Theor Appl Genet (2007) 114:961–970
DOI 10.1007/s00122-006-0481-1

ORIGINAL PAPER

Fine mapping of the *Pc* locus of *Sorghum bicolor*, a gene controlling the reaction to a fungal pathogen and its host-selective toxin

Ervin D. Nagy · Tso-Ching Lee · Wusirika Ramakrishna · Zijun Xu · Patricia E. Klein · Phillip SanMiguel · Chiu-Ping Cheng · Jingling Li · Katrien M. Devos · Keith Schertz · Larry Dunkle · Jeffrey L. Bennetzen

Received: 26 July 2006 / Accepted: 30 November 2006 / Published online: 14 March 2007
© Springer-Verlag 2007

Abstract Milo disease in sorghum is caused by isolates of the soil-borne fungus *Periconia circinata* that produce PC-toxin. Susceptibility to milo disease is conditioned by a single, semi-dominant gene, termed *Pc*. The susceptible allele (*Pc*) converts to a resistant form (*pc*) spontaneously at a gametic frequency of 10^{-3} to 10^{-4}. A high-density genetic map was constructed around the *Pc* locus using DNA markers, allowing the *Pc* gene to be delimited to a 0.9 cM region on the short arm of sorghum chromosome 9. Physically, the *Pc*-region was covered by a single BAC clone. Sequence analysis of this BAC revealed twelve gene candidates. Several of the predicted genes in the region are homologous to disease resistance loci, including one NBS-LRR resistance gene analogue that is present in multiple tandem copies. Analysis of *pc* isolines derived from *Pc/Pc* sorghum suggests that one or more members of this NBS-LRR gene family are the *Pc* genes that condition susceptibility.

Introduction

Plant responses to potential pathogens are frequently governed by specific interactions between elicitor compounds secreted by the pathogen and plant resistance gene products. Although the known pathogenic elicitors represent a wide assortment of compounds, most resistance genes can be classified into a few well-defined groups (Nimchuk et al. 2003). Many of these resistance or *R* genes provide gene-for-gene type resistance, wherein the dominant *R* gene and a specific pathogen *avirulence* (*avr*) gene are needed to initiate a signal transduction process that elicits a hypersensitive

Communicated by F. Ordon.

E. D. Nagy · J. L. Bennetzen (✉)
Department of Genetics, University of Georgia,
Athens, GA 30602, USA
e-mail: maize@uga.edu

T.-C. Lee
Department of Botany and Plant Pathology,
Purdue University, West Lafayette, IN 47907, USA

W. Ramakrishna · Z. Xu
Department of Biological Sciences,
Michigan Tech University, Houghton, MI 49931, USA

P. E. Klein
Institute for Plant Genomics and Biotechnology,
Texas A&M University, College Station, TX 77843, USA

P. SanMiguel
Agricultural Genomics Center, Purdue University,
West Lafayette, IN 47907, USA

C.-P. Cheng · J. Li
Department of Biological Sciences,
Purdue University,
West Lafayette, IN 47907, USA

K. M. Devos
Department of Crop and Soil
Sciences and Department of Plant Biology,
University of Georgia, Athens, GA 30602, USA

K. Schertz
Department of Soil and Crop Sciences,
Texas A&M University, College Station,
TX 77843, USA

L. Dunkle
Crop Production and Pest Control Research, USDA-ARS,
West Lafayette, IN 47907, USA

resistance response in the host. Many *R* genes encode NBS-LRR proteins that are composed of three major structural features: a nucleotide binding site (NBS), a leucine-rich repeat (LRR) domain and either a coiled-coil (CC) or a Toll-interleukin receptor (TIR) domain at their N-termini.

Some pathogenic fungi secrete phytotoxins that selectively damage specific plant genotypes. These host-selective toxins are structurally quite diverse, including several classes of low-molecular-weight metabolites or proteins (Wolpert et al. 2002). Very little is known about toxin resistance genes in plants. In contrast to the *R* genes, the known genes conferring toxin resistance are highly variable in their genetic and molecular properties. The *HM1* gene confers resistance against the HC-toxin produced by the pathogen *Cochliobolus carbonum* in maize (*Zea mays* L.). It encodes a carbonyl reductase that directly inactivates the toxin molecule (Johal and Briggs 1992; Meeley and Walton 1991). The fungal AAL-toxin inhibits the tomato sphinganine *N*-acyltransferase, an enzyme involved in sphingolipid metabolism. The corresponding resistance gene (*Asc-1*), probably as part of an enzyme complex, is involved in sphingolipid metabolism and is insensitive to AAL-toxin action (Brandwagt et al. 2000; Spassieva et al. 2002). The mitochondrial gene T-*Urf13* exists only in the maize lines carrying the T male-sterile cytoplasm and causes sensitivity to the T-toxin of *C. heterostrophus*. The T-*Urf13* gene product is a transmembrane protein that conditions membrane leakage upon exposure to the T-toxin (Levings et al. 1995).

In the 1920s, a new disease of sorghum (*Sorghum bicolor* (L.) Moench), termed milo disease, arose and became the major threat to sorghum production in the US (Leukel 1948). However, resistant individuals arose spontaneously and frequently in susceptible populations. Schertz and Tai (1969) established that one mutation to resistance occurred in every 7876 gametes in at least one sorghum genetic background. They also found that susceptibility is controlled by the single, semi-dominant locus designated *Pc*. The causal agent of milo disease is *Periconia circinata* (Mangin) Sacc., a soil-borne saprophytic fungus (Leukel 1948) that produces two phytotoxic chlorinated polyketide peptides, peritoxin A and B (Macko et al. 1992), referred to as PC-toxin. PC-toxin alone is sufficient to produce the symptoms of milo disease in susceptible genotypes of sorghum (Wolpert and Dunkle 1983). Typically, infection by PC-toxin-producing strains of *P. circinata* causes dark red discoloration on the roots and crown. The leaves become chlorotic and eventually necrotic. Panicles of infected susceptible plants produce little or no grain. An early sign of infection is inhibition of seedling root development. Root growth inhibition of germinating seeds has been applied as a bioassay for characterizing sorghum *Pc* genotypes (Dunkle 1979).

PC-toxin induces several physiological changes in susceptible plant cells (reviewed in Dunkle and Macko 1995), including an increase in the synthesis of a group of low molecular weight proteins, electrolyte leakage and chromosome condensation. Resistant genotypes exhibit none of the toxin-induced symptoms or biochemical responses. However, the primary mode of action for PC-toxin remains to be determined. The ultimate goal of our investigations is to characterize the *Pc* gene and determine the molecular mechanism of the spontaneous mutations from susceptibility to resistance. In the present work, we develop a high-resolution genetic map of the *Pc* locus and identify potential gene candidates in the *Pc*-region.

