

Mapping of T-DNA insertion sites using sequence capture and Illumina-Miseq

Mauricio A. Reynoso/Bailey-Serres Lab, Center for Plant Cell Biology, UC Riverside

Citation

This protocol has been published as an appendix of: Reynoso M, Pauluzzi G, Kajala K, Cabanlit S, Velasco J, Bazin J, Deal R, Sinha N, Brady SM, Bailey-Serres J. *Nuclear transcriptomes at high resolution using retooled INTACT*. Plant Physiol. DOI: <https://doi.org/10.1104/pp.17.00688>

Purpose and Background

The purpose of this method is to localize T-DNA insertion sites in transgenic plants. This procedure was tested with DNA from *Oryza sativa* L. ssp. *Japonica* (cv. Nipponbare). The protocol involves the hybridization and capture of sequences containing T-DNA borders using biotinylated oligos. This allows for an enrichment of junction reads carrying both T-DNA and genomic DNA (gDNA) sequences. Insertions in multiple (25 or more) transgenic plants can be localized in a single sequencing run. This method includes the bioinformatic analyses.

Materials

- 100 mg young leaves
- Bioruptor sonicator (Diagenode) or Covaris S2
- Agencourt AMPure XP beads (Beckman Coulter)
- Low-bind microcentrifuge tube
- Savant Speedvac
- Dynabeads M-270 Streptavidin or Dynabeads MyOne Streptavidin C1 (Life Technologies)
- Illumina library preparation reagents (required for end-repairing, dA tailing, ligation and PCR amplification).
- Qubit fluorometer (Life Technologies)
- Agilent 2100 BioAnalyzer (Agilent Technologies)
- 1.5 mL microcentrifuge tube magnet
- Rotator/nutator mixer

Solutions and buffers

- 40% PEG-8000
- 5 M NaCl
- Elution Buffer: 10 mM Tris-HCl pH 7.5
- 40X Denhart's solution: Bovine Serum Albumin 0.8%; Ficoll 400 0.8%; Polyvinylpyrrolidone 0.8%. Dissolve in water. Filter to sterilize. Aliquot and store at -20° C
- 20X Saline Sodium Citrate (SSC) buffer: 3 M NaCl; 0.3 M Na-Citrate. Adjust to pH 7.0 with HCl. Autoclave
- 1 M Sodium phosphate buffer pH 7.0: 39% (v/v) 1M NaH₂PO₄; 61% (v/v) 1M Na₂HPO₄
- 5 µg/µl Salmon Sperm DNA: use sterile forceps to collect the DNA and dissolve in water. Autoclave for 15 min to fragment DNA
- 2X Hybridization buffer: 0.5 M Sodium phosphate buffer pH 7.0; 1% SDS; 2 mM EDTA; 2X SSC, 4X Denhardt's solution
- 2X Bind and Wash Buffer: 10 mM Tris-HCl pH 7.5; 2 M NaCl; 1 mM EDTA; 0.1% Tween-20
- Wash Buffer 1: 1X SSC; 0.1% SDS
- Wash Buffer 2: 0.1X SSC; 0.1% SDS
- Wash Buffer 3: 0.2X SSC
- 0.15M NaOH
- 1 M Tris-HCl pH 8.0

Biotinylated and blocking oligos

The biotinylated oligos were designed for the T-DNA borders of H-TRAP and H-INTACT based on the experience from Soichi and Isabelle Henry from Comai lab (UC Davis). There is a pair of 70 nt oligonucleotides for each border and they overlap 35 nt with each other. The 3' end of the oligos RB1 and RB2 contain 16 nt and 51 nt matching with the right border, respectively. The 5' sequence of the oligos LB1 and LB2 contain 67 nt and 32 nt matching the left border, respectively. The biotinylated oligos are diluted to a 3 μ M working solution.

