

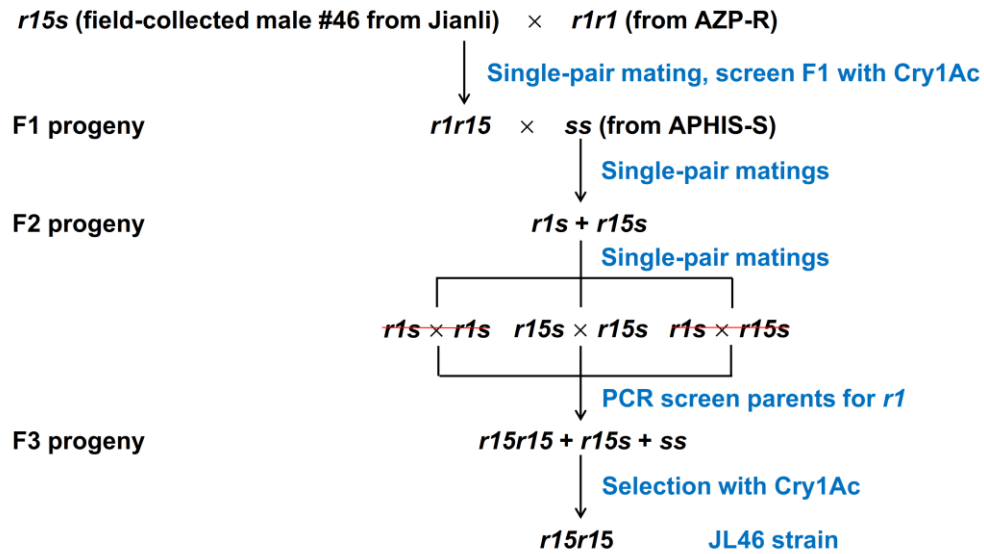
**Transposon insertion causes cadherin mis-splicing and confers  
resistance to Bt cotton in pink bollworm from China**

**Supplementary Figures S1-S8**

Ling Wang<sup>1,2</sup>, Jintao Wang<sup>1,3</sup>, Yuemin Ma<sup>4</sup>, Peng Wan<sup>1</sup>, Kaiyu Liu<sup>4</sup>, Shengbo Cong<sup>1</sup>,  
Yutao Xiao<sup>5</sup>, Dong Xu<sup>1</sup>, Kongming Wu<sup>2\*</sup>, Jeffrey A. Fabrick<sup>6</sup>, Xianchun Li<sup>7</sup> & Bruce  
E. Tabashnik<sup>7</sup>

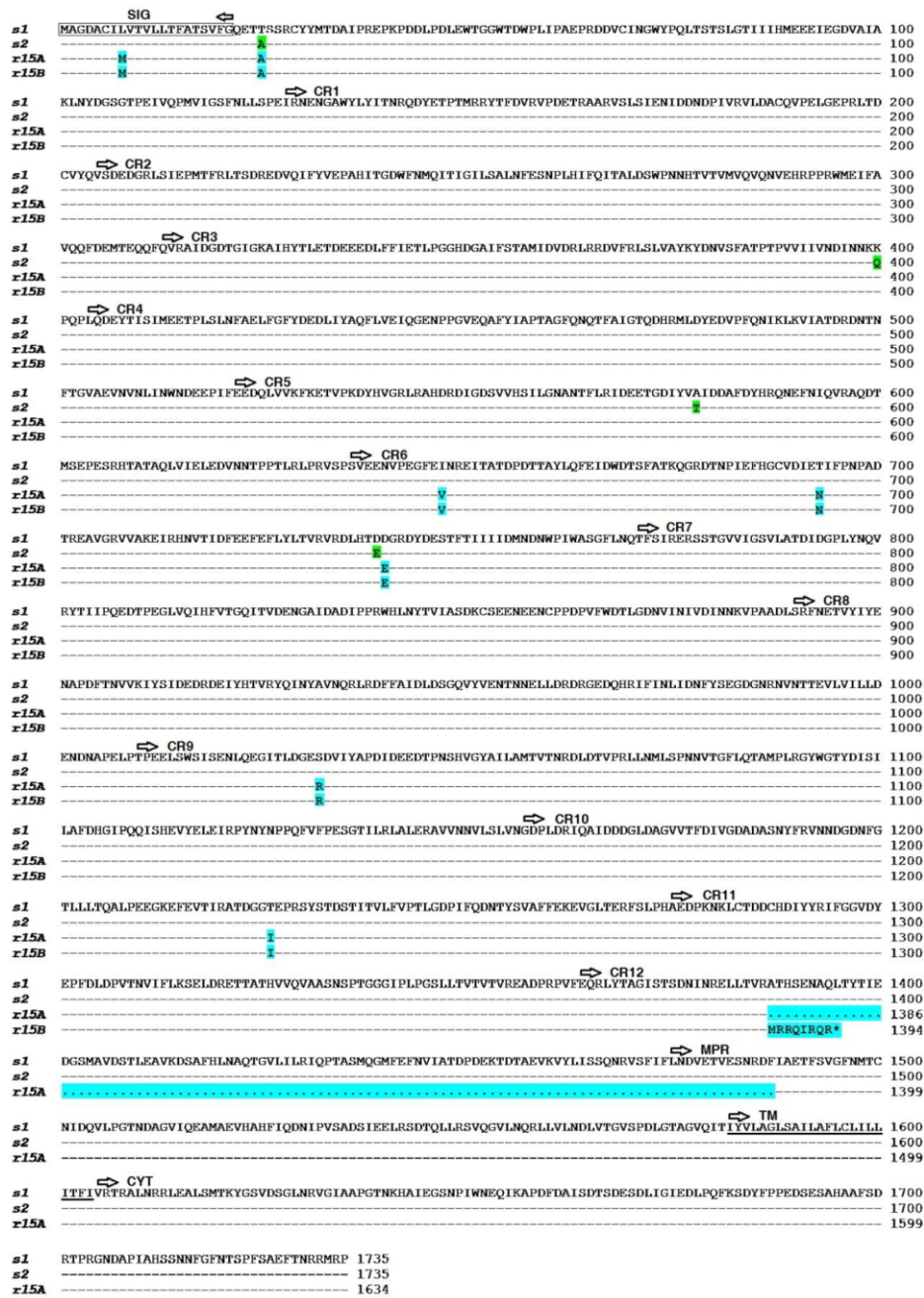
<sup>1</sup>Key Laboratory of Integrated Pest Management On Crops in Central China, Ministry of Agriculture, Hubei Key Laboratory of Crop Disease, Insect Pests and Weeds Control, Institute of Plant Protection and Soil Fertility, Hubei Academy of Agricultural Sciences, Wuhan 430064, China. <sup>2</sup>State Key Laboratory for Biology of Plant Diseases and Insect Pests, Institute of Plant Protection, Chinese Academy of Agricultural Sciences, Beijing 100193, China. <sup>3</sup>Hubei Insect Resources Utilization and Sustainable Pest Management Key Laboratory, Huazhong Agricultural University, Wuhan 430070, China. <sup>4</sup>School of Life Science, Central China Normal University, Wuhan, 430079, China. <sup>5</sup>Agricultural Genomics Institute at Shenzhen, Chinese Academy of Agricultural Sciences, Shenzhen, 518120, China. <sup>6</sup>USDA, ARS, U.S. Arid Land Agricultural Research Center, Maricopa, Arizona 85138, USA. <sup>7</sup>Department of Entomology, University of Arizona, Tucson, Arizona 85721, USA.

