Supplementary Figure Legends

Figure S1. Sequence alignment of the *TGTT1* elements

Sequence identities over 70% were shaded in blue. The TSD site, 5′LTR, PBS, PPT and 3′LTR regions are indicated using red lines and arrows.

Figure S2. Wet laboratory validation of the TGTT and AACA elements.

Each element was randomly selected and analyzed from each family. A schematic view of the LTR-RT, with the location of the region that was amplified (dashed pink boxes), a picture of the PCR amplification of the LTRs using DNA material of the four Rosaceae species (see Table S1), and the alignments of the sequenced PCR products together with the sequences that were mined from the genome sequence are provided. The positions of TSDs, beginnings and ends of LTRs, PBS and PPT are also indicated.

Figure S3. Multiple alignments of PBS sites from TGTT, AACA and normal TGCA elements of the *Ale* lineage.

A 20-bp region of each PBS site was extracted from all of the *TGTT* and *AACA* elements, as were the consensus sequences of the other four *Copia Ale* families from pear [38]. Sequence identities over 70% are shaded in blue.

Figure S4. INT phylogenetic relationships among TGTT, AACA and normal TGCA elements.

The INT domains (the integrase core domain, pfam00665) of 67 elements from the six *Copia-like TGTT* and *AACA* families and the consensus sequences of each normal *Copia* TGCA family from pear [38] were used to build the phylogenetic tree. The *TGTT* and *AACA* elements are indicated by various colored branches, and the normal TGCA families are shown in gray. The nucleotide sequence distances are indicated by the scales.

Figure S5. Evidence of transcriptional activity in five *TGTT* and *AACA* elements.

For each element, the whole-nucleotide sequences were used as queries to construct the HISAT alignments using the RNA-seq data from the NCBI species Sequence Read Archive database. Normal short pair reads are shown in the gray bar. Too long paired reads are shown in the purplish red bar. One-sided sequences are indicated by green or blue boxes.

Figure S1.

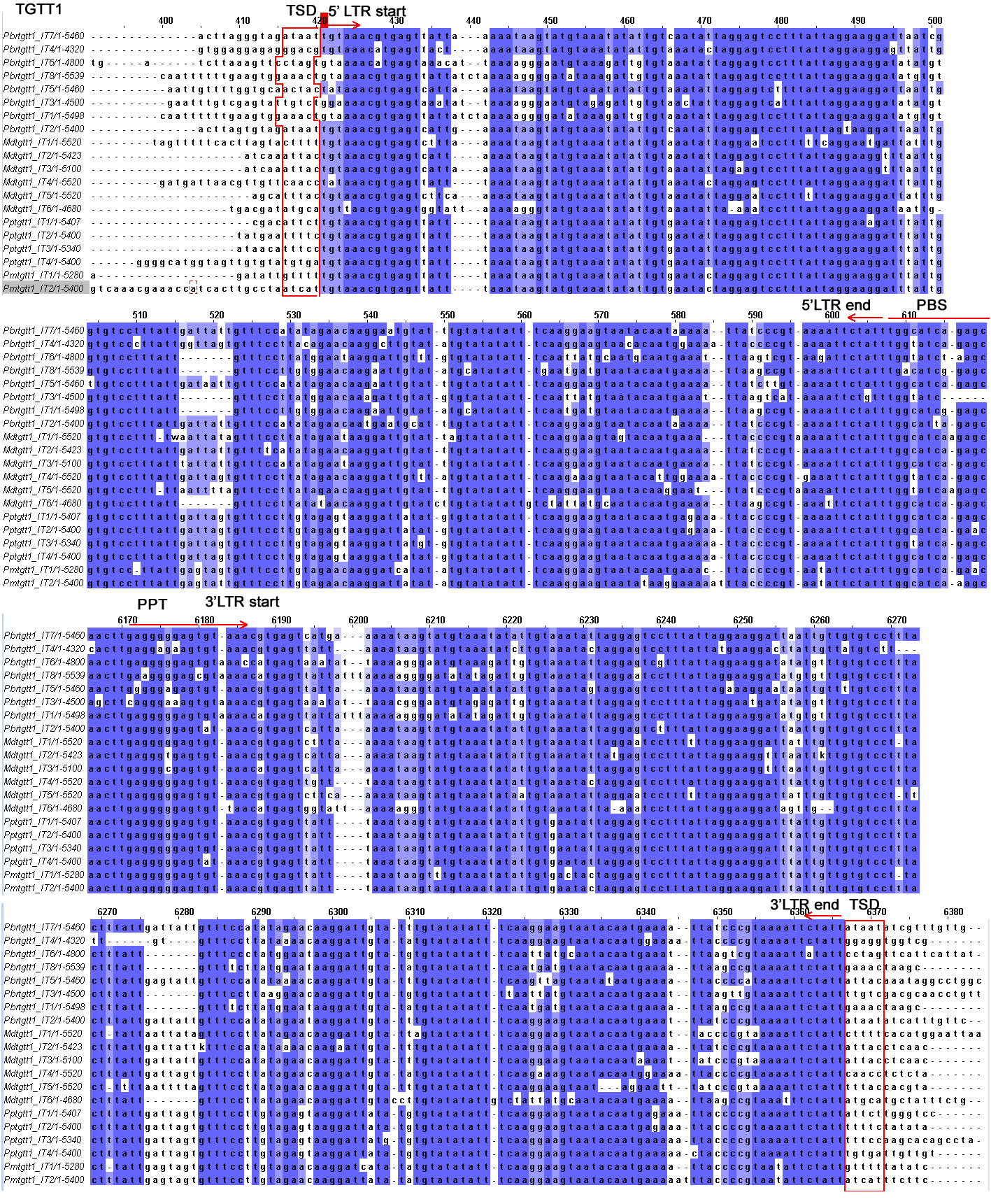


Figure S1.

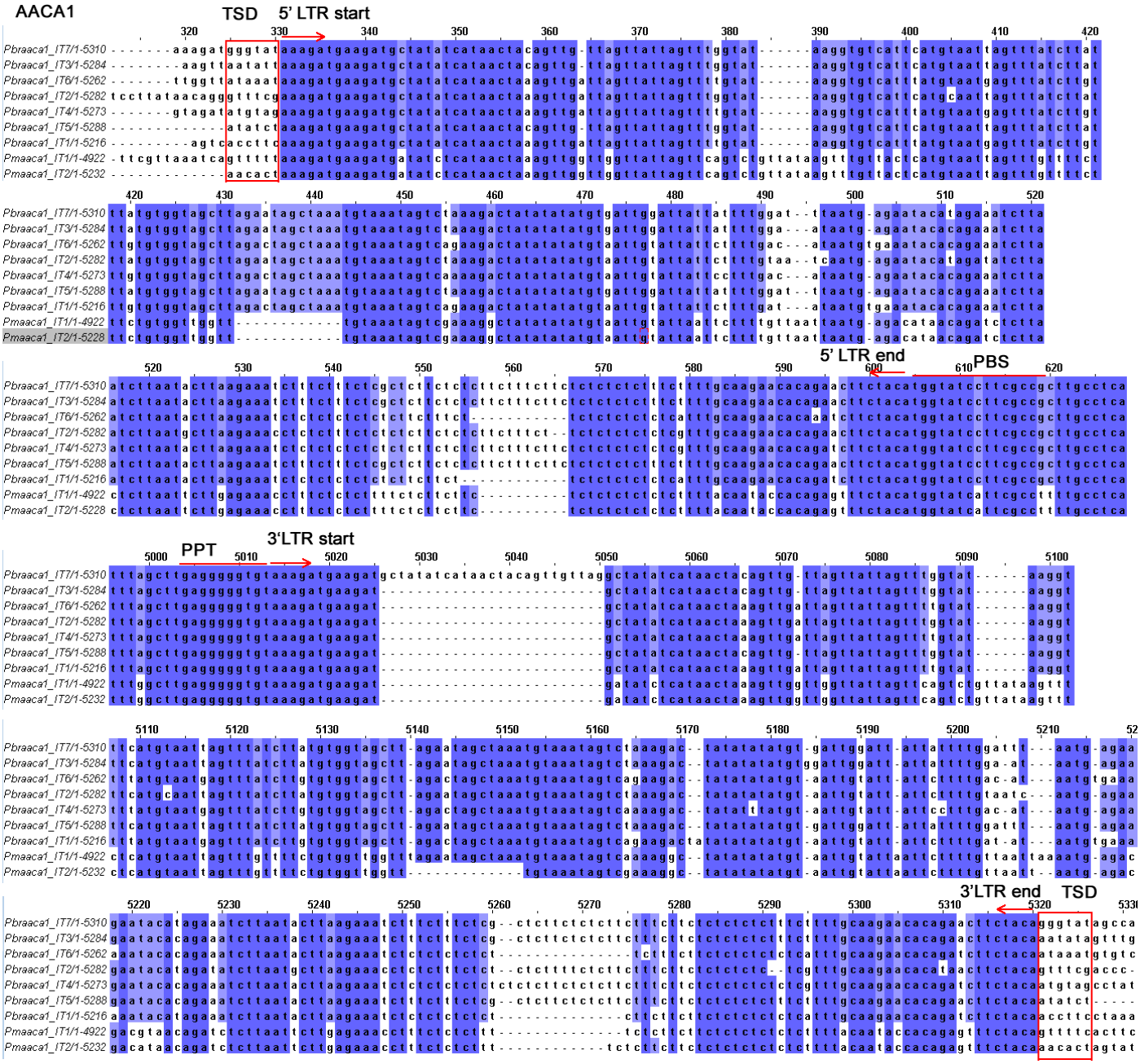
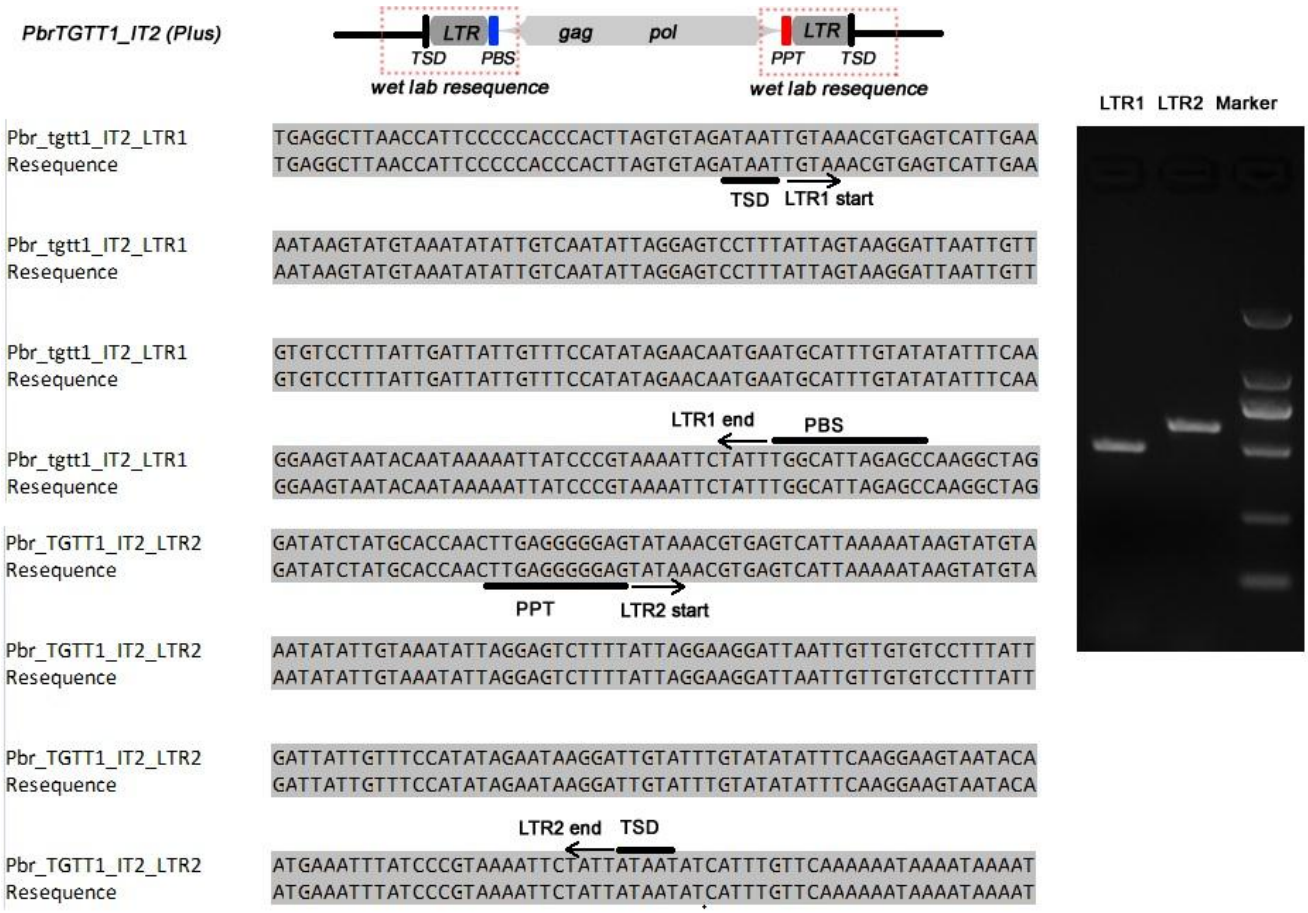


Figure S2.


PbrTGTT1_IT2 (Plus)
LTR
gag
pol
LTR
TSD
PBS
PPT
TSD
wet lab resequence
wet lab resequence
Pbr_tgtt1_IT2_LTR1
Resequence
TGAGGCTTAACCATTCCCCCACCCACTTAGTGTAGATAATTGTAAACGTGAGTCATTGAA
TGAGGCTTAACCATTCCCCCACCCACTTAGTGTAGATAATTGTAAACGTGAGTCATTGAA
TSD
LTR1 start
Pbr_tgtt1_IT2_LTR1
Resequence
AATAAGTATGTAAATATATTGTCAATATTAGGAGTCCTTTATTAGTAAGGATTAATTGTT
AATAAGTATGTAAATATATTGTCAATATTAGGAGTCCTTTATTAGTAAGGATTAATTGTT
Pbr_tgtt1_IT2_LTR1
Resequence
GTGTCCTTTATTGATTATTGTTTCCATATAGAACAATGAATGCATTTGTATATATTTCAA
GTGTCCTTTATTGATTATTGTTTCCATATAGAACAATGAATGCATTTGTATATATTTCAA
LTR1 end
PBS
Pbr_tgtt1_IT2_LTR1
Resequence
GGAAGTAATACAATAAAAATTATCCCGTAAAATTCTATTTGGCATTAGAGCCAAGGCTAG
GGAAGTAATACAATAAAAATTATCCCGTAAAATTCTATTTGGCATTAGAGCCAAGGCTAG
Pbr_TGTT1_IT2_LTR2
Resequence
GATATCTATGCACCAACTTGAGGGGGAGTATAAACGTGAGTCATTAAAAATAAGTATGTA
GATATCTATGCACCAACTTGAGGGGGAGTATAAACGTGAGTCATTAAAAATAAGTATGTA
PPT
LTR2 start
Pbr_TGTT1_IT2_LTR2
Resequence
AATATATTGTAAATATTAGGAGTCTTTTATTAGGAAGGATTAATTGTTGTGTCCTTTATT
AATATATTGTAAATATTAGGAGTCTTTTATTAGGAAGGATTAATTGTTGTGTCCTTTATT
Pbr_TGTT1_IT2_LTR2
Resequence
GATTATTGTTTCCATATAGAATAAGGATTGTATTTGTATATATTTCAAGGAAGTAATACA
GATTATTGTTTCCATATAGAATAAGGATTGTATTTGTATATATTTCAAGGAAGTAATACA
LTR2 end
TSD
Pbr_TGTT1_IT2_LTR2
Resequence
ATGAAATTTATCCCGTAAAATTCTATTATAATATCATTTGTTCAAAAAATAAAATAAAAT
ATGAAATTTATCCCGTAAAATTCTATTATAATATCATTTGTTCAAAAAATAAAATAAAAT
LTR1 LTR2 Marker

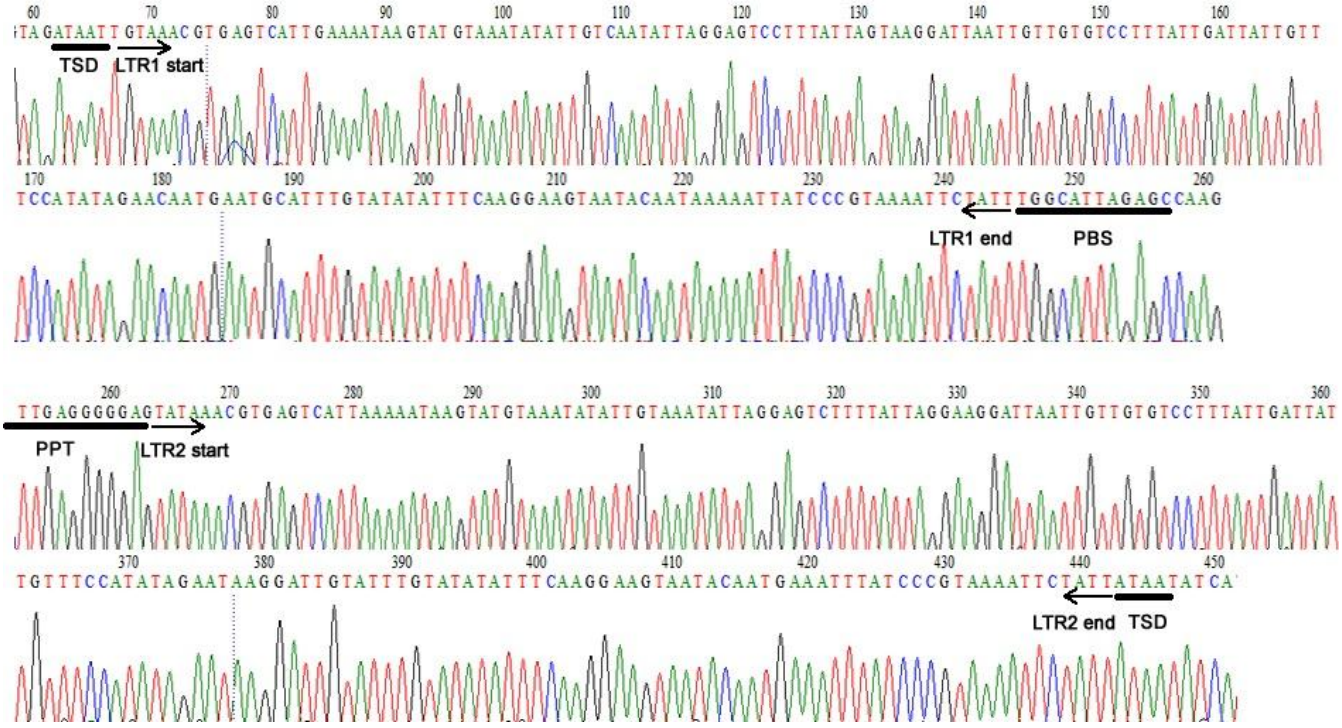

TSD
LTR1 start
LTR1 end
PBS
PPT
LTR2 start
LTR2 end
TSD

Figure S2.

