

**G**

:

**K. T** , **V.N. N** , **M.J. I** , **S. Y** , **C. T** , **D.A. L**

**Abstract:** Candidate genes were identified for two *Qr1*2 providing resistance to the leaf scorch called soybean (*Glycine max* (L.) Merr.) sudden death syndrome (SDS) and *Qr1*1 providing resistance to root infection by the causal pathogen *Fusarium solani* f.sp. *glycines*. The  $7.5 \pm 0.5$  cM region of chromosome 18 (linkage group G) was shown to encompass a cluster of resistance loci using recombination events from 4 near-isogenic line populations and 9 DNA markers. The DNA markers anchored 9 physical map contigs (7 are shown on the soybean Gbrowse, 2 are unpublished), 45 BAC end sequences (41 in Gbrowse), and contiguous DNA sequences of 315, 127, and 110 kbp. Gene density was high at 1 gene per 7 kbp only around the already sequenced regions. Three to 4 gene-rich islands were

tally dependent (Hnetkovsky et al. 1996; Njiti et al. 1996) but has a genetic component (Stevens et al. 1993; Hnetkovsky et al. 1996) that may be related to toxin synthesis (Baker and Nemeč 1994; Jin et al. 1996) or root resistance (Njiti et al. 1997). Root resistance, measured as infection severity (IS%) when the disease is severe or infection frequency (IF%) when disease is not severe, is environmentally stable and may be more consistent and reliable for determining the location of genes conferring resistance to SDS (Njiti et al. 1997, 1998; Prabhu et al. 1999).

Soybean linkage groups G, C2, I, and N contain quantitative trait loci (QTL) for resistance to SDS (Iqbal et al. 2001; Njiti et al. 2002). Linkage group G (chromosome 18) (Zou et al. 2003) contains a cluster of loci for resistance to SDS comprising *QRfs*, *QRfs1*, *QRfs2*, and *QRfs3*. Based on DNA marker maps in Meksem et al. (1999, 2001 *b*), each is located in 2- to 5-cM intervals that are mutually linked. The genes are  $0.25 \pm 0.25$  to  $30 \pm 2.5$  cM from *rhg1*, the major gene for resistance to *Heterodera glycines* (Webb et al. 1995), the soybean cyst nematode (SCN). *QRfs* is 30 cM from *rhg1* and probably corresponds to *Rfs*, a gene identified without DNA marker association (Stevens et al. 1993; D.A. Lightfoot, unpublished). *QRfs2* may be clustered closely with *rhg1*. *Rfs2* reduces the leaf DX but the effect on root resistance, where the *Rfs1* gene has the major effect, is not clear (Meksem et al. 1999). Low map resolution means that *QRfs2* may be a separate gene linked to *rhg1* or a pleiotropic effect of the gene. One goal of this study was to dissect the gene clusters by recombination and examine the effect of *Rfs2* on root resistance.

Analysis of pachytene chromosome spreads predicts that the architecture of the soybean genome consists of 31 euchromatic regions and 21 heterochromatic regions among 20 chromosomes (Singh and Hymowitz 1988). The short arm of chromosome 18 (linkage group G) is entirely heterochromatic. Euchromatin represents 50% of the chromosome and the entire long arm. The soybean genome shows

genotypes was grouped based on the interval containing the recombination events.

#### **Field assays of resistance to SDS**

The field characterization of the small populations has been described previously (Njiti et al. 1998; Meksem et al. 1999). For the large populations, 4 fields (environments) located in southern Illinois were selected for this study based on a history of uniform SDS leaf symptom expression during 1999 and 2000. All experiments were planted in a randomized complete block design with 2 replications as described in Njiti et al. (1997, 1998). SDS disease was scored at the R6 stage, and all plots were rated for disease incidence (DI) and disease severity (DS) as described earlier (Njiti et al. 1996). DX was then calculated as  $(DI \times DS)/9$ , with a possible range from 0 (no disease) to 100 (death of all plants). IS was determined by isolating fungi from taproot slices collected between R6 and R8 (harvest maturity) with a possible range of 0-100 (Njiti et al. 1997).

#### **DNA isolation and microsatellite marker scores**

Plant leaves were collected and kept frozen at  $-70^{\circ}\text{C}$ . The genomic DNA of the plants that were used for the recombinant screening was extracted following Bell-Johnson et al. (1998) with some modifications. Approximately 50 mg of

For line mean comparisons, the data were subjected to false associations is reduced in isogenic lines (Lander and ANOVA (SAS Institute Inc., Cary, North Carolina) Botstein 1989; Paterson et al. 1990). mean separation by LSD (Gomez and Gomez 1984) as described by Njiti et al. (1998). To determine the effect **DNA homology and gene prediction** genomic markers associated with resistance to SDS, the NIL DNA sequences corresponding to loci on linkage group G and recombination events were classified as Essex' and Forrest' type for each polymorphic DNA marker. Markers were compared with the trait response scores by the test in ANOVA done with SAS (SAS Institute Inc.). The probability of association of each marker with each trait was determined and a significant association was declared  $P \leq 0.05$  (unless noted otherwise in the text), since the detection of

*Arabidopsis thaliana* (TIGR, Beltsville, Maryland), and Plant ERGO July 2002 version (Integrated Genetics, Chicago, Illinois). Where the 3 search methods disagreed, the two methods that agreed were accepted. Repeat identification was done by BLAST alone, since the gene prediction packages incorporated repeat masking.

