Physical mapping of resistant and susceptible soybean genomes near the soybean cyst nematode resistance gene \( Rhg_4 \)


Abstract: The soybean cyst nematode (SCN), \( Heterodera glycines \) Ichinohe, is the foremost pest of soybean (\( Glycine \) max L. Merr.). The \( rhg_1 \) allele on linkage group (LG) G and the \( Rhg_4 \) allele on LG A2 are important in conditioning resistance. Markers closely linked to the \( Rhg_4 \) locus were used previously to screen a library of bacterial artificial chromosome (BAC) clones from susceptible ‘Williams 82’ and identified a single 150-kb BAC, Gm_Isb001_056_G02 (56G2). End-sequenced subclones positioned onto a restriction map provided landmarks for identifying the corresponding region from a BAC library from accession PI 437654 with broad resistance to SCN. Seventy-three PI 437654 BACs were assigned to contigs based upon \( HindIII \) restriction fragment profiles. Four contigs represented the PI 437654 counterpart of the ‘Williams 82’ BAC, with PCR assays connecting these contigs. Some of the markers on the PI 437654 contigs are separated by a greater physical distance than in the ‘Williams 82’ BAC and some primers amplify bands from BACs in the mid-portion of the connected PI 437654 BAC contigs that are not amplified from the ‘Williams 82’ BAC. These observations suggest that there is an insertion in the PI 437654 genome relative to the ‘Williams 82’ genome in the \( Rhg_4 \) region.

Key words: BAC, deletion, insertion, resistance gene, soybean cyst nematode.

Résumé: Le nématode à kystes des racines du soja (SCN), \( Heterodera glycines \) Ichinohe, est le principal ravageur du soja (\( Glycine \) max L. Merr.). L’allèle \( rhg_1 \) sur le groupe de liaison G et l’allèle \( Rhg_4 \) sur le groupe de liaison A2 contribuent de façon importante à la résistance. Des marqueurs étroitement liés au locus \( Rhg_4 \) avaient été employés précédemment pour cibler une banque de clones BAC (chromosomes bactériens artificiels) préparée à partir du cultivar sensible ‘Williams 82’ et avaient permis d’identifier un seul BAC de 150 kb, Gm_Isb001_056_G02 (56G2). Des sous-clones provenant des extrémités séquencées ont été placés sur une carte et ont ainsi fourni des repères pour identifier la région correspondante dans une banque de clones BAC préparée à partir de l’ADN de l’accession PI 437654, une lignée qui présente une large résistance face au SCN. Soixante-treize clones BAC parmi la banque PI 437654 ont été assemblés en contigs sur la base de leur profil de restriction à l’aide de l’enzyme \( HindIII \). Quatre contigs PI 437654 représentaient l’équivalent du BAC provenant de Williams 82 et des analyses PCR permettaient de relier ces contigs. La distance physique entre certains des marqueurs situés sur les contigs PI 437654 était plus grande qu’entre ces mêmes marqueurs chez le clone issu de Williams 82. De plus, certaines amorces produisaient des amplicons à partir de BAC situés au centre des contigs PI 437654 qui étaient liés, mais ces mêmes amorces n’amplifiaient pas une région correspondante chez le clone BAC de Williams 82. Ces observations suggèrent qu’il y a une insertion dans la région de \( Rhg_4 \) chez génome du PI 437654 par rapport au génome de ‘Williams 82’.

Mots clés: BAC, délétion, insertion, gène de résistance, nématode à kystes des racines du soja.

Introduction

The soybean cyst nematode (SCN), \( Heterodera glycines \) Ichinohe, is the foremost pest of soybean (\( Glycine \) max L. Merr.) and is responsible for the loss of more of the annual soybean crop than all other pests combined (http://aes.missouri.edu/delta/research/soyloss.stm). As an example, the loss to the 97.4 million t (metric ton) (2763 million...
Thus far, the most practical management strategy has been the use of resistant cultivars. Several sources of resistance have been reported and utilized in cultivar development (Riggs et al. 1998). Several soybean genes confer resistance to SCN (Riggs and Schmitt 1991; Caldwell et al. 1960). The Rhg4 resistance allele on linkage group (LG) G and the Rhg4 resistance allele on LG A2 of the soybean genetic map are especially important (Webb et al. 1995).