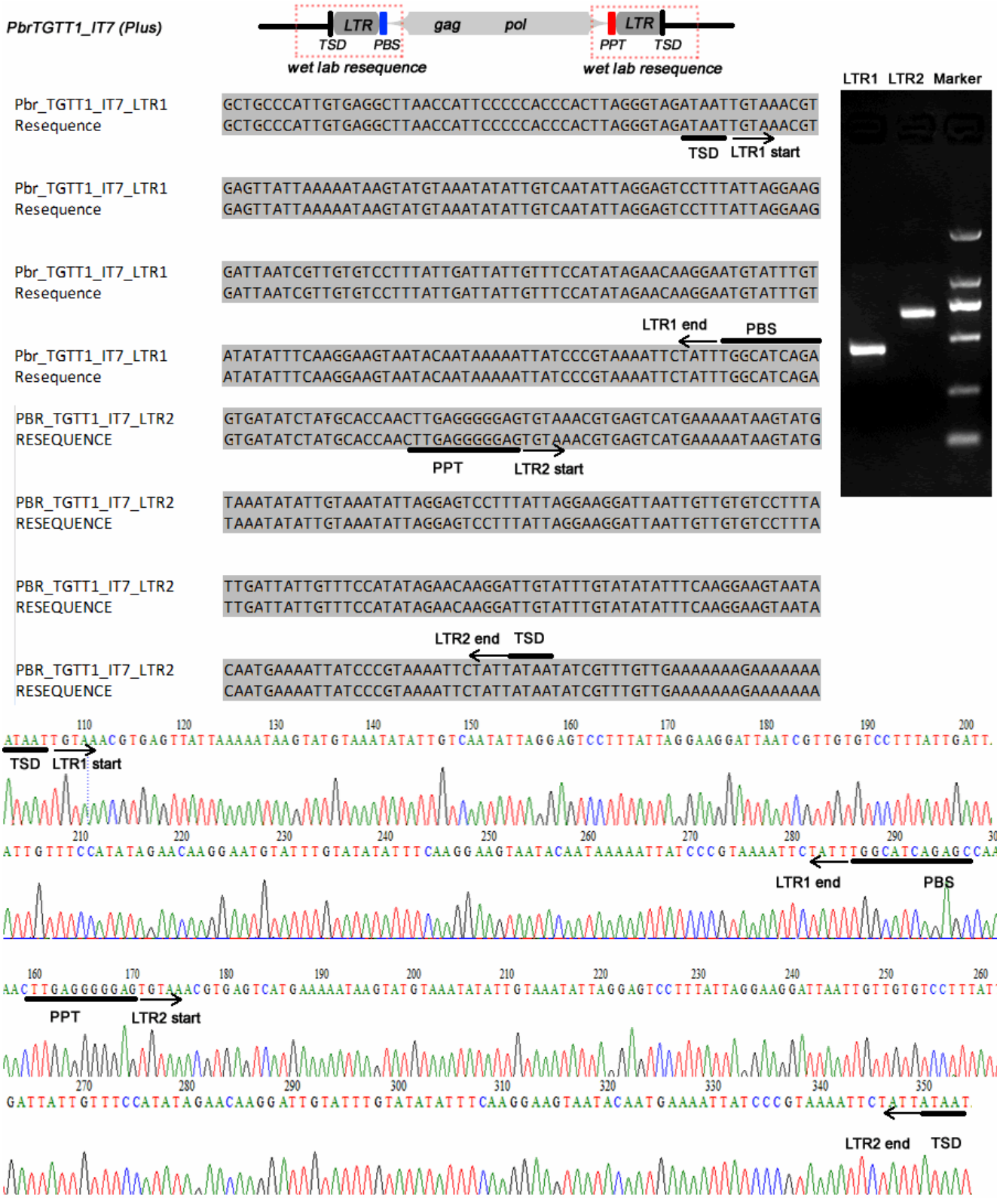

PbrTGTT1_IT7 (Plus)
LTR
gag
pol
LTR
TSD
PBS
PPT
TSD
wet lab resequence
wet lab resequence
LTR1 LTR2 Marker
Pbr_TGTT1_IT7_LTR1
Resequence
GCTGCCCATTGTGAGGCTTAACCATTCCCCCACCCACTTAGGGTAGATAATTGTAAACGT
GCTGCCCATTGTGAGGCTTAACCATTCCCCCACCCACTTAGGGTAGATAATTGTAAACGT
TSD
LTR1 start
Pbr_TGTT1_IT7_LTR1
Resequence
GAGTTATTAAAAATAAGTATGTAAATATATTGTCAATATTAGGAGTCCTTTATTAGGAAG
GAGTTATTAAAAATAAGTATGTAAATATATTGTCAATATTAGGAGTCCTTTATTAGGAAG
Pbr_TGTT1_IT7_LTR1
Resequence
GATTAATCGTTGTGTCCTTTATTGATTATTGTTTCCATATAGAACAAGGAATGTATTTGT
GATTAATCGTTGTGTCCTTTATTGATTATTGTTTCCATATAGAACAAGGAATGTATTTGT
LTR1 end
PBS
Pbr_TGTT1_IT7_LTR1
Resequence
ATATATTTCAAGGAAGTAATACAATAAAAATTATCCCGTAAAATTCTATTTGGCATCAGA
ATATATTTCAAGGAAGTAATACAATAAAAATTATCCCGTAAAATTCTATTTGGCATCAGA
PBR_TGTT1_IT7_LTR2
RESEQUENCE
GTGATATCTATGCACCAACTTGAGGGGGAGTGTAAACGTGAGTCATGAAAAATAAGTATG
GTGATATCTATGCACCAACTTGAGGGGGAGTGTAAACGTGAGTCATGAAAAATAAGTATG
PPT
LTR2 start
PBR_TGTT1_IT7_LTR2
RESEQUENCE
TAAATATATTGTAAATATTAGGAGTCCTTTATTAGGAAGGATTAATTGTTGTGTCCTTTA
TAAATATATTGTAAATATTAGGAGTCCTTTATTAGGAAGGATTAATTGTTGTGTCCTTTA
PBR_TGTT1_IT7_LTR2
RESEQUENCE
TTGATTATTGTTTCCATATAGAACAAGGATTGTATTTGTATATATTTCAAGGAAGTAATA
TTGATTATTGTTTCCATATAGAACAAGGATTGTATTTGTATATATTTCAAGGAAGTAATA
LTR2 end
TSD
PBR_TGTT1_IT7_LTR2
RESEQUENCE
CAATGAAAATTATCCCGTAAAATTCTATTATAATATCGTTTGTTGAAAAAAAGAAAAAAA
CAATGAAAATTATCCCGTAAAATTCTATTATAATATCGTTTGTTGAAAAAAAGAAAAAAA
110 120 130 140 150 160 170 180 190 200
ATAATTGTAAACGTGAGTTATTAAAAATAAGTATGTAAATATATTGTCAATATTAGGAGTCCTTTATTAGGAAGGATTAATCGTTGTGTCCTTTATTGATT
TSD
LTR1 start
210 220 230 240 250 260 270 280 290 300
ATTGTTTCCATATAGAACAAGGAATGTATTTGTATATATTTCAAGGAAGTAATACAATAAAAATTATCCCGTAAAATTCTATTTGGCATCAGAGCCAAG
LTR1 end
PBS
160 170 180 190 200 210 220 230 240 250 260
AACTTGAGGGGGAGTGTAAACGTGAGTCATGAAAAATAAGTATGTAAATATATTGTAAATATTAGGAGTCCTTTATTAGGAAGGATTAATTGTTGTGTCCTTTATT
PPT
LTR2 start
270 280 290 300 310 320 330 340 350
GATTATTGTTTCCATATAGAACAAGGATTGTATTTGTATATATTTCAAGGAAGTAATACAATGAAAATTATCCCGTAAAATTCTATTATAAT
LTR2 end
TSD

Figure S2.

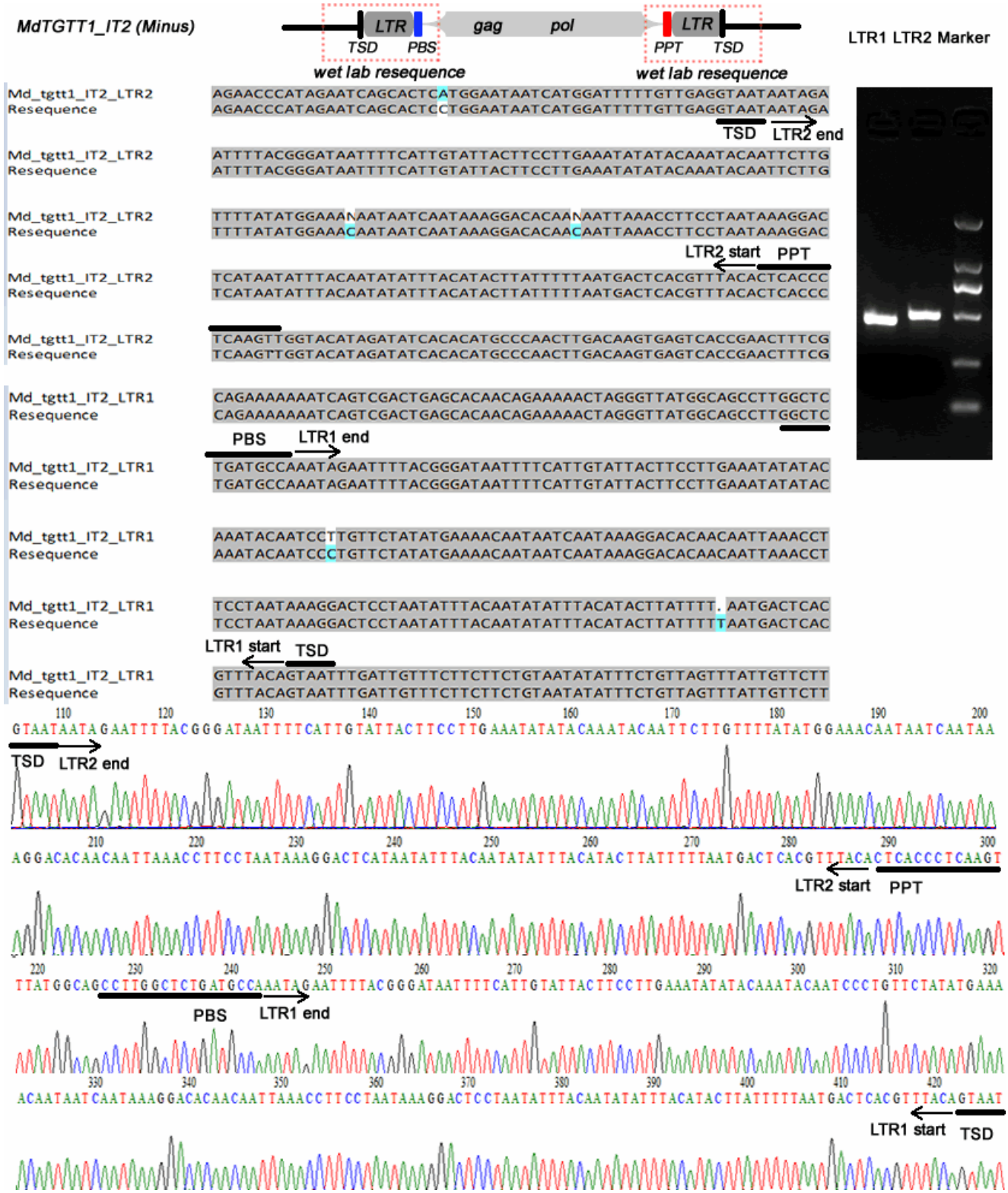

MdTGTT1_IT2 (Minus)
LTR
gag
pol
LTR
TSD
PBS
PPT
TSD
wet lab resequence
wet lab resequence
LTR1 LTR2 Marker
Md_tgtt1_IT2_LTR2 Resequence
TSD
LTR2 end
LTR2 start
PPT
Md_tgtt1_IT2_LTR1 Resequence
PBS
LTR1 end
LTR1 start
TSD
TSD
LTR2 end
LTR2 start
PPT
PBS
LTR1 end
LTR1 start
TSD

Figure S2.

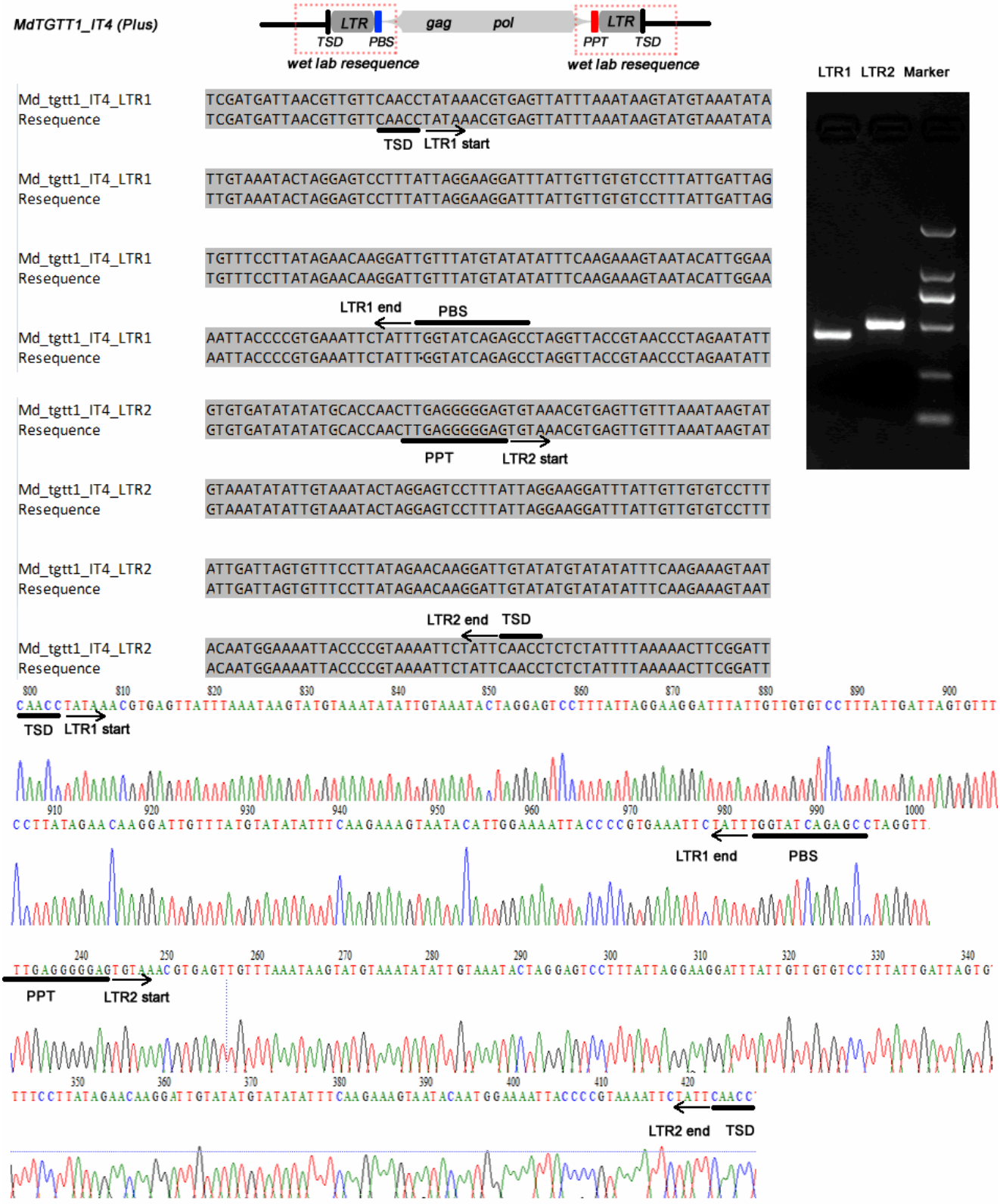
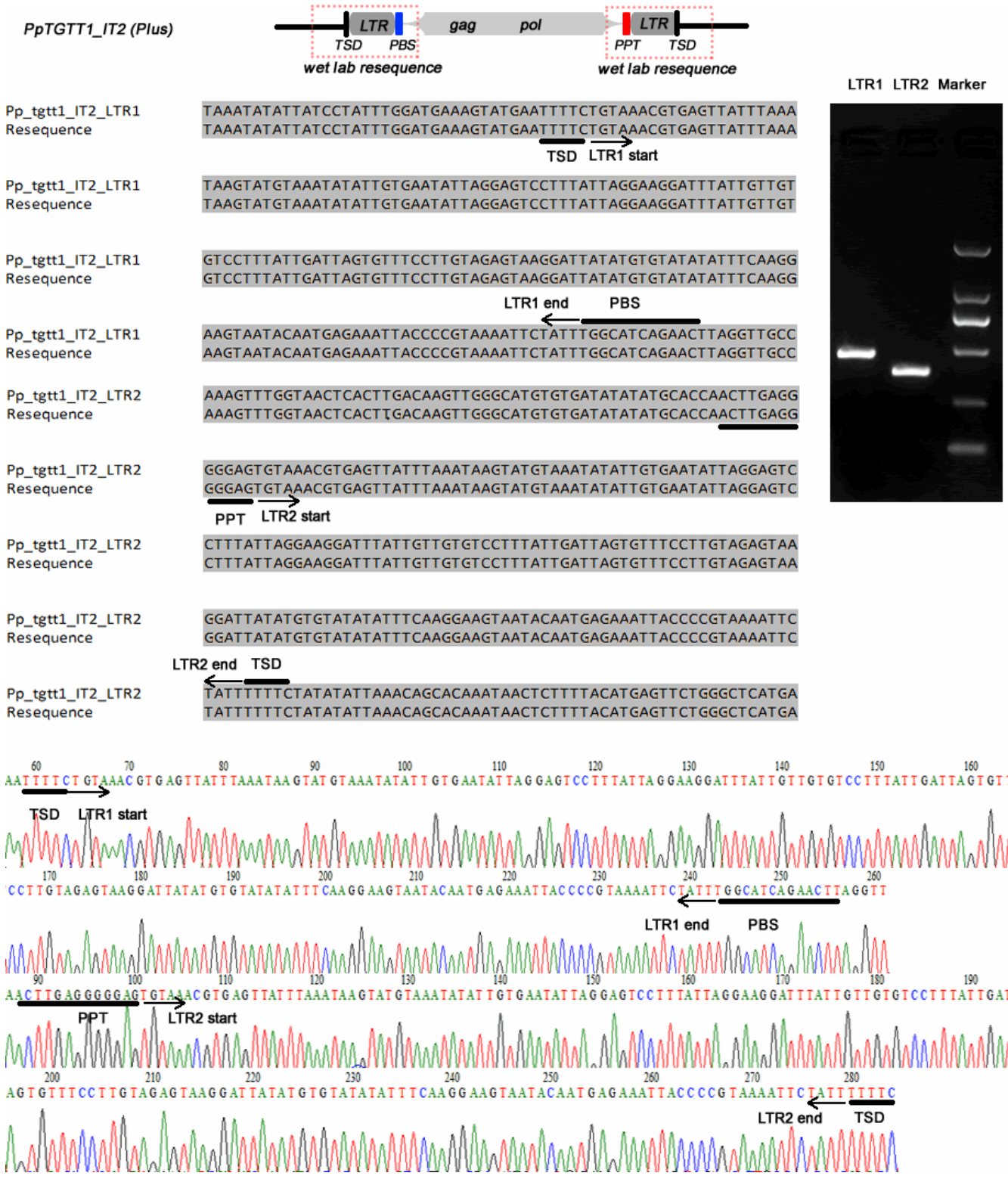

MdTGTT1_IT4 (Plus)
LTR
gag
pol
LTR
TSD
PBS
PPT
TSD
wet lab resequence
wet lab resequence
LTR1 LTR2 Marker
Md_tgtt1_IT4_LTR1
Resequence
TCGATGATTAACGTTGTTCAACCTATAAACGTGAGTTATTTAAATAAGTATGTAAATATA
TCGATGATTAACGTTGTTCAACCTATAAACGTGAGTTATTTAAATAAGTATGTAAATATA
TSD
LTR1 start
Md_tgtt1_IT4_LTR1
Resequence
TTGTAAATACTAGGAGTCCTTTATTAGGAAGGATTTATTGTTGTGTCCTTTATTGATTAG
TTGTAAATACTAGGAGTCCTTTATTAGGAAGGATTTATTGTTGTGTCCTTTATTGATTAG
Md_tgtt1_IT4_LTR1
Resequence
TGTTTCCTTATAGAACAAGGATTGTTTATGTATATATTTCAAGAAAGTAATACATTGGAA
TGTTTCCTTATAGAACAAGGATTGTTTATGTATATATTTCAAGAAAGTAATACATTGGAA
LTR1 end
PBS
Md_tgtt1_IT4_LTR1
Resequence
AATTACCCCGTGAAATTCTATTTGGTATCAGAGCCTAGGTTACCGTAACCCTAGAATATT
AATTACCCCGTGAAATTCTATTTGGTATCAGAGCCTAGGTTACCGTAACCCTAGAATATT
Md_tgtt1_IT4_LTR2
Resequence
GTGTGATATATATGCACCAACTTGAGGGGGAGTGTAAACGTGAGTTGTTTAAATAAGTAT
GTGTGATATATATGCACCAACTTGAGGGGGAGTGTAAACGTGAGTTGTTTAAATAAGTAT
PPT
LTR2 start
Md_tgtt1_IT4_LTR2
Resequence
GTAAATATATTGTAAATACTAGGAGTCCTTTATTAGGAAGGATTTATTGTTGTGTCCTTT
GTAAATATATTGTAAATACTAGGAGTCCTTTATTAGGAAGGATTTATTGTTGTGTCCTTT
Md_tgtt1_IT4_LTR2
Resequence
ATTGATTAGTGTTTCCTTATAGAACAAGGATTGTATATGTATATATTTCAAGAAAGTAAT
ATTGATTAGTGTTTCCTTATAGAACAAGGATTGTATATGTATATATTTCAAGAAAGTAAT
LTR2 end
TSD
Md_tgtt1_IT4_LTR2
Resequence
ACAATGGAAAATTACCCCGTAAAATTCTATTCAACCTCTCTATTTTAAAAACTTCGGATT
ACAATGGAAAATTACCCCGTAAAATTCTATTCAACCTCTCTATTTTAAAAACTTCGGATT
TSD
LTR1 start
LTR1 end
PBS
PPT
LTR2 start
LTR2 end
TSD

Figure S2.


