Epigenetic variation in *Arabidopsis* disease resistance

Trevor L. Stokes, Barbara N. Kunkel, and Eric J. Richards

Department of Biology, Washington University, St. Louis, Missouri 63130, USA

Plant pathogen resistance is mediated by a large repertoire of resistance (*R*) genes, which are often clustered in the genome and show a high degree of genetic variation. Here, we show that an *Arabidopsis thaliana* *R*-gene cluster is also subject to epigenetic variation. We describe a heritable but metastable epigenetic variant *bal* that overexpresses the R-like gene *At4g16890* from a gene cluster on Chromosome 4. The *bal* variant and *Arabidopsis* transgenics overexpressing the *At4g16890* gene are dwarfed and constitutively activate the salicylic acid (*SA*)-dependent defense response pathway. Overexpression of a related *R*-like gene also occurs in the *ssi1* (*suppressor of *SA* insensitivity 1*) background, suggesting that *ssi1* is mechanistically related to *bal*.

**Key Words**: Epigenetics; cytosine methylation; resistance genes; *bal*; DDM1; *Arabidopsis*

Received October 5, 2001; revised version accepted November 26, 2001.

Epigenetic variation is often overlooked as a source of phenotypic variation for natural or artificial selection. Variation in epigenetic information, encoded at the chromatin level rather than the nucleotide sequence level, is commonly thought to be transient and unlikely to underlie stable changes in phenotype. There is considerable evidence, however, that epigenetic changes, particularly those due to alterations in DNA methylation, can be inherited through meiosis and mimic traditional mutations.

Some of the earliest recognized heritable epigenetic alleles involve transposons whose differential DNA methylation and chromatin structure affect the expression of neighboring genes [McClelland 1951; Martienssen et al. 1990; Martienssen and Baron 1994; Fedoroff et al. 1995]. Another classic example of epigenetic inheritance is paramutation, wherein interactions between alleles cause chromatin-mediatedheritable changes in gene expression [Kermicle et al. 1995; Hollick et al. 1997; Walker 1998]. Several groups have recently reported defects in flower development caused by heritable epigenetic alleles [epialleles] associated with abnormal DNA methylation. Hypermethylated epialleles of *AGAMOUS* [Jacobsen et al. 2000] and *SUPERMAN* [Jacobsen and Meyerowitz 1997; Jacobsen et al. 2000], which affect flower structure, and hypomethylated epialleles of *FWA* [Soppe et al. 2000], which delay flowering time, have been recovered from both mutagenized *Arabidopsis* populations and DNA hypomethylation lines such as *ddm1*, *met1*, and antisense-cytosine methyltransferase *MET1* [Finnegan et al. 1996; Ronemus et al. 1996]. Plant epimutations associated with altered DNA methylation may also occur spontaneously. Heritable, differentially methylated *P1* alleles that condition altered kernel pigmentation were reported in maize [Das and Messing 1994]. Bender and colleagues have shown that an endogenous inverted repeat can induce methylation and silencing of homologous gene sequences (*pal2*) at an unlinked site in the *Arabidopsis* genome [Bender and Fink 1995; Jeddelloh et al. 1998; Luff et al. 1999; Melquist et al. 1999]. Recently, a naturally occurring floral variant of the plant *Linaria* was discovered to be caused by a hypermethylated, silenced *Leyc* allele [Cubas et al. 1999]. Fungi [Colot et al. 1996; Irelan and Selker 1997; Selker 1997] and animals [Michaud et al. 1994; Morgan et al. 1999] also show heritable epigenetic variation associated with differential DNA methylation.

We have been studying the effects of genomic hypomethylation using mutations in the *DDM1* (*DECREASE IN DNA METHYLATION 1*) gene of *Arabidopsis*. *ddm1* mutations lead to a reduction in cytosine methylation throughout the genome [Vongs et al. 1993]. Repeated DNA loses methylation immediately in *ddm1* mutants, but low copy sequences gradually lose methylation as *ddm1* mutants are propagated through successive generations by self-pollination [Kakutani et al. 1996]. The *DDM1* locus encodes a SWI2/SNF2-like protein that resembles chromatin remodeling enzymes in yeast, *Drosophila*, and humans [Jeddelloh et al. 1999].