### Construction of a high-resolution map on linkage group G between BARC-Satt309 and SIUC-Sat122

A fine genetic map within the interval between markers BARC-Satt309 and SIUC-Sat122 on linkage group G was constructed. The map was derived from 12 recombination events among 96 RILs and 24 recombination events among

c

B

## R

A genetic map that included 240 polymorphic microsatellite markers (Kassem et al. 2004), selected from 1012 such DNA markers, was used to identify loci that segregated in the NILs (Iqbal et al. 2001; Song et al. 2004; see Soybase 2002 at <http://soybase.agron.iastate.edu>). About 8% of the genome segregates in each NIL pair. Typically, the segregating regions are separated into 24 regions (Njiti et al. 1998; Meksem et al. 1999). The 10-cM region of linkage group G between BARC-Satt309 and SIUC-Sat122 was the only region that was associated with quantitative trait loci for SDS in ExF34-derived NILs. Markers SATT371 (C2) and SATT354 (I) that segregated in the NILs are at least 10 cM from QTL for resistance to SDS (Iqbal et al. 2001) because the resistance to SDS of the NILs was only associated with the markers on linkage group G as judged by ANOVA ( $P < 0.05$ ).

**Fig. 1** (*concluded*).

**Table 4.** Substitution mapping among recombination event lines derived from F<sub>2</sub> crosses between (ExF34-5) × (ExF34-29) and (ExF34-6) × (ExF34-29) near-isogenic lines that segregate for SDS resistance.

Group	n	SIUC-Sat1		SIUC-Sat185		CGG5-SCAR		OI03-P4		CTA13-SCAR		SIUC-Sat122		Satt570		DI		IS	
		allele	allele	allele	allele	allele	allele	allele	allele	allele	allele	allele	allele	allele	allele	allele	mean ± SEM <sup>a</sup>	Phenotype	mean ± SEM <sup>b</sup>
1b	4	E	F	F	E	E	E	E	E	E	E	E	E	E	E	1.25±1.25	R	66.75±10.75	S
2b	52	E	F	F	F	E	E	E	E	E	E	E	E	E	E	8.27±1.34	R	58.66±2.31	S
3b	4	E	F	F	F	E	E	E	E	E	E	E	F	F	2.50±1.44	R	65.65±4.35	S	
4b	4	E	F	F	F	E	E	E	E	E	E	E	F	F	2.50±1.44	R	67.95±8.55	S	
5b	8	E	E	E	E	E	E	F	F	F	F	F	E	E	34.38±12.59	S	40.25±6.81	R	
6b	8	E	E	E	E	E	E	E	E	F	F	F	E	E	6.88±2.30	R	63.08±5.08	S	
7b	36	F	F	F	F	F	F	F	F	F	F	F	E	E	7.36±1.81	R	58.56±2.71	S	

Note: E, Essex; F, Forrest; R, resistant; S, susceptible.

<sup>a</sup>A DI of 20 or less than and one of more than 20 were critical values used to distinguish between resistant and susceptible lines for DI.

<sup>b</sup>An IS of 50 or less than and one of more than 50 were used to distinguish between resistant and susceptible lines for IS.

ExF34-derived NILs were lined up to construct substitution maps (Table 4). The recombination events were located on both sides of the loci of interest.

**Substitution mapping of the *Rft/QRfs2* gene using DX**

Disease phenotypes were compared with the recombination events in genomic regions to establish the position of the disease resistance loci *QRfs1* and *QRfs2* (Table 4). All genotype groups (14, 6, and 7) carrying a Forrest' allele at SIUC-Sat1 showed resistance to leaf scorch by SDS. This marker identified the only region all resistant lines had in common. Therefore, the *Rft/QRfs2* gene is probably located in the region encompassing Satt309, TMD, and SIUC-Sat1.

**Substitution mapping of the *QRfs1* gene using IS**

Within the genotype set (Table 4), the majority of the genotype groups (14, 6, and 7) were susceptible to root infection. However, no single locus or marker model could explain the segregation observed. The consistent model was that susceptible genotypes did not carry the beneficial alleles from Forrest' for the region encompassed by OI03-P4 to SIUCSATT122. Therefore, *QRfs1* may be a locus composed of several genes located between OI03-P4 and CTA13-SCAR. In addition, *QRfs1* was present and in the absence of *Rft/QRfs2*, disease incidence was very high, suggesting that *QRfs1* may have pleiotropic effects or be linked to gene(s) that cause leaf scorch to be more evident (Njiti et al. 1998).