PpTGTT1_IT2 (Plus)
LTR
gag
pol
LTR
TSD
PBS
PPT
TSD
wet lab resequence
wet lab resequence
LTR1 LTR2 Marker
Pp_tgtt1_IT2_LTR1
Resequence
TAAATATATTATCCTATTTGGATGAAAGTATGAATTTTCTGTAAACGTGAGTTATTTAAA
TSD
LTR1 start
Pp_tgtt1_IT2_LTR1
Resequence
TAAGTATGTAAATATATTGTGAATATTAGGAGTCCTTTATTAGGAAGGATTTATTGTTGT
Pp_tgtt1_IT2_LTR1
Resequence
GTCCTTTATTGATTAGTGTTTCCTTGTAGAGTAAGGATTATATGTGTATATATTTCAAGG
LTR1 end
PBS
Pp_tgtt1_IT2_LTR1
Resequence
AAGTAATACAATGAGAAATTACCCCGTAAAATTCTATTTGGCATCAGAACTTAGGTTGCC
Pp_tgtt1_IT2_LTR2
Resequence
AAAGTTTGGTAACTCACTTGACAAGTTGGGCATGTGTGATATATATGCACCAACTTGAGG
Pp_tgtt1_IT2_LTR2
Resequence
GGGAGTGTAAACGTGAGTTATTTAAATAAGTATGTAAATATATTGTGAATATTAGGAGTC
PPT
LTR2 start
Pp_tgtt1_IT2_LTR2
Resequence
CTTTATTAGGAAGGATTTATTGTTGTGTCCTTTATTGATTAGTGTTTCCTTGTAGAGTAA
Pp_tgtt1_IT2_LTR2
Resequence
GGATTATATGTGTATATATTTCAAGGAAGTAATACAATGAGAAATTACCCCGTAAAATTC
LTR2 end
TSD
Pp_tgtt1_IT2_LTR2
Resequence
TATTTTTTCTATATATTAAACAGCACAAATAACTCTTTTACATGAGTTCTGGGCTCATGA

TSD
LTR1 start
LTR1 end
PBS
PPT
LTR2 start
LTR2 end
TSD

Figure S2.

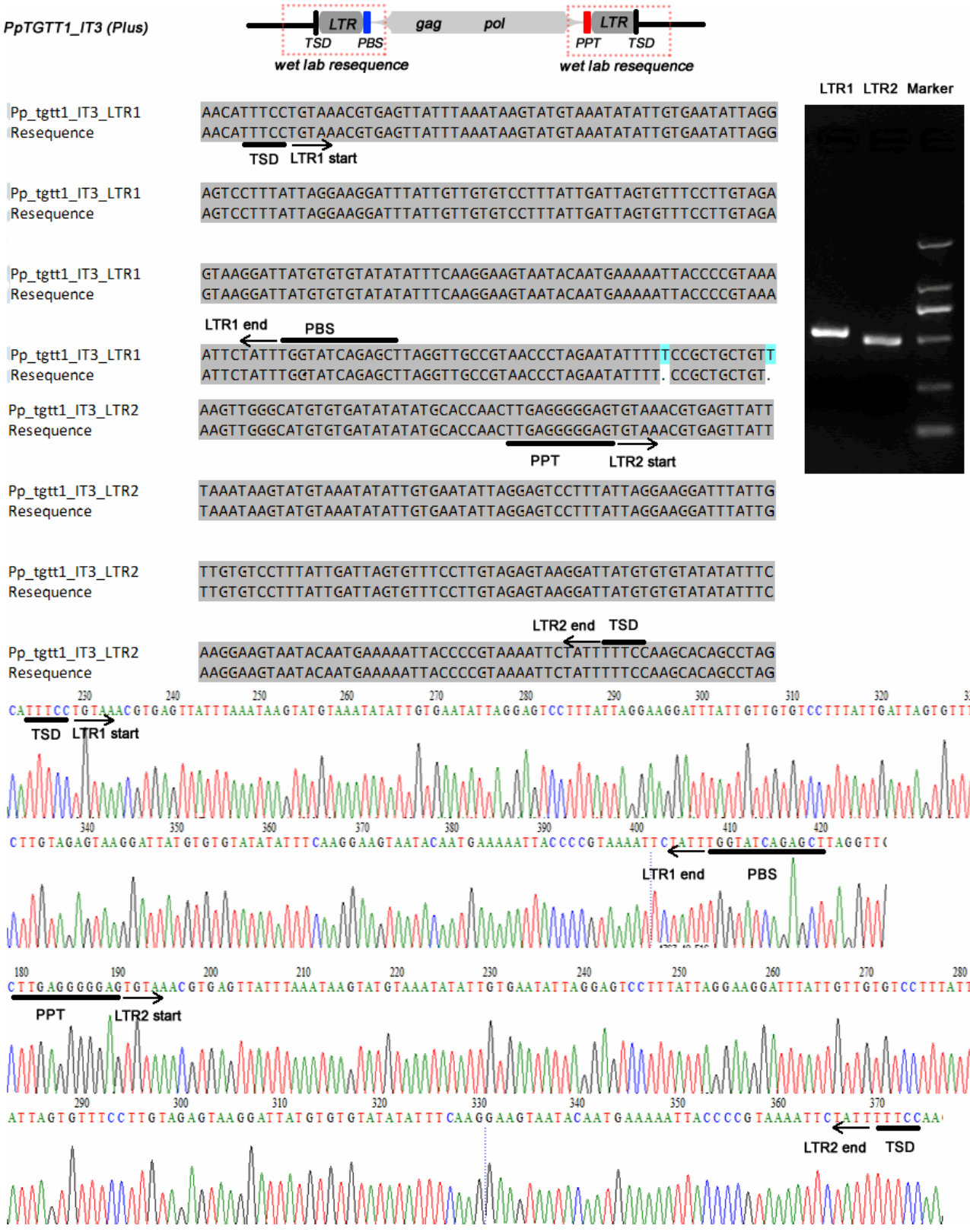

PpTGTT1_IT3 (Plus)
LTR
gag
pol
LTR
TSD
PBS
PPT
TSD
wet lab resequence
wet lab resequence
LTR1 LTR2 Marker
Pp_tgtt1_IT3_LTR1
Resequence
AACATTTCCTGTAAACGTGAGTTATTTAAATAAGTATGTAAATATATTGTGAATATTAGG
TSD
LTR1 start
AGTCCTTTATTAGGAAGGATTTATTGTTGTGTCCTTTATTGATTAGTGTTTCCTTGTAGA
GTAAGGATTATGTGTGTATATATTTCAAGGAAGTAATACAATGAAAAATTACCCCGTAAA
LTR1 end
PBS
ATTCTATTTGGTATCAGAGCTTAGGTTGCCGTAACCCTAGAATATTTTTCCGCTGCTGTT
Pp_tgtt1_IT3_LTR2
AAGTTGGGCATGTGTGATATATATGCACCAACTTGAGGGGGAGTGTAAACGTGAGTTATT
PPT
LTR2 start
TAAATAAGTATGTAAATATATTGTGAATATTAGGAGTCCTTTATTAGGAAGGATTTATTG
TTGTGTCCTTTATTGATTAGTGTTTCCTTGTAGAGTAAGGATTATGTGTGTATATATTTC
LTR2 end
TSD
AAGGAAGTAATACAATGAAAAATTACCCCGTAAAATTCTATTTTTCCAAGCACAGCCTAG
TSD
LTR1 start
LTR1 end
PBS
PPT
LTR2 start
LTR2 end
TSD

Figure S2.

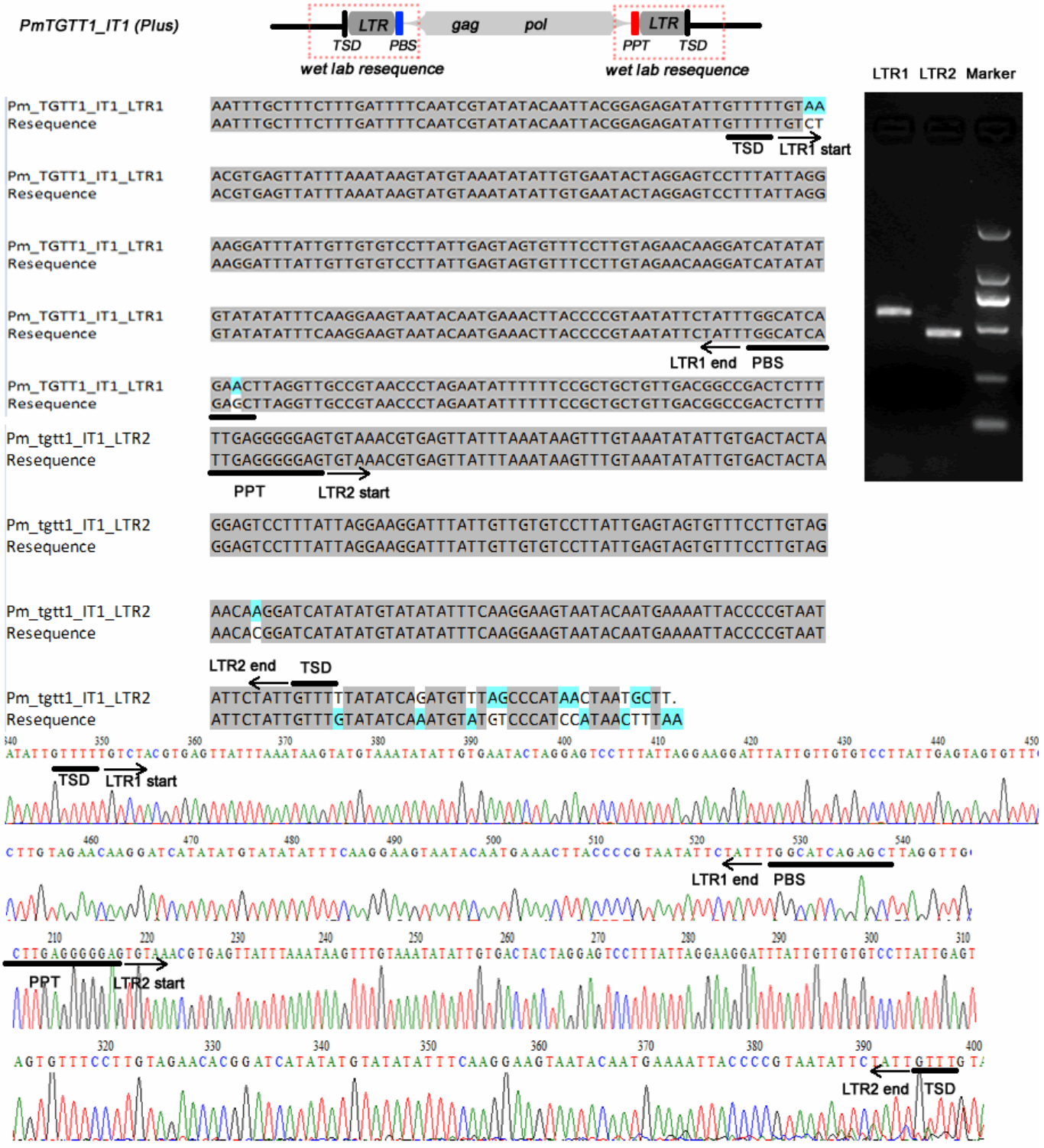

PmTGTT1_IT1 (Plus)
LTR
gag
pol
LTR
TSD
PBS
PPT
TSD
wet lab resequence
wet lab resequence
LTR1 LTR2 Marker
Pm_TGTT1_IT1_LTR1
Resequence
TSD
LTR1 start
LTR1 end
PBS
Pm_tgtt1_IT1_LTR2
Resequence
PPT
LTR2 start
LTR2 end
TSD

Figure S2.

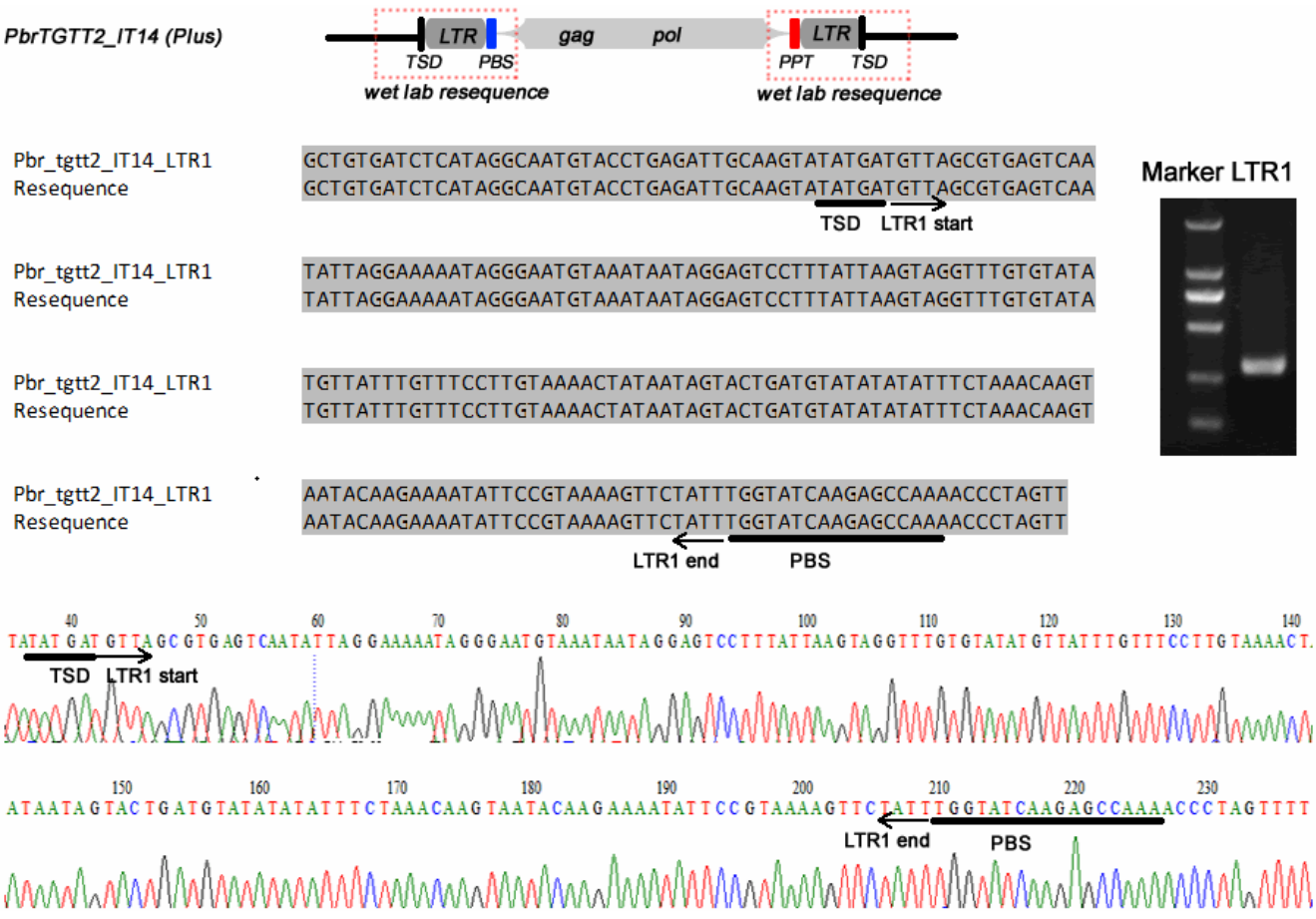

PbrTGTT2_IT14 (Plus)
LTR
gag
pol
LTR
TSD
PBS
PPT
TSD
wet lab resequence
wet lab resequence
Pbr_tgtt2_IT14_LTR1
Resequence
GCTGTGATCTCATAGGCAATGTACCTGAGATTGCAAGTATATGATGTTAGCGTGAGTCAA
GCTGTGATCTCATAGGCAATGTACCTGAGATTGCAAGTATATGATGTTAGCGTGAGTCAA
TSD
LTR1 start
Marker LTR1
Pbr_tgtt2_IT14_LTR1
Resequence
TATTAGGAAAAATAGGGAATGTAAATAATAGGAGTCCTTTATTAAGTAGGTTTGTGTATA
TATTAGGAAAAATAGGGAATGTAAATAATAGGAGTCCTTTATTAAGTAGGTTTGTGTATA
Pbr_tgtt2_IT14_LTR1
Resequence
TGTTATTTGTTTCCTTGTAAAACTATAATAGTACTGATGTATATATATTTCTAAACAAGT
TGTTATTTGTTTCCTTGTAAAACTATAATAGTACTGATGTATATATATTTCTAAACAAGT
Pbr_tgtt2_IT14_LTR1
Resequence
AATACAAGAAAATATTCCGTAAAAGTTCTATTTGGTATCAAGAGCCAAAACCCTAGTT
AATACAAGAAAATATTCCGTAAAAGTTCTATTTGGTATCAAGAGCCAAAACCCTAGTT
LTR1 end
PBS
40
50
60
70
80
90
100
110
120
130
140
TATATGATGTTAGCGTGAGTCAATATTAGGAAAAATAGGGAATGTAAATAATAGGAGTCCTTTATTAAGTAGGTTTGTGTATATGTTATTTGTTTCCTTGTAAAACT
TSD
LTR1 start
150
160
170
180
190
200
210
220
230
ATAATAGTACTGATGTATATATATTTCTAAACAAGTAATACAAGAAAATATTCCGTAAAAGTTCTATTTGGTATCAAGAGCCAAAACCCTAGTTTT
LTR1 end
PBS

Figure S2.