*ddm1* loss-of-function mutations lead to developmental abnormalities including late flowering, dwarfing, and altered floral structures [Kakutani et al. 1996; Jacobsen et al. 2000]. Typically, these defects are only observed after self-pollinating *ddm1* mutants for several generations. *Arabidopsis* normally is a self-pollinator, but *ddm1* mu-
tants display the inbreeding depression typical of many outcrossing species. The *ddm1*-induced developmental defects are caused by stable alterations at sites unlinked to the *ddm1* mutation (Kakutani et al. 1996). These alterations behave like classical Mendelian mutations and are stable when segregated from *ddm1*. Some of these heritable alterations are caused by insertion of transgenesis-related (nucleotide binding site–leucine-rich repeat) class–like proteins (Jones 1996; Glazebrook 1999). There are R genes and multiple leucine-rich repeat (LRR) motifs (Hammond-Kosack and Jones 1997). The most common class of *R* genes encode proteins that recognize, either directly or indirectly, specific signals or gene products produced by pathogens, leading to the activation of defense signal transduction pathways (Ellis et al. 2000). The most common class of *R*-gene products consists of both a nucleotide-binding site (NBS) and multiple leucine-rich repeat (LRR) motifs (Hammond-Kosack and Jones 1997). There are ~165 NBS-LRR (nucleotide binding site–leucine-rich repeat) class *R*-like genes in the Arabidopsis genome (for details, see http://www.nibr.srs.ucdavis.edu). One of the major induced pathways is dependent on salicylic acid (SA), a small aromatic compound related to aspirin (Durner et al. 1997). SA-dependent signaling induces *PR* gene expression, which is correlated with elevated resistance (Uknes et al. 1992; Maleck et al. 2000). Several modifiers of SA signaling have been identified by mutations causing either constitutive expression of *PR* genes (e.g., *cpr* [Bowling et al. 1994], *ssi* [Shah et al. 1999]) or nonexpression of *PR* genes (*npr1*; Cao et al. 1994, 1997; Ryals et al. 1997).

Our results indicate that *bal* is an epigenetic alteration mapping to a cluster of NBS-LRR-class disease-resistance genes. In the *bal* variant, overexpression of one gene in the cluster stimulates the disease response pathway and causes dwarfing and elevated disease resistance. Overexpression of a related NBS-LRR gene occurs in *ssi*1 (suppressor of SA insensitivity 1) plants, which constitutively express *PR* genes (Shah et al. 1999). The overexpression phenotype and other genetic parallels suggest that the *ssi1* mutation is mechanistically related to the *bal* variant.

Results

*bal* is a heritable *ddm1*-induced defect

Inbred *Arabidopsis* lines deficient in DNA methylation accumulate a spectrum of phenotypes, including late flowering, floral structure defects, and dwarfing (Kakutani et al. 1996; Jacobsen et al. 2000). To understand the molecular basis of such developmental abnormalities, we focused on a dwarfing variant (*bal*) generated in an inbred *ddm1* hypomethylation mutant line (strain Columbia). The *bal* phenotype is characterized by twisted leaves, dwarfed stature, and reduced fecundity (Fig. 1; Kakutani et al. 1996). We crossed a phenotypic *bal ddm1* plant to wild-type Columbia and Landsberg erecta plants and generated segregating *F2* populations. In these *F2* families, the *bal* phenotype segregated as a single Mendelian trait independent of the *ddm1* mutation (Kakutani et al. 1996). The Columbia *bal* defect behaved as a semidominant allele relative to the wild-type Columbia *BAL* allele (Fig. 1A), but acted in a recessive manner relative to the wild-type Landsberg erecta *BAL* allele (Kakutani et al. 1996).

We followed the stability of the *bal* defect in four independent *bal DDM1* lines propagated through five generations of self-pollination, and we observed no phenotypic reversion when the plants were grown under normal conditions. We conducted all subsequent phenotypic and molecular characterization of the *bal* defect using *DDM1* wild-type backgrounds from Columbia and Landsberg erecta.

The *bal* variant resembles constitutive pathogen defense mutants

The morphology of the *bal* variant resembles *Arabidopsis* mutants that constitutively express pathogen defense genes (e.g., *sis* [Shah et al. 1999]). *Sis* is a semidominant allele relative to the wild-type Columbia *SIS* allele, but acted in a recessive manner relative to the wild-type Landsberg erecta *SIS* allele (Kakutani et al. 1996).

Figure 1. Phenotype of the semidominant *bal* defect. (A) Twenty-five-day-old individuals (Columbia strain, wild-type *DDM1*) from a population segregating the *bal* alteration. Heterozygous *bal/bal* individuals show severe dwarfing and twisted leaves. Heterozygous individuals are intermediate in size and leaf twisting. (B) A higher-magnification view of *bal/bal* homozygotes compared with *cpr1-1/cpr1-1* homozygotes (Columbia strain, 4 wk old).
genes (e.g., cpr 1 [Bowling et al. 1994], mpk4 [Petersen et al. 2000], and ssi1 [Shah et al. 1999]). Figure 1B compares the morphology of bal homozygotes and plants homozygous for the cpr1 mutation. cpr1 was isolated from a chemical mutagenesis screen for plants that constitutively express PR genes (Bowling et al. 1994). The morphological similarities prompted us to determine whether the bal variant also constitutively expresses PR genes in the absence of pathogen. Figure 2A shows that PR2 transcripts are expressed in uninfected bal and cpr1 homozygotes, but are absent in wild-type siblings. PR1 and PR5 transcripts were also expressed in uninfected bal [Fig. 2B] and cpr1 plants [data not shown.] PR2 transcript is detectable in BAL/bal heterozygotes, but at significantly lower levels compared with bal homozygotes [Fig. 2B]. Therefore, the bal variant is a constitutive expresser of PR genes.