Materials and methods

Plant materials and BAC library

A mapping population comprising 178 F_3 families was established from a cross between the sorghum cultivars Shanqui Red (*pc/pc*) and Colby (*Pc/Pc*). Sorghum BAC libraries containing inserts from *pc/pc* inbred BTx623 were obtained from the Texas A&M University BAC Center (Woo et al. 1994) and the Clemson University Genomics Institute (www.genome.clemson.edu/projects/stc/sorghum).

Bioassay for PC-toxin resistance

The genetic status of the *Pc* gene in the mapping population was evaluated by the root growth inhibition bioassay described by Dunkle (1979). At least ten seeds from each F_3 family were soaked in distilled water for 4–6 h at room temperature and incubated in germination papers for 36–48 h. Seeds with ~3-mm radicles were placed into *P. circinata* culture filtrate diluted 1:100 in sterile distilled water or in purified peritoxin A at 10 ng/ml and incubated for 48 h in the dark at room temperature. The scored toxin reaction was either sensitive (>90% inhibition of root growth as compared to the control), or intermediate (50–75% inhibition) or insensitive (similar to the control), representing the homozygous susceptible (*Pc/Pc*), heterozygous intermediate (*Pc/pc*) and homozygous resistant (*pc/pc*) genotypes, respectively.

Bulk segregant analysis

Genomic DNA was isolated from the parents and 165 of the 178 F_3 families using the CTAB (cetyl-trimethyl-ammonium bromide) method as described by Murray and Thompson (1980). Leaves from 12 to 15 F_3 seedlings, representing each F_2 plant, were collected. Equal amounts of

DNA from six phenotyped F_3 families were mixed to create bulk DNA samples. Four homozygous resistant and four homozygous susceptible bulk DNA samples were analyzed. The parents and these bulks were used to screen for molecular markers that were polymorphic and linked to the *Pc* gene.

Marker development and genetic mapping

RFLP markers distributed throughout the sorghum genome were selected from maize, sorghum and rice comparative genetic maps (Gale and Devos 1998), and used to perform DNA gel blot analysis. DNA from the parental genotypes and the bulked F_3 families were digested with one of seven restriction enzymes (*Bam*HI, *Eco*RI, *Eco*RV, *Hin*dIII, *Pst*I, *Sac*I or *Xba*I) and hybridized with the probes as previously described (Hulbert et al. 1990). One RFLP probe, that was found to be tightly linked to *Pc* was used to screen the Texas A&M BTx623 BAC library (Woo et al. 1994), and the clone 99E05 from the 9S chromosome arm was identified. Chromosome designation was as in Kim et al. (2005b).

For further marker development, BAC sequence data were produced by low-redundancy shotgun sequencing of sorghum clones. Two clones (Sbb12448 and c0184G17) in the same contiguous series (contig) of BACs derived from sorghum fingerprint analysis (Klein et al. 2000) with 99E05 were partially sequenced. Preparation of shotgun libraries of BAC subclones, sequencing and analysis were as described by Dubcovsky et al. (2001). As these sequences did not produce a sufficient number of polymorphic markers, additional genomic sequence data from chromosome arm 9S were collected from GenBank and the Sorghum Genomics Sequence Database at Texas A&M University (http://www.sorgblast2.tamu.edu). PCR primers were designed using the program PRIMER3 (Rozen and Skaletsky 2000) to develop sequence-tagged site (STS) markers. For the mapping of STS markers, 15 µl PCR mixture contained 1 U of *Taq* DNA polymerase (Roche), 6 nmol of forward and reverse primers, 75 µM of each dNTP and 2 ng/µl genomic DNA. PCRs were performed in 35 cycles, using the following conditions: 94°C for 30 s, 54–56°C for 30 s and 72°C for 1 min. Single stranded conformational polymorphism (SSCP) was applied using a non-denaturing MDE^{TM} (Cambrex) gel matrix in electrophoresis. Gels (43 × 35 cm) were run overnight at room temperature at 8 W. The DNA sequences that gave no polymorphism as STS markers were used to generate sequence-specific amplified polymorphism (SSAP; Waugh et al. 1997) markers to exploit the potential sequence variation in their flanking, unknown genomic regions. DNA was digested with *Eco*RI or *Mse*I and the appropriate adaptor was ligated to the restricted ends. Sequences between the anchored primers designed to the ends of the selected sequences and the adaptor-specific primers were amplified. The adaptor sequences and PCR conditions were the same as described earlier (Nagy and Lelley 2003).

The markers selected after the bulk segregant analysis were mapped in the entire population using the program MAPMAKER Version 3.0 (Lander et al. 1987). The F_2 intercross algorithm and default linkage criteria (LOD 3.0 and 50 cM maximum distance) were applied. The Kosambi function was used to establish genetic distances.

Physical mapping and sequence analysis

The BAC contig covering the *pc* region was first assembled by restriction enzyme fingerprinting (Sorghum Genomics Sequence Database, http://sorgblast2.tamu.edu) and then further refined using the present markers. BAC clone Sbb12448 was subcloned and sequenced as described earlier (Dubcovsky et al. 2001). Base calling and assembly of the BAC end sequences were performed using the programs PHRED and PHRAP, respectively (Ewing et al. 1998), and further analyzed with the program CONSED (Gordon et al. 1998). Gene candidates were detected using the programs FGENESH (Solovyev and Salamov 1997) and GENSCAN (Burge and Karlin 1997). The sequences of the gene candidates were analyzed using the BLAST programs (Altschul et al. 1990).

SSCP analysis of *Pc*/*Pc* and *pc*/*pc* Colby cultivars

Standard Colby cultivars have the genotype *Pc/Pc* and are, hence, susceptible to *P. circinata*. As previously described (Schertz and Tai 1969), 14 independent *pc* derivatives of *Pc/Pc* Colby were isolated and brought to a homozygous *pc/pc* state. Using a large number of RFLP and other DNA markers, all but one of these *pc/pc* Colby lines were found to be isogenic at all investigated loci with *Pc*/*Pc* Colby. The one exception was highly polymorphic compared to *Pc/Pc* Colby and was not further studied because it is expected to be a contaminant. Thirteen of these *Pc* to *pc* mutant lines were analyzed at the DNA level by SSCP. Primers were designed for the gene candidates in a way that all the exons of all the predicted genes were covered by adjacent or overlapping, 200–500 bp-long PCR fragments. The PCR products were run on an SSCP gel and polymorphism was analyzed between the *Pc* and *pc*-mutant isogenic lines. The PCR and electrophoresis conditions were the same as described above for the STS marker analysis.

