Supporting Information

Morbitzer et al. 10.1073/pnas.1013133107

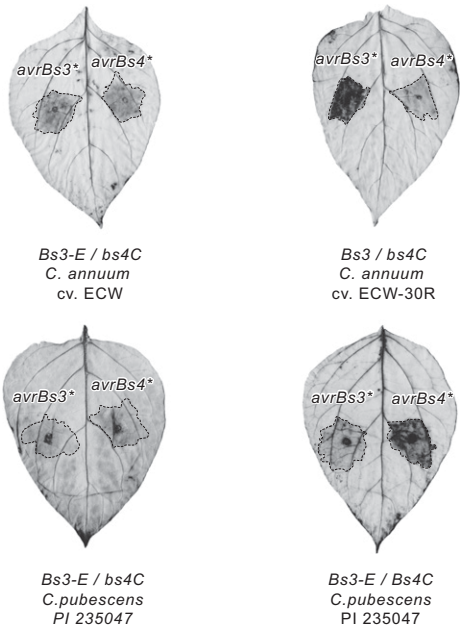

Fig. S1. Codon-optimized transcription activator-like effectors (*TALE*) genes of *avrBs3* and *avrBs4* trigger a hypersensitive response (HR) in *Bs3*- and *Bs4C*-resistant *Capsicum* species, respectively. *In planta* functional analysis of codon-optimized (*) *TALE* genes. The *35S* promoter-driven *TALE* genes indicated on the left and right sides of leaves were delivered via *Agrobacterium tumefaciens* into the four *Capsicum* genotypes shown. The allelic configuration at the *Bs3* locus [alleles: *Bs3* (recognition of AvrBs3), *Bs3-E* (recognition of AvrBs3Δrep16)] and the *Bs4C* locus [alleles: *Bs4C* (recognition of AvrBs4), *bs4C* (no recognition of AvrBs3, AvrBs3Δrep16, and AvrBs4)] is indicated for each plant genotype. Dashed lines mark the inoculated areas. Two days after infiltration, the leaves were cleared in ethanol to visualize the hypersensitive response (dark areas).