Decreased pathogen growth and dwarf morphology in the bal variant are dependent on salicylic acid signaling

Because the bal variant constitutively expresses PR genes, we sought to determine if the bal alteration affects resistance to pathogens. We investigated the growth of the bacterial pathogen, Pseudomonas syringae (pv. tomato strain DC 3000), to which Arabidopsis strain Columbia is susceptible [Whalen et al. 1991]. Figure 3A shows that this virulent P. syringae strain grows ∼3 orders of magnitude on wild-type Columbia hosts by 4 d postinoculation. Bacterial growth on bal homozygotes is attenuated; bacterial populations in bal homozygous leaf tissue are reduced 10- to 100-fold [Fig. 3A; data not shown] relative to populatons in wild-type leaves at 4 d postinoculation.

Many plant defense responses are mediated through salicylic acid (SA)-dependent signaling pathways. Therefore, we examined whether the morphological and pathogenesis defects shown by bal homozygotes could be suppressed by disrupting SA signaling. We crossed the bal variant [strain Columbia] to an nahG transgenic line [strain Landsberg erecta; Delaney et al. 1994] expressing a microbial enzyme that inactivates SA. P. syringae (DC 3000) grows to a higher density on nahG transgenic tissue relative to wild-type Columbia tissues [Fig. 3A] and Landsberg erecta [data not shown]. Bacterial growth in nahG bal homozygous tissue is also significantly higher than growth in either bal or wild-type plants. As shown in Figure 3B, nahG also significantly suppresses bal morphological defects, although the nahG bal plants retain some characteristics of the bal variant, such as mildly twisted leaves and reduced stature. These findings indicate that the full phenotypic expression of the bal variant is dependent on SA signaling.

The bal defect is metastable

Attempts to induce suppressor/revertant mutations in a bal DDM1 background resulted in the recovery of an unexpectedly high frequency of phenotypically normal plants or individuals with dramatically reduced bal phenotypes. We treated bal DDM1 seeds (M1) with either the alkylating agent ethylmethylenesulfonate (EMS) or γ-ray irradiation [see Materials and Methods]. We recovered >10% normal or phenotypically intermediate plants among the progeny [M2] of the plants grown from the EMS-treated seeds [Fig. 4, Table 1]. In the γ-ray treated experiment, ~7% of the M2 plants showed a wild-type or weakened bal phenotype. The mock-treated M2 populations contained no phenotypic revertants. The observed phenotypic reversion of bal in both experiments is at least 20-fold greater than the expected recessive knockout frequency in a comparable Arabidopsis mutagenesis (~0.1% of the M2 population; Haughn and Somerville 1987). Five phenotypically normal M2 plants were outcrossed to wild-type Columbia plants, and no segregation of phenotypically normal bal plants was seen in F2 generations among a total of 2040 F2 individuals examined. These results suggest that the restoration of the normal phenotype is caused by reversion of the bal allele rather than by the action of unlinked extragenic suppressors.
Characterization of the BAL locus

The high frequency of bal reversion is consistent with either hypermutable genetic mechanisms [e.g., transposition, DNA rearrangements] or metastable epigenetic alterations [e.g., chromatin-controlled gene expression states]. To determine the basis of the bal variant, we first sought to molecularly identify the BAL locus. We mapped the bal alteration to a small genetic interval (<0.2 cM) on the lower arm of Chromosome 4 (see Fig. 5A). This genetic interval corresponds to a 152-kb window. The region has an ∼4-fold increase in the ratio of physical to genetic distance, which in Arabidopsis normally averages ∼200 kb/cM (Copenhaver et al. 1998). The observed recombination suppression is caused by strain-specific genomic rearrangements within the region (Parker et al. 1997; Noël et al. 1999).

A large portion of the physical window containing the bal alteration is composed of a 90-kb gene cluster that contains 10 genes or gene fragments encoding NBS-LRR-class disease-resistance proteins (Parker et al. 1997; Bevan et al. 1998). Jones and colleagues discovered that the functional RPP5 [resistance to the downy mildew pathogen Peronospora parasitica] gene was embedded within this gene cluster in the Landsberg erecta strain (Parker et al. 1997). The Columbia version of the NBS-LRR gene cluster contains two apparently full-length, potentially functional disease-resistance genes. The eight remaining NBS-LRR genes are fragments, or are interrupted by mutations or retrotransposon insertions. The 152-kb window defined by our mapping contains 22 additional predicted genes (Bevan et al. 1998; http://mips.gsf.de).

Gene expression profile in the bal variant

We searched the 152-kb window for molecular changes correlated with the bal phenotype. We first screened for genomic alterations using Southern blots and hybridization probes tiling across the window. We found no DNA rearrangements in this comparison of bal and wild-type BAL samples. In particular, we did not detect structural polymorphisms involving the two retrotransposable elements inserted into the NBS-LRR gene cluster (data not shown).