\* Correspondence and requests for materials should be addressed to K.W. (E-mail address: [kmwu@ippcaas.cn](mailto:kmwu@ippcaas.cn))



**Figure S1.** Isolation of pink bollworm resistant strain JL46. Pink bollworm strain JL46 was started from a single-pair cross between a field-collected male (#46) from Jianli in Hubei province of the Yangtze River Valley and a female from a Cry1Ac-resistant strain AZP-R (cadherin genotype  $r1r1$ ) from Arizona. Their F<sub>1</sub> progeny (family #46) were screened with 10  $\mu$ g Cry1Ac protoxin per ml diet, a diagnostic concentration of Cry1Ac. Survival of the F<sub>1</sub> progeny at the diagnostic concentration was 47% and recessive resistance to Cry1Ac of pink bollworm was described before<sup>1,2</sup>, indicating that the field-collected male parent of family #46 carried only one recessive allele at the cadherin locus. Sequencing of cDNA from the resistant F<sub>1</sub> progeny showed that male #46 had two transcripts of cadherin allele, which we name  $r15A$  (Figures 1, S2 and S3) and  $r15B$  (Genbank accession numbers KY814704 and KY814705, respectively).

Survivors from family #46 were reared to adults and single-pair crosses between these survivors and adults of the opposite sex from the susceptible strain APHIS-S (cadherin genotype  $ss$ ) from Arizona were conducted separately to produce resistant strain JL46. We used PCR to identify  $r1$  allele for the parents of all single-pair F<sub>2</sub> families<sup>3</sup>, and retained only the families from pairs of parents without carrying  $r1$  allele ( $r15s$ ), then these single pairs of these F<sub>2</sub> progeny were selected by exposing larvae to diet containing the diagnostic concentration of Cry1Ac. The individuals survived on the diagnostic concentration were homozygotes  $r15r15$ , which were reared their progeny as strain JL46, and feeding larvae of JL46 with diet containing the diagnostic concentration of Cry1Ac every fifth generation to maintain the resistance to Cry1Ac.