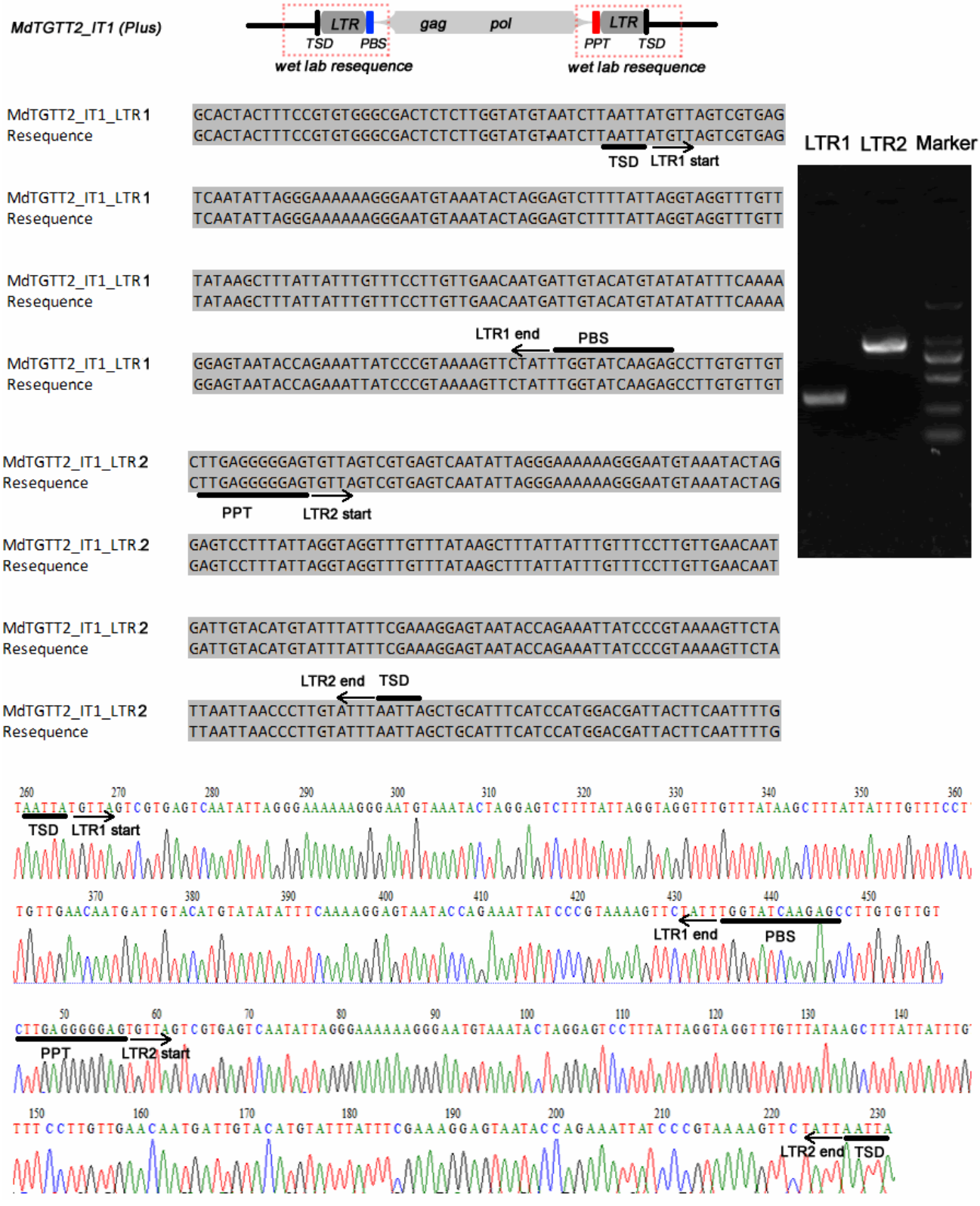
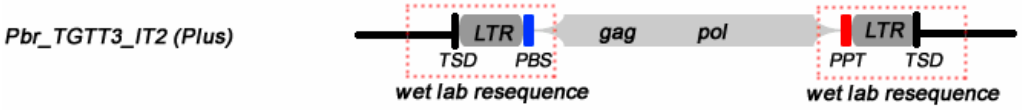
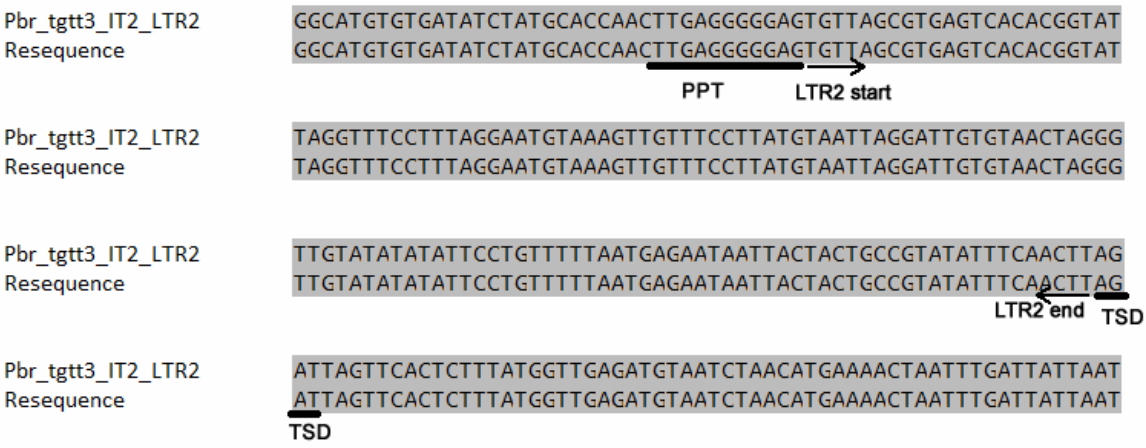

MdTGTT2_IT1 (Plus)
LTR
gag
pol
LTR
TSD
PBS
PPT
TSD
wet lab resequence
wet lab resequence
MdTGTT2_IT1_LTR1
Resequence
GCACTACTTTCCGTGTGGGCGACTCTCTTGGTATGTAATCTTAATTATGTTAGTCGTGAG
GCACTACTTTCCGTGTGGGCGACTCTCTTGGTATGTAATCTTAATTATGTTAGTCGTGAG
TSD
LTR1 start
MdTGTT2_IT1_LTR1
Resequence
TCAATATTAGGGAAAAAAGGGAATGTAAATACTAGGAGTCTTTTATTAGGTAGGTTTGTT
TCAATATTAGGGAAAAAAGGGAATGTAAATACTAGGAGTCTTTTATTAGGTAGGTTTGTT
MdTGTT2_IT1_LTR1
Resequence
TATAAGCTTTATTATTTGTTTCCTTGTTGAACAATGATTGTACATGTATATATTTCAAAA
TATAAGCTTTATTATTTGTTTCCTTGTTGAACAATGATTGTACATGTATATATTTCAAAA
LTR1 end
PBS
MdTGTT2_IT1_LTR1
Resequence
GGAGTAATACCAGAAATTATCCCGTAAAAGTTCTATTTGGTATCAAGAGCCTTGTGTTGT
GGAGTAATACCAGAAATTATCCCGTAAAAGTTCTATTTGGTATCAAGAGCCTTGTGTTGT
LTR1 LTR2 Marker
MdTGTT2_IT1_LTR2
Resequence
CTTGAGGGGGAGTGTTAGTCGTGAGTCAATATTAGGGAAAAAAGGGAATGTAAATACTAG
CTTGAGGGGGAGTGTTAGTCGTGAGTCAATATTAGGGAAAAAAGGGAATGTAAATACTAG
PPT
LTR2 start
MdTGTT2_IT1_LTR2
Resequence
GAGTCCTTTATTAGGTAGGTTTGTTTATAAGCTTTATTATTTGTTTCCTTGTTGAACAAT
GAGTCCTTTATTAGGTAGGTTTGTTTATAAGCTTTATTATTTGTTTCCTTGTTGAACAAT
MdTGTT2_IT1_LTR2
Resequence
GATTGTACATGTATTTATTTCGAAAGGAGTAATACCAGAAATTATCCCGTAAAAGTTCTA
GATTGTACATGTATTTATTTCGAAAGGAGTAATACCAGAAATTATCCCGTAAAAGTTCTA
LTR2 end
TSD
MdTGTT2_IT1_LTR2
Resequence
TTAATTAACCCTTGTATTTAATTAGCTGCATTTCATCCATGGACGATTACTTCAATTTTG
TTAATTAACCCTTGTATTTAATTAGCTGCATTTCATCCATGGACGATTACTTCAATTTTG
260 270 280 290 300 310 320 330 340 350 360
TSD
LTR1 start
370 380 390 400 410 420 430 440 450
LTR1 end
PBS
50 60 70 80 90 100 110 120 130 140
PPT
LTR2 start
150 160 170 180 190 200 210 220 230
LTR2 end
TSD

Figure S2.


Pbr_TGTT3_IT2 (Plus)
LTR
gag
pol
LTR
TSD
PBS
PPT
TSD
wet lab resequence
wet lab resequence


Pbr_tgtt3_IT2_LTR2
Resequence
GGCATGTGTGATATCTATGCACCAACTTGAGGGGGAGTGTTAGCGTGAGTCACACGGTAT
GGCATGTGTGATATCTATGCACCAACTTGAGGGGGAGTGTTAGCGTGAGTCACACGGTAT
PPT
LTR2 start
Pbr_tgtt3_IT2_LTR2
Resequence
TAGGTTTCCTTTAGGAATGTAAAGTTGTTTCCTTATGTAATTAGGATTGTGTAACTAGGG
TAGGTTTCCTTTAGGAATGTAAAGTTGTTTCCTTATGTAATTAGGATTGTGTAACTAGGG
Pbr_tgtt3_IT2_LTR2
Resequence
TTGTATATATATTCCTGTTTTTAATGAGAATAATTACTACTGCCGTATATTTCAACTTAG
TTGTATATATATTCCTGTTTTTAATGAGAATAATTACTACTGCCGTATATTTCAACTTAG
LTR2 end
TSD
Pbr_tgtt3_IT2_LTR2
Resequence
ATTAGTTCACTCTTTATGGTTGAGATGTAATCTAACATGAAAACTAATTTGATTATTAAT
ATTAGTTCACTCTTTATGGTTGAGATGTAATCTAACATGAAAACTAATTTGATTATTAAT
TSD

Marker
LTR2

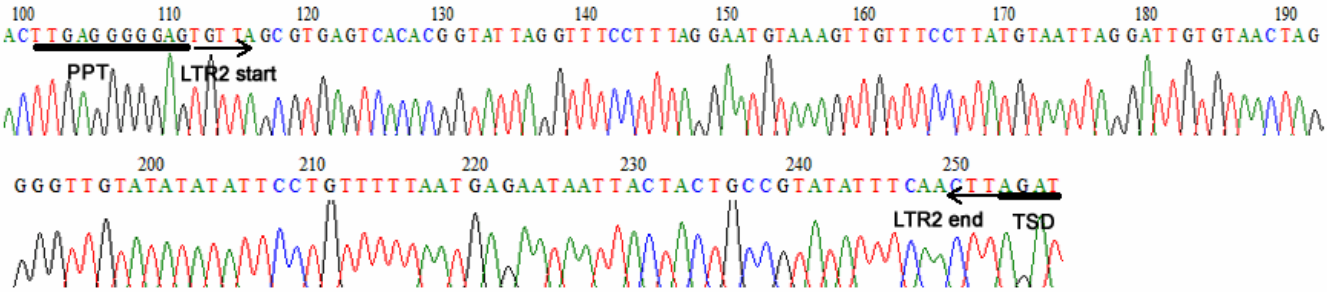

100
110
120
130
140
150
160
170
180
190
ACTTGAGGGGAGTGTTAGCGTGAGTCACACGGTATTAGGTTTCCTTTAGGAATGTAAAGTTGTTTCCTTATGTAATTAGGATTGTGTAACTAG
PPT
LTR2 start
200
210
220
230
240
250
GGGTTGTATATATATTCCTGTTTTTAATGAGAATAATTACTACTGCCGTATATTTCAACTTAGAT
LTR2 end
TSD

Figure S2.

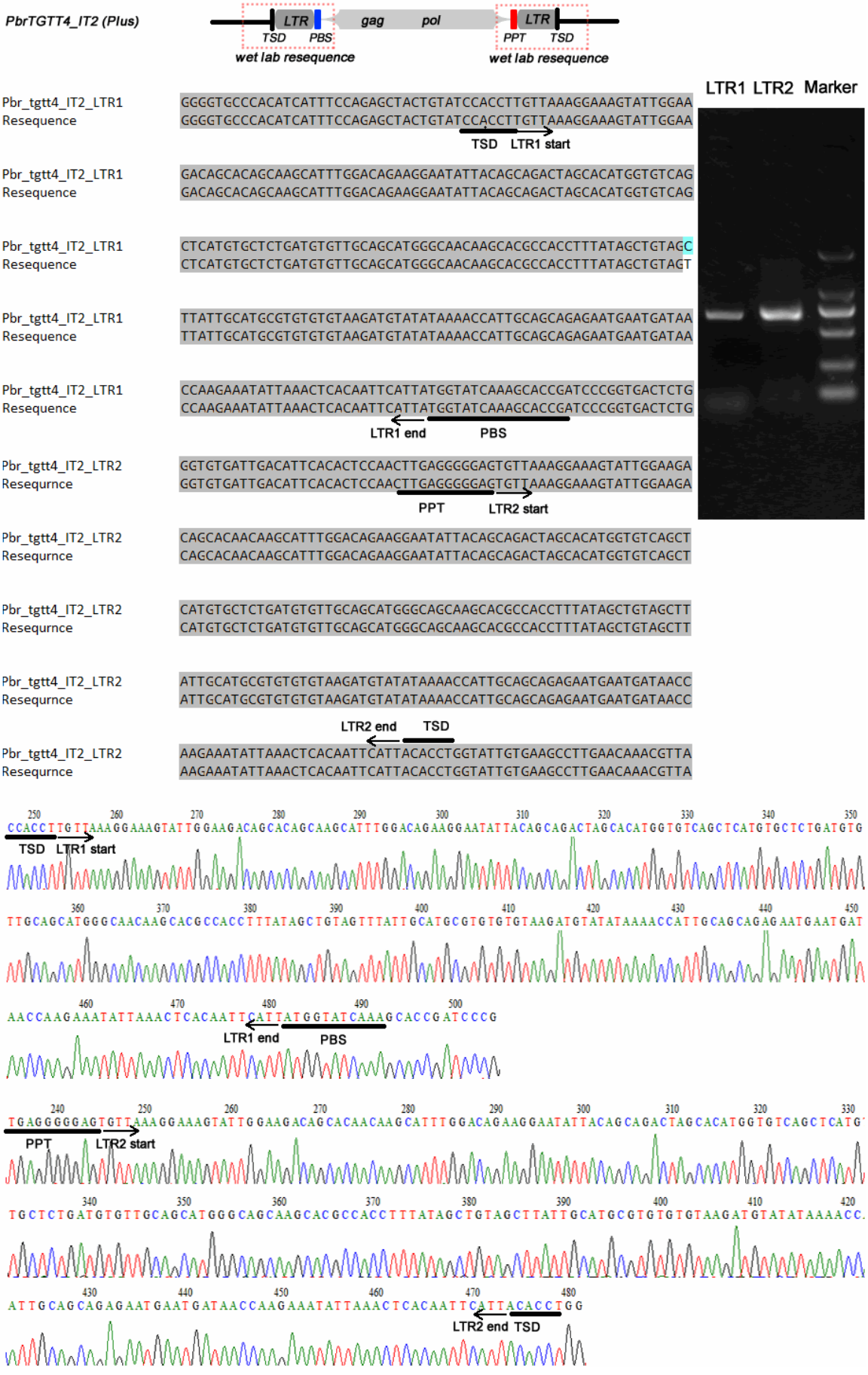

PbrTGTT4_IT2 (Plus)
LTR
gag
pol
LTR
TSD
PBS
PPT
TSD
wet lab resequence
wet lab resequence
LTR1 LTR2 Marker
Pbr_tgtt4_IT2_LTR1
Resequence
GGGGTGCCCACATCATTTCCAGAGCTACTGTATCCACCTTGTTAAAGGAAAGTATTGGAA
GGGGTGCCCACATCATTTCCAGAGCTACTGTATCCACCTTGTTAAAGGAAAGTATTGGAA
TSD
LTR1 start
Pbr_tgtt4_IT2_LTR1
Resequence
GACAGCACAGCAAGCATTTGGACAGAAGGAATATTACAGCAGACTAGCACATGGTGTCAG
GACAGCACAGCAAGCATTTGGACAGAAGGAATATTACAGCAGACTAGCACATGGTGTCAG
Pbr_tgtt4_IT2_LTR1
Resequence
CTCATGTGCTCTGATGTGTTGCAGCATGGGCAACAAGCACGCCACCTTTATAGCTGTAGC
CTCATGTGCTCTGATGTGTTGCAGCATGGGCAACAAGCACGCCACCTTTATAGCTGTAGT
Pbr_tgtt4_IT2_LTR1
Resequence
TTATTGCATGCGTGTGTGTAAGATGTATATAAAACCATTGCAGCAGAGAATGAATGATAA
TTATTGCATGCGTGTGTGTAAGATGTATATAAAACCATTGCAGCAGAGAATGAATGATAA
Pbr_tgtt4_IT2_LTR1
Resequence
CCAAGAAATATTAAACTCACAATTCATTATGGTATCAAAGCACCGATCCCGGTGACTCTG
CCAAGAAATATTAAACTCACAATTCATTATGGTATCAAAGCACCGATCCCGGTGACTCTG
LTR1 end
PBS
Pbr_tgtt4_IT2_LTR2
Resequrnce
GGTGTGATTGACATTCACACTCCAACTTGAGGGGGAGTGTTAAAGGAAAGTATTGGAAGA
GGTGTGATTGACATTCACACTCCAACTTGAGGGGGAGTGTTAAAGGAAAGTATTGGAAGA
PPT
LTR2 start
Pbr_tgtt4_IT2_LTR2
Resequrnce
CAGCACAACAAGCATTTGGACAGAAGGAATATTACAGCAGACTAGCACATGGTGTCAGCT
CAGCACAACAAGCATTTGGACAGAAGGAATATTACAGCAGACTAGCACATGGTGTCAGCT
Pbr_tgtt4_IT2_LTR2
Resequrnce
CATGTGCTCTGATGTGTTGCAGCATGGGCAGCAAGCACGCCACCTTTATAGCTGTAGCTT
CATGTGCTCTGATGTGTTGCAGCATGGGCAGCAAGCACGCCACCTTTATAGCTGTAGCTT
Pbr_tgtt4_IT2_LTR2
Resequrnce
ATTGCATGCGTGTGTGTAAGATGTATATAAAACCATTGCAGCAGAGAATGAATGATAACC
ATTGCATGCGTGTGTGTAAGATGTATATAAAACCATTGCAGCAGAGAATGAATGATAACC
LTR2 end
TSD
Pbr_tgtt4_IT2_LTR2
Resequrnce
AAGAAATATTAAACTCACAATTCATTACACCTGGTATTGTGAAGCCTTGAACAAACGTTA
AAGAAATATTAAACTCACAATTCATTACACCTGGTATTGTGAAGCCTTGAACAAACGTTA
TSD
LTR1 start
LTR1 end
PBS
PPT
LTR2 start
LTR2 end
TSD

Figure S2.

