

Identification and Sequencing of the T-DNA Insertion Region in the Genomic DNA of Transgenic Bt-cotton (cv. Bikaneri Nerma) using Genome Walking Method

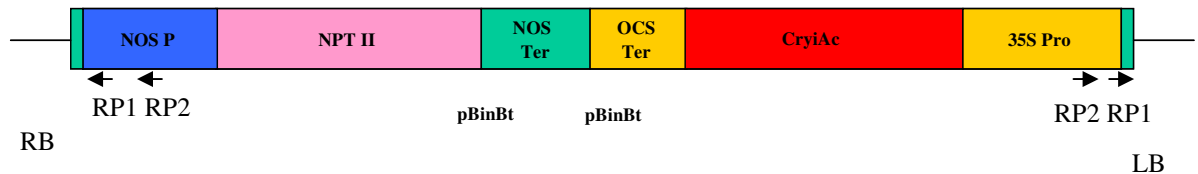
**NRC on Plant Biotechnology, New Delhi
University of Agricultural Sciences, Dharwad
Central Institute of Cotton Research, Nagpur**

INTRODUCTION:

NRC on Plant Biotechnology, New Delhi, University of Agricultural Sciences, Dharwad and Central Institute of Cotton Research, Nagpur have jointly developed transgenic cotton expressing a truncated Cry1Ac toxin of *Bacillus thuringiensis* in a popular cotton variety Bikaneri Nerma (Katageri et al., 2007). The programme was supported by the National Agricultural Technology Project of Indian Council of Agricultural Research (ICAR). The transgenic event namely 'Dharwad Event' was characterized by molecular analysis such as Southern hybridization, gene integration, and segregation and expression. The T-DNA flanking sequences in the cotton genome were also analyzed by genome walking approach, the details of which are furnished in the following account.

Identification and sequencing of the T-DNA flanking regions of transgenic plant lines by using genome walking

The genomic DNA of transgenic cotton plant was prepared. This was single copy insertions generated by *Agrobacterium* mediated transformation. The region of integration of T-DNA in the genomic DNA had to be identified. T-DNA maps and sequences close to the right border (RB) and left border (LB) regions of the T-DNA are depicted below.



EXPERIMENT

The isolation and cloning of the T-DNA flanking regions were carried out using Universal Genome Walker Kit from Clontech. The experiment was carried out by preparing a library with PvuII enzyme. The DNA was digested and the adapters were ligated. Primary PCR was carried out using NosRP1- AP1 and LBRP1-AP1.

The product from the primary PCR was used for secondary PCR using NosRP2 – AP2

Clone 36 sequence analysis Product of AP2 and NOSRP2

CTATAGGGCACGCGTGGTCGACGGCCCCGGGCTGGTCTGTCAACTACGGACAAT
ACGCTTACGGAGGCTACTTCCCCAACCGCCCAACACTAAGCCGAAGGTTTCATG
CCTGAGAAAGGCACCCCTGAGTATGCAGAGCTTGAAAAGAACCCTGAGAAGGT
CTTCTTTAGAATCATGTCTTCGCAGCTACAGTCCCTGATTGTCATTACGGTGGTC
GAAACGCTGTCAAACACGCATCGGATGAGGTGTATCTTGGACAGCGAACCCC
CAACTGGACCACTGATGCAGTTCGCTACAAGCTTCGGATGCTTTCAATAGGA
GACTTGCTGAAATCGAAGGGGAAATCTTAAAGATGAACATCGACAATTATGAC
CCCCGCCGATGACGCGGGACAAGCCGTTTTACGTTTGGAAGTACAGAACCGC
AACGTTTATC