**Figure S2.** Predicted amino acid sequence of pink bollworm cadherin protein PgCad1 for alleles *s1* (GenBank accession no. AY198374) and *s2* (GenBank accession no. MF276974) from susceptible strain APHIS-S; and *r15A* (GenBank accession no. KY814704) and *r15B* (GenBank accession no. KY814705) from resistant strain JL46). The *s1* allele was sequenced in 2003 from APHIS-S reared in Arizona<sup>1</sup> and *s2* was sequenced in 2015 from the subset of APHIS-S reared in China<sup>4</sup>. Shown are the signal sequence (SIG), cadherin repeats (CR1-CR12), membrane-proximal region (MPR), transmembrane region (TM) and cytoplasmic region (CYT). The blue shading indicates differences in *r15* compared with *s1*: deletion of 101 amino acids (1387 to 1487 for *r15A*, indicated by dots) and frameshift mutation for *r15B* and seven single amino acid substitutions (L8M, T25A, I647V, T693N, D740E, S1032R, T1226I). The green shading shows differences in *s2* versus *s1*: four single amino acid substitutions.

```

#1 ATGGCGGGTACGCGTACTACTGTCGCGTCTTCTGACCTTCGCACTAGTTCGGCCAGAAACACATCTGCGAGATGTACTACTGACTGAGCTATCCGAGAGAACCGAAACGGATGATTGCTGATTGAAGTGG 150
r15A .....A.....C.....G..... 150
r15B .....C.....G..... 150

#1 ACTGGTGTATGACCGATGCGCTTTGATCCCGCGTGGCCAGAGAGACGCTGTGCATAAAGCGGTACCCCAACTCCAGCACTCTCTCGGACCATCATCCACATGGAAGGAGATCGAGGGAGTGTGCTATCGCT 300
r15A .....C..... 300
r15B .....C..... 300

#1 AAATCTAATGATGCTGCGAACCCGAAATGTCACCGGATGTTATGAGACTTTTAACTCTAGTCCAGGATCGGATGAAAGGGGGGCTGACTTATATACCAATAGGCAAGATATTAACACCAACATG 450
r15A ..... 450
r15B ..... 450

#1 CCGCGTATACATGCTGCGAGTCCGAGACGAGACTGTCGCGGAGAGTGTCTCCATCGAAGACTGCAATGAGCACTCTCTGAGGCTGAGCGCTTCCAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGTCCGAGT 600
r15A .....C.....T..... 600
r15B .....C.....T..... 600

#1 TGCGTTTACCAAGTGTGAGAGATGGAGGCTTAGTATCGAGCCATGACATTCGCGCTCAGTACAGCCGTGAGAGCTACAGATATCTATGTGGAGCCAGCTCACTACTGGTGTGCTCAACATGCAAACTACTATCGT 750
r15A .....C.....A..... 750
r15B .....C.....A..... 750

#1 ATCTCTATGCGCTTAACTCGAAGCAACCCGCTGCACTTTCAATCACTGCTTGGATCTCGGCGCAACCATACGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 900
r15A .....T.....C..... 900
r15B .....T.....C..... 900

#1 GTCACAGCTTGTGAGATGAGAGCACTTCCAGTCCGCGCCATCGAGAGACACTGGCATCGGAAAGCTATACACTATACCTCGAGAGATGAGGAGAGATTTGCTTCAACAACTCCGCGCGCATAC 1050
r15A ..... 1050
r15B ..... 1050

#1 GGAGCCATCTTACGACTGCGCATTTGATGTGATGAGCTCGGCGAGATGTCTTCAGACTTCCTCGTGGATACAGAGTACAGCAATGTCTCCGCAACCCGAGCCGCTGATCATATGCTCAATGACATCAACCAAGAAA 1200
r15A .....T.....C..... 1200
r15B .....T.....C..... 1200

#1 CCCACCGCTGCAAGTATGACAAATCCATAATGAGAGAACTCCACTGCTGCAATTTGCTGACTTTTGGTCTATGATGAGATTTGATCAGCAAACTTCTGCGAAATCAAGCGAGACCTCCGCGTAT 1350
r15A ..... 1350
r15B ..... 1350

#1 GAGCAAGCTTTTATATGCGCCACCGAGGCTTCAGAACAGACATTCGCAATAGGACTCAAGATCCGCAATGCTGATTTAGGATGCTTCTTCCAAACATCAAGCTCAAGTAAAGCAAGCCAGCCGCTGACATCAAT 1500
r15A .....G..... 1500
r15B .....G..... 1500

#1 TTACTGGATGCGGAGTCAAGCTGAACTGATTAATGAGACGAGAGGCGGATCTTGGAGAGACCACTGTTGAAATTAAGAGACTGTACCGAGACTATCACTCGGAGACTGAGGCTCAGACCGGACATA 1650
r15A .....T.....G.....G..... 1650
r15B .....T.....G.....G..... 1650

#1 GGAGCAGCTGTGCTTCACTCTGGGAAAGCAATTTGAGATCGAGAGAACTGGGACATATAGCTATGATGAGCGCTGATTCAGACAGAAATGAAATGATCAAACTCGCGCTCAGAGACC 1800
r15A .....C..... 1800
r15B .....C..... 1800

#1 ATGCGAGCGAGATCCAGCATACAGACTCTCAGCTGCTATAGACTCGAGAGCTCAACAGCACTCTCTACTGAGCTCGCTGAGTCCGCTGTGAGAGAAATGCGCAGAGGCTTGAATCAACCGGAG 1950
r15A .....G..... 1950
r15B .....G..... 1950

#1 ATACCGCGACGAGCCCTGACCCAGCATACTGCTGAAATGATGAGTGGAGCACTATCTCCACTAAACAGGGCGTATACCAATGAGTATCCAGATCGCTGATATGAGAACCTTCCCAACCGAGAC 2100
r15A .....A..... 2100
r15B .....A..... 2100