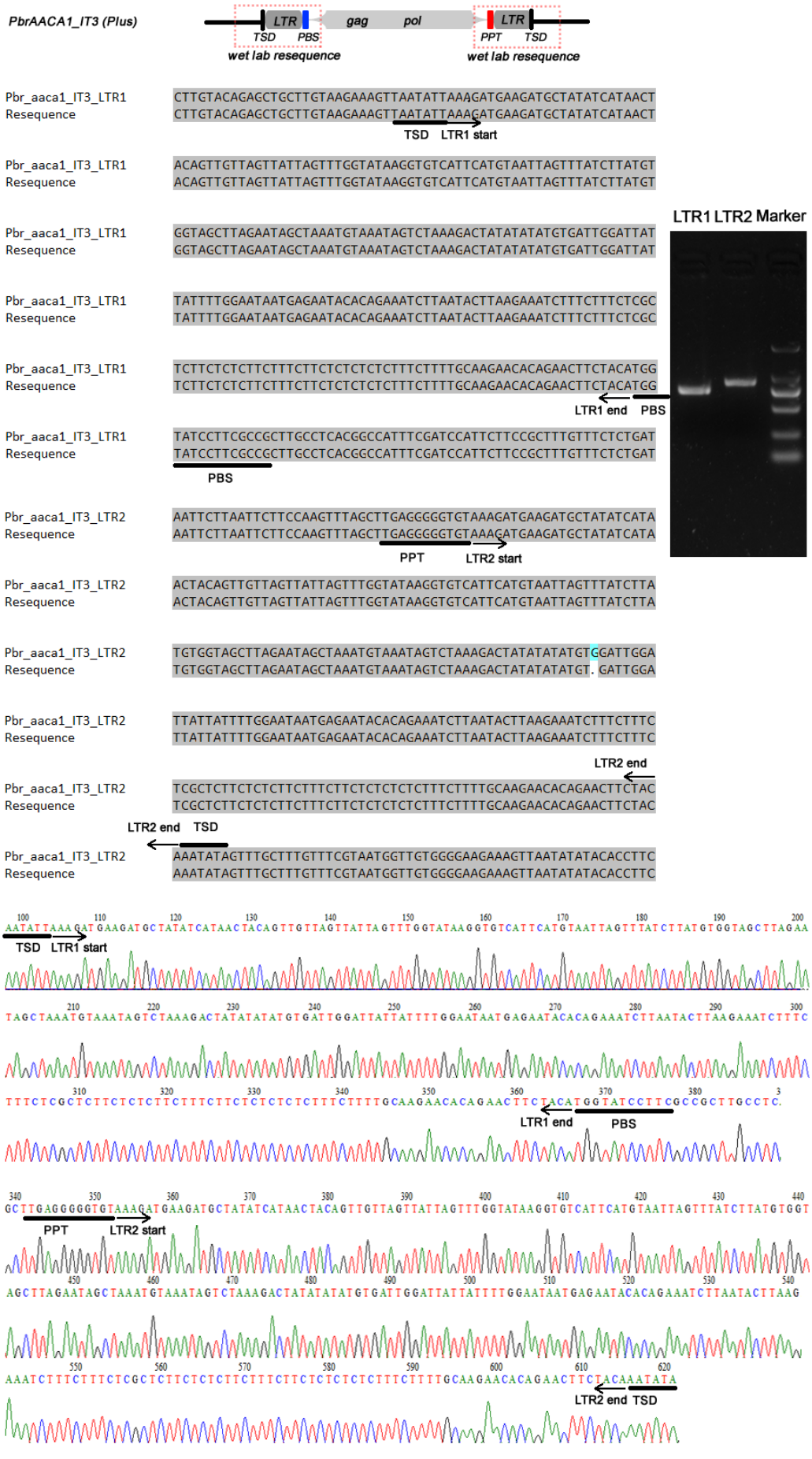

PbrAACA1_IT3 (Plus)
LTR
gag
pol
LTR
TSD
PBS
PPT
TSD
wet lab resequence
wet lab resequence
Pbr_aaca1_IT3_LTR1
Resequence
CTTGTACAGAGCTGCTTGTAAGAAAGTTAATATTAAAGATGAAGATGCTATATCATAACT
CTTGTACAGAGCTGCTTGTAAGAAAGTTAATATTAAAGATGAAGATGCTATATCATAACT
TSD
LTR1 start
ACAGTTGTTAGTTATTAGTTTGGTATAAGGTGTCATTCATGTAATTAGTTTATCTTATGT
GGTAGCTTAGAATAGCTAAATGTAAATAGTCTAAAGACTATATATATGTGATTGGATTAT
TATTTTGGAATAATGAGAATACACAGAAATCTTAATACTTAAGAAATCTTTCTTTCTCGC
TCTTCTCTCTTCTTTCTTCTCTCTCTCTTTCTTTTGCAAGAACACAGAACTTCTACATGG
LTR1 end
PBS
TATCCTTCGCCGCTTGCCTCACGGCCATTTCGATCCATTCTTCCGCTTTGTTTCTCTGAT
PBS
AATTCTTAATTCTTCCAAGTTTAGCTTGAGGGGGTGTAAAGATGAAGATGCTATATCATA
PPT
LTR2 start
ACTACAGTTGTTAGTTATTAGTTTGGTATAAGGTGTCATTCATGTAATTAGTTTATCTTA
TGTGGTAGCTTAGAATAGCTAAATGTAAATAGTCTAAAGACTATATATATGTGATTGGA
TTATTATTTTGGAATAATGAGAATACACAGAAATCTTAATACTTAAGAAATCTTTCTTTC
LTR2 end
TCGCTCTTCTCTCTTCTTTCTTCTCTCTCTCTTTCTTTTGCAAGAACACAGAACTTCTAC
LTR2 end
TSD
AAATATAGTTTGCTTTGTTTCGTAATGGTTGTGGGGAAGAAAGTTAATATATACACCTTC
LTR1 LTR2 Marker
100 110 120 130 140 150 160 170 180 190 200
AATATTAAAGATGAAGATGCTATATCATAACTACAGTTGTTAGTTATTAGTTTGGTATAAGGTGTCATTCATGTAATTAGTTTATCTTATGTGGTAGCTTAGAA
TSD
LTR1 start
210 220 230 240 250 260 270 280 290 300
TAGCTAAATGTAAATAGTCTAAAGACTATATATATGTGATTGGATTATTATTTTGGAATAATGAGAATACACAGAAATCTTAATACTTAAGAAATCTTTC
310 320 330 340 350 360 370 380 3
TTTCTCGCTCTTCTCTCTTCTTTCTTCTCTCTCTCTTTCTTTTGCAAGAACACAGAACTTCTACATGGTATCCTTCGCCGCTTGCCTC
LTR1 end
PBS
340 350 360 370 380 390 400 410 420 430 440
GCTTGAGGGGGTGTAAAGATGAAGATGCTATATCATAACTACAGTTGTTAGTTATTAGTTTGGTATAAGGTGTCATTCATGTAATTAGTTTATCTTATGTGGT
PPT
LTR2 start
450 460 470 480 490 500 510 520 530 540
AGCTTAGAATAGCTAAATGTAAATAGTCTAAAGACTATATATATGTGATTGGATTATTATTTTGGAATAATGAGAATACACAGAAATCTTAATACTTAAG
550 560 570 580 590 600 610 620
AAATCTTTCTTTCTCGCTCTTCTCTCTTCTTTCTTCTCTCTCTTTCTTTTGCAAGAACACAGAACTTCTACAAATATA
LTR2 end
TSD

Figure S2.

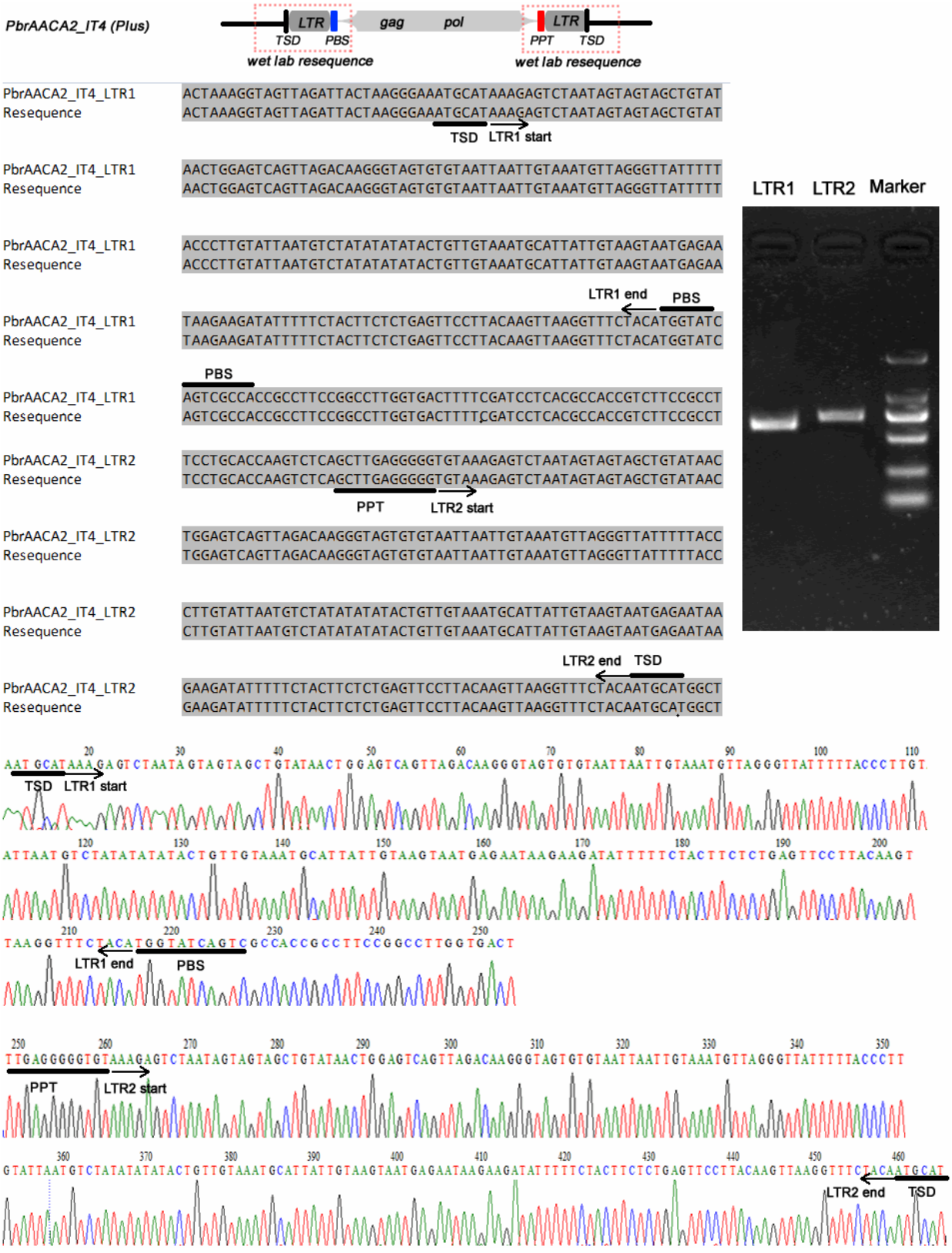
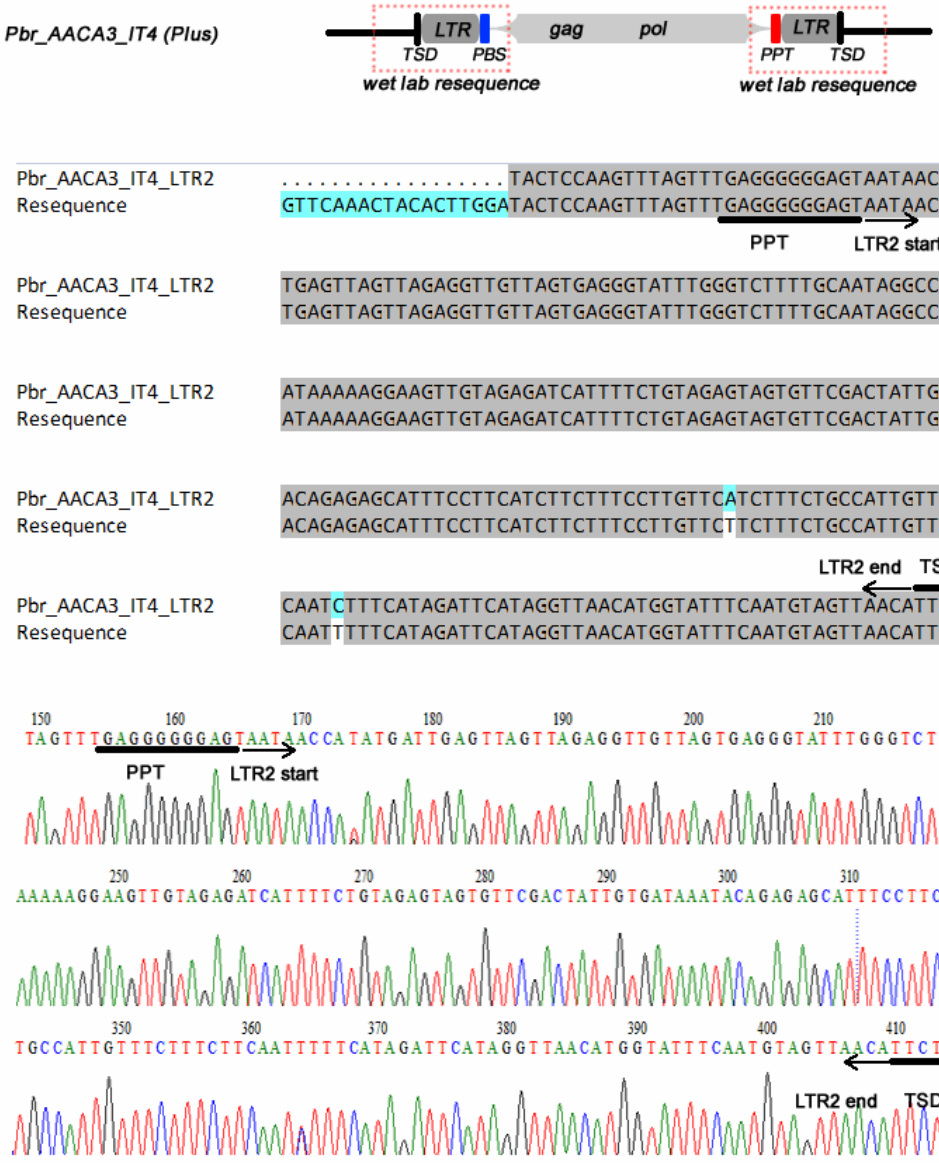

PbrAACA2_IT4 (Plus)
LTR
gag
pol
LTR
TSD
PBS
PPT
TSD
wet lab resequence
wet lab resequence
PbrAACA2_IT4_LTR1
Resequence
ACTAAAGGTAGTTAGATTACTAAGGGAAATGCATAAAGAGTCTAATAGTAGTAGCTGTAT
ACTAAAGGTAGTTAGATTACTAAGGGAAATGCATAAAGAGTCTAATAGTAGTAGCTGTAT
TSD
LTR1 start
PbrAACA2_IT4_LTR1
Resequence
AACTGGAGTCAGTTAGACAAGGGTAGTGTGTAATTAATTGTAAATGTTAGGGTTATTTTT
AACTGGAGTCAGTTAGACAAGGGTAGTGTGTAATTAATTGTAAATGTTAGGGTTATTTTT
PbrAACA2_IT4_LTR1
Resequence
ACCCTTGTATTAATGTCTATATATATACTGTTGTAAATGCATTATTGTAAGTAATGAGAA
ACCCTTGTATTAATGTCTATATATATACTGTTGTAAATGCATTATTGTAAGTAATGAGAA
LTR1 end
PBS
PbrAACA2_IT4_LTR1
Resequence
TAAGAAGATATTTTTCTACTTCTCTGAGTTCCTTACAAGTTAAGGTTTCTACATGGTATC
TAAGAAGATATTTTTCTACTTCTCTGAGTTCCTTACAAGTTAAGGTTTCTACATGGTATC
PBS
PbrAACA2_IT4_LTR1
Resequence
AGTCGCCACCGCCTTCCGGCCTTGGTGACTTTTCGATCCTCACGCCACCGTCTTCCGCCT
AGTCGCCACCGCCTTCCGGCCTTGGTGACTTTTCGATCCTCACGCCACCGTCTTCCGCCT
PbrAACA2_IT4_LTR2
Resequence
TCCTGCACCAAGTCTCAGCTTGAGGGGGTGTAAAGAGTCTAATAGTAGTAGCTGTATAAC
TCCTGCACCAAGTCTCAGCTTGAGGGGGTGTAAAGAGTCTAATAGTAGTAGCTGTATAAC
PPT
LTR2 start
PbrAACA2_IT4_LTR2
Resequence
TGGAGTCAGTTAGACAAGGGTAGTGTGTAATTAATTGTAAATGTTAGGGTTATTTTTACC
TGGAGTCAGTTAGACAAGGGTAGTGTGTAATTAATTGTAAATGTTAGGGTTATTTTTACC
PbrAACA2_IT4_LTR2
Resequence
CTTGTATTAATGTCTATATATATACTGTTGTAAATGCATTATTGTAAGTAATGAGAATAA
CTTGTATTAATGTCTATATATATACTGTTGTAAATGCATTATTGTAAGTAATGAGAATAA
LTR2 end
TSD
PbrAACA2_IT4_LTR2
Resequence
GAAGATATTTTTCTACTTCTCTGAGTTCCTTACAAGTTAAGGTTTCTACAATGCATGGCT
GAAGATATTTTTCTACTTCTCTGAGTTCCTTACAAGTTAAGGTTTCTACAATGCATGGCT
LTR1
LTR2
Marker
TSD
LTR1 start
LTR1 end
PBS
PPT
LTR2 start
LTR2 end
TSD

Figure S2.


Pbr_AACA3_IT4 (Plus)
LTR
gag
pol
LTR
TSD
PBS
PPT
TSD
wet lab resequence
wet lab resequence
Pbr_AACA3_IT4_LTR2
Resequence
TACTCCAAGTTTAGTTTGAGGGGGGAGTAATAACCATATGAT
GTTCAAACTACACTTGGATACTCCAAGTTTAGTTTGAGGGGGGAGTAATAACCATATGAT
PPT
LTR2 start
TGAGTTAGTTAGAGGTTGTTAGTGAGGGTATTTGGGTCTTTTGCAATAGGCCTAATGACT
ATAAAAAGGAAGTTGTAGAGATCATTTTCTGTAGAGTAGTGTTCGACTATTGTGATAAAT
ACAGAGAGCATTTCCTTCATCTTCTTTCCTTGTTCATCTTTCTGCCATTGTTTCTTTCTT
ACAGAGAGCATTTCCTTCATCTTCTTTCCTTGTTCTTCTTTCTGCCATTGTTTCTTTCTT
LTR2 end
TSD
CAATCTTTCATAGATTCATAGGTTAACATGGTATTTCAATGTAGTTAACATTCTTTTCTA
CAATTTTTCATAGATTCATAGGTTAACATGGTATTTCAATGTAGTTAACATTCTTTTCTA
Marker LTR2

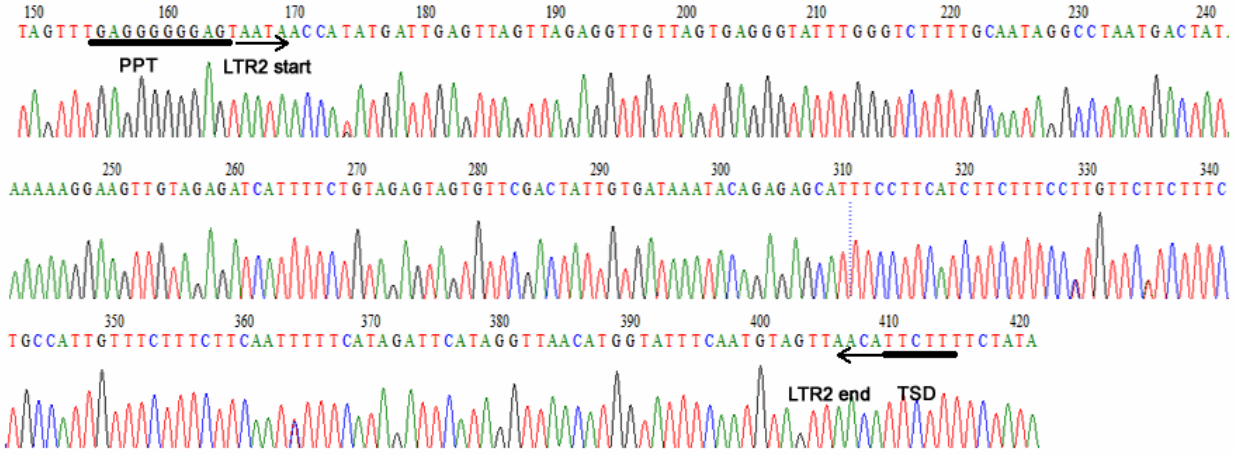

150 160 170 180 190 200 210 220 230 240
TAGTTTGAGGGGGGAGTAATAACCATATGATTGAGTTAGTTAGAGGTTGTTAGTGAGGGTATTTGGGTCTTTTGCAATAGGCCTAATGACTAT.
PPT
LTR2 start
250 260 270 280 290 300 310 320 330 340
AAAAAGGAAGTTGTAGAGATCATTTTCTGTAGAGTAGTGTTCGACTATTGTGATAAATACAGAGAGCATTTCCTTCATCTTCTTTCCTTGTTCTTCTTTC
350 360 370 380 390 400 410 420
TGCCATTGTTTCTTTCTTCAATTTTTCATAGATTCATAGGTTAACATGGTATTTCAATGTAGTTAACATTCTTTTCTATA
LTR2 end
TSD

Figure S2.