**Gene island inference and gene prediction within intervals**

Intervals containing *QRfs2* and *QRfs1* were examined for structure and gene content using the soybean genome browser (<http://bioinformatics.siu.edu>) that represents the physical map (Wu et al. 2004 a, 2004 b), the genetic map (Song et al. 2004) (Fig. 1), and the BAC end sequence database (Shultz et al. 2002). In addition, we used an annotation (Table 5) of the available sequences from 2 cultivars (Hauge et al. 2001; Meksem et al. 2001 a), contig data for *rhg1* (Meksem et al. 2005), and published hybridization data (Foster-Hartnett et al. 2002). There were 10 contigs with 45 BAC end sequence data within the 7.5-cM region examined and 4 showed significant homology to known genes. Three contig sequences identified only repeat, nongenic sequences. Therefore, the region may contain 3 or 4 gene-rich islands but they do not appear to be contiguous.

Gene density was further explored using composition and homology searches. Within both published sequences on linkage group G (the 322- and 127-kbp tracts), there were 74 predicted genes at a density of 1 gene per 7 kbp (Table 5; AX196295). Sixty-two aligned with (expressed sequence tags) ESTs; the best hits were judged identical (31), paralogous (26), or orthologous (5) by sequence. Orthologs were from *A. thaliana*, *Oryza sativa*, and *Medicago truncatula*, but there was no clear evidence of microsyntenic regions between soybean and other partially or completely sequenced genomes (Ku et al. 2000; Yan et al. 2003). Lack of conserved microsynteny may be caused by the presence of a major disease resistance gene in the region in soybean. Data suggest that 77% of soybean genes appeared to be paralogs of the existing soybean ESTs. About 83% have orthologs

**Table 5.** Results of the annotation of the 315 kbp of the finished DNA sequence from soybean cultivar Asgrow 3433 for linkage group G within 1100–172 444 bp (GenBank No. AX196295): shown are the best EST matches from plant species.

	Gene name	Gene coordinates	Predicted exons	EST
1	Predicted protein	11004023	11104023	
2	Predicted protein G31	11932019	11934540 19602019	AW830182 BG156466 AW831483
3	Predicted protein	52205605	52205385 54665605	
4	Predicted protein G30 Na/H transporter	6782–10 029	67826817 69086991 73258323 8791–10 029	
5	Predicted protein G29 Chorismate dehydrogenase Embryonic abundant protein	16 234–18 215	16 234–16 400 16 603–16 927 17 614–17 696 17 786–17 855 17 942–18 016 18 120–18 215	AW185583 TC78116 A1938433 TC77282
6	Predicted protein G26	25 760–21 294	25 760–25 081 24 711–24 582 23 752–23 225 22 204–22 097 22 008–21 460 21 368–21 294	BG550903 TC75279
7	Predicted protein G25 AT-rich putative protein	34 265–28 983	34 265–34 061 33 522–33 446 33 188–33 135 32 602–32 342 32 237–32 175 30 741–30 673 30 591–30 546 29 044–28 983	TC63131 TC63130
8	Predicted protein G24 Hypothetical protein	40 482–38 327	40 482–40 351 40 167–40 060 39 952–39 872 39 765–39 593 38 723–38 327	
9	Unnamed protein product ( <i>Rtg1</i> )	46 891–49 573	44 891–48 763 48 975–49 573	
10	Predicted protein G1 Diphenol oxidase Laccase	58 247–62 782	58 247–58 372 58 486–58 637 59 236–59 480 59 737–59 865 60 760–61 301 61 854–62 298 62 652–62 782	AW307139 AW707121 AW706547
11	Predicted protein G13 Na/H antiporter protein like (At)	67 540–64 896	67 540–66 689 66 089–64 896	
12	Predicted protein G15	74 602–69 934	74 602–74 363 73 613–73 524 72 553–71 609 71 184–69 934	AW279576
13	Predicted protein G2	75 245–75 418	75 245–75 263 75 288–75 418	
14	Predicted protein G3 DNA repair protein	80 601–84 056	80 601–81 542 81 850–81 996	BF425110 (GM)

**Table 5** (continued).

	Gene name	Gene coordinates	Predicted exons	EST
	DNA helicase protein		82 131 -84 056	
15	Predicted protein G4	94 907 -95 599	94 907 -95 559	
16	Predicted protein G16	108 228 -104 233	108 228 -108 033 107 943 -107 700 105 535 -105 364 105 254 -105 055 104 946 -104 788 104 572 -104 233	
17	Predicted protein G17	112 800 -110 429	112 800 -112 707 112 502 -112 396 112 281 -112 243 111 725 -111 658 111 542 -111 464 110 712 -110 665 110 509 -110 429	BE473522
18	Predicted protein G18	117 271 -114 083	117 271 -117 054 116 477 -116 435 114 253 -114 083	
19	Predicted protein G5	119 155 -123 583	119 155 -119 418 119 647 -119 883 120 217 -120 477 120 657 -120 736 120 826 -120 931 121 078 -121 191 121 295 -121 393 121 496 -121 585 121 976 -122 082 122 166 -122 309 123 316 -123 583	AI759877
20	Predicted protein G6	124 288 -125 524	124 288 -124 507 125 391 -125 524	TC77682
21	Predicted protein G7	130 110 -136 768	130 110 -130 277 130 955 -131 092 131 703 -131 812 131 900 -132 011 132 103 -132 210 133 048 -133 160 134 242 -134 323 134 837 -134 917 135 130 -135 179 135 636 -135 771 135 867 -135 950 136 076 -136 768	BG044009 BG881498
22	Predicted protein G19	139 055 -137 811	139 055 -137 811	
	Legumin protein			
23	Predicted protein G20	146 284 -145 489	146 284 -146 129 145 830 -145 489	
24	Hypothetical protein			
24	Predicted protein G8	151 349 -151 628	151 349 -151 395 151 454 -151 628	GM
	Ferritin			
25	Predicted protein G9	156 824 -159 344	156 824 -157 092 157 479 -157 514 157 571 -157 654 157 747 -157 807 157 943 -158 030 158 558 -158 619 158 890 -158 955 159 095 -159 158	