#1 ACCAGAGGCTGTGGGCGAGTGTAGCAAGGATCCCGCAAGCTGACTGATTTGAGAGTGTGAAATCTCTACTCCAGTATGAGTTCGAGCTCCAGACAGATGAGGAGATGATGAACTACCTCAGC 2250
r15A .....A.....G..... 2250
r15B .....A.....G..... 2250

#1 ATATGATATGATATGAGCAACTGCGCTATCTGGGCGTGGTCTTCTGCAAGCTCTGATTCGGAGGACTCATCTCGGCGCTGCTCATCGGCTGCTCATCGCTCAGACATGATGGCCCATTTACAGCAATC 2400
r15A .....T.....C..... 2400
r15B .....T.....C..... 2400

#1 CGTACACATATCCCGAGAGACTCCGAGGCTGATCGAGTACATCTGTCAGGGTCAATGAGTGTAGAGAAATGTCAGCTGATTCACCTGTCGACCTCACTACAGCTATATGAGCGAGC 2550
r15A .....A.....G..... 2550
r15B .....A.....G..... 2550

#1 AAATGTCGAGAAATGAGAGAACTGTCGCGGATCCAGTCTTGGGATCTCGGGCAATGTAATTAACATGCGGACATAAACAAGGTCGCGGACAGCTCAGTCCGATCAACGAAAGCTGATATTATGAA 2700
r15A .....C..... 2700
r15B .....C..... 2700

#1 AATGACCGCTTCAAAAGCTGCAAGATATATCCAGCGAGAGCAATATACAGACGCTCGGATCAGATCAATATGCTGATCAACCGGCTGAGACTTCTCCGATAGACTGATTCGCGAGCTG 2850
r15A .....T..... 2850
r15B .....T..... 2850

#1 TAGTGGAGAACCAACATGATCTCTGAGTGGAGAGGAGAGCAACAGAGATATCTAACCCTCATGCAACTTTATAGCGAGGATGAGAAATGAGAAATGAAACTACAGAGTGTGGTATCATTAAT 3000
r15A ..... 3000
r15B ..... 3000

#1 GAGATGACAACTGCTGAAATGCGACTCCAGAGAGCTGATGGAGCATTTCCGAAATTTACAGAGGATATAACTGATGGCGAAGCGATGATATACGACCGGATATAGCAAGAGAGACCGCAACTCTCACGTT 3150
r15A .....A.....A..... 3150
r15B .....A.....A..... 3150

#1 GCGTACGATCTCGGATCAGCACTCACTAGAGACTGCGACTGTCAGACTCTCAACTCTCGCTCAACAGTAAACCGGATCTCCAGAGCAATGCTTGGAGAGATATGGGGACTACGATATGATATA 3300
r15A .....T.....A..... 3300
r15B .....T.....A..... 3300

#1 CTGGGCTGACAGAGTATCTCAGAGATATCATGAGTGTATGAAATGCAATGCACTTACATTAACAATCCCGCACTGTTTCTGAAATCGGGAGATCTCAGACTGGCTTGGAGCGAGTGTGATTAAT 3450
r15A ..... 3450
r15B ..... 3450

#1 GTTTCTACTGTAAAGCTGACCGTATGAGCAGATACAGCAATGACAGAGCTGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 3600
r15A ..... 3600
r15B ..... 3600

#1 ACCTTTCTGACAGCGCTCTCTGAGGAGGAGGATGAGGTTACCTCGGCTACAGCGGAGAGAGACTGATCATATTAACAGACTCCATATAAGCTCTCTGCTCCGACTTGGGTATCGGACTTT 3750
r15A .....T..... 3750
r15B .....T..... 3750

#1 CAGATAACACTTACTAGTACTCTTGAAGAGGTTGGCTTACTGAGAGTCTCGCTCCACATCAGAGAGGACTTAAAGCAAACTGCACTGAGACTGACAGATATTAACAGAGATCTTGGTGTGATGAT 3900
r15A ..... 3900
r15B ..... 3900

#1 GAGCAATTTAGCTGAGCGGCTGACAGCTGATCTCTGAAATGAGACTGAGCGGAGACTCTGACAGCTGATGCTGAGTGGAGGCGGATTTCCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4050
r15A .....T..... 4050
r15B .....T..... 4050

#1 ACCTGACGTGAGAGAGGAGTCCAGCGCTGTGTGAGAGCGTGTGACAGCGCTGCAATTCACCTCCGATACATCAAGAGGAACTACTACCGTGTGCACTTCCGAAAGCGCAATGACATACCATCGAA 4200
r15A ..... 4158
r15B ..... 4158

#1 GACGGTCTATGGCGTGGACTCCACTGGAAGCGTCAAGGCTCGGCTTCCATCTAAGCGCAGACCGGCTCTCATACTAGGATACAACACTACTCCAGCATGAGGAGATGTCGAGTTCACGCTACGCTACTGATCCA 4350
r15A ..... 4158
r15B ..... 4158

#1 GATGAGAGAGAGATGCGAGAGTGAAGTCTACTTTCATCCCAAAATAGGTTGCTCTATATCTGACGATGTGGAGCGTTGAGATGAGAGAGATATCGAGAAAGCTCAGGCTTGGCTCAACATGACTTC 4500
r15A ..... 4197
r15B ..... 4197

#1 AATATAGTCAAGTCTCGCGGCAACAGCGCGGCTGATCAGAGGCAATGCGGAACTCAAGTCACTCATACAGGATACATCCCTGTGAGCGCGAGATGAGAGCTTCGAGTGCAGCTCAGCTCGCGCTCC 4650
r15A ..... 4347
r15B ..... 4657

#1 GTCAGAGTGTGAAACAGCGCTTGGGCTGAAACAGCTGAGCGGGCTCAGCTGATCGGACTCGCGGCTGCAATGACATATGCTGAGCGGGTGTGAGGACTCTGCTCTCTGCTCTATCTGCTC 4800
r15A ..... 4497
r15B ..... 4607

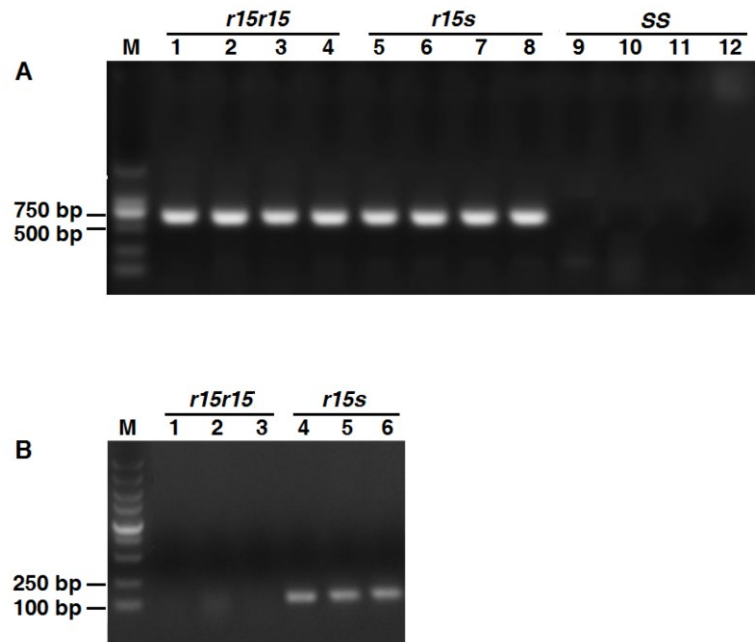
#1 ATCAATTCATGTGAGGCGGACTTGAACCGCTTGGAGCACTGTGATGAGAAATCGGCTCGGTGATTCGCGGCTGAGCGAGTGGGATAGCGGCCAGCAAGCAACAGCCATGAAAGCTCCAGCCCATC 4950
r15A .....C..... 4647
r15B .....C..... 4757

#1 TGGACAGAGAGTCAAGGCCCGGACTTGCATCCATGATGACACTGAGAGAGTGTGATGATGGATCGAGGATCAACCAATCAAGAGGAGCTATTTCCGCGTGGAGCTGGAAATCGGCTCAGCGGCTTTAGGAC 5100
r15A ..... 4797
r15B ..... 4907

#1 CGACCGCAGCGGAGAGTCCGCTTGCACAGTAGCACTTGGTTCACACAGCTTTTTCAGCGGAGTCACTACAGCGGAGTCACTACAGCGGAGTATGAGCAATG 5208
r15A .....C..... 4905
r15B .....C..... 5015

