

## TRANSPOSON INSERTION SITE VERIFICATION

Transposon and T-DNA insertion in Arabidopsis genes can be identified using the Arabidopsis thaliana Insertion Database (ATldb) (<http://atldb.org/cgi-perl/gbrowse/atibrowse>). There is, as yet, no publicly available insertion site verification data

### VERIFICATION OF INSERTIONS SITES

- Enter the gene id into the **Landmark or Region** window (in this case At2g38230) (<http://atldb.org/cgi-perl/gbrowse/atibrowse>) press the **Search** button.

Address <http://atldb.org/cgi-perl/gbrowse/atibrowse>

**Arabidopsis thaliana Insertion Database**

Home Genome Browse BLAST Gene Traps Sub-seq Contact

Arabidopsis thaliana (TIGR genome version 5)

Showing 1 bp from 3, positions 13,748,211 to 13,748,211

**Instructions:** Search using a sequence name, gene name, locus, oligonucleotide (15 bp minimum), or other landmark. The wildcard character \* is allowed. To center on a location, click the ruler. Use the Scroll/Zoom buttons to change magnification and position.  
**Examples:** 1:97,000, 117,000, 2:102,000, 112,000, At2g26800.1, apetala1, homeotic, flowering, FCA, F4N2, F4N\*, MGT\_3\_6639\*, MGT\*6639\*, CGAATTCAATAGAATG, CGAATTCAATAGAAT, RNA\*.

[Hide banner] [Hide instructions] [Bookmark this view] [Link to an image of this view] [Publication quality image] [Help]

Landmark or Region: At2g38230 Search Reset Flip Scroll Zoom: Show 1 bp

Overview of 3

Resizing small interval to 5000 bp

Insertions in Landsberg erecta

- Click on the bar in the **Gene Region** field. This links to a reference page in ATldb that give you details of the gene and the insertions;

Arabidopsis thaliana (TIGR genome version 5)

Showing 1.039 kbp from 2, positions 16,018,534 to 16,019,572

**Instructions:** Search using a sequence name, gene name, locus, oligonucleotide (15 bp minimum), or other landmark. The wildcard character \* is allowed. To center on a location, click the ruler. Use the Scroll/Zoom buttons to change magnification and position.  
**Examples:** 1:97,000, 117,000, 2:102,000, 112,000, At2g26800.1, apetala1, homeotic, flowering, FCA, F4N2, F4N\*, MGT\_3\_6639\*, MGT\*6639\*, CGAATTCAATAGAATG, CGAATTCAATAGAAT, RNA\*.

[Hide banner] [Hide instructions] [Bookmark this view] [Link to an image of this view] [Publication quality image] [Help]

Landmark or Region: At2g38230 Search Reset Flip Scroll Zoom: Show 1.039 kbp

Overview of 2

Insertions in Landsberg erecta

Insertions in Columbia

Gene Traps with GIS patterns

**Gene Region (Transcription Unit)**  
 At2g38230  
 EST:At2g38230.1 stress-responsive protein, putative, GI:10002554 (molecular\_function unknown)

Pseudogenes

Transcript  
 At2g38240.1

Select an insertion line from the **Gene Insertions** field. Click on the line number. This links in to a second reference page in ATIdb that gives you details of the bp position of the insertion relative to the psuedochromosome (in this case, SM\_3.22656 is located on Chromosome 2 position 16018847)

Transcript(s)			
Transcript name	5' coding	3' coding	Peptide Length translation
At2g38230.1	1601853	16019482	309

TargetP Results					
Transcript name	Compartment	CTP	mTP	SP	Other
At2g38230.1	UNKNOWN	0.233	0.053	0.083	0.867

Gene Insertions					
Insertion flanking sequence	Mapping Score (README)	Chromosome	Base	Position in gene	
T-DNA_LB_GK-755D10-023550	5	2	16019428	895	
T-DNA_LB_GK-755D10-023563	5	2	16019428	895	
T-DNA_LB_GK-423C02-017847	3	2	16019132	599	
SM_3.22656	1	2	16018847	314	
SM_3.22655	1	2	16018847	314	
SM_3.22656	1	2	16018847	314	
SM_3.1248	1	2	16018728	195	