Primer sequence on NOS promoter

GGAAGTACAGAACCGCAACGTT

pBIN19 vector sequence including the NOS promoter region

GATCATGAGCGGAGAATTAAGGGAGTCACGTTATGACCCCCGCCGATGACGCG
GGACAAGCCGTTTTACGTTTGGAAGTACAGAACCGCAACGTT

Adapter sequence

CTATAGGGCACGCGTGGTCGACGGCCCCGGGCTGGT

Part of Right Border repeat region

CGACAATCT

**Alignment of Sequence obtained with NOS Promoter region on pBIN 19 vector
sequence**

```

      10      20      30      40      50
Seq obtained
Gen bank sequence of Nos P+RB AACACTGATAGTTTAAACTGAAGCGGGAAACGACAATCTGATCATGAGC

      60      70      80      90     100
Seq obtained
Gen bank sequence of Nos P+RB GGAGAAATAAGGGAGTCACGTTATGACCCCGCCGATGACGCGGGACAAG
GGAGAAATAAGGGAGTCACGTTATGACCCCGCCGATGACGCGGGACAAG

      110     120     130     140     150
Seq obtained
Gen bank sequence of Nos P+RB CCGTTTACGTTTGGAACTGACAGAACCACAACGTT
CCGTTTACGTTTGGAACTGACAGAACCACAACGTTGAAGGAGCCACTCA

      160     170     180     190     200
Seq obtained
Gen bank sequence of Nos P+RB GCCGCGGGTTCTGGAGTTTAATGAGCTAAGCACATACGTCAGAAACCAT

      210     220     230     240     250
Seq obtained
Gen bank sequence of Nos P+RB TATTGCGCGTTCAAAGTCGCCTAAGGTCACATCAGCTAGCAAATATTT

      260     270     280     290     300
Seq obtained
Gen bank sequence of Nos P+RB CTTGTCAAAAATGCTCCACTGACGTTCCATAAATCCCTCGGTATCCAA

      310     320     330
Seq obtained
Gen bank sequence of Nos P+RB TTAGAGTCTCATATTCACTCTCAATCCAAATAACTGCA

```

The highlighted region is the overlap between pBin19 derived pBinAR vector sequence and the sequence obtained

Sequence flanking the pBIN19 vector sequence where the T-DNA integration has taken place


```

CTGTGTCAACTACGGACAATACGCTTACGGAGGCTACTTCCCAACCGCCCAA
CACTAAGCCGAAGGTTTCATGCCTGAGAAAGGCACCCCTGAGTATGCAGAGCTT
GAAAAGAACCCTGAGAAGGTCTTCTTTAGAATCATGTCTTCGCAGCTACAGTCC
CTGATTGTCATTACGGTGGTCGAAACGCTGTCAAACCACGCATCGGATGAGGT
GTATCTTGACAGCGAACCCCAACTGGACCACTGATGCAGTTCGCTACAAG
CTTCGGATGCTTTCAATAGGAGACTTGCTGAAATCGAAGGGGAAATCTTAAAG
ATGAACAG

```

BLAST analysis of the Sequence flanking the T-DNA right border

```

> gb|AF361893.4|  Gossypium hirsutum bacterial-induced lipoxygenase (Lox1) mRNA,
complete cds
Length=2835

```

Score = 100 bits (110), Expect = 4e-18
 Identities = 159/228 (69%), Gaps = 0/228 (0%)
 Strand=Plus/Plus

```

Query 5      GTCAACTACGGACAATACGCTTACGGAGGCTACTTCCCCAACCGCCCAACTAAGCCGA 64
          || ||||| || ||||| |||| | ||||| |||| ||||| || || ||||| ||
Sbjct 2254    GTTAACTTTGGTCAATATCCTTATGCAGGCTACCTCCCGAACCGACCGACTATAAGTCGT 2313

Query 65     AGGTTTCATGCCTGAGAAAGGCACCCCTGAGTATGCAGAGCTTGAAAAGAACCCTGAGAAG 124
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Sbjct 2314    CGGTTTCATGCCTGAGAAAGGAACACCAGAATACACAGAACTAGAGTCCAACCCGGACAAG 2373

Query 125    GTCTTCTTTAGAATCATGTCTTCGCAGCTACAGTCCCTGATTGTCATTACGGTGGTCGAA 184
          || ||||| || || || || || || || || || || || || || || || || ||
Sbjct 2374    GTTTCTTGAAAACAATAACTGCTCAACTGCAGACACTCCTGGGAATATCCTTGATAGAA 2433

Query 185    ACGCTGTCAAACCACGCATCGGATGAGGTGTATCTTGGACAGCGAACC 232
          | || ||||| || | || ||||| ||||| || || || || ||
Sbjct 2434    ATCCTATCAAGGCATTTCGTGATGAGGTTTATCTGGGGCAGAGAGCC 2481
  
