



TILLING using the ABI3730

David Baker.



**[1] Overview of platform**

**[2] Organisms**

**[3] Data**

**[4] New developments**

**[5] Other considerations**

# Capillary Electrophoresis on ABI 3730



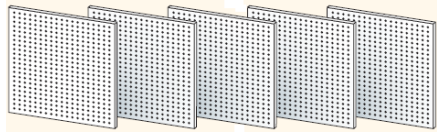
ABI 3730 & 3730xl DNA Analyzer  
provide a single HTP Capillary Electrophoresis Platform  
for combined sequencing and fragment analysis



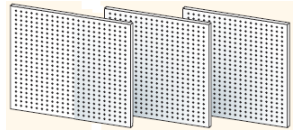
Compatible with 96 and 384 well plate formats this platform  
is readily integrated into automated sample processing with  
liquid handling robotics and integral HTP PCR

# Overview of platform

## Workflow



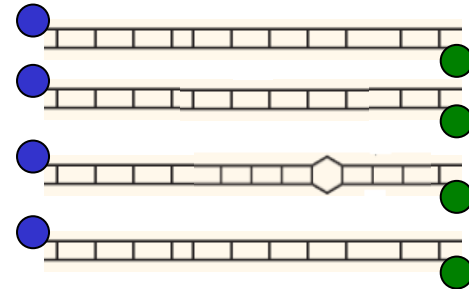
Plates of individuals  
M2 DNA



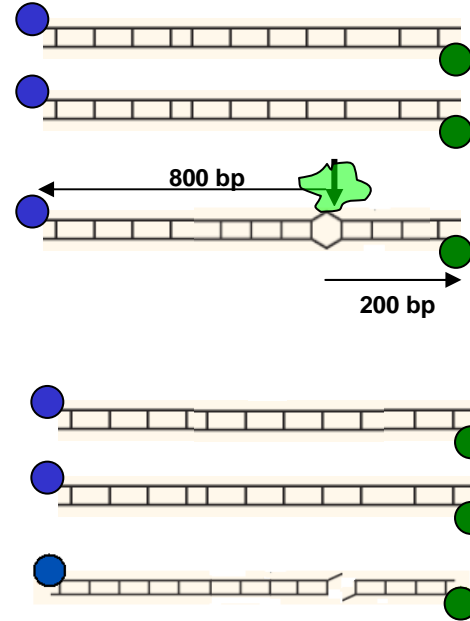
Quantification,  
normalisation  
and  
pooling 4x 8x or  
12x



PCR with  
5' labelled primers

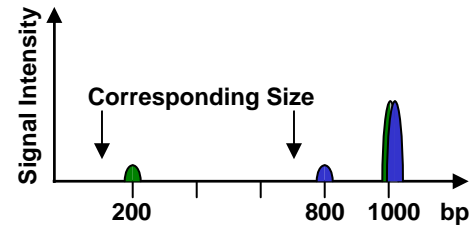


Heteroduplex  
formation

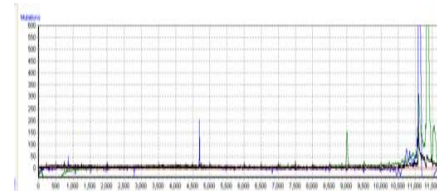


Cell Cleavage

Separation



Analysis



# 2 Day Turnover per gene region



## Primers arrive DAY 1

PCR Stage 1	4 hours	This is just for 1x 96 of an 8 pooled plate. Larger populations would take longer
Cel1 Cleavage	4.5 hours	
Cleanup	5.5 hours	However, JGL has access to three 3730's but currently uses only one which copes with all Genotyping, DNA Sequencing and TILLING that passes through
3730 Run	7 hours	
Analysis	8 hours	
PCR Stage 2	overnight	

## DAY 2

Cleanup	1 hour	With increased number of TILLING samples there is the possibility of using a single 3730 designated to run just TILLING samples
Sequencing Reactions	4 hours	
Cleanup	5 hours	
3730 run	7 hours	
Sequence analysis	8 hours	
Report of mutations		