We next examined the transcripts originating from the NBS-LRR gene cluster using Northern blot analysis. We detected no transcripts from the two retrotransposons in the NBS-LRR cluster in either wild-type Columbia or bal homozygotes. However, hybridization probes corresponding to exons of the NBS-LRR genes detected a 5-kb transcript in wild-type BAL Columbia plants. This transcript species was overexpressed fourfold in bal compared with BAL homozygotes (Fig. 6A).
We investigated the correlation between the abundance of the 5-kb transcript and the phenotypic severity of the \textit{bal} defect by measuring transcript levels in EMS-induced revertants. As shown in Figure 6A, the abundance of the 5-kb transcript was decreased to approximately wild-type levels in the EMS-induced revertants lacking \textit{bal} variant phenotypes. Similar results were seen for the \textit{\textgamma}-ray-induced phenotypic revertants (data not shown). BAL/bal heterozygotes, which displayed intermediate phenotypes, contained an intermediate level of the 5-kb transcript (Fig. 6B). The levels of the 5-kb transcript were not elevated in nonphenotypic \textit{ddm1} plants (data not shown).

**Expression of genes in the NBS-LRR gene cluster**

Because the NBS-LRR coding sequences in the region share 90\% or greater nucleotide identity, Northern analysis was unable to distinguish among the different genes in the region. We took four approaches to identify which genes are expressed from the NBS-LRR cluster in wild-type Columbia (see Fig. 5B). First, a search of EST databases identified cDNA clones originating from four genes in the region (\textit{At4gxxxxx}: 16890, 16900, 16950, and 16990). In addition, five EST matches were found for the two retrotransposons inserted into the penultimate 3' intron of \textit{At4g16860} and \textit{At4g16900}. We also screened a wild-type Columbia cDNA library (Kieber et al. 1993) and isolated six additional clones that matched three genes: 16890, 16900, and 16950. As a supplementary approach, we designed RT-PCR primers to amplify products corresponding to transcripts from the most intact NBS-LRR coding sequences in the cluster. Expression from individual NBS-LRR genes in the cluster was detected by RFLP or sequence analysis of the RT-PCR-amplified products. Sequence analysis of nine RT-PCR clones indicated that the major transcript in wild-type Columbia originates from the 16890 gene (5 of 9 clones); we also found expression from 16900 (3 of 9 clones) and 16950 (1 of 9 clones). RFLP analysis of bulk RT-PCR products supported the conclusion that the major transcript source in wild-type Columbia is the 16890 gene (data not shown). A parallel RT-PCR RFLP profile indicated that 16890 is also the major source of transcript in \textit{bal} variant plants (data not shown). These analyses indicate that the 5-kb transcript detected on Northern blots (hereafter, referred to as the \textit{BAL} transcript) originates from the 16890 gene. The 16890 gene is one of the two genes in the cluster predicted to produce a full-length NBS-LRR-class protein.

**The DNA sequence of the \textit{At4g16890} gene is unchanged in the \textit{bal} variant**

To see if the observed overexpression of the 16890 gene was caused by a mutation, we determined the nucleotide sequence of a 7-kb region from the \textit{bal} variant encompassing the coding sequence of the gene and the entire upstream region up to the next annotated gene, 16900 (data not shown). No sequence changes were found. These data support the argument that the increased expression of the 16890 gene is not caused by a nucleotide sequence change.

**Overexpression of the \textit{At4g16890} gene phenocopies the \textit{bal} defect**

We then asked if overexpression of the 16890 gene causes the \textit{bal} phenotype. Using \textit{Agrobacterium}-mediated T-DNA transformation, we constructed three independent transgenic lines (strain Columbia) containing

**Table 1.** Chemical and physical mutageneses cause \textit{bal} reversion at a high frequency

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Wild-type</th>
<th>Intermediate</th>
<th>\textit{bal}</th>
<th>\textit{n}</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 mM EMS</td>
<td>100%</td>
<td>88.7%</td>
<td>100%</td>
<td>1725</td>
</tr>
<tr>
<td>30 mM EMS</td>
<td>2.5%</td>
<td>8.8%</td>
<td>120%</td>
<td>1102</td>
</tr>
<tr>
<td>0 kRad \textit{\textgamma}-ray</td>
<td>1.6%</td>
<td>5.6%</td>
<td>92.8%</td>
<td>625</td>
</tr>
<tr>
<td>120 kRad \textit{\textgamma}-ray</td>
<td>100%</td>
<td>90%</td>
<td>525</td>
<td></td>
</tr>
</tbody>
</table>