Results

A population segregating for PC-toxin resistance was generated from a cross between sorghum inbreds Shanqui Red

(*pc/pc*) and Colby (*Pc/Pc*). The segregation of the *Pc* gene was monitored using the root growth inhibition bioassay on the mapping population. The F_1 population exhibited an intermediate root development rate. Out of 178 F_3 families, representing 178 F_2 individuals, 53 were found to be homozygous susceptible, 83 families segregated for the resistance (suggesting that the original F_2 plants were heterozygous), and 42 families were homozygous resistant. The ratio of the susceptible, segregating and resistant F_3 families fits the expected 1:2:1 ratio for a single locus ($\chi^2 = 2.168$, $P > 0.20$).

Molecular marker analysis and genetic mapping was carried out in 165 of the 178 F_3 families. A set of RFLP markers distributed across the sorghum genome was tested in a bulk segregant analysis to select *Pc*-linked markers, and these were then employed in analysis of the mapping population. One marker, a maize chitinase gene (UIU101), exhibited good linkage to *Pc* (~0.6 cM, Fig. 1). This marker was used to screen a sorghum BAC library, and it hybridized to clone 99E05. This clone is from a contig on the 9S chromosome arm. Genomic sequence data were obtained by partial sequencing in this specific contig, or collected from various databases focusing on the 9S chromosome arm. About two-hundred PCR-based marker candidates were developed using these sequence data. Following bulk segregant analysis, sixteen markers were selected for mapping. Two of them, pSB1945 (BH246322) and pSB0134*d* (BH245277), were based on previously published sequence data (Bowers et al. 2003). The marker pSB0134*d* was an SSAP marker representing the downstream flanking region of the sequence pSB0134. RFLP marker UIU101 and the 16 newly developed markers, along with two previously developed microsatellite markers (Txp410 and Txp047, P.E. Klein, unpublished), were used in the genetic fine-mapping. The *Pc* locus completely co-segregated with STS marker S12a (Fig. 1). Two of the 330 gametes showed recombination between the *Pc* gene and markers S12b, S66e, S66f and UIU101 (0.6 cM). In the opposite direction, one recombinant was found (0.3 cM) between the *Pc* gene and marker F4T7. Therefore, the *Pc* gene was delimited to the chromosome segment flanked by markers F4T7 and S12b, a region that spans ~0.9 cM. This will henceforth be referred to as the *Pc*-region.

Physically, the *Pc*-region was covered by a single BAC clone (Fig. 2). The markers UIU101, S66e and S66f, although co-segregating with S12b in the recombinational analysis, were located on two additional BAC clones, indicating that the local distribution of recombination events was quite uneven in this mapping study. The BAC clone including the *Pc*-region was sequenced completely. The total length of the BAC clone is approximately 118 kb, including about 110 kb containing the *Pc*-region (i.e., between markers S12b and F4T7). Twelve candidate genes were identified in the *Pc*-region (Table 1). Three of them are members of an NBS-LRR gene family that is more homologous to the maize rust resistance gene *Rp3* (Fig. 3) (Webb et al. 2002) than to any other DNA sequences in GenBank. Its overall similarity to *Rp3* is 47% at the amino acid level. A second gene, a putative amino acid selective channel protein, is also duplicated on this BAC, with one of the copies located in the intergenic region of the *Rp3*-homologous gene family (Fig. 2). The putative *Cf2/Cf5*-homologous gene represents another type of *R* gene (Dixon et al. 1996). Another candidate gene in the *Pc*-region, homologous to an *Xa21*-binding protein in rice, could also be functionally associated with a *Xa21*-type resistance gene in sorghum (Song et al. 1995). As revealed by PSI-Blast analysis, this gene consists of an ankyrin repeat domain and a Zn-finger domain, suggesting its involvement in both protein–protein and protein–DNA interactions. Hence, several

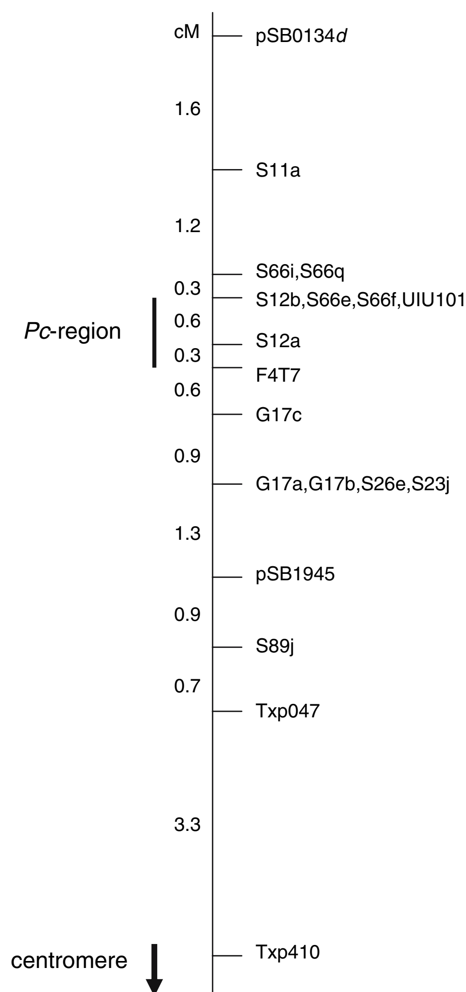

Fig. 1 Genetic map of the *Pc* locus on the 9S chromosome arm in sorghum. The position of the *Pc* gene was delimited to an ~0.9 cM region between markers S12b and F4T7

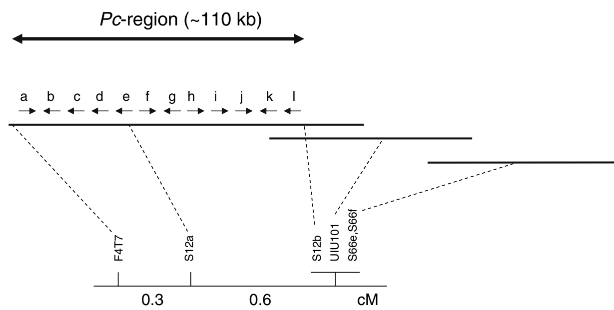

Fig. 2 Comparison of the genetic and physical maps of the *Pc*-region. The *Pc*-region flanked by markers F4T7 and S12b was assigned to an ~110 kb segment of sorghum BAC clone Sbb12448. The co-segregating markers S12b, UIU101, S66e and S66f were localized on three different BAC clones (Sbb12448, 99E05 and Sbb6675, respectively). Twelve *Pc* gene candidates were identified (*arrows*). These were **a** an *Xa21*-receptor-like kinase binding protein, **b** a pentatricopeptide (PPR) protein, **c** a pectin acetylesterase precursor, **d** a catalytic protein phosphatase type 2C, **e** an amino acid selective channel protein, **f** an *Rp3* (NBS-LRR)-like protein, **g** an amino acid selective channel protein, **h** an *Rp3* (NBS-LRR)-like protein, **i** an *Rp3* (NBS-LRR)-like protein, **j** a *Cf2*/*Cf5* disease resistance protein homologue, **k** a small nuclear ribonucleoprotein D2, and **l** a replication protein A1-like protein