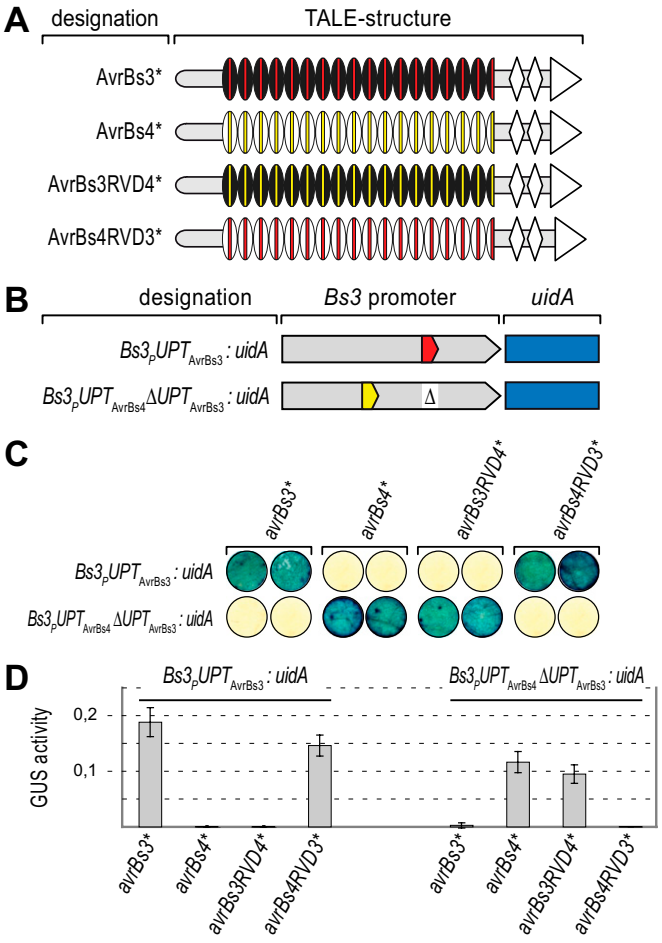

Fig. S2. TALE nonrepeat-variable diresidues (non-RVD) residues have no measurable influence on TALE specificity. (*A*) Graphic representation of chimeric TALES and corresponding *uidA* reporter constructs. TALEs are displayed as gray bars. Black and white ovals represent the repeat units of AvrBs3 and AvrBs4, respectively. Red and yellow vertical lines within the ovals represent the RVDs of AvrBs3 and AvrBs4, respectively. The chimeric TALE AvrBs3RVD4 contains the RVDs of AvrBs4 but is otherwise identical to AvrBs3. Reciprocally, the chimeric TALE AvrBs4RVD3 contains the RVDs of AvrBs3 but is otherwise identical to AvrBs4. TALE nuclear localization signals (NLSs) and activation domain (AD) are displayed as white diamonds and white arrowheads, respectively. Asterisks (*) indicate that proteins are translated from genes with optimized codon usage. (*B*) Graphic representation of *Bs3* promoter derivatives. Gray arrows represent the promoter of the pepper *Bs3* gene ($Bs3_P$). Red and yellow arrowheads depict the UPT_{AvrBs3} and the UPT_{AvrBs4} boxes, respectively. The white area within the $Bs3_PUPT_{AvrBs4}\Delta UPT_{AvrBs3}$ promoter indicates the deleted UPT_{AvrBs3} box (Δ). Blue boxes represent the *uidA* reporter gene, encoding the β-glucuronidase (GUS) reporter protein. The nucleotide sequences of the promoters are provided as *SI Text*. (*C*) RVDs of AvrBs3 and AvrBs4 retain promoter-activation specificity regardless of the repeat-array backbone. Depicted *uidA* reporter constructs were delivered into *Nicotiana benthamiana* leaves in combination with the depicted *35S*-driven *TALE* genes. Asterisks (*) indicate *TALE* genes with optimized codon usage. Leaf discs were stained with 5-bromo-4-chloro-3-indolyl-β-D-glucuronic acid, cyclohexylammonium salt (X-Gluc) to visualize activity of the GUS reporter. Samples were taken at 40 hpi. (*D*) Quantitative analysis of promoter activation via wild-type and chimeric TALEs. GUS activity (pmol 4-MU·min·g protein) was determined 27 h after *A. tumefaciens*-mediated codelivery of the depicted *uidA* reporter constructs and the *35S*-driven *TALE* gene shown below columns. Error bars denote SD. Asterisks (*) indicate *TALE* genes with optimized codon usage.