> RB1

5'/5BiotinTEG/CTTGTGATATCACTAGTGC GGCCGCTGCAGGTCGACCATATGGGAGAGCTCAAGCTTAGCTTGAGCTTG -3'

>RB2

5'/5BiotinTEG/CCATATGGGAGAGCTCAAGCTTAGCTTGAGCTTGGATCAGATTGTCGTTTCCCGCCTTCAGTTTAACTA -3'

>LB1

5'/5BiotinTEG/TTGACGCTTAGACAACCTTAATAACACATTGCGGACGTTTTTAATGTACTGAATTAACGCCGAATTGAATT -3'

>LB2

5'/5BiotinTEG/CGTTTTTAATGTACTGAATTAACGCCGAATTGAATTATCAGCTTGCATGCCGGTCGATCTAGTAACATAG -3'

DNA Libraries made from different plants will share the universal adapter and a part of the index adapter. During the hybridization these adapters can interact causing cross hybridization and non-specific binding to the probes. To prevent this effect the hybridization mix contains an excess of adapter oligos that hybridize with the adapter region of molecules of the library protecting from cross hybridization. The blocking oligos contain a modification at the 3' end prevent amplification in the PCR enrichment that follows hybridization. These oligos should be designed to match the adapters used during library preparation. The blocking oligos are used at 1 mM concentration.

>Universal Oligo Blocker

AAT GAT ACG GCG ACC ACC GAG ATC TAC ACT CTT TCC CTA CAC GAC GCT CTT CCG ATC T/3SpC3/ -3'

>Index Oligo Blocker

'- CAA GCA GAA GAC GGC ATA CGA GAT **NNN NNN** GTG ACT GGA GTT CAG ACG TGT GCT CTT CCG ATC T/3SpC3/ -3'

The six N bases correspond to the index zone. If using eight nt barcode this stretch should contain eight Ns.

Primers

Primers needed for the enrichment after hybridization:

>IlluminaP5

AATGATACGGCGACCAACCGA

>IlluminaP7

CAAGCAGAAGACGGCATACGA

Primers used to calculate efficiency of the hybridization:

>qRB_1685_fwd

GCTTTTTTGTACAACTTGTGAT

>qRB_1796_rev

CTGAAGGCGGGAAACGACA

>qLB_8408_fwd

TGACGCTTAGACAACCTTAATAACACAT

>qLB_tNos_8527_Rev

GCGCGCGGTGTCATCT

The primers are used at 10 μ M concentration

DNA extraction

1. Extract genomic DNA from transgenic plants using your regular lab protocol. 100mg of tissue should give enough material for DNA library preparation.
2. Check DNA quality by electrophoresis in 1.2% agarose gels.
3. Quantify DNA concentration using a Nanodrop spectrophotometer.

DNA fragmentation and purification

Fragmentation and purification of genomic DNA samples from transgenic plants. Follow the next steps for each DNA sample.

Note: PEG content on AMPure XP beads binding buffers can be modified to obtain a distribution of fragments required for your Illumina sequencing (see figure at the end of the document). In this case we selected DNA fragments of approximately 400bp and sequenced using Illumina MiSeq to get 250bp paired-end reads.

1. Dilute one microgram of genomic DNA to 100 μ L using DNase free water.
2. Sonicate at 4°C in a Bioruptor water bath sonicator for 25 min with power set to 'low' and sonication intervals set to 30 sec on/30 sec off. Replace the water with fresh 4°C water every 5 minutes.
3. Wash 15 μ L of AMPure XP beads. Collect to the magnet and resuspend in 15 μ L 20% PEG-8000 and 2.5 M NaCl. Collect the beads again and resuspend in 100 μ L of a solution 12% PEG-8000; 2.5 M NaCl. This buffer will allow the binding of fragments bigger than 500 bp. (Note: amount of beads can be optimized).
4. Bind fragmented DNA with the resuspended AMPure XP beads from the previous step by adding to a low-binding microcentrifuge. Mix well and incubate for 6 min at room temperature.
5. Collect beads to the magnet and transfer the supernatant to a new tube. This solution contains DNA fragments of smaller than 500 bp.
6. Wash 15 μ L of AMPure XP beads. Resuspend in 15 μ L 20% PEG-8000 and 2.5 M NaCl. Collect the beads to the magnet and resuspend in 13.33 μ L of a solution of 30% PEG-8000 and 1.25 M NaCl. (Note amount of beads can be optimized).
7. Add 13.33 μ L of resuspended beads to the supernatant of step 5. This solution will increase the concentration of PEG-8000 to 7.5% allowing for the binding of DNA fragments bigger than 400 bp.
8. Incubate for 6 min at room temperature to allow DNA binding to AMPure XP Beads. Collect the beads to the magnet. Wash twice for 30 sec with 150 μ L 80% ethanol. Evaporate the remaining ethanol for 10 min at room temperature. Elute with the volume required in your Illumina library preparation protocol at the End-Repair step of library preparation.