```

**Figure S3.** The full length cDNA sequence of pink bollworm cadherin protein PgCad1 for alleles *s1* (GenBank accession no. AY198374) from susceptible strain APHIS-S; and *r15A* (GenBank accession no. KY814704) and *r15B* (GenBank accession no. KY814705) from resistant strain JL46). The blue shading indicates differences in *r15* compared with *s1*: deletion of 303 bp for *r15A* (4159 to 4461, indicated by hyphens) and 193 bp for *r15B* (4159 to 4351, indicated by hyphens).



**Figure S4.** PCR detection of *PgCad1* genotype using primers in Table S1. (A) Primers for *r15* (*r15F* and *r15allR1*) generate a single band of 652 bp in *r15r15* from JL46 (lanes 1-4) and in *r15s* ( $F_1$  progeny of APHIS-S  $\times$  JL46; lanes 5-8); and no band in *ss* from APHIS-S (lanes 9-12). (B) Primers for *s* (*notr15F* and *r15allR2*) generate no band from *r15r15* (lanes 1-3) and a single band of 152 bp from *r15s* (lanes 4-6).

**CAGGGAACACTACTACCGGTCGT** GTAAGTTCCTCTCTCTCTCTCTATTTAAGAGCTCGCGCTTGTGCGTGGAGTAATCGCCATTCTCTCTCTCCGCCAAAACCTCACCTCCCGA 120  
TACGACACGACCTGCACCTTCTCTTTATTTGTTTCAATAATGTTACTAGGTCTAGGTCCTCTCTCCCTTAAATTTTTCTCTATGATGTTGTTATAAAGAAATCGTGTC 240  
GTATCAGGTGGCCCAATCATATTTCTCTCCGGTCTCTATAGTATCTTCAACTCTTCCAGTACTTTTCATTGACACCATTTCCGTCACGCTTTTACTTATTAATGATTAATAA 360  
TAAAAAACCCCTCTCTACCCCTCTCTCTCCCTCAATTTTCTCTCTATAATGTTGTTATAAATAAAGTGGTATCAAGTCTACTTTATTTGACCATATTCATTAGCTCC 480  
GACCATTTTACAGGGGCAATTAACCAAGAAATGAATGAATAAAGATGAGTATTTTCCGCAATAAATAAAGAAATGTTGTTGAAATCCGAAACAGACAGAATAGTAAAA 600  
ATAAGTTTACCAAAAGCAATCATCAATTTGAAAAATAATGAATCAAGATATAAAAAATGTTGATTTGGATAAAAAATAAAGAAATAAAGGACACTAAAAAGGAAATAAAGTCTA 720  
CAAGTAAATGTTACTTACCTACCTATCACTTTTATGTAAGCCCTCATCATCAATTTAAGAGCCACGCTTGTGCGGTGAGCATTTCCATCCAGTCTATCAAAAGCCAAATC 840  
CTTGACTTCCCTATAAGACAGACGTTAACATTTTCTTAATCTGTCCATGTAAGCTCTCTTGGTCTTCCCTTCTCCCTTTCCTCTAGCTTCTCTATGATGTTTTATAA 960  
TTGCTGCTGTCAAAATTTGAAGCCCTATGCTATAGTTTTTGTCTTAGTAGGATTTTGTGCTCATCACCTTTTGTCTGATCGCATTTACCATTTAATTTAAGCTTTTATAAATTC 1080  
CCTACCAACGACTGTCCTTCAAATCACTACAAACCGATTCTTGTATCGAGCACTTCCGAAAACCGCAATTTGACATATGGGCTATATTTATTTGGGTCGCATAACCTGTC 1200  
TCTCACGTTCAAACACGATTTTTATAAAGCAAAACATTTTTTAAAAATAATTACACTATGGTACATTAAGTATGATGTTTTATACTCGTGTCTTATAATCCAGTTTTTATCAAATTAAT 1320  
TGAATTTAACCAAAAACAGCGTAACCTAAGTTGACTTTTGTAGTAACCGAGTGTTCCTCAAAAATTCGACTGTCCTCATATAATTTACTCCATATAAACACACATTACTTAAAAAC 1440  
TCTTCAAAGCGCTATTAATAATTAATTTTGGCGTAATAACACGTAATTTAACTTTTAACTTTAACTTTAACTTTAACTTTAACTTTAACTTTAACTTTAACTTTAACTTTAACTTT 1560  
TGTCCCGCGGCAAAACGAAAAATACGAAACACTATGTTACCATACGAAACCTGAGTCTTAAATTTAGAAATTTGATGACATTTGCGCCAAAAGTCGATAACATATAATCAAATATTATT 1680  
GGGGCGCGGGGAGTGCCTGGCTTTGAAATTCGCGGAGTCACTCGCGAGGACGCGCGGAGTAGGTGCGTATTAAATGGGAGACCGCTGAGTACCCGCTGTGATAGGTGAGTTTT 1800  
TGAATTTGAATAATCTTTACGTTATCTTTAAAAAAGGAGCAAAATTTATTGACACCATCTGTTTTATAAAAATAAATACCATATGAGTGTATTTTGAATAACACGTTGGTTTT 1920  
TACAACGTGAGTTACCGATGCTCTGGCTTACTTGGGAAATGTAACGCAAGTTGAGCCAGTCTAAAGCAATTTGCTCAATTTTACAAAATTTAATAGTAAAAAT 2040  
AGTTTAAATTAATAATTTTCTGTACAGAGTGTGCGAGTTTATCGTCCCAAGTCCCTAGTCAACACGCTGTTGACCATCGGTAGTGTTCGAGCAGGGTAGCGTCTGATCTCAGAA 2160  
ATCATCTCGATATTTGCTGACTGTATCCAGGTATACGCTCCGACCTCGGACACCCAGTGAACGAAAGTGGAGACTATGATGGAAGATTTCTAAAGCCATGCAATAGCCATAAAAGCCA 2280  
CTACACGATTTGATGGGGGACTTTAATGCACAGTTGGGGAAGCCTTGGCGTACGAGCTGAGAGTAGGGCAATTTGGATTTGGGGTTTGGAAACCCAGAGCCGGAATGCTGGCGGACTT 2400  
TATGGAGAAGAAGAACTCTATATGATGAACCTCTTTCAGAAAGCCCGCACCCGCAATGGACCTGAATAAGTCCAGTGGAAATATAAAAAAGCAGATGATTTTATCATGTGAGC 2520  