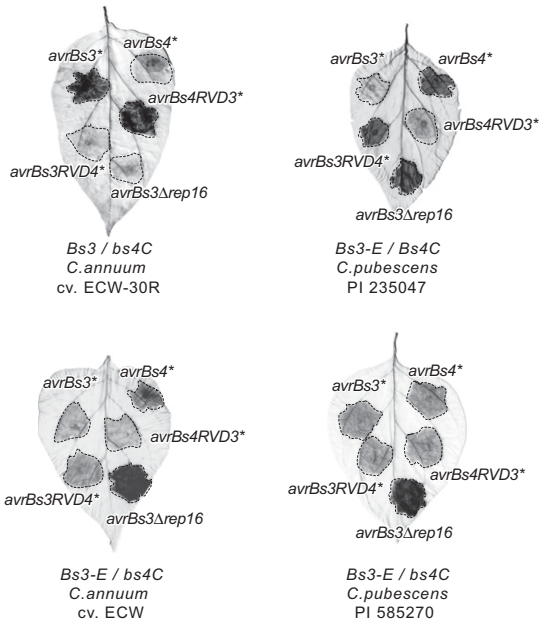

Fig. S3. *In planta* functional analysis of chimeric *TALE* genes. The *35S* promoter-driven *TALE* genes indicated on the left and right side of the leaves were delivered via *A. tumefaciens* into leaves of the *Capiscum annuum* cultivars ECW and ECW-30R or the *Capiscum pubescens* accessions PI 585270 and PI 235047. The allelic configuration of the plant genotypes at the *Bs3* locus (alleles: *Bs3* or *Bs3-E*) and the *Bs4* locus (alleles: *Bs4C* or *bs4C*) locus is shown below the leaves. Dashed lines mark the inoculated areas. Two days after inoculation, the leaves were harvested and cleared with ethanol to visualize the HR (dark areas). Asterisks (*) indicate that AvrBs3 was translated from a codon-optimized gene.

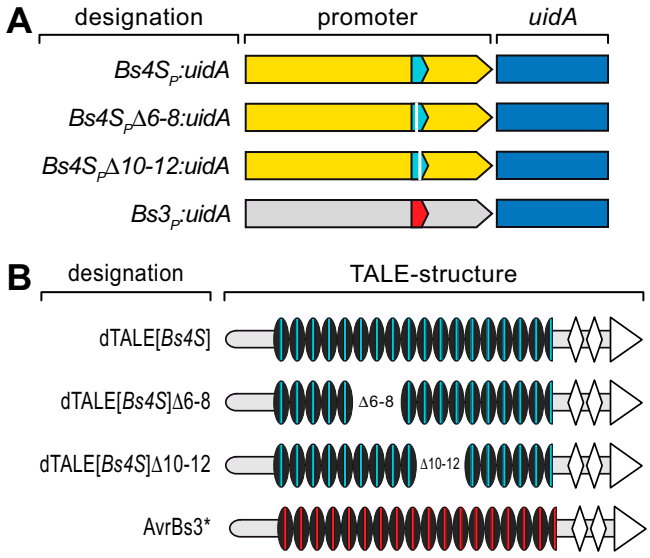

Fig. S4. Graphic display of *Bs4S* promoter:*uidA* reporter constructs and matching dTALEs. Corresponding experimental data (Fig. 1). (*A*) Arrows represent the promoters of the tomato *Bs4S* (yellow) and the pepper *Bs3* gene (gray). Cyan and red arrowheads depict the $UPT_{dTALE[Bs4S]}$ and UPT_{AvrBs3} boxes, respectively. White vertical bars within the cyan arrowheads represent deletions in the $UPT_{dTALE[Bs4S]}$ box. Blue boxes represent the *uidA* reporter gene, encoding the GUS protein. The nucleotide sequences of the promoters are provided as *SI Text*. (*B*) Graphic display of the TALEs used in this experiment. The TALEs are displayed as gray bars. Repeat units, NLSs, and AD are depicted as black ovals, white diamonds, and white arrowheads, respectively. Cyan and red colored vertical lines within the ovals represent the RVDs of designer TALES (dTALE) [*Bs4S*] and AvrBs3*, respectively. Repeat unit deletions in dTALE[*Bs4S*]Δ6–8 and dTALE[*Bs4S*]Δ 10–12 are indicated. Asterisks (*) indicate that the AvrBs3 is translated from a codon-optimized gene.