Illumina Library Preparation

Follow your library preparation protocol starting at the end repair step. Perform End-Repair and dA tailing. Ligate a different barcode or index for each DNA sample. Purify twice with AMPure XP beads to capture ligated fragments and avoid adapter dimers.

Amplify fragments with 10 cycles of PCR. Purify with AMPure XP beads.

Note: Concentration of enzymes and incubation times can be modified (increased with respect to RNA-seq protocol) to improve the efficiency of library preparation.

Quality control and library pooling

1. Measure the concentration of each library using a Qubit® Fluorometer (Life Technologies).

2. Run 1 μL of library on a bioanalyzer (Bio-Rad Experion™ or Agilent 2100) using a DNA 1K chip. Take note of concentration and the average fragment length of library.

Optionally: Perform library quantification using the KAPA Library Quantification Kit.

3. Based on the concentration and average fragment length for each library, pool the libraries to obtain a total of 500 ng of DNA.

For example, when multiplexing 25 samples, add to a microcentrifuge tube 20 ng of each DNA library to obtain 500 ng of DNA for hybridization. It is possible to increase the total amount of DNA to 1 μg when multiplexing more samples.

Note: To evaluate efficiency of the capture it is possible to pool 5% more of DNA for each sample (1 ng more in the example) and store that extra volume at -20°C to use after the capture.

4. Dry out 500 ng of pooled DNA libraries using a Speedvac concentrator. After drying the DNA, resuspend in 2 μL of water.

Hybridization of pooled libraries

This part of the protocol is modified from the **xGen® Target Capture Protocol** *Protocol for DNA Probe Hybridization and Target Capture Using an Illumina Library* IDT

Hybridize biotinylated oligos to target sequences in pooled libraries.

a. Combine the following in a low-bind 0.2 mL PCR tube:

3 pmol each biotinylated oligo (3 pmol/ μL ; 1 μL per oligo)	4.0 μL
500 ng Illumina library (dried down, and then resuspended in water to 125 ng/ μL)	2.0 μL
10 μg Salmon Sperm DNA (5 $\mu\text{g}/\mu\text{L}$)	2.0 μL
1 mM Universal Oligo Blocker	1.0 μL
1 mM Index Oligo Blocker	1.0 μL
Total Volume	10 μL

b. Add 10 μL 2X Hybridization Buffer and vortex.

c. Overlay with 40 μL mineral oil. DO NOT vortex after overlaying with oil.

IMPORTANT: *It is essential that you overlay with mineral oil to avoid significant volume loss.*

d. Denature the mixture in a thermal cycler at 95°C for 5 min, and then slowly decrease the temperature ($0.1^{\circ}\text{C}/\text{sec}$) to 62°C . DO NOT use the heated lid of the thermal cycler.

e. Incubate at 62°C for 48 hr. This reaction will be used in Step 2.e. (below).

Purification of captured DNA

1. Bind hybridized target sequences to streptavidin beads.

a. Allow Dynabeads MyOne Streptavidin C1 to equilibrate to room temperature for 30 min.

b. Pipet 50 μL C1 streptavidin beads into a fresh 1.7 mL low-bind microcentrifuge tube and wash twice with 50 μL 2X Bind and Wash Buffer. Use a pipette to remove the buffer after each wash.

c. Combine the following to make Bead Resuspension Buffer

2X Bind and Wash Buffer	50 μL
Nuclease-Free Water	30 μL
Total Volume	80 μL

d. Resuspend the washed beads in 80 μL Bead Resuspension Buffer (from 1.c.).

e. Recover the hybridization reaction (20 μL) from under the mineral oil (step e.; above), being careful not to withdraw any mineral oil, and add to the resuspended beads.