\textit{bal/bal, DDM1/DDM1} seeds were treated with either ethylmethanesulfonate (EMS) or \textit{\textgamma}-rays, or mock-treated. The treated seeds (M1) were planted and M2 progeny were examined for the \textit{bal} phenotype. \textit{n} indicates the number of M2 progeny examined.
the 16890 gene driven by the strong viral promoter, P35S. In the first generation \(T_1\), we recovered 35S::BAL transgenic plants that phenocopied the dwarfing and twisted leaves of the bal variant. Control T1 transformants containing the vector construct without the 16890 gene always displayed a wild-type phenotype. These results indicate that the 35S::BAL transgene recapitulates the bal phenotype (see Fig. 7). In T2 families generated by self-pollination of phenotypic T1 35S::BAL transgenics, plants displaying the bal phenotype were recovered at a frequency of 55% [96 bal::79 wild-type phenotype]. The T2 segregation ratios are consistent with stochastic silencing of the 35S::BAL transgene, which is frequently noted in 35S-driven transgenes in plants.

Next, we examined the expression of the 16890 gene using Northern analysis in the T2 generation comparing individuals with wild-type or bal phenotypes. As shown in Figure 8, the level of the 5-kb BAL transcript was elevated in transgenic plants showing the bal phenotype. No increase in the abundance of the BAL transcript was noted in nonphenotypic segregants. Moreover, PR1 gene expression was elevated in 35S::BAL phenotypic individuals but was absent in nonphenotypic segregants. Based on progeny tests for the transgenic kanamycin-resistance marker, some of the nonphenotypic T2 individuals seg-
regated kanamycin-resistant seedlings in the T₃ generation, consistent with the hypothesis that silencing of the 35S::BAL transgene occurs. However, all T₂ bal phenotypic plants contained the 35S::BAL transgene and overexpressed the 5-kb BAL transcript.

The transgenic data support the hypothesis that overexpression of an NBS-LRR gene causes the bal phenotype. This result was corroborated by the phenotype of bal eds1-2 double mutants [data not shown]. EDS1 distinguishes TIR (toll interleucin receptor) and LZ (leucine zipper) classes of NBS-LRR gene products; eds1 mutations block signaling downstream of TIR, but not LZ, class R-genes (Aarts et al. 1998; Clarke et al. 2001). The eds1-2 mutation suppressed the bal morphology and constitutive PR gene expression phenotype [data not shown], consistent with the 16890 gene being a TIR-class R-gene.

**Overexpression of R-gene coding sequence in suppressor of SA insensitivity 1 (ssi1) variant**

There are several parallels between the phenotype and genetics of the bal variant and the ssi1 mutant. The ssi1 mutation was isolated in the Nossen background as an extragenic suppressor of an npr1 [nonexpresser of pathogenesis-related genes] mutation [Shah et al. 1999]. Like the bal variant, the phenotype of the ssi1 mutant is characterized by dwarfing, narrow leaves, and reduced fertility. The phenotypic effects of the ssi1 mutation and the bal variant are not dependent on the npr1 mutation [data not shown]. The ssi1 mutant shows constitutive PR gene expression and decreased growth of P. syringae. The ssi1 dwarfing and constitutive PR gene expression are suppressed by the presence of the nahG transgene, indicating that these phenotypes are SA-dependent. The ssi1 mutation is semidominant to the Nossen SSI1 allele and maps to the lower arm of Chromosome 4 close to the BAL locus (Shah et al. 1999).

**Figure 6.** Overexpression of an NBS-LRR gene is associated with the bal phenotype. (A) Northern blot analysis showing overexpression of an NBS-LRR-like gene in phenotypic bal variants. [BAL] BAL/BAL; [bal] bal/bal variant; [bal mock] phenotypic plants from the mock-treated EMS M2 population; [1, 2, 3] independent nonphenotypic revertant plants from an EMS-treated M2 population. [Bottom panel] RNA loading control. (B) Northern blot analysis showing that the NBS-LRR like transcript level correlates with phenotypic severity in a population segregating the bal variant allele. The autoradiogram resulted from hybridization of the membrane used in Figure 2.

**Figure 7.** Overexpression of Atâ₉16890 in transgenic plants phenocopies the bal variant. (Top) Control plants: (BAL/BAL) wild type; (bal/bal) homozygous bal variant; (BAL/BAL vector only) wild type transformed with the T-DNA vector [pMD1]. (Middle, bottom) Segregation of the dwarfing bal phenotype is shown by representative plants from two T₂ families segregating the 35S::BAL transgene. All plants pictured [Columbia strain background] are 24-day-old plants and were grown in parallel in an environmental chamber.
The genetic and phenotypic parallels between the bal and ssi1 defects led us to investigate the ssi1 variant further. First, we examined the stability of the ssi1 phenotype in response to EMS mutagenesis. We found a high frequency of phenotypically normal or intermediate plants in the M1 generation: 14% (50 mM EMS) and 9% (120 krad of H9253-irradiation) compared with 0% in mock-treated controls. Next, we investigated the expression within the Chromosome 4 NBS-LRR gene cluster in the ssi1 variant and the appropriate wild-type background. Figure 9 shows that in the Nossen background, the level of the transcript was increased in the ssi1 variant. The metastability of the ssi1 variant and its association with the overexpression of an R-like gene further suggest that the ssi1 and bal variants may be caused by the same underlying mechanism.