Table 5 (continued).

	Gene name	Gene coordinates	Predicted exons	EST
26	Predicted protein G10 Histidine acid phosphatase	165 544 -169 870	159 292 -159 344 165 544 -165 638 165 799 -165 886 166 182 -166 475 166 765 -166 923 167 505 -167 645 167 758 -167 814 168 148 -168 286 168 929 -169 033 169 164 -169 360 169 487 -169 528 169 646 -169 729 169 811 -169 870	AW755680 BE473897 BG239199 TC64328 BE331398 BG404904 BE804031
27	Predicted protein	172 070 -175 222	172 070 -172 444 173 355 -173 411 173 485 -173 571 173 662 -173 808 173 948 -174 052 174 795 -174 893 175 070 -175 222	
28	Predicted protein	177 603 -179 938	177 603 -177 894 178 008 -178 097 178 348 -178 522 178 824 -179 061 179 430 -179 591 179 695 -179 938	
29	Predicted protein G21 RNA-binding protein	184 314 -183 637	184 314 -183 637	
30	Predicted protein G33 Putative wall-associated kinase Receptor protein kinase (At)	195 177 -198 887	195 177 -195 291 195 388 -195 526 195 830 -195 901 196 019 -196 162 196 246 -196 383 196 499 -196 908 197 167 -197 357 197 448 -197 580 198 295 -198 887	BE661144 TC71019 AW397578 BE021685 TC77189 BG239559 AW397813
31	Predicted protein G34 Putative proline-rich protein	201 848 -202 372	201 848 -202 372	
32	Predicted protein G40 Unknown protein	205 130 -202 719	205 130 -205 045 204 957 -204 813 204 644 -204 487 204 007 -203 923 203 813 -202 719	Rice
33	Predicted protein G35 Proteophosphoglycan	206 400 -210 804	206 400 -206 909 208 094 -208 198 208 310 -208 579 209 691 -210 203 210 721 -210 804	AW185645 TC64679 BF597582
34	Predicted protein G36 Unknown protein	213 752 -221 998	213 752 -213 772 214 093 -214 577 215 317 -215 422 215 513 -215 662 215 792 -215 953 216 039 -216 211 216 310 -216 376 216 818 -216 984	TC64084 TC63961 TC69770 BF425975

Table 5 (continued).

Gene name	Gene coordinates	Predicted exons	EST
		217 626 -217 728	
		219 391 -219 468	
		219 607 -219 687	
		220 175 -220 249	
		220 342 -220 433	
		220 581 -220 743	
		221 647 -221 707	
		221 886 -221 998	
35	Predicted protein G37 Unknown protein	227 781 -228 320	
36	Predicted protein	241 484 -237 452	
		241 484 -237 747	
		237 592 -237 452	
37	Predicted protein G38 Zinc finger protein	245 985 -246 512	Tc75269
38	Predicted protein G41 DNA binding protein	257 344 -256 038	BG359491 TC71751
39	Predicted protein G42 Putative protein (At)	263 929 -263 089	TC69298 BG157051
40	Predicted protein G43 Clathrin heavy chain (GM)	279 167 -267 347	TC68007 AW596783
		277 632 -277 544	
		277 448 -277 370	
		277 243 -277 116	
		276 744 -276 646	
		276 516 -276 408	
		275 720 -275 641	
		275 566 -275 339	
		275 257 -275 192	
		274 024 -273 908	
		273 716 -273 546	
		273 356 -273 240	
		273 038 -272 979	
		272 781 -272 695	
		272 483 -272 265	
		272 172 -272 110	
		272 023 -271 919	
		271 838 -271 774	
		271 614 -271 362	
		271 256 -271 120	
		271 044 -270 766	
		270 684 -269 682	
		269 427 -268 942	
		268 845 -268 675	
		268 590 -268 405	
		267 985 -267 863	
		267 737 -267 627	
		267 493 -267 347	
41	Predicted protein G44 Ribosomal protein (60s) (At)	282 698 -282 016	BG52844 TC65646 TC61913 TC61914 TC65646 TC61915 BF423807
		282 698 -282 522	
		282 216 -282 016	
42	Predicted protein G45 Unknown protein	290 895 -289 298	TC66532 A I496112
		290 895 -290 414	
		290 290 -290 119	