Hint: *A gel-loading pipette tip makes it easy to remove the liquid from under the mineral oil. To avoid trapping any mineral oil in the tip, leave a small amount of positive pressure in the tip until reaching the*

mineral oil/hybridization solution interface. Expel the small amount of oil in the tip before removing the hybridization mixture.

f. Place the tube on tube rotator/nutator mixer for 30 min at room temperature to bind the biotinylated capture probes to the streptavidin beads.

2. Wash the streptavidin beads to remove unbound DNA.

a. Briefly spin down the tube from step 1.f. (above) and place on a magnetic separation rack. Allow beads to separate from the supernatant. Using a pipette, remove the supernatant containing unbound DNA and discard.

b. Wash the beads with the following solutions, sequentially:

i. 1000 µL Wash Buffer 1 for 5 min, 62°C with rotation

ii. 1000 µL Wash Buffer 2 for 5 min, 62°C with rotation

iii. 1000 µL Wash Buffer 2 for 5 min, 62°C with rotation

iv. 1000 µL Wash Buffer 2 for 5 min, RT with rotation

v. 1000 µL Wash Buffer 3 for 30 sec, RT with tube on magnet; keep tube on magnet

Ensure that each solution has been pre-equilibrated to the appropriate wash temperature.

Add the stated volume of each wash buffer to the tube and place on a bench top tube rotator/nutator mixer for the indicated amount of time.

Pulse spin (1–3 sec, 800 x g) to collect the liquid without pelleting the beads.

Place tube on magnetic holder to attract the magnetic beads.

Remove wash buffer using a pipette.

IMPORTANT! After the final wash (v), do not remove tube from magnet!

Remove ALL of the Wash Buffer 3 from the tube after the final wash. SSC present in the buffer may react with NaOH used in subsequent steps to produce a precipitate that will interfere with downstream manipulations.

c. Add 50 µL 125 mM NaOH (freshly diluted from a more concentrated stock solution) and incubate at RT for 10 min. Vortex every 2 min to keep beads in suspension.

d. Return the tube to the magnet for 1 min. Leave the tube on the magnet while you perform the next steps.

e. Add 50 µL 1 M Tris-HCl, pH 8.8, to a new 1.7 mL low-bind tube.

f. Keeping the tube on the magnet, transfer the supernatant (d) to the tube of 1 M Tris-HCl (e). This neutralizes the NaOH added to the hybridization mixture.

g. Add 1.5X volumes (150 µL) Agencourt® AMPure® XP beads (Beckman-Coulter). Proceed according to the manufacturer's protocol, eluting in 22 µL Elution Buffer.

h. Transfer 20 µL of eluted product into a new 1.7 mL low-bind tube, ensuring no beads are carried over.

Quantification of the efficiency of hybridization

This step is recommended but can be omitted.

Perform quantitative PCR on pooled libraries from before and after hybridization.

a. Take 1 µL of eluted product and dilute to 10 µL in Elution buffer. This DNA corresponds to 5% fraction of the total eluted product.

b. Take 5% of the DNA from pooled libraries from input of the hybridization. In the example that is 25 ng of pooled libraries as $1/20$ (fraction of output used for quantification) * 500 ng (input) = 25 ng.

c. Quantify the abundance of RB and LB ends for both samples. Perform qPCR with the primers for the RB and LB indicated in material section.

d. Use dilutions of a vector containing RB and LB to calculate primer efficiency in qPCR.

e. Normalize Ct values using a primer efficiency curve and calculate the ratios of efficiency of RB and LB capture. This can be obtained by dividing the normalized values for the captured samples to those obtained for the input sample.

PCR enrichment and quality control**1. Post-Capture PCR**

The eluate from step h allows retention of a small amount of captured library for troubleshooting, if necessary. Otherwise, the library can be eluted in a smaller volume and the entire amount used in the final PCR enrichment.