**Discussion**

Here, we show that a ddm1 mutation induced a stable alteration, bal, which activates plant defense responses in the absence of pathogens. Several lines of evidence suggest that bal is an epigenetic variant. First, the bal alteration arose in an inbred ddm1 hypomethylation background known to produce epigenetic variation (e.g., sup [Jacobsen et al. 2000], ag [Jacobsen et al. 2000], and fwa [Soppe et al. 2000] epialleles). Second, the bal alteration causes a change in gene expression that is not associated with large-scale genomic alterations or changes in nucleotide sequence of the implicated At4g16890 gene. Third, the bal alteration can be destabilized in response to DNA-damaging agents, consistent with previous demonstrations that mutagens can alter DNA methylation and produce epigenetic alleles that are either hypomethylated or hypermethylated (Riggs and Jones 1983; Wilson and Jones 1983; Jacobsen and Meyerowitz 1997; Soppe et al. 2000). Furthermore, our preliminary results indicate that the Chromosome 4 R-like gene cluster and associated LTR retrotransposons are methylated in wild-type Columbia, and that this region is hypomethylated in ddm1-inbred backgrounds (T.L. Stokes and E.J. Richards, unpubl.). We hypothesize that overexpression of At4g16890 in the bal variant is caused by heritable hypomethylation of critical sites in the region or is triggered by hypomethylation below a threshold level. Although we have found hypomethylated sites in the R-gene cluster of the bal variant, we have not yet identified methylation changes that are strictly correlated with At4g16890 overexpression.

In our study, overexpression of the At4g16890 R-like gene in both the bal variant and transgenic Arabidopsis induces defense responses in the absence of pathogen. Previous studies show that overexpression of a non-LRR-class resistance gene Pto or an LRR-NBS resistance gene Prf induces defense responses and enhances pathogen resistance in transgenic tomato (Oldroyd and Staskawicz 1998; Tang et al. 1999). In contrast to the bal variant and the 35S::BAL Arabidopsis transgenics, tomato plants overexpressing Pto or Prf fail to show dramatic whole-
plant defects. Recently, Xiao et al. (2001) reported that the overexpression of *RPW8*, a novel *R*-gene locus containing two genes, in *Arabidopsis* leads to broad-spectrum fungal resistance without obvious phenotypic abnormalities. The different phenotypic consequences of *R*-gene overexpression may reflect peculiarities of the different plant species and/or specific interactions of the different resistance genes. We note that some signaling mutants that affect steps downstream of *R*-genes show dwarfing and twisted leaf phenotypes (e.g., *cpr1* [Bowling et al. 1994], *mpk4* [Petersen et al. 2000]), which closely resemble the *bal* variant phenotype. *R*-gene products act to recognize, either directly or indirectly, specific pathogen strains. After recognition, *R*-gene products signal through one or more defense pathways (Glazebrook 1999). The *BAL* gene is transcribed in the absence of pathogen, and the level of the transcript is not increased by pathogen attack (data not shown). Most characterized NBS-LRR *R*-genes are expressed at low constitutive levels (e.g., Meyers et al. 1999) similar to that seen for the *BAL* transcript. Mild developmental phenotypes were seen in *BAL/bal* heterozygotes that express ~2.5-fold higher levels of the *BAL* transcript. Severe morphological abnormalities resulted from a fourfold overexpression of the *BAL* transcript, levels similar to the highest seen in recovered 35S::*BAL* transgenics. These results suggest that the expression of the *BAL* transcript may be constrained to an upper limit of expression.

Our results also address the mechanism of the *ssi1* variant, which was originally identified in the Nossen background as a suppressor of *npr1-5* (Shah et al. 1999). Several pieces of evidence indicate that *bal* and *ssi1* variants are caused by the same mechanism. First, both *bal* and *ssi1* are semidominant alterations that cause dwarfing and activation of defense pathways independent of *NPR1* function. Second, the *ssi1* mutation maps to the same genetic region (<5 cM) as the *BAL* locus. Third, both *bal* and *ssi1* overexpress an *R*-like gene from the cluster mapping to the bottom arm of Chromosome 4. Finally, both the *bal* and *ssi1* variants are destabilized by mutants. Our results suggest that the *ssi1* phenotype is caused by overexpression of an LRR-NBS *R*-like gene, which may be tied to a metastable epigenetic alteration.