Tight genetic linkage of 0.35 cM was first reported between SCN resistance at the Rhg4 locus and black seed coat at the I (inhibitor) locus (Matson and Williams 1965). A deletion of a portion of the promoter of a gene encoding chalcone synthase (CHS4) allows the production of black pigment in the seed coat (Todd and Vodkin 1996). CHS4 resides in a cluster of three genes encoding CHS isoforms, CHS1, CHS3, and CHS4 (Akada and Dube 1995). Primers were designed from the sequence of the region between the genes encoding CHS3 and CHS4 for use in a PCR assay.

Also closely linked to the I locus, and hence to the Rhg4 locus, is the restriction fragment length polymorphism (RFLP) molecular marker pBLT65 (Weiseman et al. 1992; Webb et al. 1995). Marker pBLT65 (GenBank accession No. AF049706) is a partial cDNA clone from a gene encoding the bifunctional enzyme aspartokinase – homoserine dehydrogenase (AK–HSDH) in soybean cultivar Century (Gebhardt et al. 1999). Primers 548 and 563 were designed from a portion of a genomic clone from Century of a gene encoding AK–HSDH (GenBank accession No. AF049708) for a PCR assay to use in marker-assisted selection for cultivar development (Heer et al. 1998; Matthews et al. 1998). It was demonstrated that at least two loci defined by these primers mapped near the I locus in crosses of SCN-resistant genotypes with susceptible genotypes No1 (PI 290136) and BARC-2 (PI 547895) (Matthews et al. 1998).

Primers 548 and 563 were used in PCR assays to screen a bacterial artificial chromosome (BAC) library of genomic clones of ‘Williams 82’ (Marek and Shoemaker 1997) to isolate BACs associated with the Rhg4 locus (Lewers et al. in press). BACs identified by this assay were screened with primers designed from the sequence of the region between the genes encoding CHS3 and CHS4. Only 150-kb BAC Gm_1S8001_056_G02 (56G2) was positive for both assays. Simple sequence repeat markers (SSRs; Sat157 and Sat162) were developed from BAC 56G2 and cosegregated with the I locus in a population of 240 recombinant inbred lines (RILs) (Cregan et al. 1999). Several EcoRI subclones located to Apol fragments of a restriction map of BAC 56G2 were end sequenced to provide additional landmarks for use in identifying the same region in a resistant soybean genome (Lewers et al. in press). A library of BAC genomic clones from the soybean accession PI 437654 was made (Tomkins et al. 1999) because of broad resistance to SCN in PI 437654, and because PI 437654 was used to genetically map resistance to SCN as a quantitatively inherited trait (Webb et al. 1995). The objective of this research is to use the landmarks derived from ‘Williams 82’ BAC 56G2 to identify BACs in the PI 437654 library that may contain the Rhg4 gene.

Materials and methods

The BAC library used in this work consists of 73 728 clones of genomic DNA from PI 437654 and represents a nine haploid genome equivalent, resulting in a 99% probability of recovering any targeted sequence (Tomkins et al. 1999). The PI 437654 BAC library is maintained at Clemson University Genomics Institute (www.genome.clemson.edu). The library was arrayed on membranes to facilitate hybridization with probes. Landmarks derived from ‘Williams 82’ BAC 56G2 were used as probes to screen the library to identify a similar region in PI 437654. DNA probes were made from amplification products of PCRs using genomic DNA from PI 437654 as template rather than using subclones from the ‘Williams 82’ BAC because of possible sequence divergence between ‘Williams 82’ and PI 437654. PCR products were made using several primer pairs (Table 1) including those used to identify ‘Williams 82’ BAC 56G2 and primers designed from end sequences of ‘Williams 82’ BAC 56G2 EcoRI subclones 915, 916, and 917. Also, because several genes encoding CHS isoforms are located near the I locus (Todd and Vodkin 1996) and in BAC 56G2 (Lewers et al. in press), primers for a PCR assay were designed from exon II of a consensus sequence of the genes encoding CHS1, CHS2, CHS3, CHS4, CHS5, CHS6, and CHS7 to identify all PI 437654 BACs containing genes encoding CHS.