Table 1 Predicted genes in the *Pc*-region of BTx623

Homology	Copy no.	*E* value
Xa21-receptor-like kinase binding protein (*Oryza*)	1	0.0
Pentatricopeptide (PPR) protein (*Oryza*)	1	7×10^{-51}
Pectin acetylesterase precursor (*Oryza*)	1	2×10^{-166}
Catalytic protein phosphatase type 2C (*Arabidopsis*)	1	2×10^{-106}
Amino acid selective channel protein (*Oryza*)	2	4×10^{-53}
Rp3 (NBS-LRR) protein (*Zea*)	3	0.0
Cf2/*Cf5* disease resistance protein (*Oryza*)	1	0.0
Small nuclear ribonucleoprotein D2 (*Oryza*)	1	7×10^{-47}
Replication protein A1-like (*Arabidopsis*)	1	1×10^{-179}

genes in the *Pc*-region appear to be likely candidates for the *Pc* locus.

From *Pc/Pc* Colby genotypes, numerous independent *pc* mutations were selected (Schertz and Tai 1969), and we chose 13 of these for further analysis. The derived *pc/pc* Colby isolines were confirmed to be isogenic across the sorghum genome by the use of several hundred RFLP and other DNA markers (data not shown). The instability of the *Pc* locus is suggestive of a tandem gene family that undergoes unequal recombination to remove some gene family members, as originally defined by Sturtevant (1925) and subsequently observed at many dozens of loci, including the *Rp1* disease resistance gene cluster in maize (Bennetzen et al. 1988; Sudupak et al. 1993; Webb et al. 2002). Hence, we decided to investigate whether the *Pc* to *pc* mutations were associated with changes in the copy number or other organizational feature(s) of the genes in the *Pc*-region. For this purpose, several PCR oligonucleotide primers were designed in order to completely cover each of the candidate genes identified in the *Pc*-region by the BAC sequence analysis. Nine primer-pairs were used to analyze the NBS-LRR genes. Six of them showed multiple band profiles. This suggests that the paralogous gene copies harbored sequence variations in these regions that was detected by SSCP. With these six primer pairs, some of the bands in the susceptible (*Pc/Pc*) Colby plants were missing in the resistant (*pc/pc*) mutant lines, suggesting that some of the NBS-LRR gene copies of the susceptible Colby plants were deleted in their mutant, resistant derivatives (Fig. 4). All other gene candidates were identical between the *Pc* and *pc*-mutant Colby isolines (data not shown).

Discussion

Fine-mapping of the *pc* locus

In the present study, we have fine-mapped a genomic region in sorghum that contains the *pc* toxin resistance locus. The plants heterozygous for the *Pc* gene (*Pc/pc*) provided an intermediate resistance response to the peritoxins. The resistance bioassay used can reliably discriminate between heterozygous and homozygous plants (Dunkle et al. 1979). However, the rate of germination and seedling root growth can also be affected by a number of other genetic and physiological factors that may obscure the effects of the PC-toxin. Therefore, to obtain more accurate resistance scores for the genetic mapping, F_3 families were tested for their PC-toxin response to assess the *Pc* allele genotypes in the F_2 progeny.

After a broad coverage screen of the sorghum genome with a few RFLP markers, one *Pc*-linked marker, a chitinase gene called RFLP probe UIU101, was identified. Many additional RFLP probes mapping in the colinear regions in maize and rice were then tested for their *Pc* linkage in sorghum. Despite the overall high degree of polymorphism between sorghum parent Shanqui Red and US sorghums like Colby (Deu et al. 2006; Oliveira et al. 1996), none of these additional RFLP markers were polymorphic in the sorghum parents that were utilized in the *Pc* mapping cross (data not shown). Thus, the apparent low level of sequence polymorphism in the area around *Pc* indicated the need for a more sensitive detection technique. For this reason, the majority of the DNA markers developed in the present