The semidominant nature of the *bal* and *ssi1* alterations is consistent with the proposed overexpression mechanism. *BAL/bal* heterozygotes expressed ~2.5 times the level of *BAL* transcript compared to wild-type *BAL/BAL* homozygotes in a Columbia background. Interestingly, the Columbia *bal* alteration behaved as a recessive allele when crossed into a wild-type Landsberg erecta background. The *RPP5* NBS-LRR haplotype in Landsberg erecta is rearranged relative to Columbia, and we have not detected a corresponding *BAL* transcript from the Landsberg locus (data not shown). Combination of a 4x-overexpressing Columbia *bal* allele with a null Landsberg allele is not sufficient to exceed a threshold for phenotypic expression. On the other hand, combination of a 4x-overexpressing Columbia *bal* allele with a 1x-expressing Columbia *BAL* allele exceeds this thresh-

old, leading to mild phenotypic consequences. We note that the Columbia *bal* allele in combination with a wild-type Nossen allele acted semidominantly, consistent with the idea that the wild-type Nossen allele can specify some *R*-like gene function (data not shown). The overexpression model of *ssi1* is also consistent with recent findings that *SSI1/SSI1/ssi1* triploid progeny (tetraploid Columbia × Nossen *ssi1* diploid) fail to show morphological phenotypes (Greenberg 2000). In this case, the presence of the two wild-type Columbia loci in a triploid background would dilute the *R*-like gene function below the critical threshold.

Resistance genes are highly reiterated in plants, and our results point toward the importance of epigenetic regulation of these gene clusters prevalent in plant genomes. For example, the *Arabidopsis* genome, once believed to be relatively simple in structure, contains ~37% of the predicted protein coding sequences in families of five members or more, and 17% of all predicted *Arabidopsis* genes are multicopy and tandemly arrayed (The *Arabidopsis* Genome Initiative 2000). Our results suggest that epigenetic modification may play an important role in regulating gene clusters by cementing silent or intermediate expression states.

Recent results from several groups suggest that epigenetic alterations may play an important role in controlling phenotypic variation in both laboratory and natural populations. In *Arabidopsis* laboratory strains, epigenetic variation at several loci has been induced by both chemical mutagenesis and DNA hypomethylation (Jacobsen and Meyerowitz 1997; Jacobsen et al. 2000; Soppe et al. 2000). The induced epialleles described to date fall into two categories: metastable, silenced hypermethylated alleles (*ag, sup, and pai2*) and stable, hypomethylated epialleles associated with ectopic gene expression (*fwa*). The stability of the *bal* alteration is consistent with the hypothesis that DNA hypomethylation is involved. The recent discovery of hypermethylated silenced *Lcyc* alleles in *Linaria* shows that epigenetic variation can affect plant morphology and fitness in natural populations as well (Cubas et al. 1999). The induced genomic alterations characterized here (*bal* and *ssi1*) can modulate pathogen–plant interactions, raising the possibility that similar variants might arise and be selected in natural populations.

Materials and methods

Plant growth

Plants were grown in a mixture of Redi-Earth (Scotts):vermiculite (60%:40%) in environmental growth chambers (16 h of illumination [fluorescent + incandescent]/day, 70% relative humidity, 22°C) or under similar conditions in a greenhouse. Under short day length (8 h of illumination/day) and high humidity (95%), the morphological phenotypes of *bal* are attenuated. *bal* plants grown under short day length were used for the *P. syringae* inoculations (see Fig. 3).

Genetic mapping of the *bal* defect

The *bal* variant was generated from a *ddm1-2* line in the Columbia strain background, self-pollinated through six genera-
tions. We mapped bal in an F2 segregating population generated from a cross between a bal/bal; DDM1/DDM1 [Columbia] plant and a BAL/BAL; DDM1/DDM1 [Landsberg erecta] individual (Kakutani et al. 1996). Phenotypic bal/bal F2 plants were geno-
typed initially using the flanking CAPS markers, AG and SC5 (Konieczny and Ausubel 1993; http://www.arabidopsis.org). We
generated new CAPS markers from the available genomic se-
quence (Bevan et al. 1998) to establish a genetic and physical
interval containing the bal alteration. This window is defined
by a centromere-proximal marker: forward primer, 5'-AGACG
CTGGAGTATCTTCCAC-3'; reverse primer, 5'-CAGGAGGT
GAGTTCCATCCTC-3'; polymorphic TaqI site and a telomere-
proximal marker: forward primer, 5'-AATCATTTGCACCGAT
CAC-3'; reverse primer, 5'-TGTTACGCCGGCTCTGCTAC-
3'; polymorphic HaeIII site.