```

Conclusions

Right Border

The T-DNA has been integrated into the genomic DNA. The right border region of the T-DNA is integrated and we were able to obtain about 331bp sequence upstream of the sequence provided. This sequence is the putative region of RB integration in the genome. Blast analysis of the above sequence showed partial homology to a bacterial induced Lipoxygenase1 from *Gossypium hirsutum*.

Left Border Region

This experiment was carried out by preparing a new library with EcoRV enzyme (Blunt cutter with recognition sequence **GATATC**). The DNA was digested and the adapters were ligated. Primary PCR was carried out using 35SRP1- AP1 and LBRP1-AP1. The product from the primary PCR was used for secondary PCR using 35SRP2 – AP2. For further confirmation PCR was carried out using 35SRP3-AP2 and the product was cloned into pJET cloning vector and sequenced.

Sequence obtained

CATTAAAAACGTCCGCAATTTGTTATCAAAGTCAAAGTCCAGACTCGGGAAAGATGAATGTGCTTTTTGTC
 ATGAGAAATGCCAGTGGAAAGAAAATTGTCCAAAGCTGAAGAATAAGGGAAAAGCTGCTGTAGATGCTTGT
 GTTGCTAAGCATGATACTAGTACTCTGAACTATCACTGGTTGCATCATCATCGTTCGTTCCATTGATGA
 GTGGATATTGGATGCATATTGAT**ACCAGCCCGGGCCGTCGACCACGCGTGCCCTATAG**

Vector sequence pBin19 (pBinAR) at the LB region

CATTAAAAACGTCCGCAATTTGTT

Adapter sequence

ACCAGCCCGGGCCGTCGACCACGCGTGCCCTATAG

Probable Cotton Genomic DNA sequence

ATCAAAGTCAAAGTCCAGACTCGGGAATGATGAATGTGCTTTTTGTCATGAG
AATTGCCAGTGGAAGAAAAATTGTTCAAAGCTGTAGAATAAGGGTAAAGATG
CTGTGAATGCTTGTGTTGCTAAGGATGGTACTAGTGACTCTGAACTATCACTG
ATTGCATAGTCATCGACGTTCCAGTCAGATGAGTGGATATTGGATGCCTATTG
AT



Alignment of the sequence obtained with the LB region of pBIN19 (pBinAR) vector

```

      10      20      30      40      50      60
pBIN19 LB region  .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Sequence obtained TTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGGCGGTGAAGGGCAATCAGCTGTTGCC
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
      70      80      90      100     110     120
pBIN19 LB region  .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Sequence obtained GTCTCACTGGTGAAAAGAAAAACCACCCAGTACATTAAAAACGTCGGCAATTTGTT
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
      130     140     150     160     170     180
pBIN19 LB region  .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Sequence obtained AAAGTCAAAGTCCAGACTCGGGAAGATGAATGTGCTTTTTGTTCATGAGAAATGCCAGTG
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
      190     200     210     220     230     240
pBIN19 LB region  .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Sequence obtained GAAGAAAAATTGTCCAAAGCTGAAGAATAAGGGAAAAGCTGCTGTAGATGCTTGTGTTGC
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
      250     260     270     280     290     300
pBIN19 LB region  .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Sequence obtained TAAGCATGATACTAGTGACTCTGAACATCACTGGTTGCATCATCGTTCGTTCCATTC
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
      310     320     330     340     350     360
pBIN19 LB region  .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Sequence obtained AGATGAGTGGATATTGGATGCATATTGATACCAGCCCGGGCCGTCGACCACGCGTGCCCT
.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
.....
pBIN19 LB region  .....
Sequence obtained  ATAG
```

The Highlighted region is the overlap between the sequence obtained and the pBIN 19 vector sequence.