PCR components included 40 ng of PI 437654 genomic DNA, 1× PCR buffer, 2.5 mM MgC12, 0.1 mM dNTPs, 2 pmol of each primer, and 1 U Taq recombinant polymerase in a volume of 25 μL. PCRs were subjected to 1 min at 94°C, 30 cycles of 30 s at 94°C, 30 s at 57°C, and 30 s at 72°C; followed by 4 min at 72°C and storage at 4°C. PCR products were separated from other reaction components by electrophoresis through an agarose gel (1 g low melting point (LMP) agarase – 100 mL of 1× Tris-acetate–EDTA (TAE) at 100 cV and 4°C and visualized using ethidium bromide staining techniques. The piece of gel containing the PCR band was cut out and melted in a microcentrifuge tube at 65°C. One microlitre of the melted gel was used in a second PCR in a volume of 100 μL with the same primers used to generate the template band, and thermocycling conditions were as described previously. PCR products were quantitated by comparison with Low DNA Mass™ Ladder (Life Technologies Inc., Gaithersburg, Md.) on a 1 g agarose/100 mL of 1× Tris-borate–EDTA (TBE) gel before being used as template in random priming reactions to make probes. Random priming reaction, hybridization, exposure, and visualization conditions were as described in Tomkins et al. (1999).

The PI 437654 BAC DNA, restricted with NcoI, was electrophoresed in a 1 g agarose/100 mL 0.5× TBE gel via clamped homogenous electric field (CHEF) for 14 h at 6.0 V/cm and 14°C. In Southern analysis, CHEF filters were hybridized with DNA probes to confirm BAC selection by that probe on the BAC library filters. CHEF filters were also hybridized with sonicated total genomic DNA probe for characterization of repetitive DNA content as previously described in Hanson et al. (1995).

PI 437654 BACs identified by the probes derived from ‘Williams 82’ BAC 56G2 were analyzed by HindIII restriction fragment pattern similarity and assigned to contigs. DNA was isolated from the selected BACs using a QIAprep 8 Miniprep Kit (Qiagen, Valencia, Calif.). BAC DNA was digested to completion overnight with restriction enzyme HindIII and electrophotothermally separated on a 1 g Fisher Biotech agarose/100 mL of 1× TAE gel at 60 cV and 12°C for 16 h. Hi-Lo marker DNA (Minnesota Molecular,
of molecular weight. Gels were stained with SYBR Gold nucleic acid gel stain (Molecular Probes, Eugene, Oreg.). Separated HindIII fragments were scanned by a Bio-Rad Fluor-S Multi-imager (Bio-Rad Laboratories, Hercules, Calif.) and assigned a molecular weight using Image software (available at http://www.sanger.ac.uk/Software/Image). HindIII restriction fragment patterns were compared using FPC v.4.5 (Sonderlund et al. 1999; available at http://www.genome.clemson.edu/fpc/) for unambiguous assignments of BACs into contigs. BACs were grouped into contigs according to stringency levels ranging from $e^{-5}$ through $e^{-12}$ (probability of a false association between BACS is from 10^{-5} through 10^{-12}).