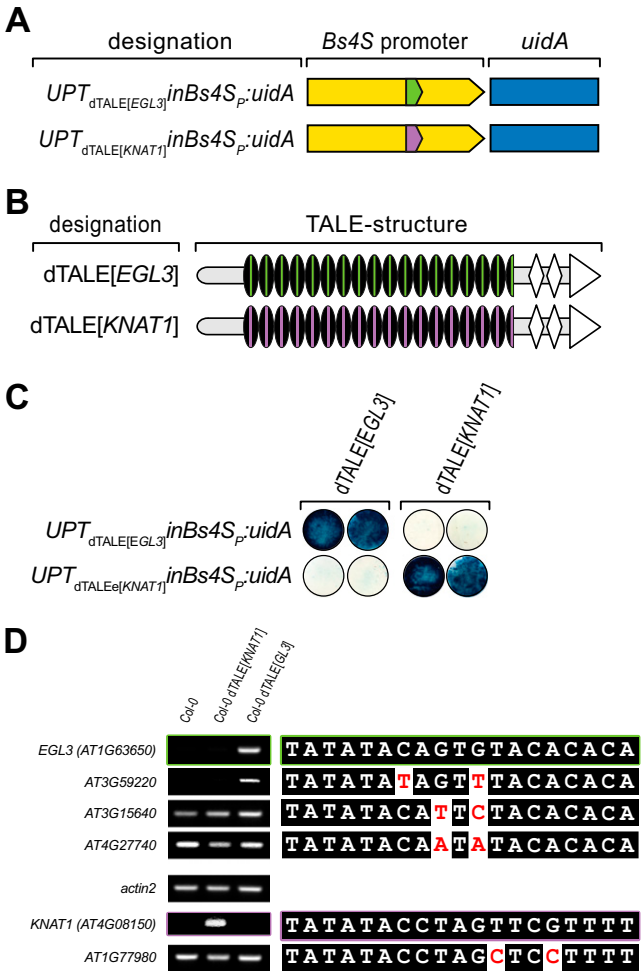

Fig. S5. Designer TALEs activate promoters of user-defined target genes in *Arabidopsis thaliana*. (*A*) Graphic display of promoter:*uidA* reporter constructs used in this experiment. Yellow arrows represent the promoter of the tomato *Bs4S* gene. Green and purple arrowheads depict the inserted $UPT_{\text{dTALE}[EGL3]}$ and $UPT_{\text{dTALE}[KNAT1]}$ boxes, respectively. Blue boxes represent the *uidA* reporter gene encoding the GUS protein. The nucleotide sequences of the promoters are provided as *SI Text*. (*B*) Graphic display of engineered dTALEs. The dTALEs are displayed as gray bars. Repeat units, NLSs, and AD are displayed as black ovals, white diamonds, and white arrowheads, respectively. Green and purple colored vertical lines within the ovals represent the RVDs of dTALE[*EGL3*] and dTALE[*KNAT1*], which direct binding to promoter elements of the *Arabidopsis EGL3* and *KNAT1* genes, respectively. Although not displayed graphically, both dTALEs contain a C-terminal GFP-tag that facilitates analysis of expression in vivo. (*C*) *In planta* functional analysis of different TALE-promoter combinations. The *uidA* reporter constructs under transcriptional control of the promoters indicated at left were delivered via *A. tumefaciens* into *N. benthamiana* leaves in combination with the *35S* promoter-driven *TALE* genes shown above leaf discs. GUS assays were carried out 40 hpi as described in Fig. 1. (*D*) Analysis of potential dTALE[*EGL3*] and dTALE[*KNAT1*] off-targets in the *Arabidopsis* genome. Semiquantitative RT-PCR was carried out on RNA from the *Arabidopsis* ecotype Columbia (Col-0) and corresponding lines containing the *35S* promoter-driven *dTALE[EGL3]*, or the *dTALE[KNAT1]* transgene. RNA was isolated from leaves and used for cDNA synthesis. Semiquantitative RT-PCR was performed with locus-specific primers (*SI Text*). The constitutively expressed gene *actin2* served as an internal normalization control. The boxes targeted by dTALE[*EGL3*] and dTALE[*KNAT1*] are shown on the right side and are highlighted with green and purple frames, respectively. Sequences of potential off-targets are displayed below the two distinct, user-defined boxes. Differences between the targeted box and the potential off-targets are displayed in boldface red font. Details on the genomic context of the targeted boxes and potential off targets are provided in the *SI Text*.