Nucleic acid isolation and analysis

Genomic DNA samples were purified using QIAGEN protocols
and columns, or by the Urea Lysis miniprep protocol [Coccio-
lone and Cone 1993]. Southern analyses were performed as de-
scribed previously [Jeddeloh et al. 1998]. Total RNA samples
were isolated using either QIAGEN RNaseasy or Bio-Rad Aqua-
Pure protocols. RNA was size-fractionated by electrophoresis
through 1.5% agarose formaldehyde gels and blotted to Gene-
Screen (NEN DuPont) nylon membranes using capillary action
and 50 mM sodium phosphate buffer (pH 6.5). All hybridiza-
tions were done following the protocol of Church and Gilbert
(1984), and membranes were washed at 60°C in 0.2× SSC, 0.1%
sodium citrate. Hybridization probes were radiolabeled using the random
priming protocol [Feinberg and Vogelstein 1983], and unincor-
porated radionucleotides were removed by size-filtration col-
umns. We used the following hybridization probes: Arabidopsis
actin 2 gene [GenBank U41998]; PR1 [Uknes et al. 1992]; PR2
[Dong et al. 1997]; R-like gene [ColE; Parker et al. 1997]; 25S rRNA gene (pARR17; Kakutani et al. 1996). DNA sequencing was performed using linear double-
stranded templates generated by genomic amplification and Big-
Dye Terminator Cycle Sequencing [Perkin-Elmer] protocols/re-
agents.

bal reversion analysis

bal/bal; DDM1/DDM1 seeds were hydrated overnight in water
and then treated with different concentrations of EMS (0, 10, 20,
30, 40, and 50 mM; Sigma M0880) for 8 h. Seeds were then
washed overnight in water, dried, and sown on soil. The result-
ing M1 plants were allowed to grow to maturity, and self-pol-
linated M2 seeds were harvested in independent batches (~20
M1 plants per M2 pool). A similar protocol was followed for the
γ-ray mutagenesis experiments except that dried seeds were

either sequenced using Big-Dye Terminator Cycle Sequencing
[Perkin Elmer] protocols/reagents or were digested with HaeIII
and/or HphI [New England Biolabs], which distinguished among
Atdg16860, 16890, 16900, 16920, and 16950.

Generation of transgenic plants

The genomic region encompassing the Atdg16890 gene was am-
plified using the following primers: forward primer, 5'-CTAG
TCATGACGGCAAAATTTTGCTGCCCATCTCGTA-3'; reverse primer, 5'-CCCGGATCCAAAAATGTGAAAAAGATAG
ATGTTAGTGCAC-3'. The ends of the resulting 5-kb fragment
were cut with Xbal and BamHI [New England Biolabs], and the
fragment was cloned into the pMDI vector [a derivative of
pBluescript II; Mark Dixon, unpubl.] downstream of a strong viral pro-
moter, P35S. The amplified 355::16890 gene encodes a full-
length protein with one missense mutation at position 950 in the
LRR domain [S 950 L]. Only 8 of the 21 LRRs in RPP5 [the
least relative in the Arabidopsis genome] contain a leucine
adjacent [toward the N terminus] to the altered position, which
is occupied by either serine or isoleucine. The poor conserva-
tion at this position and the similar structure of leucine and isoleu-
cine suggest that the change will not alter the function of the
gene product. The construct was electroporated into Agrobac-
terium tumefaciens strain GV 3101 [Koncz and Schell 1986].

Plants were transformed using the floral dip protocol [Clough
and Bent 1998]. Harvested seeds were ethanol- and bleach-ster-
ilized, and transformants were selected on plates with 4.6%
Murashige and Skoog salts [GIBCO-BRL], Gamborg’s vitamins
[Sigma], and 50 mg/L kanamycin [Sigma].

Characterization of pathogen defense response

Five-week-old Arabidopsis plants were inoculated by vacuum
infiltration with Pseudomonas syringae p.v. tomato strain
DC3000 [Whalen et al. 1991] at an initial density of 1 × 10^5
CFU/mL. The concentration of bacteria in the plant leaves was
assayed after 0, 1, and 4 d as described previously [Whalen et al.
1991]. Data points represent the means of three independent
determinations ± standard error of the mean. nahG bal double
mutants were identified by testing for kanamycin resistance
[linked to the nahG transgene] and scoring a CAPS marker
closely linked to bal: forward primer, 5'-AGACGCTGGAG
ATCTTCCAC-3'; reverse primer, 5'-CAGGAGGTGAGTT
ATCCCTC-3'; polymorphic TaqI site. Northern blots were done
done on double mutants to verify that Atdg16890 and nahG
were expressed at expected levels (data not shown). Crosses were
done between bal/bal and npr1-1/npr1-1 plants and double mu-
tants in the F2 generation were identified by first scoring for
dwarfing and then genotyping at the NRP1 locus using a mo-
lecular marker. A similar approach was used to generate bal
ed5-2 double mutants.

Acknowledgments

This work was supported by grants from the National Sci-
cience Foundation to E.J.R. [MCB9306266, MCB9604972, and
MCB9985348] and from the National Institutes of Health to
B.N.K. [GMS5236]. T.L.S. was supported in part by a predoctoral
fellowship from the Monsanto Company. We thank X. Dong
cprl], D. Klessig [ssi1], and J. Parker [eds1-2] for generously
providing seeds. We gratefully acknowledge William D. Wright
[Washington University, Department of Radiology] for assis-
tance with the γ-irradiation, Katherine Mason for technical as-
sistance, and Mike Dyer for greenhouse management.
The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 USC section 1734 solely to indicate this fact.

References