AAAAAGCACAATTAATGATGCTCTGTGATCAATGCCGTTAAGACGGAAGCAACGAAACGGGTTTCCGAAAGGCTTTAGCACCATAGCCACATCCATACGATCGCGGAGTTATA 2640  
CAGAAGCCGGAAGATATACTGCCACTTGGCTTGGCGTTGTGGACTATGAGAAAGCCTTTGATTCGATCGAGACCTGGCGGTTTGCAGTCTCTTACAGAGTTCGCAAGTGAACACT 2760  
AGGTACATCGAAGTCTAAGGACCTCTACCAAAATGACACTATGTCAAGTTCGGAACAGAACGAGCTCTAATCCGATCCAAGTCCAGCGAGGAGTGAACAGGGAGATGTTATCTCT 2880  
CCGAACCTGTTCACTGCTGCATTTGAGGATATTTTAAAGCTTGTGACTGGAAAGGATTTGGCATCAACATCAACGGCAGTACCTGACGCCCTTCTGATTTCCGATGATATAGTCTC 3000  
ATGGCTGAGACCTGGAGACCTGGAACACATGCTCGAGCATCTCAGTAACGGATCCCAATGAGTGGGTCTTAGCAGGAACATGGCAAAAGCAAAAATATGATCCAAAGACCATGTCGCA 3120  
CCCACTCCAGTTCAGTTGGGAACTGACACTCGAAGTTGAGACCAAGTATTTACTAGGACAAATAAATCCAGTTAGGTAAGTCCCACTCTCGAAGAAAGAGATCTCTGCTGGATCCAA 3240  
CTCGGATGGGACGCTTGGGAACTGCGGAAATCTTCTCGTCCAAAATACCTCAGTGTCTCAAGGAAAGTCTTTGACCAGTGCCTTTGACCAGTGAATTTACAAAATTTAATAGTAAAAAT 3360  
TGGCGCTTACTATGGTCTCGTGAAGAGGCTCGGTGTGCTCAGCGGGCAATGGAGAGAGCTATGCTCGGATTTTCCCTGCTCGATCAAAATCAGAAATGAGGAGATCCGGCAGGAGAACT 3480  
AAAGTACCGACATAGCTCGGCAAAATGCAAAAGTGAAGTGGCAGTGGGACAGACATAGCGAGGAAACCGGTGAAGTTCGCGGAAAGCCGCTGGATCGCGGACGAGCAGCCGACT 3600  
GTCCTGGAGATCTTGGGGAGGCTATGCCACGAGTGGGCGCTGTAAGGCTGTAGAGCATGATGACATACATACATAAATAAAGTCAATGATGATGATGATGATGATGATGATGATGAT 3720  
TAGAATCCATTTGCTGAAAACGAATTTGAATGAATTTGATGCTCCATTAATAATATTTGACTCAGACAGAACTAATGATGATTAATTTGAGAAAGTTTGGATTTATTTTT 3840  
ATTTAAGAGATTTGTTTAAATTTGTTATGTTTCAATTTAATTTGATTTCTGTTTTTAAATGTTTTCAGTAAAGTAAACACCATAGATGCGGAAAGT(TTACCATTATGTTACTTAC)TAA 3960  
ACTATGCGTATCTTTTTGTCGCCCATTTTTATGGAGACTTTTCAGACTTAGAACTAAAACTACCTAAGTATATTTAAAAATTTTGAATTTTATGATTTTAAATATATAAATG 4080  
ACCTCCATAAATAAATAATGCTGAACTGCTATTGTTTCTATTTTAAAAAAGTTTATCTCAATAAATAAATAAATAGGTAACGAAAGTGAACCTAACCAATAAATAAAGTATTA 4200  
TGAATGTTTTCGTGATTTAATAATAAAGTCAAAAATCTTTATCTTAATTTAAAAAAGAAAGTTTGAATTTAATGTAACGTTCCCAATAAATAAATAAATAAATAAATAAATAA 4320  
ACACAGTTGTTGTTTTCTGCCACCTCCCTAGTATAA 4440  
AGTAAGAAATATTAACAAACAATAGCATAA 4560  
TGGACTCCACTCTGGAAGCCCTCAAGACTCGGCTCCATCTGAACCGCAGACCCGGCTCTCATACTGAGGATACAACCTACTGCCAGCATGCGAGGCTGTCGAGTTCAACGCTCA 4680  
TCGCTACTGATCCAGTAAATGGGGCACAGCTACGCAAGGACTTTGAGCAAGGGGTAACAGTGTGTTGTAATAAAGGGGATGATCGGCTCCTCACACTGAGGATACAGTCAACAGGA 4800  
TTTTGTTCCACGTCATCAGTGATTCAGAGAGGCGTGCCTATAGAATCATAACATAAACAACGCTATATACGTCCTGCTGGGACAGGCTCCCTCAATCAACCGGAGG 4920  
GGTATGGAGTACTCCAGCTCAATGCGGGTGGTGGAGGTTTACAGTTTGTAGAACGCTTGCCTCTCCTTTGAGTGCAGGTTGCAATCCAGCACAGGCTAAACCAA 5040  
TTATTTGCGAATTTGTTTTCGAATTTATGTTTGGATCATAAATGATATCACGTGCTCAACAAAACATCTGAGGAAACCCGAGAAATGCATTTCCGAAAGGATGTGACCTAACCTGTA 5160  
TTGGGCTGGTTTTCCCTGGTGGAAAGTGAAGCATGCTGCTTACCTTTCGCTCTTTAAAAACCGGACTGTGCAATCTTCAGGTTAGGCAAGCGGACCTGTGAAATACGGATTAAG 5280  
GCTAGGGAGATGATGATGATGAGTCACTTCAAAGTTTTTAAACAAAAGGATTGGATTCCGATTCTGTAATGTAATAATAGACCTCTGTAACATAAATAGCCACATAGCTAGGTGATCAAG 5400  
GAACCAAGCAATACTTACTACCTTATTCAAGATTAATTTCTGAAGAAATGAAAAATTTCCAAACACAAAATAACAGTGAATCTTTTATTTATTTCTTACTTTCCAGATGAGAAAGC 5520  
AGATACCGCAGAGTGAAGCTACCTCAATTTCACTCCAAAATAGGGTCTCTCATATTTCTGAACGATGTTGGAGACGGT 5601

**Figure S5.** Sequences of the three nested TEs inserted into Exon 28 of r15 gDNA sequence. Blue=MITE-1\_PGo, Green=RTE-5\_PGo, Gray=SINE-1\_PGo. TSD= italic and underlined letters. TIR=underlined letters. Microsatellite repeats of (AYG)<sub>4</sub>(ACAT)<sub>4</sub> for RTE-5\_PGo and of (TTAY)<sub>4</sub> for SINE-1\_PGo are indicated with a pair of parentheses.