Perform final PCR enrichment. This reaction can be done using the same PCR reagents used for PCR enrichment during the preparation of the libraries.

a. Prepare the following reaction mix in a low-bind 0.2 mL PCR tube:

2X KAPA HiFi™ HotStart ReadyMix	25 µL
10 µM Illumina P5 Primer	2.5 µL
10 µM Illumina P7 Primer	2.5 µL
Nuclease-Free Water	4 µL
Eluted capture library	16 µL
Total Volume	50 µL

b. Briefly vortex the tube, and then spin briefly in a microcentrifuge to collect the reaction mixture at the bottom of the tube.

c. Place reactions in a thermal cycler and run the following program:

1. 98°C	5 min
2. 98°C	20 sec
3. 60°C	15 sec
4. 72°C	30 sec
5. Cycle 14 times to step 2	
6. 72°C	5 min
7. 4°C	Hold

d. Purify the fragments using 1.5X volume (75 µL) AMPure XP beads. Proceed according to the manufacturer's protocol, eluting in 22 µL Elution Buffer.

e. Transfer 20 µL of eluted product into a fresh 1.7 mL low-bind tube, ensuring no beads are carried over.

Note: some protocols use a second step of hybridization with the PCR enriched capture DNA, this was not tested but can be considered an alternative when obtaining a low efficiency of capture.

2. Quality control

a. Measure the concentration of captured library using a Qubit® Fluorometer (Life Technologies).

b. Run 1 µL of library on a bioanalyzer (Bio-Rad Experion™ or Agilent 2100) using a DNA 1K chip. Take note of the average fragment length.

c. Submit sample for Illumina MiSeq. (We used 250 bp paired-end reads).

Sequencing and Bioinformatic Analyses

Two approaches were taken to analyze the 250 bp paired-end reads, obtaining similar results. Jérémie Bazin from Bailey-Serres lab helped on the design of the analysis. For both cases paired-end reads that overlapped were collapsed to one longer read.

The first method selects sequences that contain a fraction of the left or right border. These reads are locally mapped to the genome, i. e. a part of the sequence align to the genome. Then, regions with a significant coverage of reads are selected to indicate potential insertions. Reads aligned to candidate regions are aligned back to the T-DNA to determine which portion of it was incorporated in the genome (indicating the site of junction between T-DNA and gDNA).

1. Merge paired-end reads containing overlapping sequences to obtain longer reads

```
### MERGE PAIRS IN R using FLASH module###
mytargetsfq <- read.delim("your_directory/targets_lane1_fastq.txt", sep="\t")
mypair1=mytargetsfq$FileName1
mypair2=mytargetsfq$FileName2
sample=mytargetsfq$SampleName
zip=paste0("flash ",mypair1[1]," ",mypair2[1]," -m 12 -M 250 -o ",sample[1]," --suffix=flash")
for(i in seq(along=mypair2)) {
  zip=paste0("flash ",mypair1[i]," ",mypair2[i]," -m 12 -M 250 -o ", sample[i])
  system(zip)
}
```

2. Filter reads containing border sequences

```
#### RUN IN UNIX using GREP module###
for i in *extendedFrgs.fastq ;
do      grep      --no-group-separator      -B      1      -A      2      -E
      "CTAGTGC GGCCGCTGCAGGTCGACCATATGGGAGAGCTC|GAGCTCTCCCATATGGTCGACCTGCAGGCGG
      CCGCACTAG|TTATCAGCTTGCATGCCGGTCGATCTAGTAAC|GTTACTAGATCGACCGGCATGCAAGCTGATA
      A" $i > $i.grep ; done
```