The sequence obtained was further analysed by blast analysis.

```
>  gb|AY632360.1  Gossypium hirsutum BAC 106I22, complete sequence
Length=135862

Score = 297 bits (328), Expect = 2e-81
Identities = 187/202 (92%), Gaps = 0/202 (0%)
Strand=Plus/Minus

Query 1 ATCAAAGTCAAAGTCCAGACTCGGGAATGATGAATGTGCTTTTGTGCATGAGAATTGCCA 60
      |||
Sbjct 14619 ATCAAAGTCAAAGTCCAGACTCGGGAAGATGAATGTGCTTTTGTGCATGAGAAATGCCA
14560

Query 61 GTGGAAGAAAAATTGTTCAAAGCTGTAGAATAAGGGTAAAGATGCTGTGAATGCTTGTGT 120
      |||
Sbjct 14559 GTGGAAGAAAAATTGTCCAAAGCTGAAGAATAAGGGAAAAGCTGCTGTAGATGCTTGTGT
14500

Query 121 TGCTAAGGATGGTACTAGTACTCTGAACATCACTGATTGCATAGTCATCGACGTTCCA 180
      |||
Sbjct 14499 TGCTAAGCATGATACTAGTACTCTGAACATCACTGGTTGCATCATCATCGTCGTTCCA
14440

Query 181 GTCAGATGAGTGGATATTGGAT 202
      |||
Sbjct 14439 TTCAGATGAGTGGATATTGGAT 14418
```

Conclusions

Left Border

The T-DNA has been integrated into the genomic DNA. The Left border region of the T-DNA is shown to be integrated and we were able to obtain about 212bp sequence upstream of the LB region. This sequence is the putative region of LB integration in the genome. Blast analysis of the above sequence showed partial homology to a BAC clone from *Gossypium hirsutum*.

Information for LOD test for BN Bt cotton

Flanking sequence **Right** border

CTGTGTCAACTACGGACAATACGCTTACGGAGGCTACTTCCCCAACCGCCCAA
CACTAAGCCGAAGGTTTCATGCCTGAGAAAGGCACCCCTGAGTATGCAGAGCTT
GAAAAGAACCCTGAGAAGGTCTTCTTTAGAATCATGTCTTCGCAGCTACAGTCC
CTGATTGTCATTACGGTGGTCGAAACGCTGTCAAACCACGCATCGGATGAGGT
GTATCTTGGACAGCGAACCCCAACTGGACCACTGATGCAGTTCGGCTACAAG
CTTCGGATGCTTTCAATAGGAGACTTGCTGAAATCGAAGGGGAAATCTTAAAG
ATGAACAG

Flanking sequence **Left** border

ATCAAAGTCAAAGTCCAGACTCGGGAATGATGAATGTGCTTTTTGTCATGAG
AATTGCCAGTGGAAGAAAAATTGTTCAAAGCTGTAGAATAAGGGTAAAGATG
CTGTGAATGCTTGTGTTGCTAAGGATGGTACTAGTGACTCTGAACTATCACTG
ATTGCATAGTCATCGACGTTCCAGTCAGATGAGTGGATATTGGATGCCTATTG
AT

Gene specific primers (Cry1Ac) used for PCR

Cry1Ac FP-----CCAGAAAGTTGAAGTACTTGGTGG

Cry1Ac RP-----CCGATATTGAAGGGTCTTCTGTAC