**Table 5** (concluded).

	Gene name	Gene coordinates	Predicted exons	EST
43	Predicted protein G46	294 688 –293 716	289 501 –289 298	AW317196
	Phosphorobosylanthranilate isomerase		294 688 –294 284	TC65409
	Anthranilate synthase		294 208 –294 045	BE612235
	Anthranilate synthase complex II amidotransferase		293 941 –293 716	AW186079
	Indol-3-glycerol phosphate synthase (5Phosphoribosyl)anthranilate isomerase			
44	Predicted protein G39	303 473 –304 812	303 473 –303 978	BF425173
	Dioxygenase (rice)		304 467 –304 812	TC72744
	Stress response SRG1 protein			TC66319
	Flavonol synthase			TC75936
	Gibberellin oxidase			TC72744
	Naringenin-2-oxoglutarate-3-dioxygenase			
	3-Hydrolase dioxygenase (rice)			
45	Predicted protein G47	321 532 –316 447	321 532 –320 370	TC71518
	Hydroxyproline-rich glycoprotein		319 725 –319 476	BG882316
	Heat shock protein like (At)		318 264 –316 447	BF020035
				TB69148 A I 988139

among plant ESTs. Identification of the remaining 1723% of genes will be dependent on further legume genome sequencing.

The mean gene coding region was 2.56 kbp (range 6 bp to 10 kbp). The mean intron number was 3.4 (range 0-24).

Intergenic distance was 5309 bp but no large intergenic distance were observed (all <12 kbp). Genetic markers matched the sequenced regions at 7 coordinates. About 20 BACs were expected to overlap the sequenced region based on current databases. However, it may represent an uncharacterized physical map anchor to a BAC fingerprint derived physical map (Shultz et al. 2003; Wu et al. 2004 a, 2004 b; Meksem et al. 2005; see <http://bioinformatics.siu.edu>). BAC end sequences matched for all those tested. Clones with large inserts and minimal overlap with the sequence regions were identified. Therefore, the region contained genes at a 10- to 20-fold higher frequency than expected for the entire soybean genome. The annotation pipelines developed at *A. thaliana* (TIGR) and Plant ESTs (ERGO) are effective for soybean DNA and orthologs are rarely the only evidence for annotation. Taken together, the gene prediction data suggest that the gene-rich island extends for 7.5 cM or 3.5 Mbp. Therefore, the sequencing of the soybean genome may be highly efficient if a rigorously edited physical map is used to select BAC clones (Shultz et al. 2003; Wu et al. 2004 b).

Candidates for the *QRfs2* gene (genes 17) clustered with the *Rhgl* gene (gene 8) for resistance to SCN (Table 5). Laccase (gene 9) was also considered a strong candidate for *QRfs2* because it is polymorphic between resistant and susceptible cultivars and may be involved in cell wall strengthening (Iqbal et al. 2002). Independently derived resistant cultivars Forrest, Pyramid, and Hartwig (Prabhu et al. 1999; Njiti et al. 2002) all share the same haplotype for laccase (Ahsan 2004). Independently derived susceptible cultivars Essex, Douglas, Flyer, and Asgrow 3244 all share the second haplotype. These cultivars represent 5 of 8 possible

haplotypes at *Rhgl* (Hague et al. 2001; Lightfoot and Meksem 2001; Afzal et al. 2004). Therefore, selection may have favored recombination, mutation, or gene conversion in this gene pair.

A single-nucleotide polymorphism within the laccase gene alters amino acid sequence within a sequence motif that is also found in *Rhgl* (Hague et al. 2001; Lightfoot and Meksem 2001). The sequence is not a recognized motif in current databases. However, it may represent an uncharacterized allosteric binding site that contributes to the interaction between *Rhgl* and *QRfs2* and therefore SDS and SCN resistance (Chang et al. 1997). Finally, laccase is a diphenol oxidase and many of the toxins produced by *Fsg* are phenolic in structure (e.g., the toxin Monorden; Baker and Nemeč 1994; Jin et al. 1996). RNAi (small RNAs that inhibit laccase gene expression) plants are being used to test the hypothesis that laccase gene silencing in collaboration with Dr. C. Taylor, Danforth Center, St. Louis, Mo.) than transformation and protein activity assays will be pursued in future (Ahsan 2004).

### Conclusions

Based on the fine map, the QRfs genes within separate loci appear to interact. Therefore, MAS for resistance to *Fsg* using *QRfs2* alone may not be effective unless TMD and SIUC-Sat1 are included. Using all 3 markers together will make prediction for resistance to SDS more precise. Moreover, selection for *QRfs1* for resistance to *Fsg* root infection by *OI03-P4* or SAT122 alone will predict the phenotype of about 66.9% because this marker is not mapped exactly on the gene. Therefore, more markers derived from contigs between *OI03-P4* and *SIUC-Sat122* region may provide better selection tools for this gene. The gene-rich regions are extensive and well localized, the contigs available are robust, and the conserved

repeat regions are well defined so that continued clone by Denny, R., Penuela, S., and Young, N.D. 2002. Comparative clone sequencing of this region will be most informative. A genomic analysis of sequences sampled from a small region on complete sequence of this chromosome arm would provide information on genome structure and resistance loci architecture. Isolation and unequivocal identification of the *ORfs1* and *ORfs2* gene sequences will provide perfect markers for selection and provide basic information addressing the molecular nature of partial resistance in plants.