To identify the contigs most likely to contain the Rhs$_G$ gene, a subset of 47 PI 437654 BACs was selected for further analysis. Selection was based on fingerprinting and hybridization of probes derived from the ‘Williams 82’ BAC DNA as template. Probes were amplified from genomic PI 437654 probe library. The probes were end sequenced using the ABI Prism dRhodamine Terminator Cycle Sequencing Kit (Perkin-Elmer, Applied Biosystems, Foster City, Calif.). Sequencing primers T7 (5'-TGTAAATCCACCTTGGA-3') and SP6 (5'-GCCTAGTTGACACTTAAAT-3') were derived from the sequence of the pBAC vector (Shizuya et al. 1992).

Primers were designed from the BAC-end sequences (Table 1) and used in duplicated PCRs using 10^{-1} dilutions of the DNA of ‘Williams 82’ BAC 56G2 and the subset of 47 PI 437654 BACs as templates. PCRs contained 1 µL template in a 20-µL volume of reaction components as described previously for non-SSR PCRs. Amounts of template used were determined empirically. Thermocycling and visualization procedures were also as described previously.

### Results

The landmarks associated with BAC 56G2 from the SCN-susceptible ‘Williams 82’ identified 87 BACs from the BAC library of 73,728 clones of the SCN-resistant PI 437654. Eighteen of these BACs were identified by hybridization with the probe 812 derived from the PCR product amplified from genomic PI 437654 using primers 548 and 563 (Table 1), similar to RFLP molecular marker pBLT65 (AK-HSDH). Several of these BACs hybridized with the probes derived from ‘Williams 82’ BAC 56G2 EcoRI subclones 915 and 917. Fifty-three of the 87 BACs were identified by hybridization with the probe 864 (derived from PCR product amplified from genomic PI 437654) using primers 849 and 850 (Table 1) (designed

### Table 1. Primers used in PCRs to assay soybean genomic BAC clones or to generate DNA for labeling as probes for hybridization.

<table>
<thead>
<tr>
<th>Primer designation</th>
<th>Primer</th>
<th>Primer sequence, 5'-3'</th>
<th>Probe</th>
</tr>
</thead>
<tbody>
<tr>
<td>AK–HSDH F (pBLT65)</td>
<td>548</td>
<td>GCAGATACACAGTGTGCAGAC</td>
<td>812</td>
</tr>
<tr>
<td>AK–HSDH R (pBLT65)</td>
<td>563</td>
<td>GGATGGACGTCGTAAGCC</td>
<td>812</td>
</tr>
<tr>
<td>AK–HSDH R</td>
<td>803</td>
<td>TGGAGACAGCTGCACTGAGAG</td>
<td>812</td>
</tr>
<tr>
<td>CHS3–CHS4 intergenic F</td>
<td>849</td>
<td>ATGTTCCTTATACCAAGACTCT</td>
<td>864</td>
</tr>
<tr>
<td>CHS3–CHS4 intergenic R</td>
<td>850</td>
<td>GTGTTGCCGATTACTTACAGCG</td>
<td>864</td>
</tr>
<tr>
<td>EcoRI subclone 915 F</td>
<td>1057</td>
<td>CAAACGTAGTCGCACTCCTCT</td>
<td>915</td>
</tr>
<tr>
<td>EcoRI subclone 915 R</td>
<td>1058</td>
<td>CACCGAGCAAGAAAGCGTTGC</td>
<td>915</td>
</tr>
<tr>
<td>EcoRI subclone 916 F</td>
<td>1042</td>
<td>CTACGGTGCAAGCAATCC</td>
<td>917</td>
</tr>
<tr>
<td>EcoRI subclone 916 R</td>
<td>1043</td>
<td>GTTCCACATGAGAACAGAG</td>
<td>916</td>
</tr>
<tr>
<td>EcoRI subclone 917 F</td>
<td>1048</td>
<td>CAGTGATGTGTTGTTGCGTGCT</td>
<td>917</td>
</tr>
<tr>
<td>EcoRI subclone 917 R</td>
<td>1047</td>
<td>GATGGCTGACAGAGGGTCG</td>
<td>917</td>
</tr>
<tr>
<td>CHS exon II, nonspecific F</td>
<td>1051</td>
<td>CGTGAGATTGCTTGCTGCGT</td>
<td>CHSn</td>
</tr>
<tr>
<td>CHS exon II, nonspecific R</td>
<td>1052</td>
<td>CCACCCGGGTGGAATTCGCAATCC</td>
<td>CHSn</td>
</tr>
<tr>
<td>EcoRI subclone 1006 F</td>
<td>1078</td>
<td>GCCAGAGCAGCACTGCGT</td>
<td>864</td>
</tr>
<tr>
<td>EcoRI subclone 1006 R</td>
<td>1079</td>
<td>GCGTTGTGACTGTTGCTCGT</td>
<td>864</td>
</tr>
<tr>
<td>BAC 35K19 T7 F</td>
<td>1095</td>
<td>CGCACTGTCTAAGCCTCAAG</td>
<td>864</td>
</tr>
<tr>
<td>BAC 35K19 T7 R</td>
<td>1096</td>
<td>CGCGGATATGCAGGCGGAGGA</td>
<td>864</td>
</tr>
<tr>
<td>BAC 76P24 T7 F</td>
<td>1097</td>
<td>GCCGCTAGCTGCTTCTCTAA</td>
<td>864</td>
</tr>
<tr>
<td>BAC 76P24 T7 R</td>
<td>1098</td>
<td>CGCGGAGTACCTCTCTCAGAT</td>
<td>864</td>
</tr>
</tbody>
</table>