Fig. 3 Sequence alignment between one of the putative *Rp3*-homologues found in the *Pc*-region (Sbrp3) and the *Rp3* protein of maize (Zmrp3). Identical and similar residue pairs are shaded with *black* and *gray*, respectively

```
Sbrp3 ---CFSIWSVKAAGDKLVSLLATEFAAITGVKRDLCQLQDIHADITGWLSAVHDRAIQSE
Zmrp3 MEVALVSTVLKVLGTKLAPLALKELSSKAGVAKDLQELQDLVEEINNWLQTVGDKGRS--

Sbrp3 TQSHWVVKLKDVAYDIDDILQEVQLEAEKQKMERDDDKSGIAACFCAKPKSFAFRYKMAH
Zmrp3 --SKWLKKLKEVAYDADDLVHEFHIEAEKQDREITGGKNTLVKYFITKPKATVTEFKIAH

Sbrp3 KIKAIKVRFAAIVKQRSDFNTLVPTR--DQHVGTRYKTVGEMTWLSKVPESKIPLRDQEK
Zmrp3 KIKKIKNRFDEIVKGRSDYSTIANSMPVDYPVQHTRKTIGEVPLYTIVDETSIFGRDQAK

Sbrp3 DEIISKLVECNAGENNMIVSIIGLGGSGKTTLAKHICHDVKIKEHFGGEIFWVHVSQEFD
Zmrp3 NQIISKLIETDSQQR--IVSVIGLGGSGKTTLAKQVFNDGNIINHF-EVLLWVHVSREFA

Sbrp3 VQKLIGKLFETIVGDNSDRHPPQHMVQKISEKLSNKKFLLILDDAWHEDRHDWEQFMVQL
Zmrp3 VEKLVEKLFEAIAGDMSDHLPLQHVSRTISDKLVGKRFLAVLDDVWTEDRVEWERFMVHL

Sbrp3 KCGAPETRIVLTTRDRKVAQAVESRYTFELAFLSESESWNLFLKGSGLAEQDLSCDEVQV
Zmrp3 KSGAPGSSILLTTRSRKVAEAVDSSYAYDLPFLSKEDSWKVFQQCFGIALKALDPEFLQA

Sbrp3 GKEIIKGCGGVPLAIQTLGAVLCDKKQISTWRAIRENNLWKVQSIKDRVFASLKLSYIHL
Zmrp3 GIEIVEKCGGVPLAIKVIAGVLHGIKGIEEWRYICNSNLLDVQDDEHRVFACLLLSFVHL

Sbrp3 ADELKQCFTFCSIFPKGYGIQKDRLIAQWIAHGFINAMNGEQLEDVGRDYLDSLVKVRFL
Zmrp3 PDHLKPCFLHCSIFPRGYEINRCHLISQWIAHGFVPTNQARQAEDVGIDYFDSLLKVGFL

Sbrp3 QEAYGSRNTDIYNMHDLIHDLTRQILKDELVTCVPIHTTEEFTHRYRYLSLTSFTENVDK
Zmrp3 QIWS-TWGEVTCKMHDLVHDLARQILRDEFVSEIETN---KQIKRCRYLSLTSCTGKLDN

Sbrp3 GVFDKVRALYISDSKTSFDTTVKSSCCMRSVVLDYAIDTPFSLFILKFEYLGYLEIHNVS
Zmrp3 KLCGKVRALYVCGRELEFDKTMNKQCCVRTIILKYITADSLPLFVSKFEYLGYLEISSVN

Sbrp3 CTTVPEAISRCWNLQSLHFVNCKGFVTLPESVGKLRKLRTLELHRITDLESLPQSIGDCY
Zmrp3 C---------------------------------VQKLRTLELNGVSSIKSLPQSIGDCD

Sbrp3 VLQCLQLYKCRKQREIPSSLGRIGNLCVLDFNGCTGLQDLPSTLSCPTLRTLNLSETKVT
Zmrp3 NLRRLYLEGCHGIEDIPNSLGKLENLRILNIVHCISLQKLPPSDSFGKLLNLQTMAFKLC

Sbrp3 ----MLPQWVTSIDTLECIDLKGCNELRELPKGIANLKRLTVLNIERCSKLCCLPSGLGQ
Zmrp3 YDLRNLPQCMTSLIHLESVDLGHCPKLVELPEGIGNLRNLKVLNLKKCKKLRGLPAGCGQ

Sbrp3 LTRLRKLGLFVVGCGADDARISELENLDMIGGHLEITNLKYLKDPSEAEKACLKRKSHMQ
Zmrp3 LTRLQQLSLFVIGDNTKHARISELENLDKLDGELQIKNIRYVKDPSDADKVRLKKKIGIR

Sbrp3 RLELNWS------LSDAEEELVSDMEHDWGVLNALEPPSQIERLDIYGYRGPCLPGWMMK
Zmrp3 KLLLDCYSRLEVQPDDVEEELS--LNKEFHLLDSLEPPSKIEKLRIRGYRGSQMPRWMTK

Sbrp3 QNDSSY-CEGGIMLKQTIASHFLCLTLLTLVRFPNLRHMRGFVELPSLKTLELLEMPNLE
Zmrp3 QSDCCCPADDTHIVMQRNPSEFSHLIELVLDNLPNLEHLGELVELPLIKILKLKILPKLG

Sbrp3 ELWTTSSGFETGEKELAAQHLFPVLSSLHIYGCPKLNVSPYFPPSLERMTLGRTNGQLLS
Zmrp3 ELLTTT----TGEEGVEVQCCFHHVSTLVIIDCPKLVVKPYFPPSLQRLTLEGNNMQLVS

Sbrp3 TGRFSHQLPSMHALVP----------------RLKSLWLSEVTGSSSGWELLQHLTELKE
Zmrp3 SGCFFHPRHHHASHAHGDESSSSSYFADETGTHLERLELRRLTGSSSGWEVLQHLTGLHT

Sbrp3 LYIDTCNDLTQLPESMRN 1066
Zmrp3 LEIYMCTDLTHLPESIH- 1047
```

study utilized single stranded conformational polymorphism (SSCP) technology to scan for polymorphism. SSCP can detect as little as a single nucleotide change if it modifies the secondary structure of a single-stranded DNA molecule (Orita et al. 1989; Martins-Lopes et al. 2001).

In the mapping population employed, providing 330 gametes for possible recombination events, we mapped the *Pc*-region to ~0.9 cM between SSCP markers F4T7 and S12b. These markers were found to be about 110 kb apart on a single BAC that is derived from a region on the short arm of sorghum chromosome 9. This suggests a recombination rate of about 120 kb/cM in this region, a somewhat higher rate of meiotic recombination than predicted for the euchromatin (250 kb/cM) and much higher than for the heterochromatin (8.7 Mb/cM) in the sorghum genome (Kim et al. 2005a). This difference is likely associated with the high gene density found in this 110 kb region, as it is known that gene-rich regions are often associated with high rates of recombination (Sandhu and Gill 2002; Kim et al. 2005a).

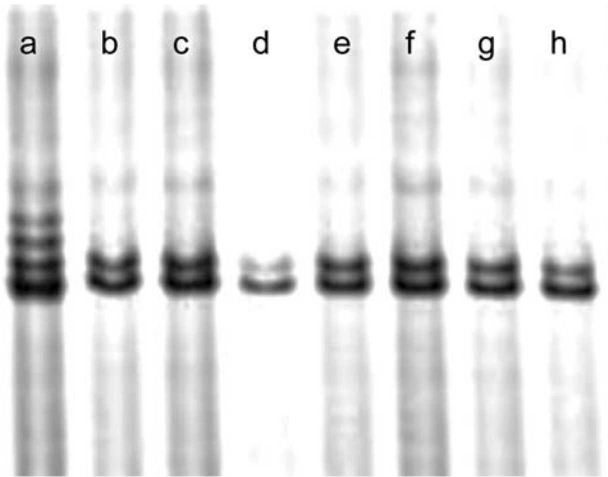

Fig. 4 Copy number change of the NBS-LRR gene in the *Pc*-region between the wild type, susceptible (*Pc/Pc*) Colby cultivar (**a**) and its resistant (*pc/pc*) derivatives (**b–h**) as revealed by SSCP

The structure of the *Pc*-region

The *Pc*-region on BAC Sbb12448 is about 110 kb and contains 12 gene candidates. Hence, on average, there is one putative gene every 9.2 kb in this region. Similar gene densities (10.7 kb/gene, 8.6 kb/gene) have been observed on other sorghum BACs (Ramakrishna et al. 2002a, b). These BACs were all selected because they had genes on them and, thus, are likely to represent gene-enriched regions. Overall, the sorghum genome is predicted to have about 50,000 genes (Bedell et al. 2005), or about one gene every 15 kb.