3. Map reads containing border sequences against the genome

in UNIX make genome file accessible for BOWTIE2####

bowtie2-build **genome.con** **genome.fa**

ALIGN SELECTED READS FROM PREVIOUS STEP In R using BOWTIE2 module####

```
setwd("/your_directory")
library(ShortRead);library(rtracklayer); library(GenomicRanges);
library(Rsamtools);library(GenomicAlignments);
library(DESeq); library(edgeR); library(ShortRead);library(chipseq); library(lattice);library(ggbio)
targets=read.delim("target.txt",header=T)
input=paste0(targets$path,".grep")
setwd("/your_directory")
for(i in seq(along=input)) {
  bowtie=paste0("bowtie2 --local -p 8 --very-sensitive -x /your_directory/genome.fa -U ", input[i]," |
  samtools view -bS - > ",paste0(targets$SampleName[i],".bam"))
  system(bowtie) # -g: ignore all alginments with >g matches; -p: number of threads to use for alignment
  step; -i/-l: min/max intron lengths; --segment-length: length of split reads (25 is default)
}
```

4. Find regions with candidate insertions

EXTRACT THE PEAKS FOR EACH SAMPLE AND WRITE A TABLE in R

```
result=data.frame()
for(i in seq(along=input)) {
  alignedReads <- readGAlignments(paste0(targets$SampleName[i],".bam"))
  coverage <- coverage(alignedReads) # calculate genomic coverage and returns an IRanges Rle object
  peakCutoff(coverage, fdr = 0.0001)
  peaks <- slice(coverage, lower = 100)
```

```
peaks.sum <- peakSummary(peaks) # Summarizes the peak data in a RangedData object.
df=as.data.frame(peaks.sum)
df$sample=targets$SampleName[i]
result=rbind(result,df)
write.table(result, "insertions.xls", sep="\t")
}
```

5. Align reads of each candidate region to the vector to define the site of junction between T-DNA and gDNA

```
##EXTRACT BAM FILES aligned to insertions with SAMTOOLS
sam_input <- read.delim("insertions.xls", sep="\t")
chr=sam_input$space
from=sam_input$start
to=sam_input$end
sample=sam_input$sample
for(i in seq(along=chr)) {
  sam=paste0("samtools view -b ",sample[i],".bam ",chr[i],":",from[i],"-",to[i],
  ",sample[i],"_ins",i,".bam")
  system(sam)}
```

```
##CONVERT BAM TO FASTQ in UNIX using BEDTOOLS modules##
targets=read.delim("insertions.xls",sep="\t")
input=paste0(targets$sample,"_ins",targets$n,".bam")
output=paste0(targets$sample,"_ins",targets$n,".fastq")
for(i in seq(along=input)) {
  bam2fastq=paste0("bedtools bamtofastq -i ",input[i]," -fq ", output[i])
  system(bam2fastq)}
```

```
#in UNIX make vector file accessible to BOWTIE2 aligner##
bowtie2-build vector.fa TDNA.fa
```

```
### ALIGN FASTQ READS FROM MAPPED in R using BOWTIE2 module####
setwd("/bigdata/mreynoso/data/illumina.bioinfo.ucr.edu/illumina_runs/299")
library(ShortRead);library(rtracklayer); library(GenomicRanges);
library(Rsamtools);library(GenomicAlignments);
library(DESeq); library(edgeR); library(ShortRead);library(chipseq); library(lattice);library(ggbio)
targets=read.delim("insertions.xls",sep="\t")
input=paste0(targets$sample,"_ins",targets$n,".fastq")
output=paste0(targets$sample,"_ins",targets$n,"_vec.bam")
setwd("/bigdata/mreynoso/data/illumina.bioinfo.ucr.edu/illumina_runs/299/")
for(i in seq(along=input)) {
  bowtie=paste0("bowtie2 --local -p 8 --very-sensitive -x TDNA.fa -U ", input[i]," | samtools view -bS ->
  ",output[i])
  system(bowtie) # -g: ignore all alignments with >g matches; -p: number of threads to use for alignment
  step; -i/-l: min/max intron lengths; --segment-length: length of split reads (25 is default)
}
```

```
#### EXTRACT THE PEAKS FOR EACH SAMPLES AND WRITE A TABLE in R ####
targets=read.delim("insertions.xls",sep="\t")
input=paste0(targets$sample,"_ins",targets$n,".fastq")
output=paste0(targets$sample,"_ins",targets$n,"_vec.bam")
```



```

name=paste0(targets$sample,"_ins",targets$n)
result=data.frame()
for(i in seq(along=input)) {
  alignedReads <- readGAlignments(output[i])
  coverage <- coverage(alignedReads) # calculate genomic coverage and returns an IRanges Rle object
  peaks <- slice(coverage, lower = 100)
  peaks.sum <- peakSummary(peaks) # Summarizes the peak data in a RangedData object.
  df=as.data.frame(peaks.sum)
  df$sample=name[i]
  result=rbind(result,df)
  write.table(result, "insertions_in_vector.xls", sep="\t")
}