Note: EcoRI subclones were derived from ‘Williams 82’ BAC 56G2. Primers were used in PCRs with BACs and genomic DNA as template. Probes were amplified from PI 437654 genomic DNA labeled, and hybridized to the PI 437654 BAC library.
from sequence of the region between genes encoding CHS3 and CHS4), and 10 of these also hybridized with probe CHSn. Although the ‘Williams 82’ BAC library contained a single BAC that was positive for the PCR assay with primers 548 and 563, and also for a second PCR assay with primers 849 and 850, no single BAC in the PI 437654 BAC library hybridized with both probes derived from the two sets of primers. This result differentiates the PI 437654 genomic library from that of ‘Williams 82’.

Complete HindIII digests of the 87 BACs were separated by gel electrophoresis and analyzed by FPC software to identify similarities among the BAC restriction fragment banding patterns. In initial analysis of the HindIII restriction fragment banding patterns, 14 BACs were eliminated because of insufficient overlap scores. At tolerance 7 and cut-off $e^{-5}$, 73 BACs were assembled into eight contigs. Probes 812, 864, 915, 917, and CHSn (Table 1) hybridized to the CHEF-gel separated BAC NotI restriction digests transferred to nylon filters by Southern hybridization, confirming that these clones were related. Based on similarities among HindIII restriction fragment profiles, the BACs were grouped into eight contigs at the lower stringency level of $e^{-5}$ and 14 contigs at the higher stringency level of $e^{-12}$. (Fig. 1). At the lower stringency level of $e^{-5}$ the ‘Williams 82’ BAC 56G2 was included in a large contig that separated into two contigs at the next higher stringency level of $e^{-6}$. BAC 56G2 remained with one of these contigs until analysis at the much higher stringency level of $e^{-10}$ separated the contig into three contigs. These three contigs remained intact at stringency level $e^{-12}$ and were designated contigs 4, 5, and 6 (Fig. 1). In addition, contig 4 contained BACs hybridizing to probe 812, derived from a gene encoding AK–HSDH, and to probes 915 and 917, derived from ‘Williams 82’ BAC 56G2 subclones 915 and 917. This association suggested that this contig represented at least part of the PI 437654 genomic counterpart to ‘Williams 82’ BAC 56G2. However, at all but the lowest stringency level of $e^{-5}$, the PI 437654 contig containing the ‘Williams 82’ BAC was not joined with any contig containing BACs hybridizing strongly to probe 864 derived from the region between genes encoding CHS3 and CHS4.