The NBS-LRR gene family found in the *Pc*-region is most similar to the maize *Rp3* gene. However, maize *Rp3* and this sorghum NBS-LRR gene family are not in colinear chromosomal positions. *Rp3* is located on chromosome 3 in maize (Webb et al. 2002), whereas the 9S chromosome arm of sorghum shows large-scale synteny with maize chromosomes 6 and 8 (Whitkus et al. 1992; Gale and Devos 1998). It has been noted previously that plant disease resistance genes, along with nucleolar organizers (Dubcovsky and Dvorak 1995), are less likely to exhibit colinearity than any other plant genes (Leister et al. 1998).

Candidate genes for *pc*

The sequenced BAC Sbb12448 contains numerous genes that might be targets for PC-toxin action. The *Rp3*- and *Cf2/Cf5*-homologues identified in the *Pc*-region have structural features of typical *R* genes (Webb et al. 2002; Dixon et al. 1996). This is in agreement with previous observations that *R* genes of different families often are organized in mega-clusters (Leister et al. 1998). The *Xa21*-binding protein homologue on sorghum BAC Sbb12448 may also be associated with disease resistance or susceptibility, but its function in sorghum has not yet been determined. The activity of the pectin acetylesterase gene product can loosen the pectin backbone in plant cell walls, making it also a possible target for pathogen gene action. Pathogenic nematodes can upregulate host pectin acetylesterase genes, thereby making the root tissues more accessible for invasion (Vercauteren et al. 2002). The 2C type protein phosphatases constitute the largest protein phosphatase family in plants (Schweighofer et al. 2004). They are involved in several signal transduction processes, including abcisic acid-mediated responses to a number of biotic and abiotic stresses.

Given that so little is known about the activities of fungal toxins against susceptible plant hosts, even genes with no previously identified association with plant disease susceptibility or resistance cannot be immediately excluded as possible PC-toxin targets. The pentatricopeptide proteins are encoded by many gene families in plants. Typically, they participate in post-transcriptional modification of organellar RNAs (Lurin et al. 2004) and usually are transported into either mitochondria or chloroplasts. The specific pentatricopeptide protein found in the *Pc*-region has a mitochondrial target sequence. The amino acid selective channel protein was discovered in connection with an oxidative stress response in plants. Its expression is highly enhanced in cold-treated barley (*Hordeum vulgare* L.) leaves (Baldi et al. 1999). The functional protein is located in the chloroplast membrane. The small nuclear ribonucleoprotein D2 is homologous to genes that are essential for the biogenesis and stability of snRNAs and, therefore, play a major role in pre-mRNA splicing (He and Parker 2000). Replication protein A is a key component of DNA replication, repair and recombination (Iftode et al. 1999). Disruption of the function of any of the proteins encoded by these genes could lead to a debilitated plant host that is more available to pathogenesis by *P. circinata.*

As the primary mode of action of PC-toxin is unknown, the proposed physiological functions of the genes in the *Pc*-region did not and could not reveal the identity of *Pc*. The semi-dominant susceptible allele (*Pc*) has been seen to mutate to the resistant *pc* allele during meiosis at a rate of once per 7876 gametes (Schertz and Tai 1969). Meiotic instability has also been observed in other plant resistance genes. Various alleles of the *Rp1* resistance gene in maize mutate to susceptibility at a rate between 1/535 and 1/20729 gametes (Bennetzen et al. 1988). Another rust resistance gene in maize, *Rp3*, also mutates spontaneously at a rate comparable to that of *Pc* and some of the *Rp1* alleles. A heterozygous (*Rp3/rp3*) population was screened for homozygous susceptible (*rp3/rp3*) mutant plants. One mutation occurred in every 4236 seedlings analyzed (Webb

et al. 2002). Both *Rp1* and *Rp3* are tandemly duplicated *R* genes, and the underlying molecular mechanism of their mutations is unequal recombination, often between different *Rp* gene copies (Bennetzen et al. 1988; Sudupak et al. 1993; Webb et al. 2002). Unequal crossing over has been found to be a major source of genetic variation in many tandemly arranged gene families, such as the genes that specify ribosomal RNA, the major histocompatibility complex in animals, and disease resistance genes in plants (Williams et al. 1990; Venkateswarlu et al. 1991; Ohta 1991; Bennetzen and Hulbert 1992). The mutation rate of the *Pc* gene supports the idea that the *Pc* locus is a tandemly duplicated gene family that undergoes frequent unequal recombination events.

Two tandem duplicated gene families were identified in the *Pc* region, the NBS-LRR genes and the putative amino acid selective channel protein genes. However, the genotype that was the source of BAC Sbb12448, BTx623, is of the *pc/pc* genotype. Hence, genes that may have been duplicated in the presumed *Pc* ancestor of this current *pc* genotype could have been reduced to single copies by an earlier unequal recombination event to give rise to a resistant genotype. Thus, all genes in the sequenced *Pc*-region were considered to be reasonable *Pc* candidates.

The SSCP analysis of parental *Pc* and derived *pc* isolines used amplification products from every gene identified in the *Pc*-region. In every case except the NBS-LRR gene family, the SSCP results were identical for *Pc* and *pc* isolines. This suggests that none of the other genes could be the *Pc* locus, unless the *Pc* gene that was lost in the *Pc* to *pc* mutation carried no sequence polymorphism that could be detected by SSCP. The lower number of gene copies identified for the NBS-LRR genes in all of the *pc* isolines suggests that it is the loss of one of these genes by unequal crossing over that is responsible for the *Pc* to *pc* transition. This analysis falls short of a proof; however, because there may be one or more additional genes that our sequencing and/or annotation have not detected in the *Pc*-region, and these might also have been lost in a *Pc* to *pc* mutation that included the loss of one or more NBS-LRR gene copies. Confirmational analysis will require complete sequencing of the *Pc*-region of the parental *Pc* and the *pc* isolines derived from *Pc/Pc* Colby.

The toxin-resistance genes characterized so far in plants do not fall into *R* gene classes. However, some properties of other, yet unidentified resistance genes suggest that classical *R* genes can act as toxin susceptibility genes. The fungus *C. victoriae* causes the disease Victoria blight in oats through production of the toxin victorin. The host gene conferring susceptibility to victorin seems to be identical to the *Pc-2* resistance gene that acts against crown rust caused by the fungus *Puccinia coronata* (Wolpert et al. 2002). All of the characterized rust resistance genes in various plant species are classical *R* genes, mainly NBS-LRR genes (e.g., Keller et al. 2005; Collins et al. 1999; Webb et al. 2002). This suggests that the *Pc-2* gene is an *R* gene with dual pathogen specificity. Future investigation of the *Pc* gene in sorghum and the *Pc-2* gene in oat will provide important results about the role of *R* genes in toxin susceptibility and resistance.