```

The second method selects reads that do not align completely (end-to-end) to the genome to do the analysis. Those reads are then locally mapped to the genome, i. e. align part of the read to the genome. Then, regions with a significant coverage of reads are extracted. And finally, reads mapping to each candidate regions are aligned to the T-DNA to determine the presence of vector sequences and the site of junction between T-DNA and gDNA.

1. Select reads that do not align completely to the genome

```

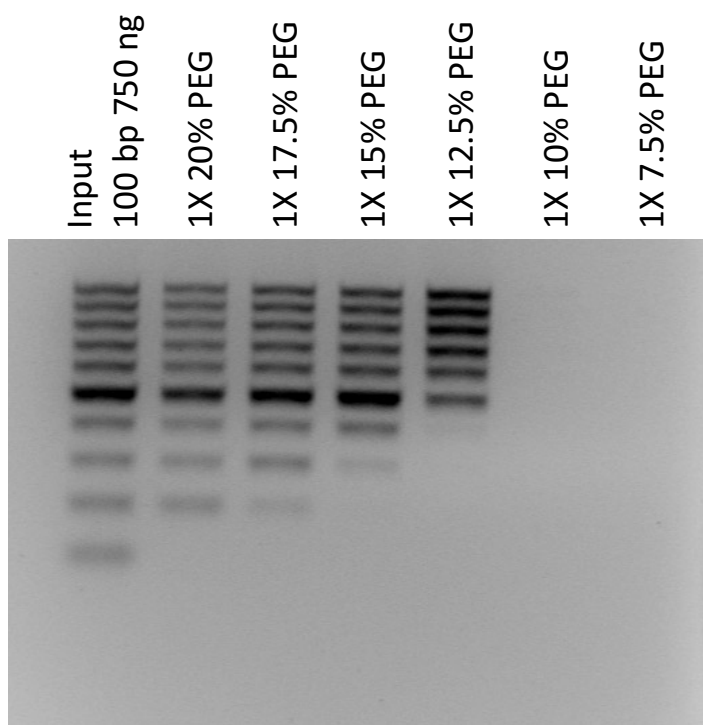
### SELECT READS THAT DO NOT MAP COMPLETELY TO THE GENOME in R using BOWTIE2 module####
setwd("/your_directory")
library(ShortRead);library(rtracklayer); library(GenomicRanges);
library(Rsamtools);library(GenomicAlignments);
library(DESeq); library(edgeR); library(ShortRead);library(chipseq); library(lattice);library(ggbio)
targets=read.delim("target.txt",header=T)
input=targets$path
for(i in seq(along=input)) {
  bowtie=paste0("bowtie2 --end-to-end -p 8 -D 20 -R 3 -N 1 -L 20 -i S,1,0.50 -x
/your_directory/genome.fa", " --un ", input[i], ".unaligned ","-U ", input[i], " | samtools view -bS - >
",paste0(input[i],".bam"))
  system(bowtie)
}

```

After this point the analysis continues with steps 3, 4 and 5 described for the previous approach using the unaligned fastQ file produced by the script above.

Note: It is important to verify that putative insertions do not map to sequences present in the transformation vector that map to the genome (in our case OsRPL18 and promoters).

The insertions should be tested by PCR using primers that amplify the junction between T-DNA and gDNA. Primers qRB_1685_fwd and qLB_tNos_8527_Rev can be used for this purpose together with a primer designed for each side of insertion, RB and LB respectively.



Elution from beads