Distribution of repetitive sequences in this region seems to be restricted to specific BAC clones (Fig. 1). In contigs 4 and 6, BACs did not give a strong signal when hybridized with total genomic DNA, suggesting that they may have low content of repetitive DNA that can interfere with attempts at chromosome walking. Similarly, BACs included in contig 13 gave low hybridization signal with the total genomic DNA probe. Contig 5, which also originated at $e^{-9}$ from the PI 437654 contig containing ‘Williams 82’ BAC, had moderate levels of repetitive DNA content.

Primers 849 and 850 derived from the region between genes encoding CHS3 and CHS4 were very specific and amplified product from only four BACs, all of which were grouped in a single contig (contig 13). Contig 13 was assembled at $e^{-3}$ and the BAC assignment to this contig was maintained at all higher stringency levels. Therefore, PI 437654 BAC contigs 4, 5, 6, and 13 may represent portions orthologous to ‘Williams 82’ BAC 56G2.

Primers 1042 and 1043, derived from EcoRI subclone 916, amplified three bands from the ‘Williams 82’ BAC 56G2 and from most of the PI 437654 BACs. An extra band was amplified from BACs 11K09 and 15G19 in contig 6 and from BAC 26M10 in contig 2. Observations of multiple bands within a BAC and the amplification of a band from most of the BACs implicate multiple primer sites in the genome that could represent sequence duplications.

Because SSR markers are designed to be locus specific, SSR markers derived from ‘Williams 82’ BACs (Cregan et al. 1999) identified with primers 548 and 563, amplifying a region of a gene encoding AK–HSDH, were expected to be ideal for confirming exclusively and unequivocally those contigs representing part of the PI 437654 genome orthologous to ‘Williams 82’ BAC 56G2. Indeed, Sat157 confirmed the similarity between ‘Williams 82’ BAC 56G2 and PI 437654 contigs by amplifying product from six BACs in contigs 4, 5, and 6, and not amplifying product from any other PI 437654 BACs. In addition, amplification from these six BACs confirms the joining of contigs 4, 5, and 6 (Fig. 2). These BACs were in the same contig at the $e^{-9}$ stringency level but were in separate contigs at the $e^{-10}$ level (Fig. 1).

Primers 849 and 850, no single BAC in the PI 437654 BAC library was included in a large contig that separated into two contigs at the next higher stringency level of $e^{-6}$. BAC 56G2 was included in a large contig that separated into two contigs at the next higher stringency level of $e^{-6}$. BAC 56G2 remained with one of these contigs until analysis at the much higher stringency level of $e^{-10}$ separated the contig into three contigs. These three contigs remained intact at stringency level $e^{-12}$ and were designated contigs 4, 5, and 6 (Fig. 1). In addition, contig 4 contained BACs hybridizing to probe 812, derived from a gene encoding AK–HSDH, and to probes 915 and 917, derived from ‘Williams 82’ BAC 56G2 subclones 915 and 917. This association suggested that this contig represented at least part of the PI 437654 genomic counterpart to ‘Williams 82’ BAC 56G2. However, at all but the lowest stringency level of $e^{-5}$, the PI 437654 contig containing the ‘Williams 82’ BAC was not joined with any contig containing BACs hybridizing strongly to probe 864 derived from the region between genes encoding CHS3 and CHS4.

Sat162 primers amplified a band of only one size from PI 437654 BACs, but amplified two bands from ‘Williams 82’ BAC 56G2, one of higher and one of lower molecular weight than the band amplified from the PI 437654 BACs. Interestingly, the band amplified from the PI 437654 BACs was amplified to some degree from each PI 437654 BAC in the 47-BAC subset. A strong band was amplified from BACs in contigs 4 and 5, as illustrated in Fig. 2, and a weak band was amplified from most BACs indicating multiple annealing sites with high enough sequence similarity to function as template. Although Sat162 primers were not as useful as Sat157 in determining which PI 437654 BACs were associated with this region, it was possible to determine the approximate placement of the Sat162 locus relative to other linked markers.