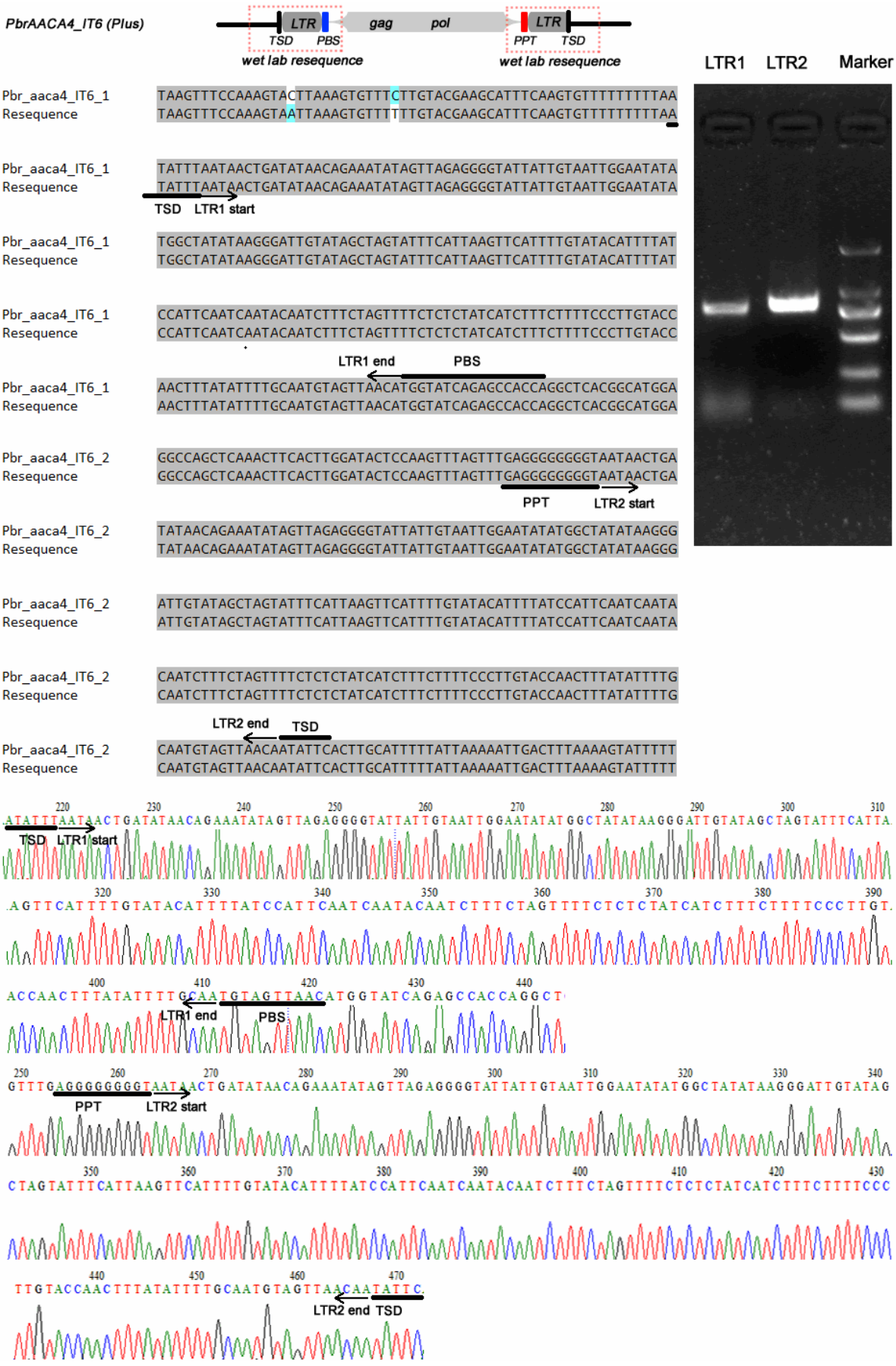
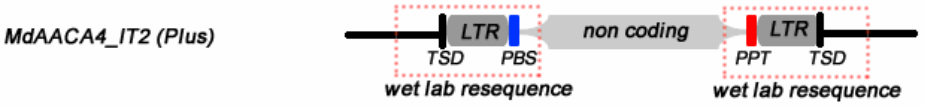
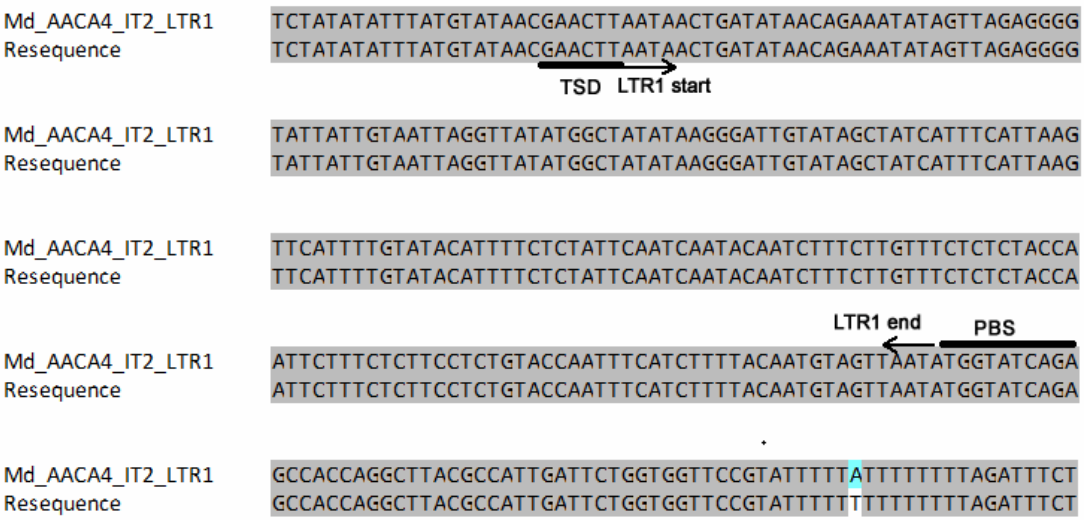
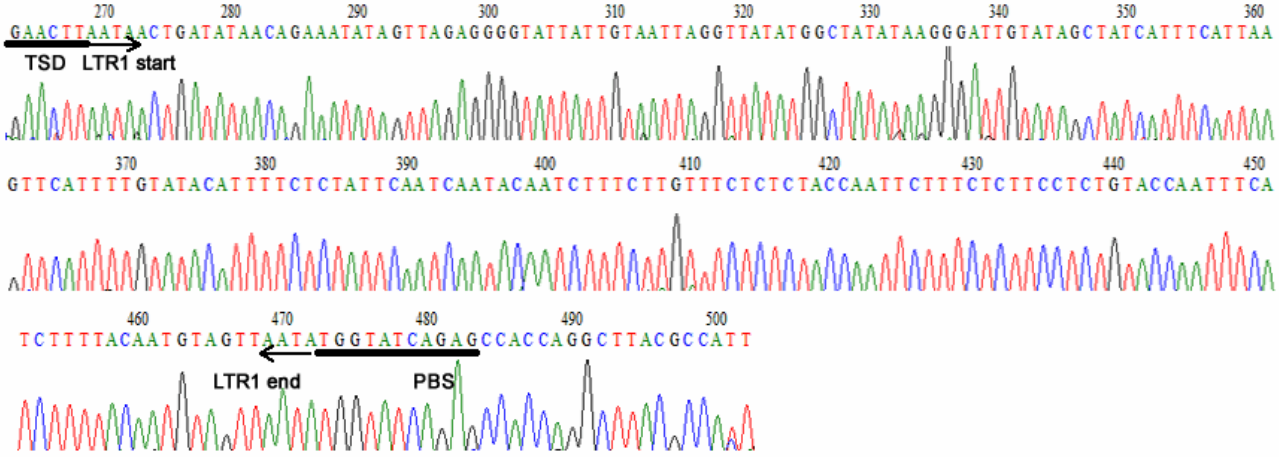

PbrAACA4_IT6 (Plus)
LTR
gag
pol
LTR
TSD
PBS
PPT
TSD
wet lab resequence
wet lab resequence
LTR1
LTR2
Marker
Pbr_aaca4_IT6_1
Resequence
TAAGTTTCCAAAGTACTTAAAGTGTTTCTTGTACGAAGCATTTCAAGTGTTTTTTTTTAA
TAAGTTTCCAAAGTAATTAAAGTGTTTTTTGTACGAAGCATTTCAAGTGTTTTTTTTTAA
Pbr_aaca4_IT6_1
Resequence
TATTTAATAACTGATATAACAGAAATATAGTTAGAGGGGTATTATTGTAATTGGAATATA
TATTTAATAACTGATATAACAGAAATATAGTTAGAGGGGTATTATTGTAATTGGAATATA
TSD
LTR1 start
Pbr_aaca4_IT6_1
Resequence
TGGCTATATAAGGGATTGTATAGCTAGTATTTCATTAAGTTCATTTTGTATACATTTTAT
TGGCTATATAAGGGATTGTATAGCTAGTATTTCATTAAGTTCATTTTGTATACATTTTAT
Pbr_aaca4_IT6_1
Resequence
CCATTCAATCAATACAATCTTTCTAGTTTTCTCTCTATCATCTTTCTTTTCCCTTGTACC
CCATTCAATCAATACAATCTTTCTAGTTTTCTCTCTATCATCTTTCTTTTCCCTTGTACC
LTR1 end
PBS
Pbr_aaca4_IT6_1
Resequence
AACTTTATATTTTGCAATGTAGTTAACATGGTATCAGAGCCACCAGGCTCACGGCATGGA
AACTTTATATTTTGCAATGTAGTTAACATGGTATCAGAGCCACCAGGCTCACGGCATGGA
Pbr_aaca4_IT6_2
Resequence
GGCCAGCTCAAACTTCACTTGGATACTCCAAGTTTAGTTTGAGGGGGGGGTAATAACTGA
GGCCAGCTCAAACTTCACTTGGATACTCCAAGTTTAGTTTGAGGGGGGGGTAATAACTGA
PPT
LTR2 start
Pbr_aaca4_IT6_2
Resequence
TATAACAGAAATATAGTTAGAGGGGTATTATTGTAATTGGAATATATGGCTATATAAGGG
TATAACAGAAATATAGTTAGAGGGGTATTATTGTAATTGGAATATATGGCTATATAAGGG
Pbr_aaca4_IT6_2
Resequence
ATTGTATAGCTAGTATTTCATTAAGTTCATTTTGTATACATTTTATCCATTCAATCAATA
ATTGTATAGCTAGTATTTCATTAAGTTCATTTTGTATACATTTTATCCATTCAATCAATA
Pbr_aaca4_IT6_2
Resequence
CAATCTTTCTAGTTTTCTCTCTATCATCTTTCTTTTCCCTTGTACCAACTTTATATTTTG
CAATCTTTCTAGTTTTCTCTCTATCATCTTTCTTTTCCCTTGTACCAACTTTATATTTTG
LTR2 end
TSD
Pbr_aaca4_IT6_2
Resequence
CAATGTAGTTAACAATATTCACTTGCATTTTTATTAAAAATTGACTTTAAAAGTATTTTT
CAATGTAGTTAACAATATTCACTTGCATTTTTATTAAAAATTGACTTTAAAAGTATTTTT
220 230 240 250 260 270 280 290 300 310
ATATTTAATAACTGATATAACAGAAATATAGTTAGAGGGGTATTATTGTAATTGGAATATATGGCTATATAAGGGATTGTATAGCTAGTATTTCATTA
TSD
LTR1 start
320 330 340 350 360 370 380 390
AGTTCATTTTGTATACATTTTATCCATTCAATCAATACAATCTTTCTAGTTTTCTCTCTATCATCTTTCTTTTCCCTTGT
400 410 420 430 440
ACCAACTTTATATTTTGCAATGTAGTTAACATGGTATCAGAGCCACCAGGCT
LTR1 end
PBS
250 260 270 280 290 300 310 320 330 340
GTTTGAGGGGGGGGTAATAACTGATATAACAGAAATATAGTTAGAGGGGTATTATTGTAATTGGAATATATGGCTATATAAGGGATTGTATAG
PPT
LTR2 start
350 360 370 380 390 400 410 420 430
CTAGTATTTCATTAAGTTCATTTTGTATACATTTTATCCATTCAATCAATACAATCTTTCTAGTTTTCTCTCTATCATCTTTCTTTTCCC
440 450 460 470
TTGTACCAACTTTATATTTTGCAATGTAGTTAACAATATTC
LTR2 end
TSD

Figure S2.


MdAACA4_IT2 (Plus)
LTR
TSD
PBS
non coding
PPT
LTR
TSD
wet lab resequence
wet lab resequence


Md_AACA4_IT2_LTR1
Resequence
TCTATATATTTATGTATAACGAACTTAATAACTGATATAACAGAAATATAGTTAGAGGGG
TCTATATATTTATGTATAACGAACTTAATAACTGATATAACAGAAATATAGTTAGAGGGG
TSD
LTR1 start
Md_AACA4_IT2_LTR1
Resequence
TATTATTGTAATTAGGTTATATGGCTATATAAGGGATTGTATAGCTATCATTTCATTAAG
TATTATTGTAATTAGGTTATATGGCTATATAAGGGATTGTATAGCTATCATTTCATTAAG
Md_AACA4_IT2_LTR1
Resequence
TTCATTTTGTATACATTTTCTCTATTCAATCAATACAATCTTTCTTGTTTCTCTCTACCA
TTCATTTTGTATACATTTTCTCTATTCAATCAATACAATCTTTCTTGTTTCTCTCTACCA
LTR1 end
PBS
Md_AACA4_IT2_LTR1
Resequence
ATTCTTTCTCTTCCTCTGTACCAATTTCATCTTTTACAATGTAGTTAATATGGTATCAGA
ATTCTTTCTCTTCCTCTGTACCAATTTCATCTTTTACAATGTAGTTAATATGGTATCAGA
Md_AACA4_IT2_LTR1
Resequence
GCCACCAGGCTTACGCCATTGATTCTGGTGGTTCCGTATTTTTATTTTTTTTAGATTTCT
GCCACCAGGCTTACGCCATTGATTCTGGTGGTTCCGTATTTTTTTTTTTTTTAGATTTCT


270 280 290 300 310 320 330 340 350 360
GAACTTAATAACTGATATAACAGAAATATAGTTAGAGGGGTATTATTGTAATTAGGTTATATGGCTATATAAGGGATTGTATAGCTATCATTTCATTAA
TSD
LTR1 start
370 380 390 400 410 420 430 440 450
GTTCATTTTGTATACATTTTCTCTATTCAATCAATACAATCTTTCTTGTTTCTCTCTACCAATTTCTTTCTCTTCCTCTGTACCAATTTCA
460 470 480 490 500
TCTTTTACAATGTAGTTAATATGGTATCAGAGCCACCAGGCTTACGCCATT
LTR1 end
PBS

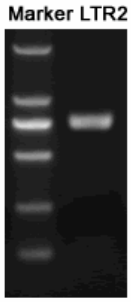

Marker LTR2

Figure S3.

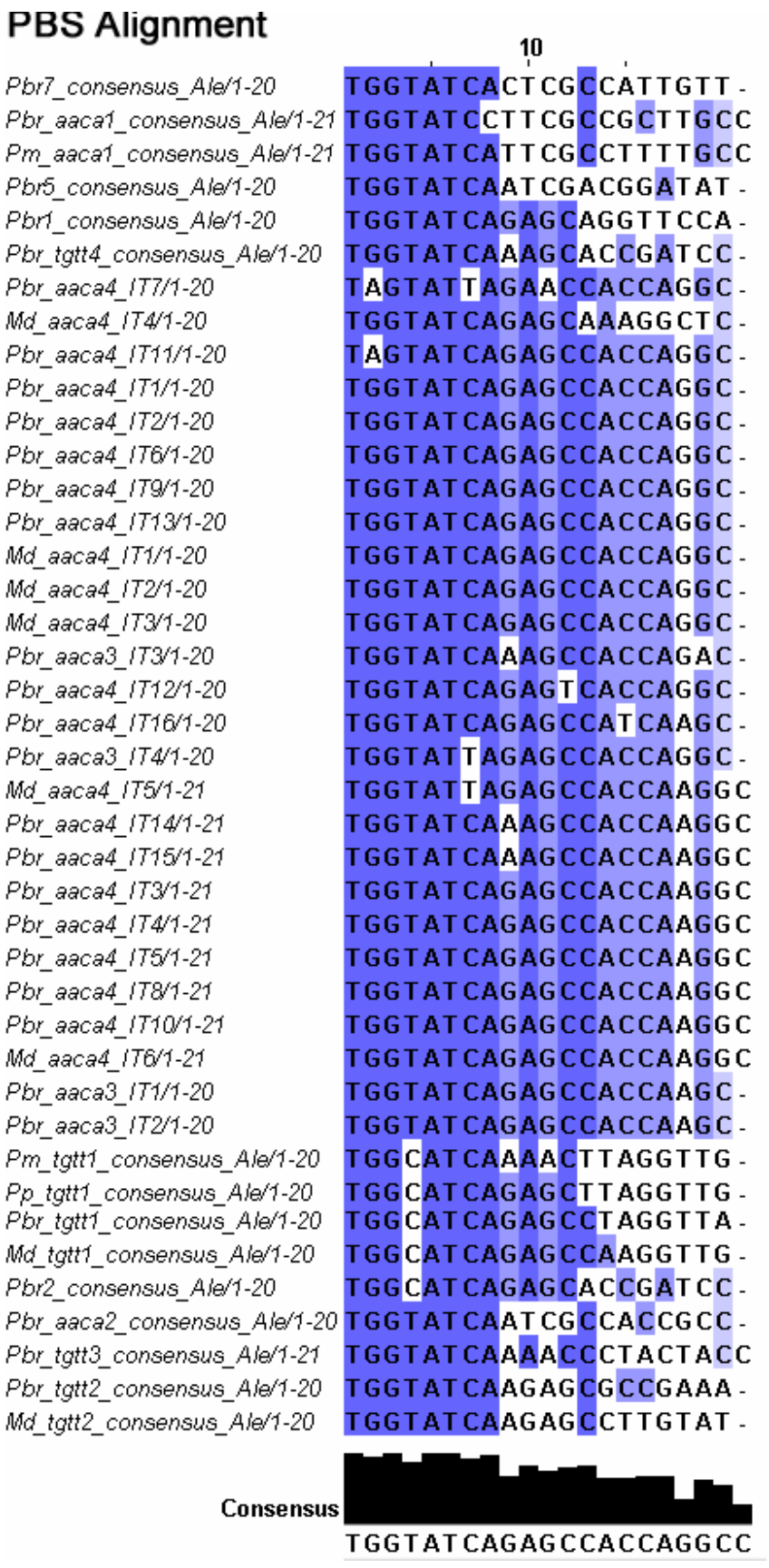
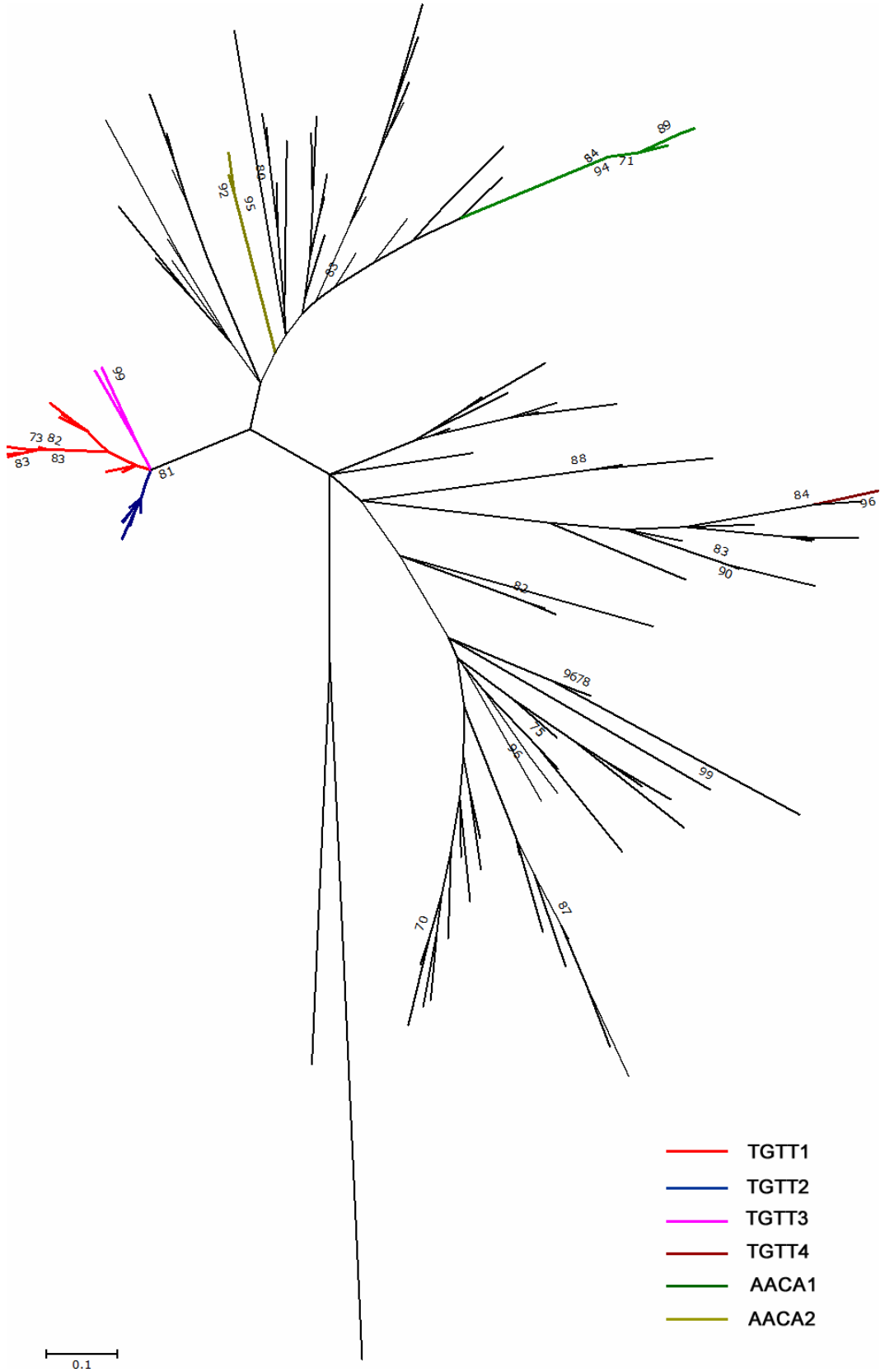

