate SCN resistance breeding programs.

1. Introduction

Soybean cyst nematode (SCN, *Heterodera glycines* Ichinohe) is the most economically important soybean [*Glycine max* (L.) Merr.] pathogen in the United States because it causes more yield loss than any other disease [1]. Management of this pest is limited to crop rotation and the use of resistant cultivars. Soybean resistance against SCN is complex because of the involvement of more than one gene and the structural changes resulting from copy number variation (CNV). To date, two major genes, *Rhg1* and *Rhg4*, and many quantitative trait loci (QTL) harboring minor genes have been identified for SCN resistance [2–7]. Three copy number classes of the *Rhg1* gene have been observed in soybean; these are generally categorized as the Plant Introduction (PI) 88788-type, carrying the highest number of copies (>6 copies); Peking-type (2 to 4 copies); and Williams 82-type (single copy) [8]. The highest magnitude of resistance has been observed within PI 88788, which has approximately nine copies [2]. In the case of *Rhg4*, the highest level of resistance is observed for the Peking-type allele because of non-synonymous variation [3]. These major genes are valuable sources of resistance and have been widely employed for the development of soybean cultivars. However, major gene resistance is not often durable, and there is an immense need for the effective utilization of more horizontal (quantitative) resistance that is derived from minor genes or QTL. Recently, Vuong et al. [4] identified a novel QTL on chromosome (Chr.) 10 (qSCN10) from an exotic accession, PI 567516C [9]. Interestingly, PI 567516C is SCN resistant and lacks the two major genes, *Rhg1* and *Rhg4* [4]. There are several other minor QTL that have been reported for SCN resistance. Among these, QTL on Chr. 11 (qSCN11) has been consistently identified from PI 437654, PI 90763, and PI 404198B [5,10].

A total of 19,652 cultivated soybean (*Glycine max*) and wild soybean (*Glycine soja* Sieb. and Zucc) accessions have been characterized using the SoySNP50K iSelect BeadChip [11]. These valuable

Abbreviations: SCN, Soybean cyst nematode; QTL, quantitative trait loci; CNV, copy number variation; SNP, single nucleotide polymorphism.

* Corresponding author.
E-mail addresses: kadamsu@missouri.edu (S. Kadam), vuongt@missouri.edu (T.D. Vuong), qiud@missouri.edu (D. Qiu), meinhardtc@missouri.edu (C.G. Meinhardt), songli@missouri.edu (L. Song), deshmukhr@missouri.edu (R. Deshmukh), patilg@missouri.edu (G. Patil), wanj@missouri.edu (J. Wan), valliyodanb@missouri.edu (B. Valliyodan), scabooa@missouri.edu (A.M. Scaboo), shannong@missouri.edu (J.Grover Shannon), nguyenhenry@missouri.edu (H.T. Nguyen).

http://dx.doi.org/10.1016/j.plantsci.2015.08.015
0168-9452/

resources can be used to find existing genetic variations in soybean germplasm. Lee et al. [12] analyzed the soybean population using SoySNP50K for the 1.5-Mb region that is centered on *Rhg1*. The identification of molecular markers that are associated with disease resistance would be helpful for developing resistant varieties because phenotyping soybean for SCN resistance is time-consuming and costly. Among the molecular markers, single nucleotide polymorphisms (SNPs) present several advantages over other genetic marker types [13]. At present, only a few simple sequence repeat (SSR) markers have been reported for the identification of the *Rhg1* gene. For instance, the marker Satt309 has been widely employed in several soybean breeding programs [14,15]. However, it has limited use because the Satt309 marker is able to detect only three copies of the *Rhg1* gene. SSR markers are not practical for a high level of multiplexing. High-throughput genotyping and the development of multiplex marker panels are more feasible with SNP markers. Some studies have reported SNP markers for the *Rhg1* and *Rhg4* genes [16–18]. Several cost effective platforms are available for SNP identification and subsequent genotyping [19,20]. Recently, next-generation KASPar assays (KBiosience, Hoddesdon, UK) have become the new SNP genotyping method because they have high-throughput, low error rates, and are cost-effective [21].

Cook et al. [2] showed that the soybean cultivar Fayette, that was developed from the *Rhg1* resistant source PI 88788 has ten copies of the *Rhg1* gene compared to nine copies in PI 88788. Recent medical science reports highlighted how CNV can be enumerated using digital polymerase chain reaction (dPCR) [22,23]. The determination of CNV with dPCR is costly and time-consuming and is not suitable for plant breeding applications; however, it is feasible to identify a haplotype that represents a particular CNV that can then be utilized for marker development. Such an assay will help breeders to select for high copy numbers of the *Rhg1* gene.

In this study, over 19,000 soybean accessions from the USDA Germplasm Collection were utilized for phylogenetic analysis of the major SCN resistant loci (*Rhg1* and *Rhg4*) using the SoySNP50K molecular marker data. The development of a panel of breeder-friendly genetic markers representing the major genes (*Rhg1* and *Rhg4*) and QTL (qSCN10 and qSCN11) is also reported. The phylogenetic information of SCN resistance loci in a large set of soybean accessions and gene/QTL specific marker resources that were developed in this study will be helpful in accelerating SCN resistance breeding in soybean.