## A

This research was funded in part by grants from the National Science Foundation (9872635), the Illinois Soybean Program Operating Board (02-127-03), and the United Soybean Board (2228). This report is based on work supported by the National Science Foundation under grant No. 9872635. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation, ISPOB, or USB. The continued support of the College of Agriculture, Southern Illinois University at Carbondale, and Office of the Vice Chancellor for Research to M.J.I. is highly appreciated. The authors thank Dr. Yakov Kogan (Integrated Genomics, Chicago, Ill.) for assistance with gene prediction and annotation. Dr. A. Jamai, Dr. K. Meksem, R. Ahsan, and J. Afzal are thanked for sharing unpublished results. J. Shultz is thanked for technical assistance. Chet Langin and Jyothi Lavu are thanked for instituting Gbrowse. For assistance with SDS scoring, members of the Southern Illinois University at Carbondale field teams from 1999 to 2003 are thanked.

## R

Afzal, J., Ahsan, R., Yaegashi, S., Kazi, S., Gillum, J., and Lightfoot, D.A. 2004. Sequence of *g1* haplotypes in 112 SCN resistant PI accessions from China. GenBank No. AY597055-AY597126, AY598406, AY598413, and AY618844, AY618867.

Ahsan, R. 2004. A analysis of haplotypes and expression in yeast of variant laccase gene candidate *lcc2* from soybean. M.Sc. thesis, Southern Illinois University at Carbondale, Carbondale, IL.

Baker, R.A., and Nene, S. 1994. Soybean sudden death syndrome: isolation and identification of a new phytotoxin from cultures of the causal agent, *Fusarium solani*. *Phytopathology*, **84**: 1144. [Abstr.]

Bell-Johnson, B.B., Garvey, G., Johnson, J., Meksem, K., and Lightfoot, D.A. 1998. Methods for high-throughput marker assisted selection for soybean. *Soybean Genet. Newslett.* **25**: 115-118.

Chang, S.J.C., Doubler, T.W., Kilo, V., Suttner, R., Klein, Schmidt, M.E., Gibson, P.T., and Lightfoot, D.A. 1996. Two additional loci underlying durable field resistance to soybean sudden death syndrome (SDS). *Crop Sci.* **36**: 1684-1688.

Chang, S.J.C., Doubler, T.W., Kilo, V., Suttner, R., Klein, Schmidt, M.E., Gibson, P.T., and Lightfoot, D.A. 1997. Association of field resistance to soybean sudden death syndrome (SDS) and cyst nematode (SCN). *Crop Sci.* **37**: 965-971.

Danesh, D., Penuela, S., Mudge, J., Denny, R.L., Nordstrom, Martinez, J.P., and Young, N.D. 1998. A bacterial artificial chromosome library for soybean and identification of clones near a major cyst nematode resistance gene. *Theor. Appl. Genet.* **96**: 1962-1970.

Foster-Hartnett, D., Mudge, J., Larsen, D., Danesh, D., Yan, Denny, R., Penuela, S., and Young, N.D. 2002. Comparative genomic analysis of sequences sampled from a small region on soybean (*Glycine max*) molecular linkage group G. *Genome*, **45**: 634-645.

Gomez, A.K., and Gomez, A.A. 1984. Statistical procedures for agricultural research. 2nd ed. John Wiley & Sons, New York.

Hague, B.M., Wang, M.L., Parsons, J., and Parnell, L.D. 2001. Nucleic acid molecules and other molecules associated with soybean cyst nematode resistance. Patent pending WO No. 20030005491.

Hague, B.M., and Effertz, R.J. 2001. Soybean SSRs and methods of genotyping. Patent pending WO No. 20020133852.

Haley, C.S., Knott, S.A., and Elson, J.M. 1994. Mapping quantitative trait loci in crosses between outbred lines using least squares. *Genetics* **136**: 1195-1207.

Hartwig, E.E., and Epps, J.M. 1973. Registration of Forrest' soybeans. *Crop Sci.* **13**: 287.

Hinetkovsky, N., Chang, S.J.C., Doubler, T.W., Gibson, P.T., and Lightfoot, D.A. 1996. Genetic mapping of loci underlying field resistance to soybean sudden death syndrome (SDS). *Crop Sci.* **36**: 392-400.

Iqbal, M.J., Meksem, K., Njiti, V., Kassen, A.M., and Lightfoot, D.A. 2001. Microsatellite markers identify three additional quantitative trait loci for resistance to sudden death syndrome (SDS) in Essex' xForrest' RILs. *Theor. Appl. Genet.* **102**: 187-192.

Iqbal, M.J., Yaegashi, S., Njiti, V.N., Ahsan, R., Cryder, K.L., and Lightfoot, D.A. 2002. Resistance locus pyramids alter transcript abundance in soybean roots inoculated with *Fusarium solani* f.sp. *glycines*. *Mol. Genet. Genomics* **268**: 407-417.