Primers 548 and 563, derived from a portion of a gene encoding AK–HSDH, amplify three bands of genomic DNA from either resistant or susceptible soybean genotypes (Matthews et al. 1999). A 996-bp band designated band 3 was amplified from both resistant PI 437654 and susceptible ‘Williams 82’ genomic DNA. The DNA sequence of this
band was most similar ($E$ value = 0.0) to GenBank accession No. AF049706, the sequence of pBLT65, a partial cDNA clone also from a gene encoding AK–HSDH in ‘Century’. Band 4 was also amplified from ‘Williams 82’ BAC 56G2, indicating some degree of difficulty with the primer annealing to the BAC template. Other primer pairs derived from BAC ends were uninformative.

### Discussion

The genetic distance between the $I$ locus and the locus defined by RFLP probe pBLT65, derived from a portion of a gene encoding AK–HSDH, has been reported at approximately 1 cM in a cross between resistant PI 437654 and susceptible BSR 101 (Webb et al. 1995). Because ‘Williams 82’ BAC 56G2 contains both these markers as determined by PCR (Lewers et al. in press), an estimate of the physical distance between these markers is approximately the size of the BAC (150 kb). Therefore, an estimate of the physical to genetic distance ratio in this region is approximately
Fig. 2. Alignment of soybean BACs from PI 437654 (SCN resistant) with BAC 56G2 from ‘Williams 82’ (SCN susceptible). PI 437654 BACs are drawn to scale. Estimated physical locations of BAC regions hybridizing to probes and PCR primer sites are indicated as vertical bars. Rectangles indicate sequenced BAC ends from which PCR primers were designed and used to identify genomic regions in common. Horizontal brackets over locus designations indicate loci with undetermined position relative to each other. (●) Probes hybridizing to BAC 56G2. (▲) Primers amplifying the same size product from either PI 437654 BACs or from ‘Williams 82’ BAC 56G2 (grey indicates same size but faint band amplified from the ‘Williams 82’ BAC 56G2). (○) Primers amplifying a product of different size from BAC 56G2 than from PI 437654 BACs. (□) Primers amplifying a product from PI 437654 BACs but not from BAC 56G2.

150 kb/cM. The genetic distance between I and Rhg4 is only 0.35 cM in a cross between resistant ‘Peking’ and a susceptible breeding line (S62-1050) (Matson and Williams 1965). Therefore, it can be expected that a BAC from a resistant cultivar containing the I locus should also contain the Rhg4 gene.

However, another possibility is that the Rhg4 locus is associated with an insertion–deletion event. The Rhg4 gene may be located in an insertion in the resistant genome relative to the susceptible genome, or conversely, the susceptible genome may have lost the Rhg4 gene through a deletion event. Two observations in this study indicate that this may be the case. The first is that although the ‘Williams 82’ BAC library contained a single BAC that was positive for the PCR assay with primers 548 and 563 and also for a second assay with primers 849 and 850, no single BAC in the PI 437654 BAC library hybridized with both probes derived from the two sets of primers. Instead, these two markers hybridized to BACs in contigs that remained separate in FPC analysis.

The second observation is that PCR with primers 1097 and 1098, derived from end-sequencing of BAC 76P24, did not amplify a product from the ‘Williams 82’ BAC. Also, primers 1042 and 1043, derived from EcoRI subclone 916, amplified an extra band on BACs 11K09 and 15G19 that was not amplified from the ‘Williams 82’ BAC. Although this could be explained by primer mispairing owing to small sequence changes, it should be noted that the primer sequences in the second case were derived from the ‘Williams 82’ BAC.