PBS Alignment
10
Pbr7_consensus_Ale/1-20 TGGTATCACTCGCCATTGTT-
Pbr_aaca1_consensus_Ale/1-21 TGGTATCCTTCGCCGCTTGCC
Pm_aaca1_consensus_Ale/1-21 TGGTATCATTCGCCTTTTGCC
Pbr5_consensus_Ale/1-20 TGGTATCAATCGACGGATAT-
Pbr1_consensus_Ale/1-20 TGGTATCAGAGCAGGTTCCA-
Pbr_tgtt4_consensus_Ale/1-20 TGGTATCAAAGCACCGATCC-
Pbr_aaca4_IT7/1-20 TAGTATTAGAACCACCAGGC-
Md_aaca4_IT4/1-20 TGGTATCAGAGCAAAGGCTC-
Pbr_aaca4_IT11/1-20 TAGTATCAGAGCCACCAGGC-
Pbr_aaca4_IT1/1-20 TGGTATCAGAGCCACCAGGC-
Pbr_aaca4_IT2/1-20 TGGTATCAGAGCCACCAGGC-
Pbr_aaca4_IT6/1-20 TGGTATCAGAGCCACCAGGC-
Pbr_aaca4_IT9/1-20 TGGTATCAGAGCCACCAGGC-
Pbr_aaca4_IT13/1-20 TGGTATCAGAGCCACCAGGC-
Md_aaca4_IT1/1-20 TGGTATCAGAGCCACCAGGC-
Md_aaca4_IT2/1-20 TGGTATCAGAGCCACCAGGC-
Md_aaca4_IT3/1-20 TGGTATCAGAGCCACCAGGC-
Pbr_aaca3_IT3/1-20 TGGTATCAAAGCCACCAGAC-
Pbr_aaca4_IT12/1-20 TGGTATCAGAGTCACCAGGC-
Pbr_aaca4_IT16/1-20 TGGTATCAGAGCCATCAAGC-
Pbr_aaca3_IT4/1-20 TGGTATTAGAGCCACCAGGC-
Md_aaca4_IT5/1-21 TGGTATTAGAGCCACCAAGGC
Pbr_aaca4_IT14/1-21 TGGTATCAAAGCCACCAAGGC
Pbr_aaca4_IT15/1-21 TGGTATCAAAGCCACCAAGGC
Pbr_aaca4_IT3/1-21 TGGTATCAGAGCCACCAAGGC
Pbr_aaca4_IT4/1-21 TGGTATCAGAGCCACCAAGGC
Pbr_aaca4_IT5/1-21 TGGTATCAGAGCCACCAAGGC
Pbr_aaca4_IT8/1-21 TGGTATCAGAGCCACCAAGGC
Pbr_aaca4_IT10/1-21 TGGTATCAGAGCCACCAAGGC
Md_aaca4_IT6/1-21 TGGTATCAGAGCCACCAAGGC
Pbr_aaca3_IT1/1-20 TGGTATCAGAGCCACCAAGC-
Pbr_aaca3_IT2/1-20 TGGTATCAGAGCCACCAAGC-
Pm_tgtt1_consensus_Ale/1-20 TGGCATCAAAACTTAGGTTG-
Pp_tgtt1_consensus_Ale/1-20 TGGCATCAGAGCTTAGGTTG-
Pbr_tgtt1_consensus_Ale/1-20 TGGCATCAGAGCCTAGGTTA-
Md_tgtt1_consensus_Ale/1-20 TGGCATCAGAGCCAAGGTTG-
Pbr2_consensus_Ale/1-20 TGGCATCAGAGCACCGATCC-
Pbr_aaca2_consensus_Ale/1-20 TGGTATCAATCGCCACCGCC-
Pbr_tgtt3_consensus_Ale/1-21 TGGTATCAAAACCCTACTACC
Pbr_tgtt2_consensus_Ale/1-20 TGGTATCAAGAGCGCCGAAA-
Md_tgtt2_consensus_Ale/1-20 TGGTATCAAGAGCCTTGTAT-
Consensus
TGGTATCAGAGCCACCAGGCC

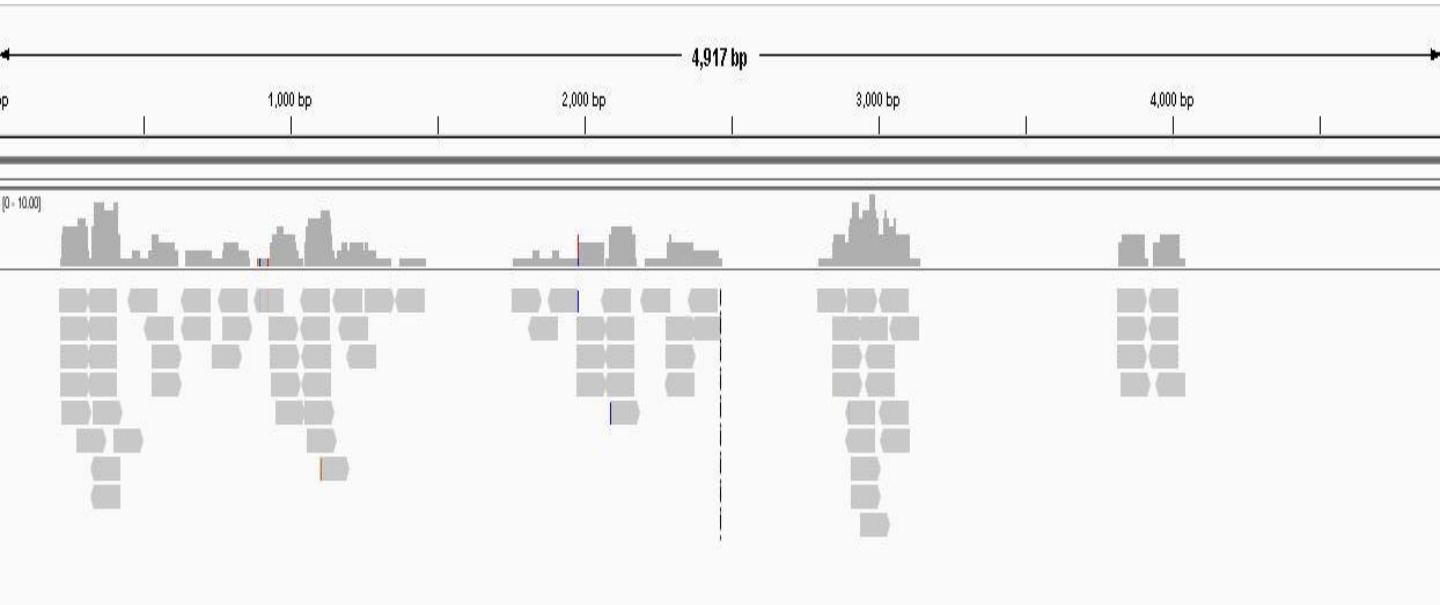
Figure S4.



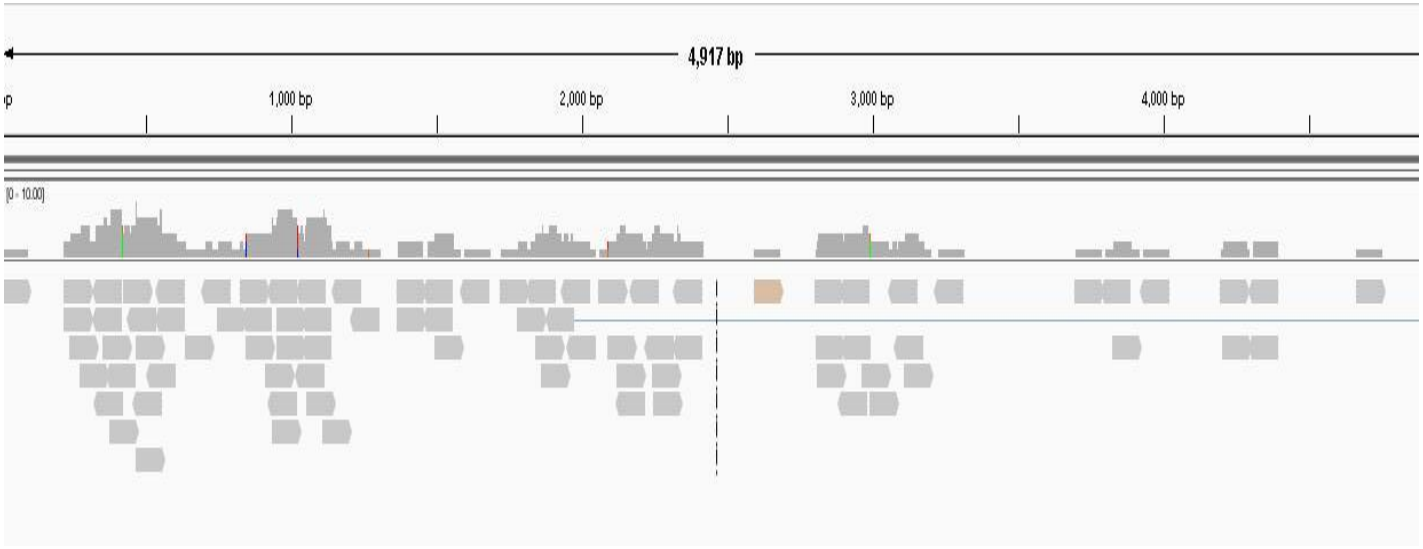
89
84
94
71
92
95
80
85
99
73
82
83
83
81
88
84
96
83
90
82
9678
75
96
99
70
87
TGTT1
TGTT2
TGTT3
TGTT4
AACA1
AACA2
0.1

Figure S5.

PbrAACA1_IT3



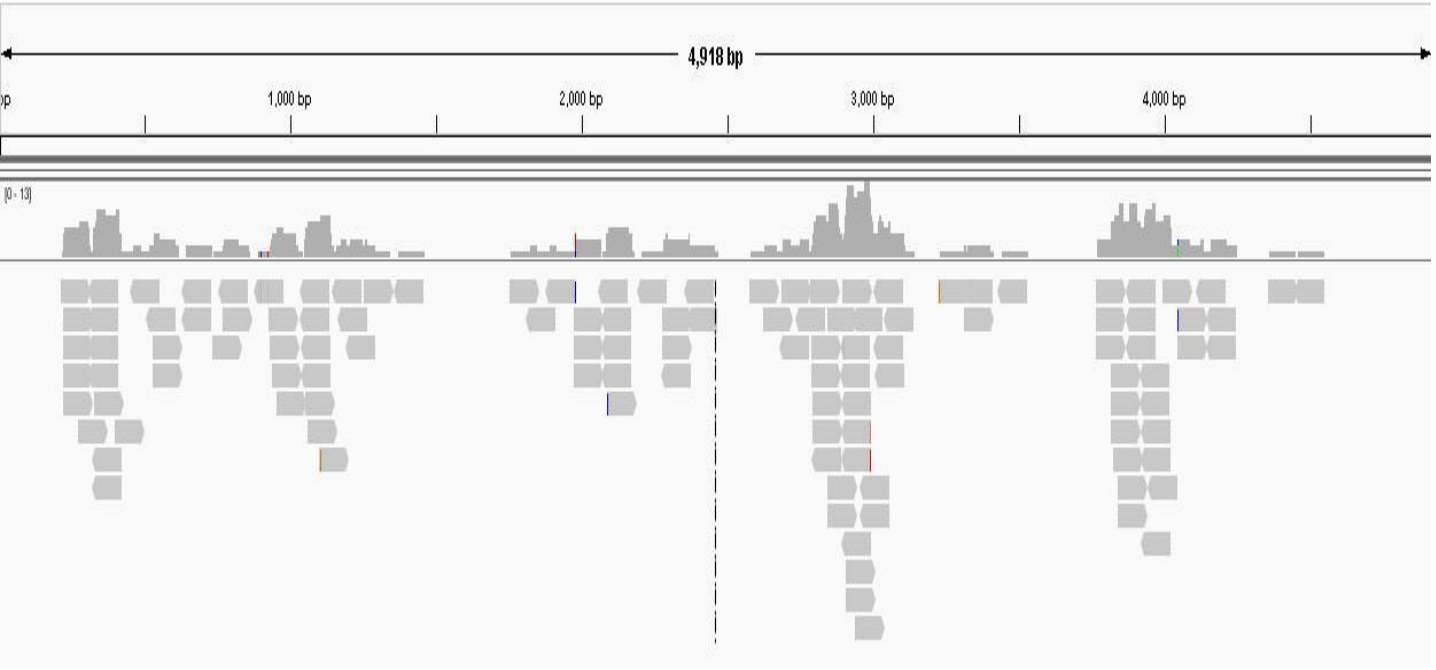
NCBI SRA database: SRX147917



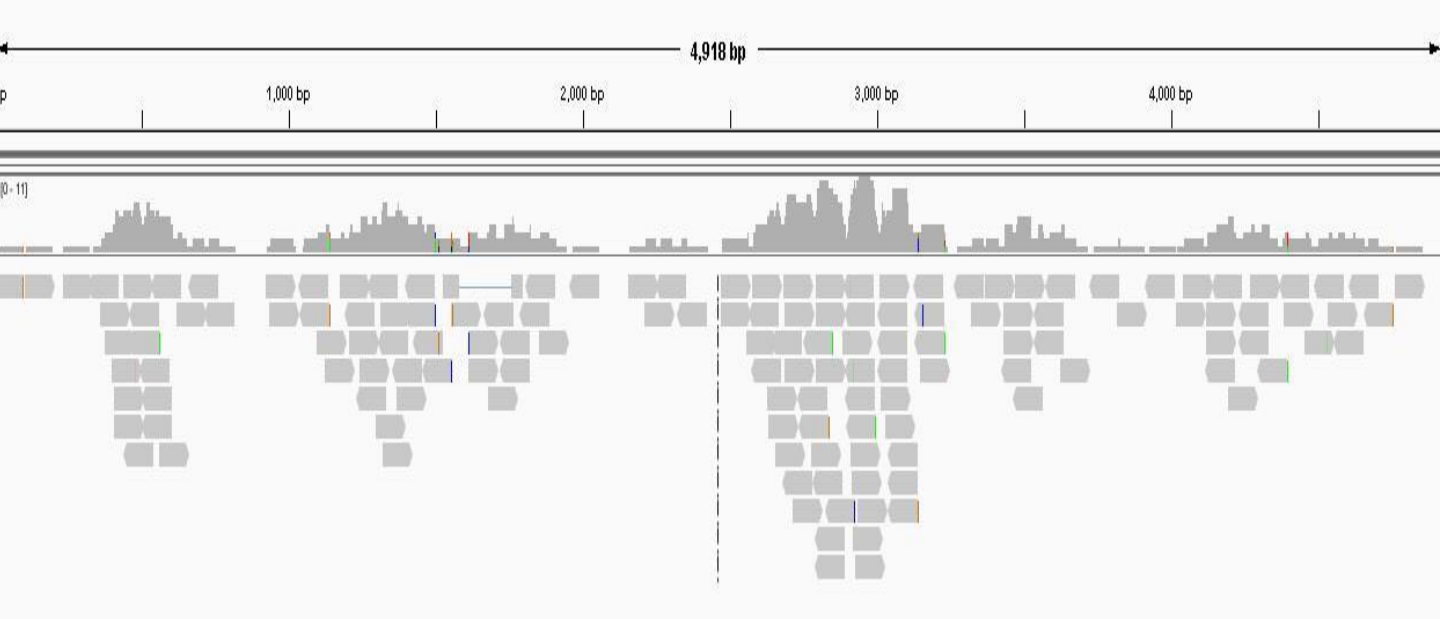
NCBI SRA database: SRX298075

Figure S5.

PbrAACA1_IT5



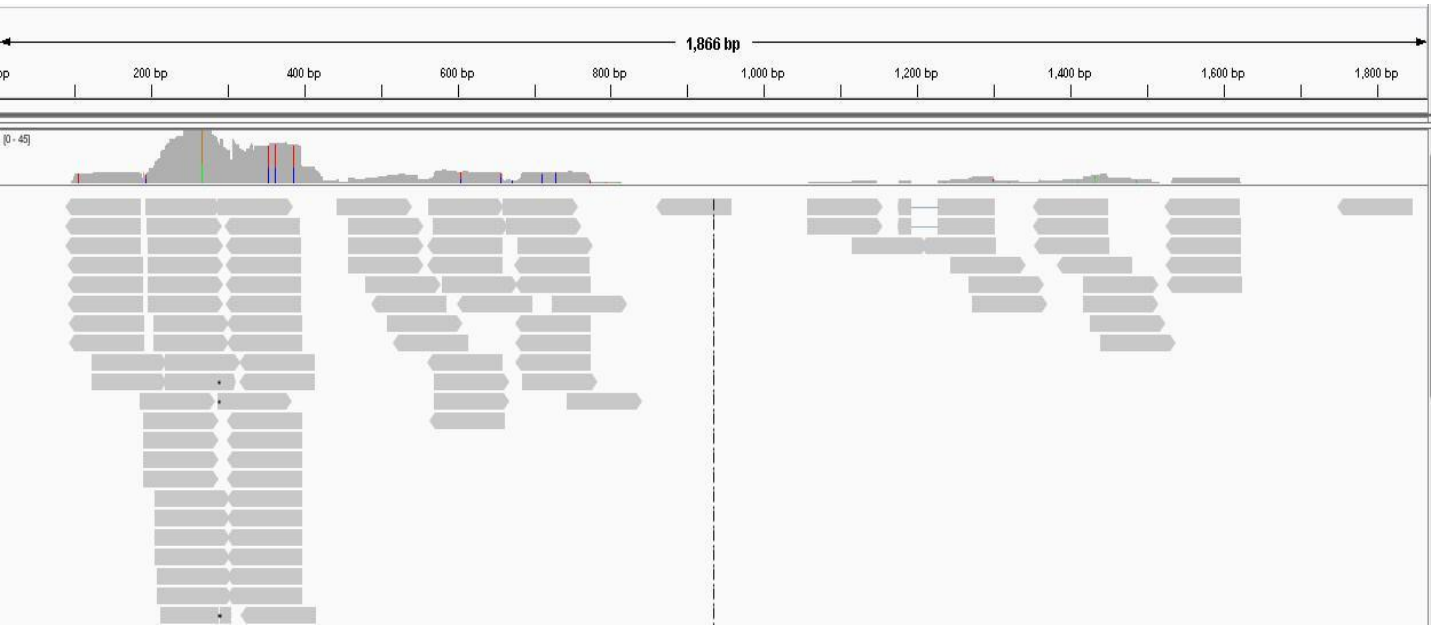
NCBI SRA database: SRX147917



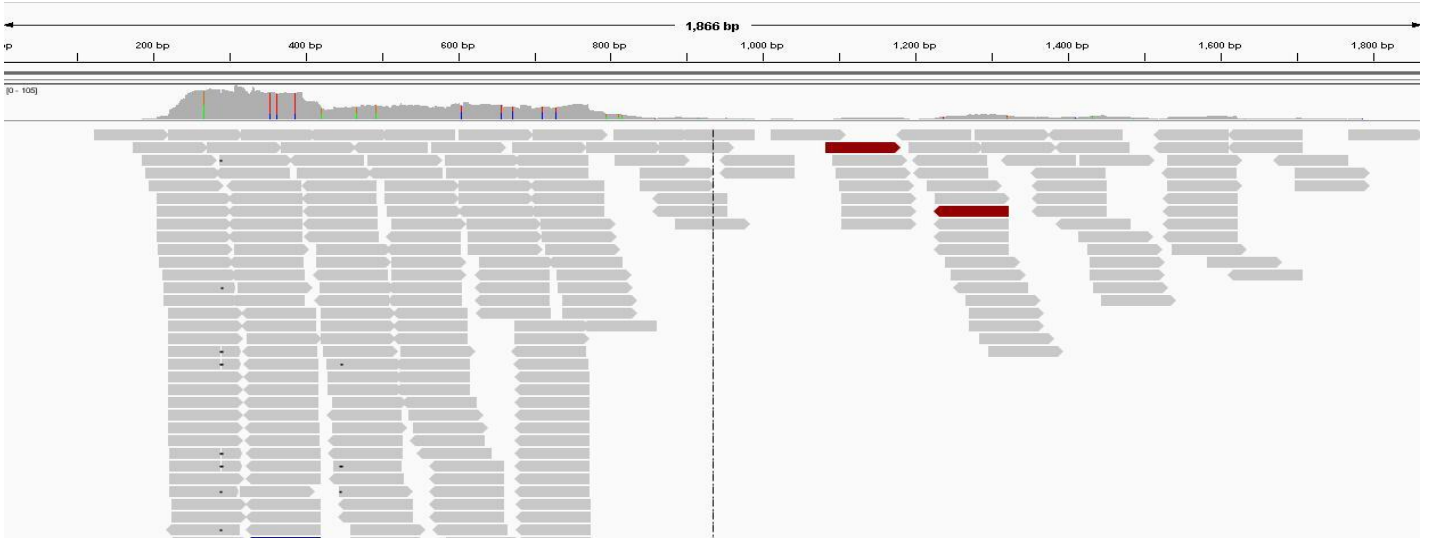
NCBI SRA database: SRX298075

Figure S5.