Iqbal, M.J., Afzal, J., Yaegashi, S., Ruben, E., Triwitayakorn, T., Njiti, V.N., Ahsan, R., Wood, A.J., and Lightfoot, D.A. 2002. A pyramid of loci for partial resistance to *Fusarium solani* f.sp. *glycines* maintains myo-inositol-1-phosphate synthase expression in soybean roots. *Theor. Appl. Genet.* **105**: 1115-1123.

Jiang, H., Hartman, G.L., Nickell, C.D., and Widholm, J.M. 1996. Characterization and purification of a phytotoxin produced by *Fusarium solani*, the causal agent of soybean sudden death syndrome. *Phytopathology*, **86**: 277-282.

Kassen, A., Meksem, K., Iqbal, M.J., Njiti, V., Banz, W.J.; Winters, T.A., Wood, A.J., and Lightfoot, D.A. 2004. Identification of three additional loci conditioning phytoestrogen content in soybean. *J. Biotechnol. Bioinform.* **2**: 52-60.

Kearsey, M.J., and Farquhar, A.G. 1998. QTL analysis in plants: where are we now? *Heredity* **80**: 137-142.

Ku, H.M., Vision, T., Liu, J., and Tanksley, S.D. 2000. Comparing sequenced segments of the tomato and *Arabidopsis* genomes: large-scale duplication followed by selective gene loss creates a network of synteny. *Proc. Natl. Acad. Sci. USA* **97**: A9121-9126.

Lander, E., and Botstein, D. 1989. Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics*, **121**: 85-99.

Lander, E., Green, P., Abrahamson, J., Barlow, A., Daley, M., Lincoln, S., and Newburg, L. 1987. MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. *Genomics*, **1**: 17-25.

Lightfoot, D.A., and Meksem, K. 2001. Novel polynucleotides and polypeptides relating to loci underlying resistance to soybean cyst nematode and methods of use thereof. Patent pending No. 09/772,134; filing date 29 January 2001.