In the present paper, we mapped and characterized a genomic region in sorghum that contains the PC-toxin resistance locus. Twelve *Pc* gene candidates were identified, and *pc* isoline analysis by SSCP suggested that an NBS-LRR gene family member is the *Pc* gene. In the future, the *Pc*-regions from isogenic *Pc* and *pc* lines will need to be cloned and structurally analyzed. Comparison of the wild type and mutated gene sequences will reveal the molecular nature of the frequent mutation events in the *Pc* locus and the precise toxin-sensitivity component of the *Pc* genotype.

Acknowledgments This article is dedicated to the memory of the late Dr. Keith Schertz, a much-valued and -missed colleague, mentor and friend. This work was supported by Umbarger and Giles Chair endowments to JLB. The technical assistance of Trudi Thomas and Matt Estep is gratefully acknowledged.

References

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410

Baldi P, Grossi M, Pecchioni N, Vale G, Cattivelli L (1999) High expression level of a gene coding for a chloroplastic amino acid selective channel protein is correlated to cold acclimation in cereals. Plant Mol Biol 41:233–243

Bedell JA, Budiman MA, Nunberg A, Citek RW, Robbins D, Jones J, Flick E, Rohlfing T, Fries J, Bradford K, McMenamy J, Smith M, Holeman H, Roe BA, Wiley G, Korf IF, Rabinowicz PD, Lakey N, McCombie WR, Jeddeloh JA, Martienssen RA (2005) Sorghum genome sequencing by methylation filtration. PLoS Biol 3:103–115

Bennetzen JL, Hulbert SH (1992) Organization, instability and evolution of plant disease resistance genes. Plant Mol Biol 20:575–578

Bennetzen JL, Qin MM, Ingels S, Ellingboe AH (1988) Allele-specific and *Mutator*-associated instability at the *Rpl* disease-resistance locus of maize. Nature 332:369–370

Bowers JE, Abbey C, Anderson S, Chang C, Draye X, Hoppe AH, Jessup R, Lemke C, Lennington J, Li Z, Lin Y, Liu S, Luo L, Marler BS, Ming R, Mitchell SE, Qiang D, Reischmann K, Schulze SR, Skinner DN, Wang Y, Kresovich S, Schertz KF, Paterson AH (2003) A high-density genetic recombination map of sequence-tagged sites for *Sorghum*, as a framework for comparative structural and evolutionary genomics of tropical grains and grasses. Genetics 165:367–386

Brandwagt BF, Mesbah LA, Takken FLW, Laurent PL, Kneppers TJA, Hille J, Nijkamp HJJ (2000) A longevity assurance gene homolog of tomato mediates resistance to *Alternaria alternata* f. sp. *lycopersici* toxins and fumonisin B_1. Proc Natl Acad Sci USA 97:4961–4966

Burge C, Karlin S (1997) Prediction of complete gene structures in human genomic DNA. J Mol Biol 268:78–94

Collins N, Drake J, Ayliffe M, Sun Q, Ellis J, Hulbert S, Pryor T (1999) Molecular characterization of the maize *Rp1-D* rust resistance haplotype and its mutants. Plant Cell 11:1365–1376

Deu M, Rattunde F, Chantereau J (2006) A global view of genetic diversity in cultivated sorghums using a core collection. Genome 49:168–180

Dixon MS, Jones DA, Keddie JS, Thomas CM, Harrison K, Jones JDG (1996) The tomato *Cf-2* disease resistance locus comprises two functional genes encoding leucine-rich repeat proteins. Cell 84:451–459

Dubcovsky J, Dvorak J (1995) Ribosomal RNA multigene loci: nomads of the triticeae genomes. Genetics 140:1367–1377

Dubcovsky J, Ramakrishna W, SanMiguel PJ, Busso CS, Yan L, Shiloff BA, Bennetzen JL (2001) Comparative sequence analysis of colinear barley and rice bacterial artificial chromosomes. Plant Physiol 125:1342–1353

Dunkle LD (1979) Heterogeneous reaction of shattercane to *Periconia circinata* and its host-specific toxin. Phytopathology 69:260–262

Dunkle LD, Macko V (1995) Peritoxins and their effects on sorghum. Can J Bot 73:S444–S452

Ewing B, Hillier L, Wendl M, Green P (1998) Basecalling of automated sequencer traces using phred. I. Accuracy assessment. Genome Res 8:175–185

Gale MD, Devos KM (1998) Plant comparative genetics after 10 years. Science 282:656–659

Gordon D, Abajian C, Green P (1998) Consed: a graphical tool for sequence finishing. Genome Res 8:195–202

He W, Parker R (2000) Functions of lsm proteins in mRNA degradation and splicing. Curr Opin Cell Biol 12:346–350

Hulbert SH, Richter TE, Axtell JD, Bennetzen JL (1990) Genetic mapping and characterization of sorghum and related crops by means of maize DNA probes. Proc Natl Acad Sci USA 87:4251–4255

Iftode C, Daniely Y, Borowiec JA (1999) Replication protein A (RPA): the eukaryotic SSB. Crit Rev Biochem Mol Biol 34:141–180

Johal GS, Briggs SP (1992) Reductase activity encoded by the *HM1* disease resistance gene in maize. Science 258:985–987

Keller B, Feuillet C, Yahiaoui N (2005) Map-based isolation of disease resistance genes from bread wheat: cloning in a supersize genome. Genet Res 85:93–100

Kim JS, Islam-Faridi MN, Klein PE, Stelly DM, Price HJ, Klein RR, Mullet JE (2005a) Comprehensive molecular cytogenetic analysis of sorghum genome architecture: distribution of euchromatin, heterochromatin, genes and recombination in comparison to rice. Genetics 171:1963–1976

Kim JS, Klein PE, Klein RR, Price HJ, Mullet JE, Stelly DM (2005b) Chromosome identification and nomenclature of *Sorghum bicolor*. Genetics 169:1169–1173

Klein PE, Klein RR, Cartinhour SW, Ulanch PE, Dong JM, Obert JA, Morishige DT, Schlueter SD, Childs KL, Ale M, Mullet JE (2000) A high-throughput AFLP-based method for constructing integrated genetic and physical maps: progress toward a sorghum genome map. Genome Res 10:789–807

Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ, Lincoln SE, Newburg L (1987) MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1:174–181