PbrAACA3_IT4



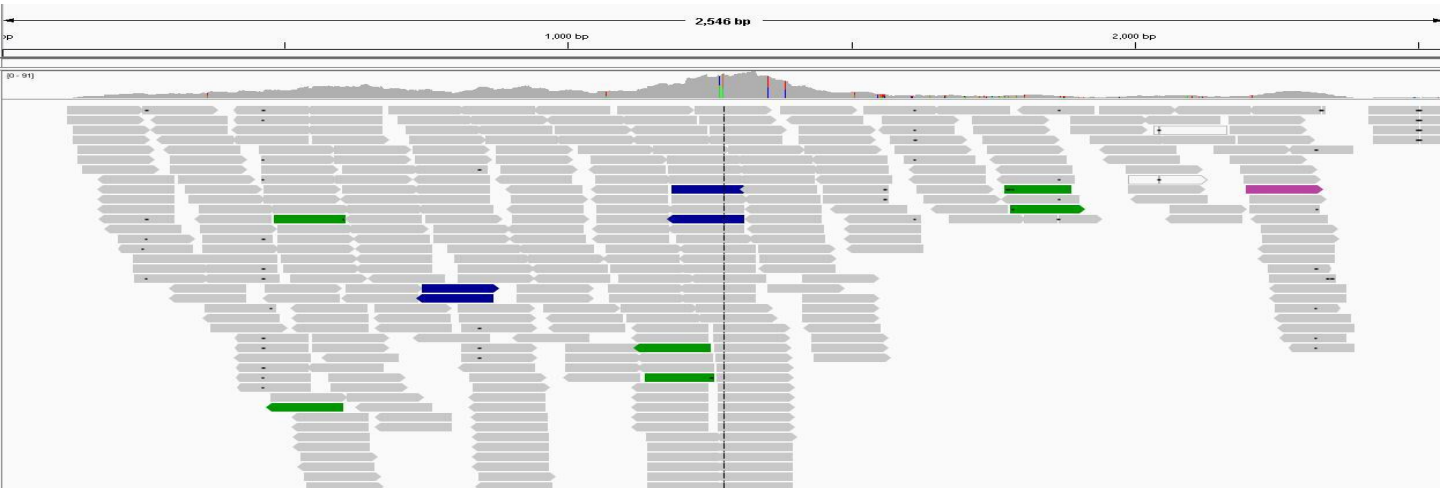
NCBI SRA database: SRX147917



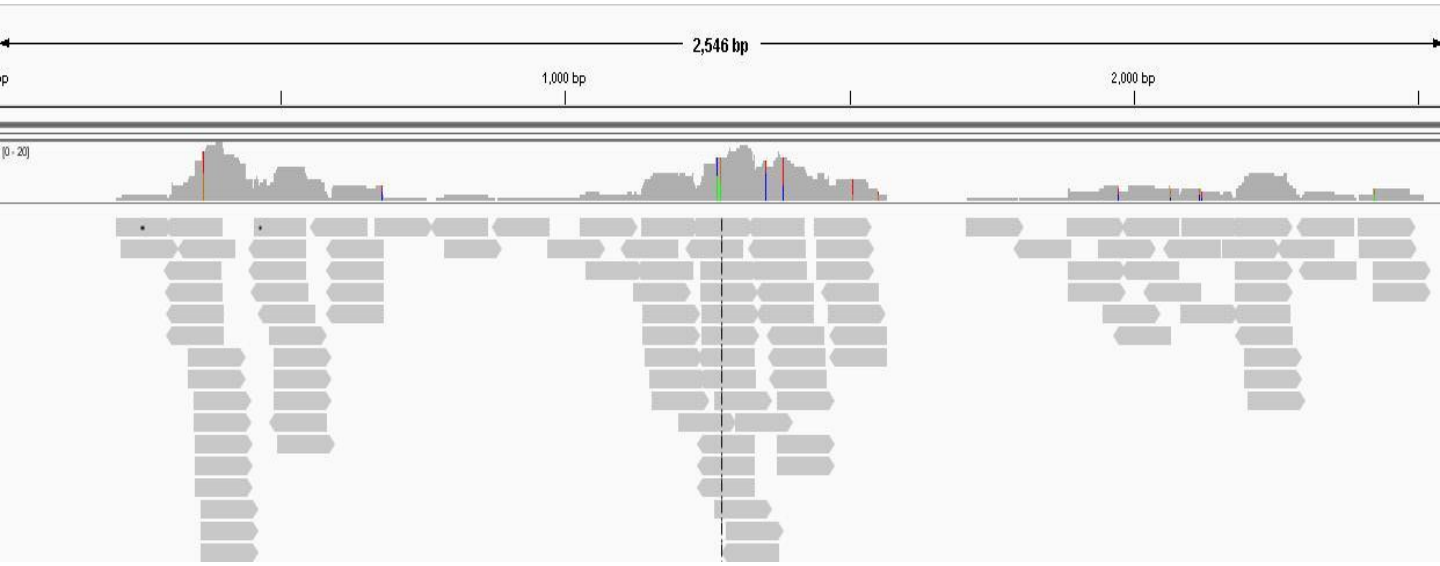
NCBI SRA database: : SRX298075

Figure S5.

PbrAACA4_IT11



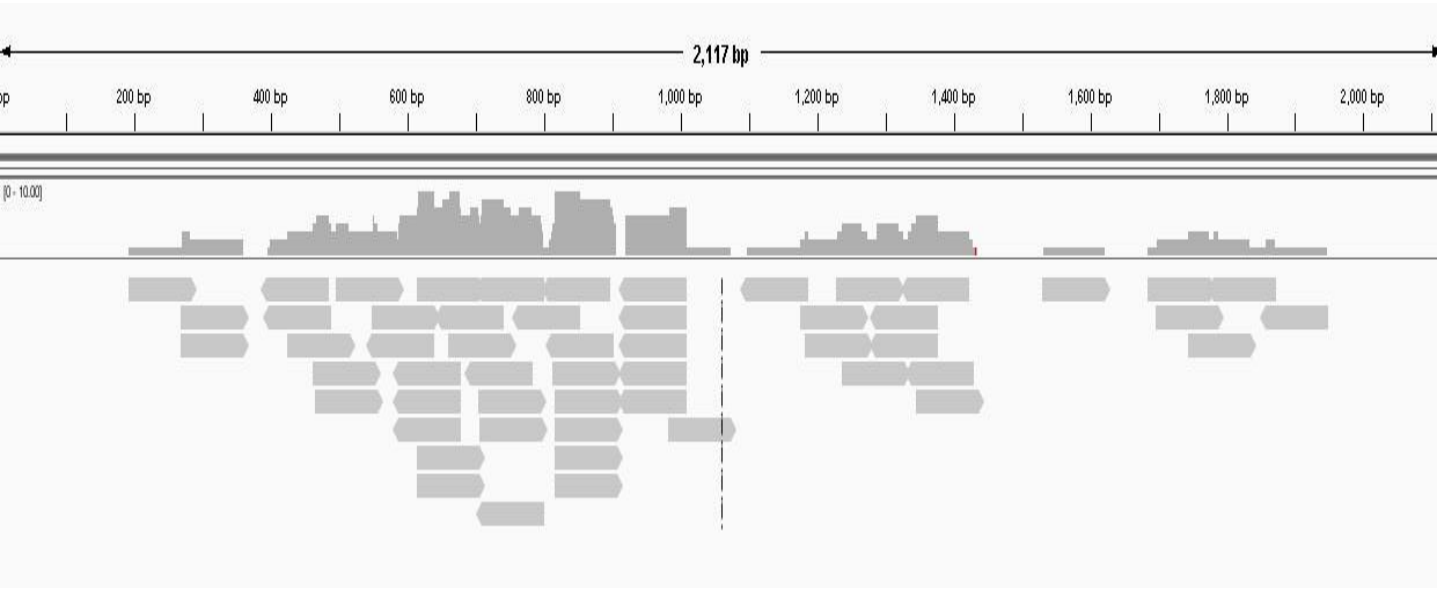
NCBI SRA database: SRX147917



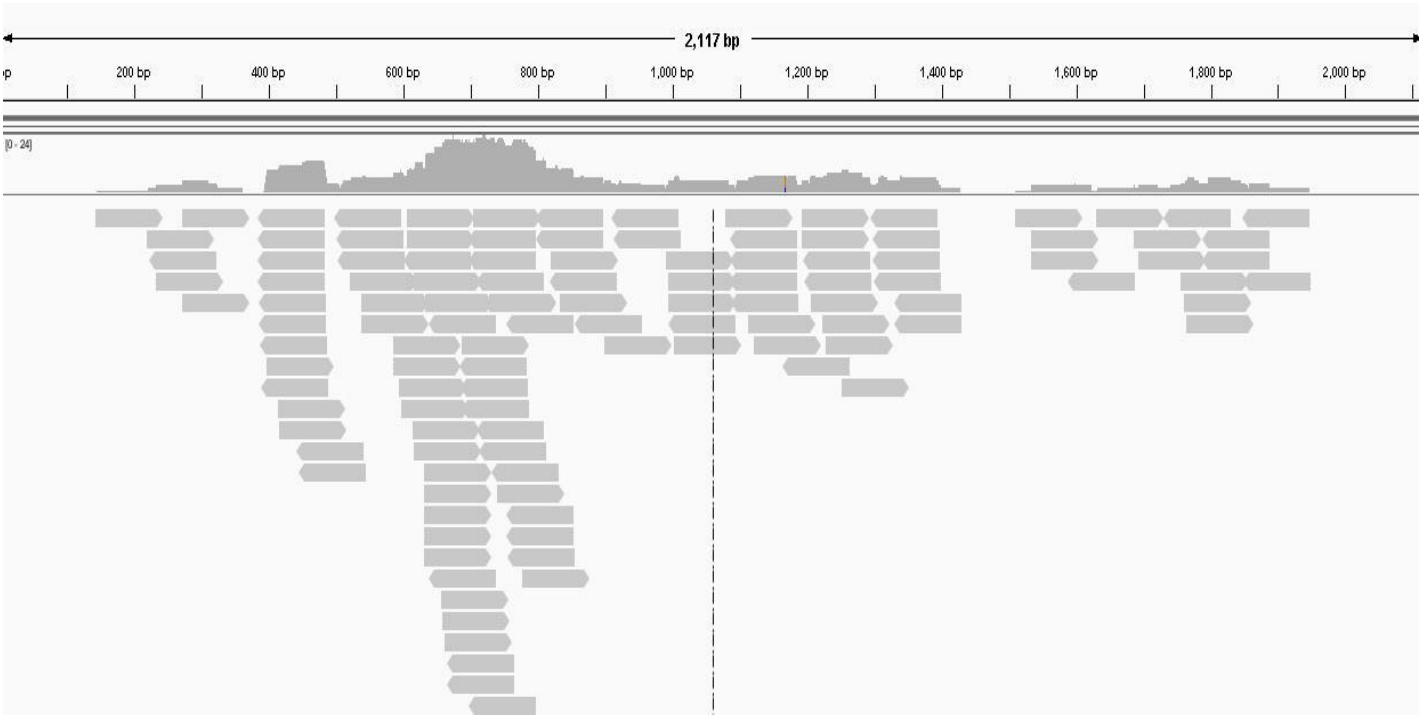
NCBI SRA database: SRX298075

Figure S5.

PbrAACA4_IT14



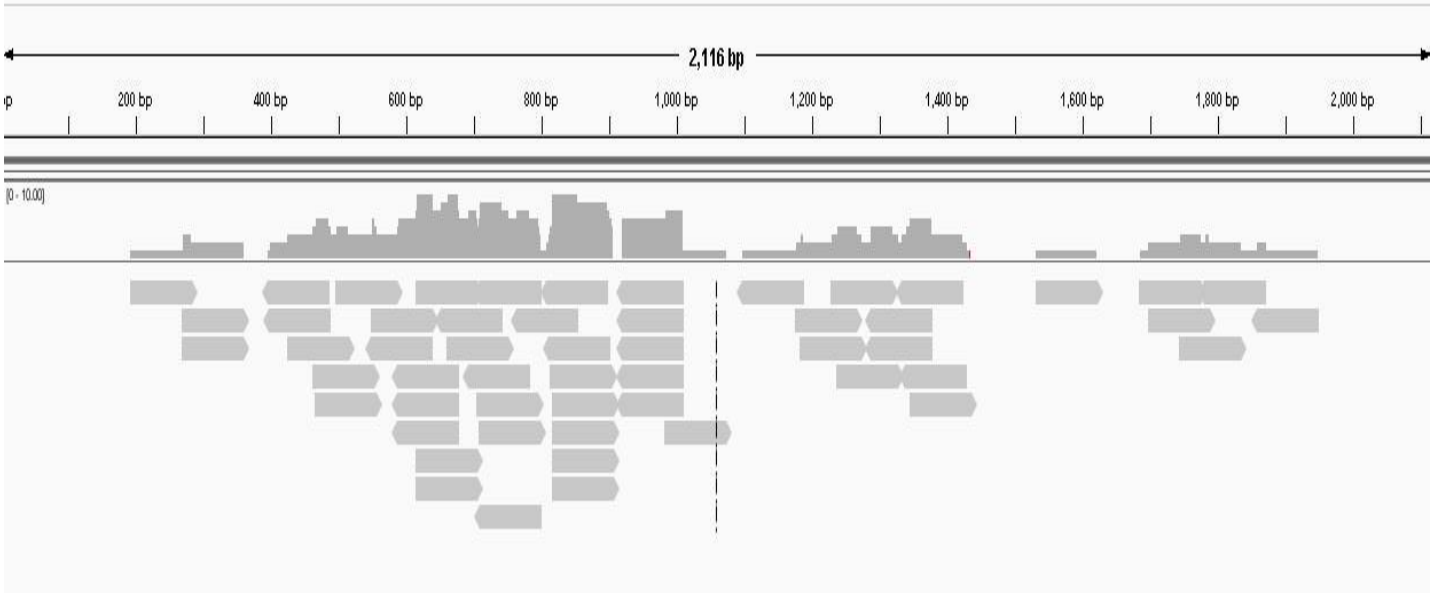
NCBI SRA database: SRX147917



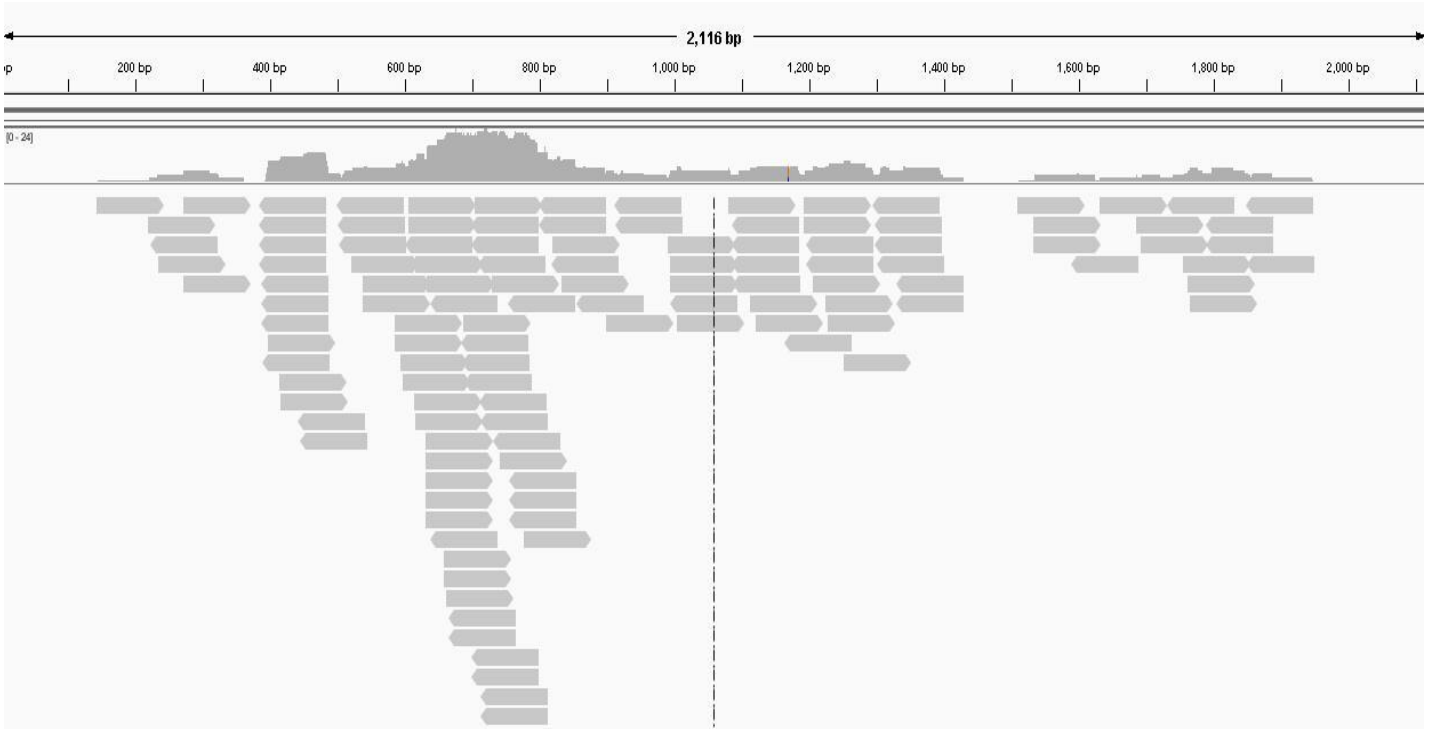
NCBI SRA database: SRX298075

Figure S5.

PbrAACA4_IT15



NCBI SRA database: SRX147917



NCBI SRA database: SRX298075