Range 1: 1719 to 2714 [Graphics](#) [Next Match](#) [Previous Match](#)

Alignment statistics for match #1

Score	Expect	Identities	Gaps	Strand
672 bits(744)	0.0	750/1001(75%)	28/1001(2%)	Plus/Plus
Query 634	CGAACGACGGGTTCCGAAAAGGCTT	AGCACCATGACACACATCATAGATGCGGCA		693
Sbjct 1719	CGAACGACGGGTTCCGAAAAGGCTATAGTACCGTAGACACACATACATACGCTGCGCCA		1778	
Query 694	GTTTATACAGAACCCGAGAGTATAATCTGCCACTTTCCTGGCGTTTGTGGACTATGA		753	
Sbjct 1779	GGTTATACAGAACCCGAGAGTATAACTTGGCCCTCTCTAGCGTTTGTGGACTACGA		1838	
Query 754	GAAAGCCTTGTATCGATCGAGACTGGGCGGTGGCAGTCTCTCAGAGGTGCCAAGT		813	
Sbjct 1839	AAAAGCCTTGTATCGGTGAAACATGGCGGTACTGGAGTCTCTCCAGCGATGCCATAT		1898	
Query 814	GGACCATAGTACATCGAAGTCTAAGGACCTCTACAAAATGACACTATGTCAGTTGG		873	
Sbjct 1899	TGACTATCGGTATATCGAAGTCTGAAAGTATTGTATAGAAAACGCCACTATGTCGGTCCG		1958	
Query 874	AGAAGCAGAACGAGCTCTAATCCGATCCACTGACGAGGAGTGAGACAGGAGATGG		933	
Sbjct 1959	ACTTCAGAAAACAGGTTTCGAGGCGGATCCCTTAAAAGAGGCGTAAAGACAGGAGAGCTG		2018	
Query 934	TATCTCTCCGAACTGTCACTGCGATGGAGGATATTTTAAAGCTCTGAGCTGGAA		993	
Sbjct 2019	CATATCTCCGAACTGTCACTGCGCGCTTGGAAAGCGCTCAAGCTCTTGGAAATGGGA		2078	
Query 994	AGGATTTGGCATCAACATCAACGGCGAGTACTGACGACCTTCGATTTGCCGATGATAT		1053	
Sbjct 2079	AGGACTGGCATCAATTAACGGCGAATACATTACTACCTTCGTTCCGCGCAGATAT		2138	
Query 1054	AGTCCTAATGCTGAGACCTGGAGACCTGAACACATGCTCGAGCATCTCAGTAACCG		1113	
Sbjct 2139	CGTAGCTAGGCAAGTGGATGGAGACTCAGTACGATGCTCGAAGGCTCAATCGAGT		2198	
Query 1114	ATCCCAATGAGTGGGTCTTAGCACGAACATGGCAAGACAAAATATATGTCATGACCA		1173	
Sbjct 2199	CTCCACAGGCTAGCTTAAATGAAAGCGGCAAGACGAACTCATGTCTAATGTCCA		2258	
Query 1174	TGTCCACCCCTCCAGTTCAGGTGGGAAAGTTACACTCGAAGTTGTAGACAGTATAT		1233	
Sbjct 2259	TGTCCACCCCTCCAGTTCAGGTGGGAAAGTTACACTCGAAGTTGTAGACAGTATAT		2318	
Query 1234	TTACCTAGGACAAATAATCCAGTTAGGTAAAGTCCAACTTCGAAAAGAGATCTCTCGTGG		1293	
Sbjct 2319	ATACCTGGACAGACAGTCCAGTTAGGAAGTCCAACTTCGAAAAGAGGTCAATGCGCG		2378	
Query 1294	GATCCAACTCGATGGGACGCTTGGGAAAGTTCGGAATATCTCTGTCGCAAAATACC		1353	
Sbjct 2379	AATCCAGCTCGATGGGACGCTTGGGAAAGTTCGGAATATCTCTGTCGCAAAATACC		2438	
Query 1354	TCAGTCTCTCAAGAGAAAGTCTTTGACAGTGGCTGTCCAGTGTGATGATCTACGGGAG		1413	
Sbjct 2439	TCAGTCTCTCAAGAGAAAGTCTTTGACAGTGGCTGTCCAGTGTGATGATCTACGGGAG		2498	
Query 1414	TGACAGTGGCGGCTT-ACTATGGTCTCTGAGGAGGCTCGGTGCTCGCTCAGCGGCAA		1472	
Sbjct 2499	CGAAACGTGG-TGCTTCCACAAAGGGCTTATGAACAGACTCAGAGTCTCAAGAGCTA		2557	
Query 1473	TGG-----AGA-----GAGCTATGCTGGTATTTCCCTGCTCGATCAATCA		1514	
Sbjct 2558	TGGCTCTCCATATCTCCGGCTCTAGTATGCTCGGAGTTTCTTACGGGATCGAATCA		2617	
Query 1515	GAAATGAGGATCCCGAGGAGAACTAAAGTACCAGACATAGCTCGGCAAAATGCAAGC		1574	
Sbjct 2618	GAAATGATGAGATACGTAGAAGAACGAGAGTACCAGACATAGCT---AAACGGATTAGC		2673	
Query 1575	---TGAAGTGGCAGTGGGACGACATAGCGAGGAGAAC		1611	
Sbjct 2674	TCTTTGAGTGGCAGTGGGACGATCATATAGCCCGGAGAAC		2714	

Range 2: 361 to 731 [Graphics](#) [Next Match](#) [Previous Match](#) [First Match](#)