Insertions Upstream within 800 b.p					
Insertion flanking sequence	Mapping Score (README)	Chromosome	Base	Position in gene	
T-DNA_LB-T-DNA_SAIL_619_G03.v3	5	2	16017790	-743	
T-DNA_LB_GK-077E08-016172	5	2	16018420	-113	
T-DNA_LB_SALK_024245.53.70.x	1	2	16018452	-81	
T-DNA_LB_GK-742E06-023520	5	2	16018480	-53	

- Select **Sub-Seq** from the toolbar (<http://atidb.org/cgi-perl/subseq>). This links to a third page in ATIdb which runs the Sub-Seq tool which allows you to extract the genomic sequence surrounding the insertion;



- Enter the details of the insertion site into the fields within Sub-Seq. (In this case Chromosome 2; 16018347 to 16019347 (500bp 5' + 500bp 3' to the insertion site, at position 16018847). Click the **Get DNA Sequence** button;

Address <http://atidb.org/cgi-perl/subseq> Go

- Sub-Seq will retrieve the DNA sequence in a FASTA Format suitable for using in sequence analysis packages;

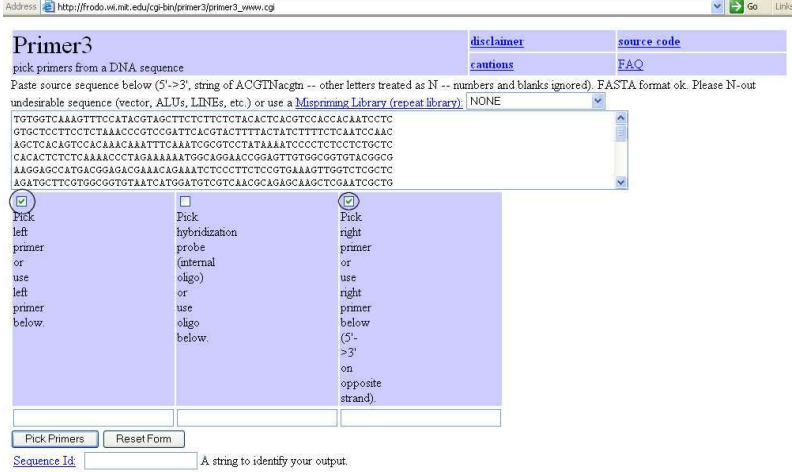
Address <http://jatidb.org/cgi-perl/subseq> Go Link

### Arabidopsis thaliana Insertion Database

[Home](#) [Genome Browse](#) [BLAST](#) [Gene Traps](#) [Sub-seq](#) [Contact](#)

```
>Arabidopsis thaliana chromosome 2 (TIGR, version v01212004, length 19705359), position 16019347 to 16019347, length 1001 b
TGTGGTCAAAGTTTCATACGCTAGCTTCTCTCTCTACACTCAGCTCCACCAATCCCTC
GTGGCTCTCTCTTAAACCCGTCGATTCACGTACTTTTACTACTTTTCTCAATCCAAAC
AGCTCAGCTCCAAACAAATTTCAAATCCGCTCTATAAAATCCCTCTCTCTCTCTC
CACACTCTCTCAAACCCCTAGAAAATGGAGGAAACCGGAGTTGTGGGGTGTACGGGG
AAGGACCATGACGGAGACGAAACAGAAATCTCCCTTCTCCGTGAAAGTTGGTCTCGCTC
AGATGCTTCGTGGGGGTGTAATCATGGATGTCCTCAACGAGAGCAAGCTCGAATCGCTG
AAGAAGCTGGGCGATGCGCGGTGATGGCTCTTGAACGCTGTCGCCCGGATATTCGAGCTC
AAGCGGGTGTGCTGGAATGAGCGATCCAGAGATGATCAAAGAAATCAAACGCGCTGA
CGATTCGGGTGATGGCGAAAGCTAGAAATTGGTCAATTCGTTGAAAGCTCAGATCCGGAAG
CAATCGGAGTTGATTACGTCGACGAGAGTGAAGTTCTCACTCTCGCGGACGAAGATAATC
ACATCAACAACATTAATTTCAAATCCCTTTTGTGTGGATGAGAAATCTCGGTGAAG
CTTTAAGCGGATCGTGAAGGAGCGCGCATGATTAAGAACCAAGGCTGAGGCTGAACTG
GTAACGTTGTAAGCCGTTAGGCACTGAGGAGTGTGALCGAGCTATTGCTTACTTA
GAAGCATGAGCATGAGGAGGTTTCACTTACGCGAAAAGATCGCTGCGCGGTATGATT
TGGTTGTCAGACTAAGGAGCTTGGGAGTTACCGTGGTTCAGTTGCTGCTGGAGGAG
TGGCAGCGCGCGGATGCGGGCTTGTATGATGAGTTGGGATGATGAGTGTGTTGTTG
GTCGGGTGTTTCAAAGTGGAGATCCGGTAAAGAGGGCT
```