2. Materials and methods

2.1. Phylogenetic tree of soybean accessions

A complete data set of 19,652 *G. max* and *G. soja* accessions that were genotyped with 52,041 SNPs was downloaded from the Soybase website [11, www.soybase.org]. The SNP information from the 0.5-Mb region flanking the *Rhg1* and *Rhg4* loci was selected to analyze phylogenetic diversity. The 0.5-Mb region flanking the genes was selected for three reasons: (1) to ensure a sufficient number of SNPs for phylogenetic analysis because there is no SNP present in the 50K data set located in the *Rhg4* or *Rhg1* genes; (2) Sonah et al. [24] and Lam et al. [25] suggested that there is up to a 1 Mb linkage disequilibrium decay in soybean; and (3) as reported in many previous studies that in the evolution major genes always carry the linked flanking region. To develop the phylogenetic tree, a single representative line from each haplotype was considered. The maximum likelihood (ML) model implemented in the MEGA6 (MEGA Inc., Englewood, NJ) program was used to construct the phylogenetic tree [26].

2.2. Plant materials

Several recombinant inbred line (RIL) populations, parental lines, and diverse PI lines, including known sources of SCN resistance, were used to identify and validate genetic markers. A subset of 95 diverse germplasm lines (64 soybean PIs, eight cultivars, and 23 elite breeding lines) and three RIL populations (Pana x PI 567690, Essex x PI 567690, Essex x PI 437654, and Magellan x PI 567305) were used for marker validation. Among the RIL populations, a first subset of 92 RILs that were derived from a Pana x PI 567690 cross, was used to validate markers linked to the *Rhg1* gene. A second subset of 44 RILs, from a Essex x PI 437654 cross, was used to test *Rhg1*, *Rhg4* and the minor QTL qSCN11 [5]. A third subset of 92 RILs, from a Magellan x PI 567305 cross, was used to validate markers linked to the novel QTL qSCN10.

2.3. Phenotyping

A subset of 95 soybean lines and three different RIL populations (Pana x PI 567690, Essex x PI 567690, Essex x PI 437654, and Magellan x PI 567305) were evaluated for SCN resistance following a well-established greenhouse bioassay [27,28] at the University of Missouri, Columbia, Missouri. Briefly, five plants from each test line, indicator lines for HG Type test (PI 548402, PI 88788, PI 90763, PI 437654, PI 209332, PI 89772, PI 548316), and susceptible checks (cv. Lee 74 and cv. Hutcheson) [29] were arranged in a randomized complete block design. Two days after transplantation, seedlings were inoculated with 2000 ± 25 eggs from near-homogenous SCN isolates (HG Types 2.5.7 (PA1), 1.2.5.7 (PA2), 0 (PA3), 2.5.7 (PA5), or 1.3.6.7 (PA14)). The experiments were maintained at $27 \pm 1\,^{\circ}C$ and were watered daily. Thirty days post-inoculation, nematode cysts were washed from the roots of each plant and were counted using a fluorescence-based imaging system [30]. The Female Index (FI) was calculated to evaluate the response of each line to SCN using the following formula: FI = (average number of female cyst nematodes on a test soybean line/average number of female nematodes on the susceptible check) $\times$ 100.

2.4. SNP identification and development of KASPar genotyping assays

For developing a KASPar assay, sequence information of *Rhg1*, *Rhg4*, the novel QTL qSCN10, and the minor QTL qSCN11 was retrieved from the available sources for diverse soybean lines [8,31]. The SNPs identified at these loci using GATK (Genome Analysis Toolkit, www.broadinstitute.org) software [32] were reconfirmed by examining read alignments in the Integrative Genomics Viewer (IGV) tool [33]. Four SNPs of *Rhg1* and three SNPs of *Rhg4* were selected. Similarly, four SNPs were selected from each of the qSCN10 and qSCN11 QTL regions. The selected SNP set was targeted for the development of KASPar assays (Table 1). Two allele-specific forward primers, along with tail sequences and one common reverse primer, were synthesized for SNP genotyping assays. The reaction mixture was prepared according to the protocol described by KBiosciences (Herts, UK) (http://www.ksre.ksu.edu/igenomics). The following cycling conditions were used: 15 min at 95 °C, followed by 10 touchdown cycles of 20 s at 94 °C, 1 min at 65–57 °C (dropping 0.8 °C per cycle) and then 23 cycles of 20 s at 94 °C, 1 min at 57 °C. The fluorescent end-point genotyping method was carried out using a Roche LightCycler (LC) 480 instrument (Roche Applied Sciences, Indianapolis, IN, USA).

2.5. Development of a Taqman assay and CNV determination

The conserved sequences of the *Rhg1* gene were selected to develop a Taqman assay to run on dPCR (Table 2). The FAM™ dye-

Please cite this article in press as: S. Kadam, et al., Genomic-assisted phylogenetic analysis and marker development for next generation soybean cyst nematode resistance breeding, Plant Sci. (2015), http://dx.doi.org/10.1016/j.plantsci.2015.08.015