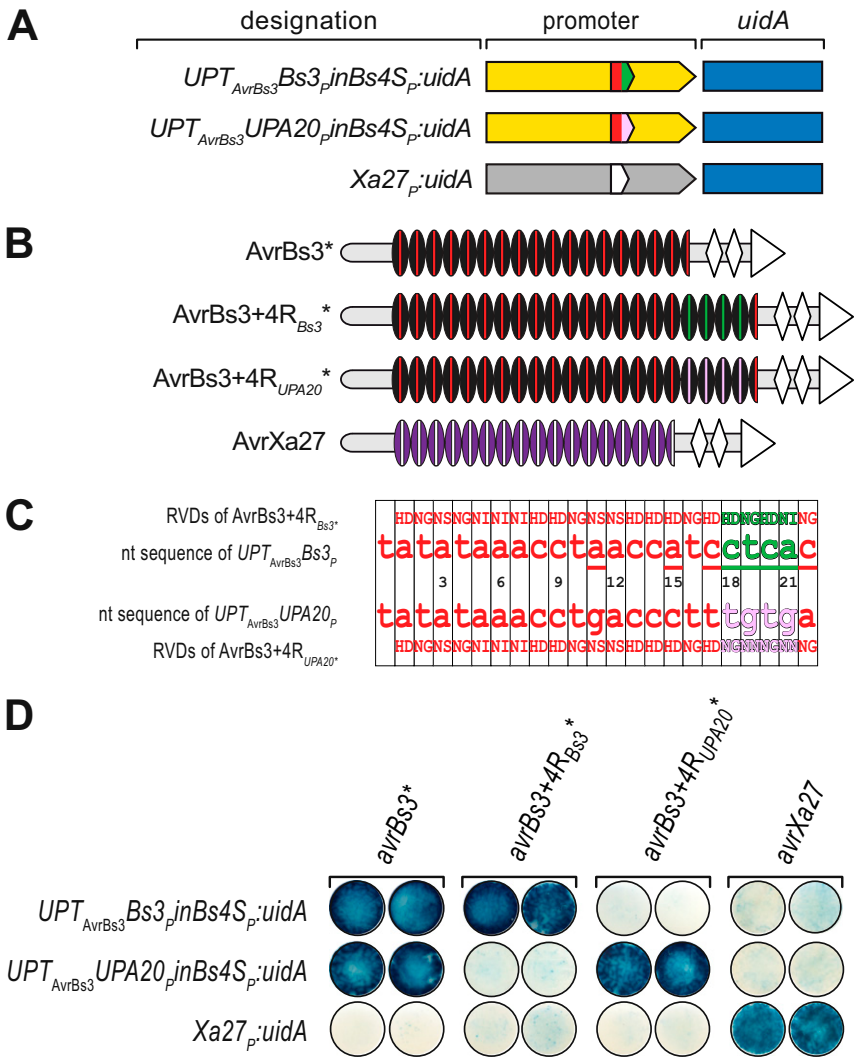

Fig. S6. Graphic display of AvrBs3 derivatives with additional repeat units and corresponding *uidA* reporter constructs. For corresponding experimental data, see also Fig. 3. (*A*) Arrows represent the promoters of the tomato *Bs4S* (yellow) and the rice *Xa27* gene (gray). Arrowheads depict *UPT* boxes. AvrBs3-targeted *UPT* boxes from the pepper *Bs3* and *UPA20* are displayed in red. *Bs3*- and *UPA20*-derived promoter regions 3′ of their UPT_{AvrBs3} boxes are displayed in green and pink, respectively. The $UPT_{AvrXa27}$ box from the rice *Xa27* gene is displayed in white. A blue box represents the *uidA* reporter gene, encoding the GUS protein. The nucleotide sequences of the promoters are provided in the *SI Text*. (*B*) Graphic display of TALEs used in this experiment. The TALEs are displayed as gray bars. Repeat units, NLSs, and AD are displayed as ovals, white diamonds, and white arrowheads, respectively. Black- and purple-colored ovals represent the repeat units of AvrBs3 and AvrXa27, respectively. Red and white vertical lines represent RVDs of the TALEs AvrBs3 and AvrXa27, respectively. Green and pink vertical lines represent C-terminal RVDs that are specific to AvrBs3+4R$_{Bs3}$* and AvrBs3+4R$_{UPA20}$*. (*C*) Alignment of RVDs and *UPT*-boxes of the depicted TALEs and promoter-reporter constructs. RVDs are shown as uppercase letters using the single-letter code. Nucleotides are shown as lowercase letters. Horizontal boxes mark repeat units, corresponding RVDs, and aligned nucleotides. Numbers mark TALE repeat units and aligned nucleotides. Lowercase red letters represent nucleotides of the naturally occurring *UPT* boxes present in the pepper *Bs3* ($UPT_{AvrBs3}Bs3_P$) and the *UPA20* ($UPT_{AvrBs3}UPA20_P$) promoters. Lowercase green and pink letters represent nucleotides that are located 3′ of the UPT_{AvrBs3} boxes in the *Bs3* and *UPA20* promoters, respectively. Nucleotides that differ between the *Bs3* and *UPA20* promoter are underlined. Uppercase red letters represent RVDs of