Alignment statistics for match #2

Score	Expect	Identities	Gaps	Strand
163 bits(180)	3e-43	265/374(71%)	7/374(1%)	Plus/Plus
Query 243	GTCATCCAGGTATACGCTCCGACCTGGGACACCACGATGACGAAAGTGGAGACTATGAT		302	
Sbjct 361	GTCGTTCAAGTATATGGCCGACTTCGTCACACAGATGAGGAAAGTGGAGTCTATGAT		420	
Query 303	GTTGAGATTTTAAGCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT		362	
Sbjct 421	GAGGACATTAGTAGGGCCATACACTAGCAAACTCACTCAATGTTGTTATGGGAGAC		480	
Query 363	TTAATGCACAGTTGGGAAAGCTTCCGTAACGAGCTGAGAGTGGGCAATTTGGATT		422	
Sbjct 481	TTAATGCAGAGCTAGGCAAGAGGCTTACGAGTTGAAAGTGGGCAATTTGGGCAAT		540	
Query 423	GGGTTTGAACACAGAGCCGATGCTGGCGGACTTTATGGAG-AGAAGAACTCTA		481	
Sbjct 541	GGCAACGGAACTTAGAGGACAGAGGCTGGCGGATTTCTGGAGAAAGAGGAC-TCTT		599	
Query 482	TATGATGAATCTCTTCCAGAAAGCCGAC-CCGAAATGGACCTGAATAAGTCCAG		540	
Sbjct 600	TGTGATGAATCTCTTCCAGAAAGCCGACTCAGCAAAATGGACTGGATGAGGCCGA		659	
Query 541	TGGAATTTAATAAAGCAGATCGATTATCATGTGACGA-AAAAG-CACATATCAAT		598	
Sbjct 660	TGTTCCACGAAATGAGATAGACTCATATAGC--CGTAGAGACGAAATATTCAT		717	
Query 599	GATGCTCTGTGAT	612		
Sbjct 718	GATGCTCAAGTAT	731		

**Figure S6.** Nucleotide sequence alignment between RTE-5\_PGo and RTE-5\_DPI. Query=RTE-5\_PGo, Sbjct = RTE-5\_DPI.

Range 1: 436 to 770 [Graphics](#) [Next Match](#) [Previous Match](#)

Alignment statistics for match #1

Score	Expect	Method	Identities	Positives	Gaps
516 bits(1328)	0.0	Compositional matrix adjust.	249/336(74%)	287/336(85%)	14/336(4%)
Query 100	ERTGFRKGFSTIDHIHTMRQVIQKTEEYNLPLCLAFVDYEKAFDSIETWAVLQSLQRCQV				159
Sbjct 436	E+ GFRKG+ST+DHIHT+RQVIQKTEEYNLPLCLAFVDYEKAFDS+ETWAVL+SLQRC +				495
Query 160	DHRYIEVLRDLYQNDTMSVREQNQSSNPIQLQRGVRQGDGISPFLFTAALEDIFKLLDVK				219
Sbjct 496	D+RYIEVL+ LY+N TMSVR Q QSS I L+RGVRQGD ISPFLFTAALED FKLL+W+				555
Query 220	GFGININGEYLTHLRFADDIVLMAETLEDLNTMLEHLSNASQMA-----KTKIMSNDH				272
Sbjct 556	G GININGEY+THLRFADDIV+MA+++E+L+TMLE L+ SQ KTK+MSN H				615
Query 273	VAPTPVQVGNVTLEVDQYIYLGQIIQLGKSNFEKEISRRIQLGWAAFGLRNIFSSKIP				332
Sbjct 616	VAPTPV V N LEVVD Y+YLGQ +QLG+SNFEKE++RRIQLGWAAFGLRN+FSSKIP				675
Query 333	QCLKRKVFDQCVLPVMIYGETWPLTMGLVRRRLGVAQRAM-----ERAMLGISLLDQIR				386
Sbjct 676	QCLK KVF+QCVLPVM YGETW T GL+ RL VAQRAM AMLG+SL D+IR				735
Query 387	NEEIRRRTKVTDIARQIAKLKQWAGHIARRTGEGR		422		
Sbjct 736	N+EIRRRT+VTDIA++I+ LKWQWAGHIARRT +GR		770		

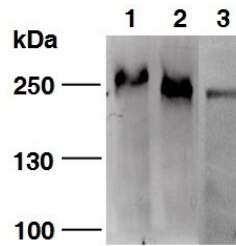
Range 2: 1 to 114 [Graphics](#) [Next Match](#) [Previous Match](#) [First Match](#)

Alignment statistics for match #2

Score	Expect	Method	Identities	Positives	Gaps
114 bits(285)	3e-31	Compositional matrix adjust.	57/114(50%)	78/114(68%)	14/114(12%)
Query 1	MYVEISKAMHSHKSHYTIVMGDFNAQLGKRCGNELRVQFGFGVWNTGRMLADFMEKKN				60
Sbjct 1	MY +IS+A+H+ K+H+ +VMGDFNA+LGKR EL+VG FG G N RG+ LADF+EK+				60
Query 61	FYMNSSFRRKPAPAN----GPE-----MSTKKHIFNDVSVINAVKTESE			100	
Sbjct 61	++MNSFF+KP+ P+ M+ K+ IFNDVSVI+ VKT S+				114

**Figure S7.** Amino acid sequence alignment between RTE-5\_PGo and RTE-5\_DPI. Query=RTE-5\_PGo, Sbjct = RT-5\_DPI. Range 2 is the EP domain. Range 1 is the RT domain.





**Figure S8.** Western blot of cadherin fusion proteins sPgCad1-GFP (lane 1) and r15APgCad1-GFP (lane 2) and r15BPgCad1-GFP (lane 3) produced in Hi5 cells transfected with vectors containing the *s* and *r15A* and *r15B* alleles, respectively.

1. Morin, S. *et al.* Three cadherin alleles associated with resistance to *Bacillus thuringiensis* in pink bollworm. *Proc Natl Acad Sci U S A* **100**, 5004-5009 (2003).
2. Tabashnik, B. E. *et al.* Sustained susceptibility of pink bollworm to Bt cotton in the United States. *GM Crops & Food* **3**, 194-200 (2012).
3. Morin, S. *et al.* DNA-based detection of Bt resistance alleles in pink bollworm. *Insect Biochem Mol Biol* **34**, 1225-1233 (2004).
4. Wang, L. *et al.* Resistance to *Bacillus thuringiensis* linked with a cadherin transmembrane mutation affecting cellular trafficking in pink bollworm from China. *Insect Biochem Mol Biol* **94**, 28-35 (2018).