- Copy this sequence and paste it into a primer design package (we recommend Primer 3: [http://frodo.wi.mit.edu/cgi-bin/primer3/primer3\\_www.cgi](http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi)). Ensure that you tick the **select left primer** and **select right primer** boxes;



Primer3

pick primers from a DNA sequence

Paste source sequence below (5'→3', string of ACGTINacgtm -- other letters treated as N -- numbers and blanks ignored). FASTA format ok. Please N-underscore undesirable sequence (vector, ALUs, LINEs, etc.) or use a [Mispriming Library](#) (repeat library): NONE

TGTGGTCAAAGTTTCATACGCTAGCTTCTCTCTCTACACTCAGCTCCACCAATCCCTC  
GTGGCTCTCTCTTAAACCCGTCGATTCACGTACTTTTACTACTTTTCTCAATCCAAAC  
AGCTCAGCTCCAAACAAATTTCAAATCCGCTCTATAAAATCCCTCTCTCTCTCTC  
CACACTCTCTCAAACCCCTAGAAAATGGAGGAAACCGGAGTTGTGGGGTGTACGGGG  
AAGGACCATGACGGAGACGAAACAGAAATCTCCCTTCTCCGTGAAAGTTGGTCTCGCTC  
AGATGCTTCGTGGGGGTGTAATCATGGATGTCCTCAACGAGAGCAAGCTCGAATCGCTG

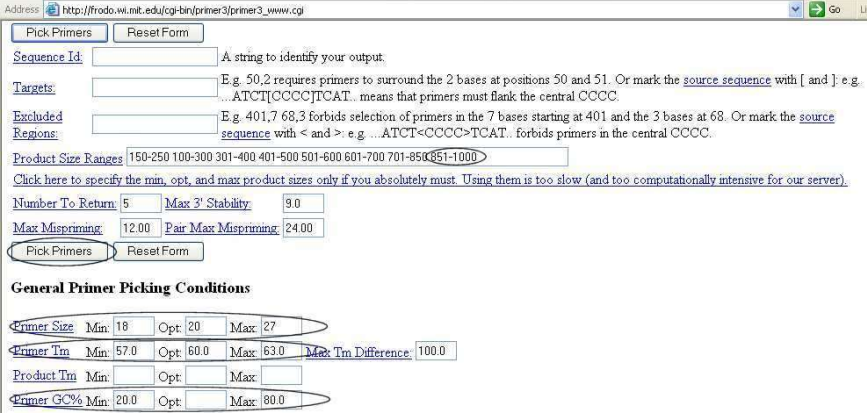
Pick left primer or use left primer below

Pick hybridization probe (internal oligo) or use oligo below

Pick right primer or use right primer below (5'→3' on opposite strand)

Sequence Id:  A string to identify your output.

- Select the product size range 850-1000 from the **select product size range** field (Delete the other size ranges). Enter the following parameters:
- **Primer size:** 27 min 29 opt 31 max
- **Primer tm:** 60 min 65opt 70 max
- **Primer GC content:** 30 min 35 opt 40 max



Product Size Ranges: 150-250 100-300 301-400 401-500 501-600 601-700 701-850 **851-1000**

Click here to specify the min, opt, and max product sizes only if you absolutely must. Using them is too slow (and too computationally intensive for our server).