the TALE AvrBs3. Green and pink big letters represent C-terminal RVDs that are specific to the dTALEs AvrBs3+4R$_{Bs3}$* and AvrBs3+4R$_{UPA20}$*, respectively and matching nucleotides of the *Bs3* and *UPA20* promoters. (*D*) AvrBs3 derivatives with additional repeat units discriminate between *UPA20* and *Bs3* promoter-derived *UPT* boxes with extensive 3′ homology. The *uidA* reporter constructs under transcriptional control of the promoters shown at left were delivered via *A. tumefaciens* into *N. benthamiana* leaves in combination with the *35S* promoter-driven *TALE* genes indicated above leaf discs. GUS assays were carried out 40 hpi, as described in Fig. 1. Asterisks (*) indicate that TALEs are translated from codon-optimized genes.

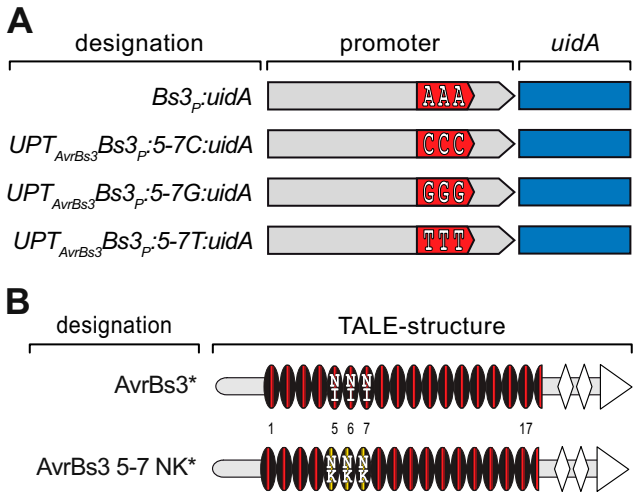

Fig. S7. Graphic display of AvrBs3 derivatives with three NK repeat units and corresponding *uidA* reporter constructs. For corresponding experimental data, see Fig. S8. (*A*) Gray arrows represent the pepper *Bs3* promoter. The red arrowhead with three tandem-arranged A nucleotides (AAA) depicts the wild-type UPT_{AvrBs3} box of the *Bs3* promoter. *Bs3*-promoter derivatives in which three tandem-arranged A nucleotides of the UPT_{AvrBs3} box are replaced by three C, G, or T nucleotides are depicted (CCC, GGG, and TTT). The nucleotide sequences of the promoters are provided in the *SI Text*. Blue boxes represent the *uidA* reporter gene encoding the GUS protein. (*B*) Graphic display of the TALEs used in this experiment. The TALEs are displayed as gray bars. Repeat units, NLSs, and AD are depicted as black ovals, white diamonds, and white arrowheads, respectively. Red vertical lines within black ovals represent the RVDs of AvrBs3. Three tandem-arranged NI type RVDs that are present in repeat-unit residues 5 to 7 of the wild-type AvrBs3 protein are depicted. Yellow vertical lines within black ovals represent RVD residues within the AvrBs3 scaffold that were changed to NK-type residues. Asterisks (*) indicate that TALEs are translated from codon-optimized genes.