Several examples of resistance genes residing in donor-specific genomic regions have been reported. A 650-kb insertion from Lycopersicon peruvianum, a wild relative of cultivated tomato Lycopersicon esculentum, has been introgressed into L. esculentum to provide resistance to root knot nematode and potato aphid. The L. peruvianum insertion contains the Mi gene, conferring resistance to root knot nematode and potato aphid, and six other gene family members (Milligan et al. 1998). The Tm-2 gene from L. peruvianum confers resistance to tomato mosaic virus in otherwise susceptible L. esculentum. Random amplified polymorphic DNA (RAPD) markers genetically linked with Tm-2 were hybridized to DNA from a heterozygote using fluorescence in situ hybridization (FISH) and demonstrated that the Tm-2 gene was in a large insertion of L. peruvianum DNA (Motoyoshi et al. 1996). In sugar beet (Beta vulgaris L.) the Hs1lpro-1 gene from Beta procumbens confers resistance to beet cyst nematode. A satellite marker specific to B. procumbens in a translocation region lead to the cloning of a cDNA from B. procumbens that was determined to be transcribed from the resistance gene (Cai et al. 1997). The previous three situations involved a dominant resistance gene residing in genomic regions exclusively from the donor genotype. In contrast, the dominant allele at the Va locus of cultivated tobacco (Nicotianum tabacum) confers susceptibility to potato virus Y. Spontaneous deletions of DNA ranging from a few bases to an estimated 1 Mb are responsible for a series of resistant tobacco varieties (Noguchi et al. 1999).

A proposed model for the evolution of clusters of resistance genes includes insertion and deletion events as part of an ongoing process (Michelmore and Meyers 1998). According to the model, the process can be initiated by the duplicating movement of a transposable element to flank a resistance gene. The region is amplified by one of several possible mechanisms such as unequal crossing over between the duplicated sequences, a rolling circle mechanism, or
slippage replication. Additional amplification is then also possible intragenically. Concurrently, the duplicated resistance genes are able to diverge in expression or function without detriment to viability, leading to clusters of resistance-gene family members. The high level of duplicated sequences causes instability because of frequent unequal crossing-over and results in deletions and additional duplications.

A similar model was proposed for the evolution of heterochromatin in plant genomes (Flavell 1986). In this model, amplifications, deletions, and transpositions occur continually in repeat regions, and genes in the region evolve as a part of the surrounding genomic region. Interestingly, repeat DNA is associated with the sugar beet Hs1pro-1 gene and the tomato Tm-2 gene. Based on these two models, it is likely that the process involves genes other than resistance genes and probably occurs in much of the plant genome.

Analysis of ‘Williams 82’ BAC 56G2 provided evidence of gene duplication on both the regional and gene level. BAC 56G2 contains several duplications including genes for CHS, protein kinases, RNA-binding proteins, and possibly membrane-associated salt-inducible proteins and glucosyl transferases (Lewers et al. in press). In PI 437654 contigs 4, 5, 6, and 13, four loci associated by PCR with AK–HSDH sequence indicate that there may be four genes encoding AK–HSDH isoform in this region in PI 437654. The high level of gene duplication in this region combined with the physical distance and PCR results are suggestive of an insertion in PI 437654 relative to ‘Williams 82’ but are not conclusive. Therefore, ‘Williams 82’ BAC 56G2 and the two PI 437654 BACs that appear to contain the putative insertion have been subcloned for complete sequencing of the region for full characterization.

Acknowledgements

We gratefully acknowledge grant support to B.F. Matthews and H.T. Knap from the United Soybean Board. In addition, we thank Ms. Margaret MacDonald for help with BAC preps and Mr. Mike Reinsel for help with BAC end sequencing. The authors wish to thank Drs. P. Cregan, D. Grant, and the anonymous reviewers for their reviews of this manuscript. Technical contribution No. 4713 of the South Carolina Agricultural Experimental Station.

References