Number To Return: 5 Max 3' Stability: 9.0

Max Mispriming: 12.00 Pair Max Mispriming: 24.00

**General Primer Picking Conditions**

Primer Size Min: 18 Opt: 20 Max: 27

Primer Tm Min: 57.0 Opt: 60.0 Max: 63.0 Max Tm Difference: 100.0

Product Tm Min: Opt: Max:

Primer GC% Min: 20.0 Opt: Max: 80.0

- Press the **Pick Primers** button





## INSERTION SITE PCR

Design a pair of insertion site specific primers (SMF/SMR) that flank the insertion site and amplify a fragment of approx 900bp using the parameters defined:

### Primer design features

- 29 mer  $\pm$  2
- tm 65°C  $\pm$  5
- GC% 35  $\pm$  5

You will need an additional primer specific to the transposon (dSpm or Ds)

### For SM lines Transposon specific primer (3' dSpm)

Spm32 ( 29-mer )

5'-TAC GAA TAA GAG CGT CCA TTT TAG AGT GA-3'

### For AT, ET, GT, MET, MGT, MT (GMT) lines Transposon specific primer (3' Ds)

Ds3-1 (20 mer)

5'-ACC CGA CCG GAT CGT ATC GGT-3'

## 3 PRIMER PCR REACTION

SMF, SMR & Spm32 or Ds3-1

### PCR REACTION.

5.0 $\mu$ l DNA

2.0 $\mu$ l x10 PCR BUFFER

2.5 $\mu$ l 2mM dNTP's

0.5 $\mu$ l 10 $\mu$ M SMF

0.5 $\mu$ l 10 $\mu$ M SMR

0.5 $\mu$ l 10 $\mu$ M Spm32 or Ds3-1

0.5 $\mu$ l TAQ

8.5 $\mu$ l dH<sub>2</sub>O

For the negative control the PCR reaction was carried out as above but with wild-type DNA.

A positive control is also carried out as above using wild-type DNA but used both the left and right primers with no Transposon specific primer (Spm32/Ds3-1).

### PCR CONDITIONS.

94°C for 2mins

94°C for 30sec

55°C for 45sec

72°C for 1mins

72°C for 5mins

4°C forever

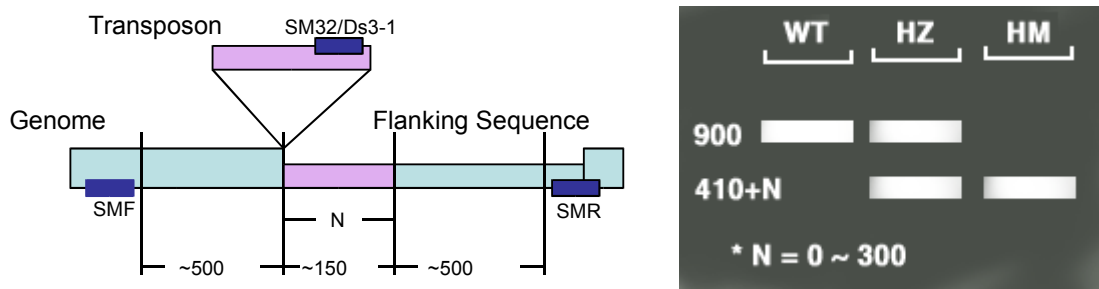
}  
} X 30 CYCLES  
}

Run out the PCR products from all reactions on a 1% agarose gel and photograph.



SMF with SMR & Spm32/Ds3-1 will generate:

1. an approx 900bp product (homozygous wild-type **WT**);
2. an approx 900bp product (wild-type) & a <500bp product (heterozygous for insertion **HZ**);
3. a single <500bp product (homozygous for insertion **HM**).



N – Transposon sequence amplified in addition to flanking sequence.

Please let me know how you get on. We are trying to collate information on confirmed insertions sites in the SM collections to define a confirmed unigene set. The JGL will update ATIdb and NASC with information on which lines have been verified (your details will be treated in confidence).

### INSERTION SITE VERIFICATION FOR T-DNA LINES

**For Salk T-DNA** lines we recommend you use the protocols from SIGnAL (<http://signal.salk.edu/tdnaprimers.html>)

**For SAIL T-DNA** lines we recommend you use the protocols from SIGnAL (<http://signal.salk.edu/tdnaprimers.html>)