Leister D, Kurth J, Laurie DA, Yano M, Sasaki T, Devos K, Graner A, Schulze-Lefert P (1998) Rapid reorganization of resistance gene homologues in cereal genomes. Proc Natl Acad Sci USA 95:370–375

Leukel RW (1948) *Periconia circinata* and its relation to milo disease. J Agric Res 77:201–222

Levings CS, Rhoads DM, Siedow JN (1995) Molecular interactions of *Bipolaris maydis* T-toxin and maize. Can J Bot 73:S483–S489

Lurin C, Andres C, Aubourg S, Bellaoui M, Bitton F, Bruyere C, Caboche M, Debast C, Gualberto J, Hoffmann B, Lecharny A, Le Ret M, Martin-Magniette ML, Mireau H, Peeters N, Renou JP, Szurek B, Taconnat L, Small I (2004) Genome-wide analysis of Arabidopsis pentatricopeptide repeat proteins reveals their essential role in organelle biogenesis. Plant Cell 16:2089–2103

Macko V, Stimmel MB, Wolpert TJ, Dunkle LD, Acklin W, Bänteli R, Jaun B, Arigoni D (1992) Structure of the host-specific toxins produced by the fungal pathogen *Periconia circinata*. Proc Natl Acad Sci USA 89:9574–9578

Martins-Lopes P, Zhang H, Koebner R (2001) Detection of single nucleotide mutations in wheat using single strand conformation polymorphism gels. Plant Mol Biol Rep 19:159–162

Meeley RB, Walton JD (1991) Enzymatic detoxification of HC-toxin, the host-selective cyclic peptide from *Cochliobolus carbonum*. Plant Physiol 97:1080–1086

Murray MG, Thomson WF (1980) Rapid isolation of high molecular weight plant DNA. Nucliec Acids Res 8:4321–4325

Nagy ED, Lelley T (2003) Genetic and physical mapping of sequence-specific amplified polymorphic (SSAP) markers on the 1rs chromosome arm of rye in a wheat background. Theor Appl Genet 107:1271–1277

Nimchuk Z, Eulgem T, Holt III BF, Dangl JL (2003) Recognition and response in the plant immune system. Annu Rev Genet 37:579–609

Ohta T (1991) Multigene families and the evolution of complexity. J Mol Evol 33:34–41

Oliveira AC, Richter T, Bennetzen JL (1996) Regional and racial specificities in sorghum germplasm assessed with DNA markers. Genome 39:579–587

Orita M, Suzuki Y, Sekiya T, Hayashi K (1989) Rapid and sensitive detection of point mutations and DNA polymorphisms using the polymerase chain reaction. Genomics 5:874–879

Ramakrishna W, Dubcovsky J, Park YJ, Busso C, Emberton J, SanMiguel P, Bennetzen JL (2002a) Different types and rates of genome evolution detected by comparative sequence analysis of orthologous segments from four cereal genomes. Genetics 162:1389–1400

Ramakrishna W, Emberton J, SanMiguel P, Ogden M, Llaca V, Messing J, Bennetzen JL (2002b) Comparative sequence analysis of the sorghum *Rph* region and the maize *Rp1* resistance gene complex. Plant Physiol 130:1728–1738

Rozen S, Skaletsky HJ (2000) Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S (eds) Bioinformatics methods and protocols: methods in molecular biology. Humana Press, Totowa, NJ, pp 365–386

Schertz KF, Tai YP (1969) Inheritance of reaction of *Sorghum bicolor* (L.) Moench to toxin produced by *Periconia circinata* (Mang.) sacc. Crop Sci 9:621–624

Schweighofer A, Hirt H, Meskiene I (2004) Plant PP2C phosphatases: emerging functions in stress signaling. Trends Plant Sci 9:236–243

Sandhu D, Gill KS (2002) Gene-containing regions of wheat and the other grass genomes. Plant Physiol 128:803–811

Solovyev VV, Salamov AA (1997) The gene-finder computer tools for analysis of human and model organisms genome sequences. Intell Syst Mol Biol 5:294–302

Song WY, Wang GL, Chen LL, Kim HS, Pi LY, Holsten T, Gardner J, Wang B, Zhai WX, Zhu LH, Fauquet C, Ronald P (1995) A receptor kinase-like protein encoded by the rice disease resistance gene, *Xa21*. Science 270:1804–1806

Spassieva SD, Markham JE, Hille J (2002) The plant disease resistance gene *Asc-1* prevents disruption of sphingolipid metabolism during AAL-toxin-induced programmed cell death. Plant J 32:561–572

Sturtevant AH (1925) The effects of unequal crossing over at the bar locus in Drosophila. Genetics 10:117–147

Sudupak MA, Bennetzen JL, Hulbert SH (1993) Unequal exchange and meiotic instability of disease-resistance genes in the *Rp1* region of maize. Genetics 133:119–125

Venkateswarlu K, Lee SW, Nazar RN (1991) Conserved upstream sequence elements in plant 5S ribosomal RNA-encoding genes. Gene 105:249–253

Vercauteren I, de Almeida Engler J, De Groodt R, Gheysen G (2002) An *Arabidopsis thaliana* pectin acetylesterase gene is upregulated in nematode feeding sites induced by root-knot and cyst nematodes. Mol Plant Microbe Interact 15:404–407

Waugh R, McLean K, Flavell AJ, Pearce SR, Kumar A, Thomas BBT, Powell W (1997) Genetic distribution of BARE-1-like retrotransposable elements in the barley genome revealed by sequence-specific amplification polymorphisms (S-SAP). Mol Gen Genet 253:687–694

Webb CA, Richter TE, Collins NC, Nicolas M, Trick HN, Pryor T, Hulbert SH (2002) Genetic and molecular characterization of the maize *rp3* rust resistance locus. Genetics 162:381–394

Whitkus R, Doebley J, Lee M (1992) Comparative genome mapping of sorghum and maize. Genetics 132:1119–1130

Williams SM, Robbins LG, Cluster PD, Allard RW, Strobeck C (1990) Superstructure of the Drosophila ribosomal gene family. Proc Natl Acad Sci USA 87:3156–3160

Wolpert TJ, Dunkle LD (1983) Alterations in gene expression in sorghum induced by the host-specific toxin from *Periconia circinata*. Proc Natl Acad Sci USA 80:6576–6580

Wolpert TJ, Dunkle LD, Ciuffetti LM (2002) Host-selective toxins and avirulence: what's in a name? Annu Rev Phytopathol 40:251–285

Woo SS, Jiang J, Gill BS, Paterson AH, Wing RA (1994) Construction and characterization of a bacterial artificial chromosome library of *Sorghum bicolor*. Nucleic Acids Res 22:4922–4931