Mek, L.F., Mudge, J., Darnielle, L., Grant, D., Hanson, N., Paz,

- M., Huihuang, Y., Denny, R., Larson, K., Foster-Hartnett, E., and Cooper, A., Danesh, D., Larsen, D., Schmidt, T., Staggs, R., Ring, A., Crow, J.A., Retzel, E., Young, N.D., and Shoemaker, R.C. 2001. Soybean genomic survey: BAC-end sequences near RFLP and SSR markers. *Genome*, **44**: 572581.
- Meksem, K., Doubler, T.W., Chanchaoenchai, K., Njiti, V., Chang, S.J.C., Rao Arelli, A.P., Cregan, P.E., Gray, L.E., Gibson, J.P., and Lightfoot, D.A. 1999. Clustering among loci underlying soybean resistance to *Fusarium solani*, SDS and SCN in near-isogenic lines. *Theor. Appl. Genet* **99**: 11314142.
- Meksem, K., Zobrist, K., Ruben, E., Tao, Q., Zhang, H.B., and Lightfoot, D.A. 2000. Two transformation ready large insert clone libraries for soybean: physical mapping of resistance to soybean cyst nematode and sudden death syndrome. *Theor. Appl. Genet* **101**: 747755.
- Meksem, K., Ruben, E., Hyten, D., Triwitayakorn, K., and Lightfoot, D.A. 2001. Conversion of AFLP bands to high-throughput DNA markers. *Mol. Genet. Genomics*, **265**: 207214.
- Meksem, K., Pantazopoulos, P., Njiti, V.N., Hyten, L.D., Arelli, P.R., and Lightfoot, D.A. 2001. Forrester' resistance to soybean cyst nematode is bigenic: saturation mapping of the *Rhg1* and *Rhg4* loci. *Theor. Appl. Genet* **103**: 710717.
- Meksem, K., Jamai, A., Zobrist, K., Ruben, E., Tao, Q., Zhang, H.B., and Lightfoot, D.A. 2005. A high resolution physical map of the *rhg1*, *Rfs2* locus for resistance to soybean cyst nematode and sudden death syndrome. *Theor. Appl. Genet.* In press.
- Njiti, V.N., Shenaut, M.A., Suttner, R.J., Schmidt, M.E., and Johnson, P.T. 1996. Inheritance of soybean response to soybean sudden death syndrome (SDS) response as influenced by cyst nematode (SCN) resistance in progeny of 'Pyramid', Douglas'. *Crop Sci* **36**: 11654170.
- Njiti, V.N., Suttner, R.J., Gray, L.E., Gibson, P.T., and Lightfoot, D.A. 1997. Rate reducing resistance to *Fusarium solani* f.sp. *phaseoli* underlies field resistance to soybean sudden death syndrome (SDS). *Crop Sci* **37**: 132438.
- Njiti, V., Doubler, T.W., Suttner, R.J., Gray, L., Gibson, P.T., and Lightfoot, D.A. 1998. Resistance to soybean sudden-death syndrome and *Fusarium solani* f.sp. *glycine* in near-isogenic lines. *Crop Sci* **38**: 472477.
- Njiti, V.N., Meksem, K., Iqbal, M.J., Johnson, J.E., Zobrist, K., Kilo, V.Y., and Lightfoot, D.A. 2002. Common loci underlying field resistance to soybean sudden death syndrome in Forrester, Pyramid, Essex, and Douglas. *Theor. Appl. Genet* **104**: 294300.
- Paterson, A., DeVerna, J.W., Lanini, B., and Tanksley, S.D. 1990. Fine mapping of quantitative trait loci using selected overlapping recombination-event chromosomes in an interspecific cross of tomato. *Genetics*, **124**: 735742.
- Prabhu, R., Njiti, V.N., Bell-Johnson, B., Johnson, J.E., Schmidt, M.E., Klein, J.H., and Lightfoot, D.A. 1999. Selecting soybean cultivars for dual resistance to soybean cyst nematode and sudden death syndrome using two DNA markers. *Crop Sci* **39**: 982987.
- Rupe, J.C.D., Sable, W.E., Robbins, R.T., and Gbur, E.E. 1993. Soil and plant factors associated with sudden death syndrome of soybean. *J. Prod. Agri* **6**: 218222.
- Shultz, J., Meksem, K., Shetty, J., Town, C.D., Koo, H., Potter, J., Wakefield, K., Zhang, H.B., Wu, C., and Lightfoot, D.A. 2003. End sequencing of BACs comprising a provisional minimal til ring path from a fingerprint physical map of soybean (*Glycine max*) cultivar Forrester'. GenBank CG826126CG812653 (13 473 sequences).
- Shultz, J., Meksem, K., and Lightfoot, D.A. 2003. Evaluating physical maps by clone location comparisons. *Genome Lett* **2**: 109118.
- Singh, R.J., and Hymowitz, T. 1988. The genomic relationship between *Glycine max* (L.) Merr. and *G. soja* Sieb. and Zucc. as revealed by pachytene chromosome analysis. *Theor. Appl. Genet.* **76**: 705711.
- Smith, T.J., and Camper, H.M. 1973. Registration of Essex-soybeans. *Crop Sci* **13**: 495.
- Song, Q.J., Shoemaker, R.C., Lark, K.G., Concibido, V.C., Delaney, X., Specht, J.E., and Cregan, P.B. 2004. A new integrated genetic linkage map of the soybean. *Theor. Appl. Genet.* **109**: 122428.
- Stevens, P.A., Nickell, C.D., and Kolb, F.L. 1993. Genetic analysis of resistance to *Fusarium solani* in soybean. *Crop Sci* **33**: 929-930.
- The Arabidopsis Genome Initiative. 2000. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* (London) **408**: 796815.
- Triwitayakorn, K. 2002. Positional cloning of *Rfh1* loci. Ph.D. thesis, Southern Illinois University at Carbondale, Carbondale, Ill.
- Walters, D.M., Baltazar, B.M., Rao-Areli, A.P., Schupp, J., Clayton, K., Kein, P., and Beavis, W.D. 1995. Genetic mapping of soybean cyst-nematode race 3 resistance loci in soybean PI 437.654. *Theor. Appl. Genet* **91**: 574581.
- Wu, C., Sun, S., Nimmakayala, P., Santos, F.A., Springman, R., Schupp, C., Meksem, K., Lightfoot, D.A., and Zhang, H.B. 2004. Construction and characterization of a bacterial artificial chromosome library and use of multiple complementary libraries for genome physical mapping. *Theor. Appl. Genet* **109**: 1041-1051.
- Wu, C., Sun, S., Nimmakayala, P., Santos, F.A., Springman, R., Ding, K., Meksem, K., Lightfoot, D.A., and Zhang, H.B. 2004. A BAC and BIBAC-based physical map of the soybean genome. *Genome Res.* **14**: 319326.
- Yan, H.H., Mudge, J., Kim, D.J., Larsen, D., Shoemaker, R.C., Cook, D.R., and Young, N.D. 2003. Estimates of conserved microsynteny among the genomes of *Glycine max*, *Medicago sativa* and *Arabidopsis thaliana*. *Theor. Appl. Genet* **106**: 12564265.
- Yuan, Z., Njiti, V., Meksem, K., Iqbal, M.J., Triwitayakorn, K., Kassem, M.A., Davis, G.T., Schmidt, M.E., and Lightfoot, D.A. 2003. Identification of yield loci in soybean populations that segregate for disease resistance. *Crop Sci* **42**(1): 271277.
- Zobrist, K., Meksem, K., Wu, C., Tao, Q., Zhang, H., and Lightfoot, D.A. 2000. Integrated physical mapping of the-soybean genome: a tool for rapid identification of economically important genes. *Soybean Genet. Newsl.* **27**: 1045.
- Zou, J.J., Singh, R.J., Lee, J., Xu, S.J., Cregan, P.B., and Hymowitz, T. 2003. Assignment of molecular linkage groups to soybean chromosomes by primary trisomics. *Theor. Appl. Genet.* **107**: 745750.