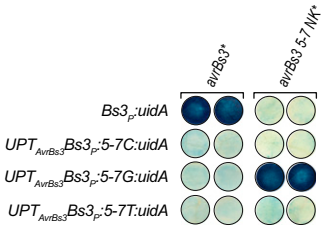

Fig. S8. The TALE RVD NK mediates specific targeting of G nucleotides. *In planta* functional analysis of different TALE-promoter combinations. The *uidA* reporter constructs under transcriptional control of the promoters indicated at left were codelivered via *A. tumefaciens* into *N. benthamiana* leaves with the *35S* promoter-driven *TALE* genes indicated above leaf discs. GUS assays were carried out 40 hpi, as described in Fig. 1. Asterisks (*) indicate *TALE* genes with optimized codon usage and corresponding gene products.

Table S1. Nucleotide sequences of primers used in this study

Number	Designation	Nucleotide sequence
RM 1	AvrBs3*-CACC-F/RobM	CACCATGGATCCCATTAGGTCTAGG
RM 2	AvrBs3*-ns-R/RobM	CTGTGGAAGCAATTCCATAAGCC
RM 3	AvrBs3*-N-term-R/RobM	CCCGCATGCAATATTCAGCGGTGCCCCCGTCAAAA
RM 4	AvrBs3*-C-term-F/RobM	GGGAGTCATGGCCGGCCAGCACTTGAGTCGATTGTT
RM 5	AvrBs3*-R6-*Xho*I-R/RobM	CGGTCTCGAGAGCCTGTTTACCTCCG
RM 6	TAL[Bs4S]-R5-R/RobM	GCCATGAGCCTGGCATAGTACAGGTAGAAGCGC
RM 7	TAL[Bs4S]-R9-F/RobM	CTTACACCAGAGCAAGTAGTTGCTATTGCATC
RM 8	TAL[Bs4S]-R9-R/RobM	TCCATGTGCCTGGCATAGAACTGGCAGGAGC
RM 9	TAL[Bs4S]-R13-F/RobM	CTAACCCCCGAACAGGTGGTGGCTATTGCATCCC
RM 10	AvrBs3*-R17-*Xho*I-F/RobM	GCACTCGAGACAGTCCAGAGATTGCTGCCG
RM 11	AvrBs3*-R17-*Xho*I-R/RobM	CCCCTCGAGTGCCTGCTTTCCACCATCATGCGAAGCGATAGCTACTACC
RM 12	$4R_{Bs3}$XhoI-F/RobM	CAAGCACTCGAGACCGTGCAGAGACTTCTG
RM 13	$4R_{Bs3}$XhoI-R/RobM	CGTCTCGAGAGCAGGTCTTCCTCCACCG
RM 14	$4R_{upa20}$XhoI-F/RobM	GGCAAGCAGGCACTCGAGACAGTTCAGAGACTG
RM 15	$4R_{upa20}$XhoI-R/RobM	CCCCTCGAGTGCTTGTTTGCCTCCGTTGTTGCTTGCGATAGCC
RM 16	$UPT_{TAL[Bs4S]}$D6-8-R/RobM	GACAATATATACAAGAAAGAAGAATTAAACAAGTACC
RM 17	$UPT_{TAL[Bs4S]}$D10-12-R/RobM	GAAGTTATATACAAGAAAGAAGAATTAAACAAGTACC
RM 18	$UPT_{TAL[Bs4S]}$F/RobM	CAAAATATCATCAATTGATCTCATCCATAC
RM 19	UPT_{AvrBs4}in3-F/PR	TATAATTAATAATCCACTTCTGGTTAAACAATGAACACGTTTGC
RM 20	UPT_{AvrBs4}in3-R/PR	GGTGTGCAAATTGTGGTTTAACCC
RM 21	DUPT_{AvrBs3}inBs3$_P$-F/RobM	TCACAACTTCAAGTTATCATCCCC
RM 22	DUPT_{AvrBs3}inBs3$_P$-R/RobM	ATAAAATTGGTCAGGCAAACGTGTTC
RM 23	$UPT_{AvrBs3/Bs3}$inBs4S$_P$-F/PR	CAATTTTATTATATAAACCTAACCATCCTCACAACGTTTCAAGTGGTACTTGT
RM 24	$UPT_{AvrBs3/upa20}$inBs4S$_P$-F/RobM	CTTCATCTTTATATAAACCTGACCCTTTGTGACATGTTTCAAGTGGTACTTGT
RM 25	$UPT_{dTAL[EGL3]}$inBs4S$_P$-F/RobM	TTTTGTGATTATATACAGTGTACACACACCTAAAAGTTTCAAGTGGTACTTGT
RM 26	$UPT_{dTAL[KNAT1]}$inBs4S$_P$-F/RobM	GACGAATTCTATATACCTAGTTCGTTTTTTCTTCCGTTTCAAGTGGTACTTGT
RM 27	Bs4p-R/PR	GTGAAAGCTTGTATTAACATTCGCTTTG
RM 28	Bs4 RT1 F/RobM	CAAGATGATAAAAGACTAGAGCATGGAG
RM 29	Bs4 RT1 R/RobM	TGGAAGAGATGCAATCTACGATCTGCT
RM 30	RS-EFrt-F1	AGTCAACTACCACTGGTCAC
RM 31	RS-EFrt-R1	GTGCAGTAGTACTTAGTGGTC
RM 32	KNAT1 RT1 F	CAAGCTTACTTGGACTGCCAAAAG
RM 33	KNAT1 RT1 R	TTAACCAACATGTCACAGTATGCTTC
RM 34	EGL3 RT1 F/RobM	CGCCGCGAGAAATTGAATGAACG
RM 35	EGL3 RT1 R/RobM	CCATGCAACCCTTTGAAGTGCC
RM 36	actin2 RT F/RobM	GACTACGAGCAGATGGAAACC
RM 37	actin2 RT R/RobM	CTGGACCTGCCTCATCATACTCGG
RM 38	UPA20 RT F/RobM	GCACTACTACATCACAAGCAG
RM 39	UPA20 RT R/RobM	CTAGAGGCTAGCTATAGTCAG
RM 40	Cand-7–01-fwd	ATGAATCAGAATTGCTTTAATTCTTGTTCA
RM 41	Cand-7–01-rev	TGATTCTTGTGCTACATTTGTTCTTTCC
RM 42	AT5G20270.1 K F/RobM	ACAACCAACGTGGTCACGCCAGTCC
RM 43	AT5G20270.1 K R/RobM	GCAAAGAGTCCAGGGAATGTAACAAGATGG
RM 44	AT3G15640 G F/RobM	AGAATCGTCTCCTCTCAGCTCAAAACC
RM 45	AT3G15640 G R/RobM	ATCCGGAGGACCACCAGGACCAACC
RM 46	AT3G59220 G F/RobM	GAAGGTGAAGGAGCTGTCGTTAGAAATGG
RM 47	AT3G59220 G R/RobM	CAAACCCGTTCTTAGCATTCTGATAGTCATC
RM 48	AT4G27740 G F/RobM	ATGGCAGCAAACAAAACTCTTCCTACG
RM 49	AT4G27740 G R/RobM	TCTTTTGGTCAACTTGAGCTTCTCGATGAC

Other Supporting Information Files

SI Text (DOC)