# Identification

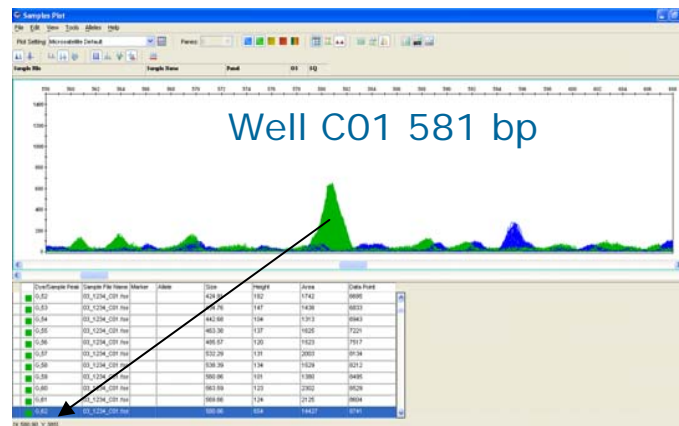
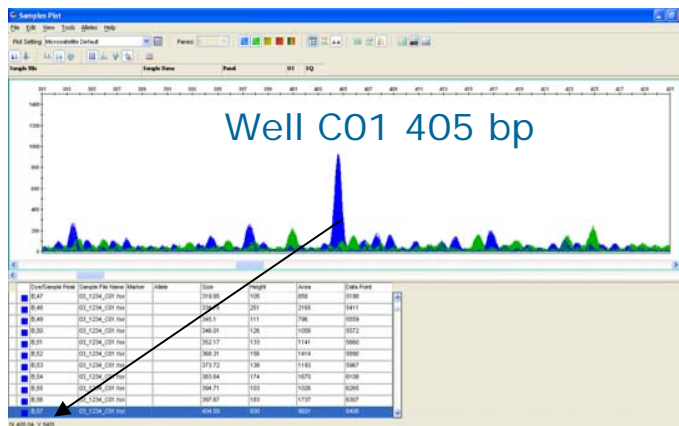
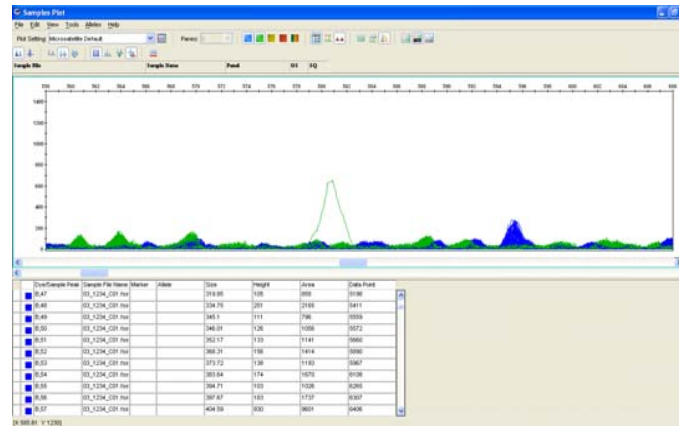
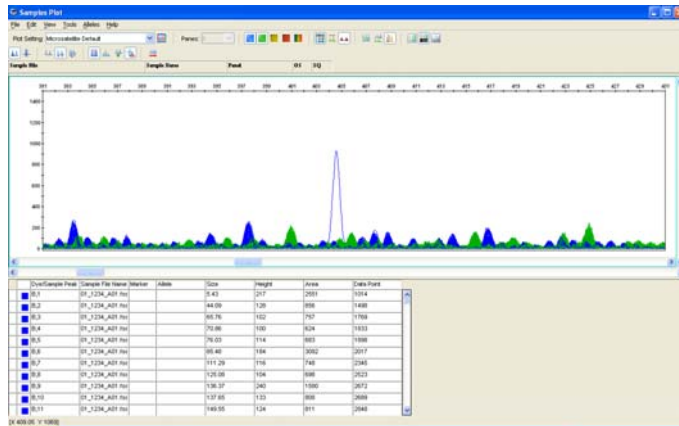
## GeneMapper v4.0 Software

[1] Select & overlay all traces & zoom in

[2] Scroll along the overlaid trace

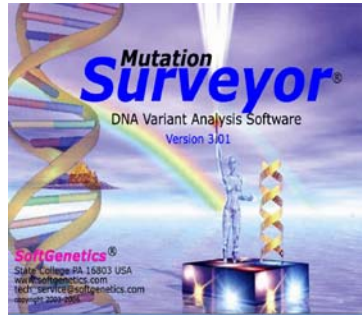
[3] Identify unique hollow peaks (no fill)

[4] Click on peak to identify pool sample



# Sequence Confirmation

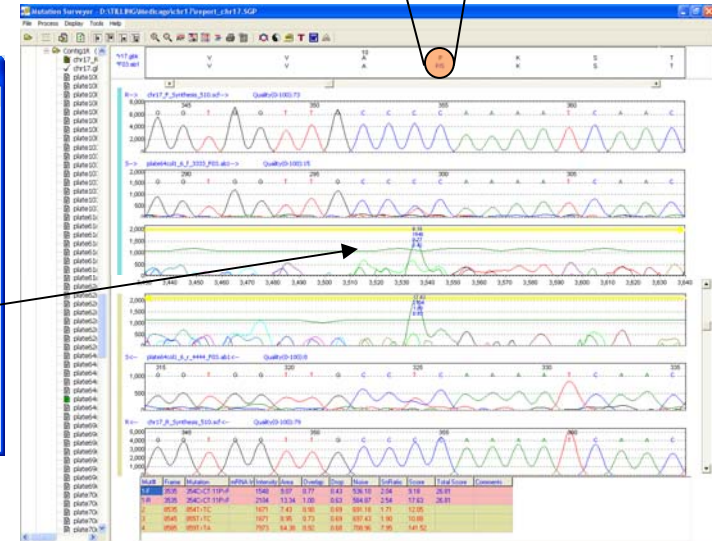
Sequence Confirmation using



Uses trace subtraction method

Software creates pseudo trace to compare sequences to, using GenBank file

No.	Sample File	Reference File	Dir	Lane	Gene	Exon	RF	Start	End	Size	Quali	Mult	Mutation1
30	plate62col10_6_f_3333_F02.ab1	chr17_F_Synthesis_510.scf	1-F		Noname-Gene1		1	82	946	865	30	0	
31	plate62col10_7_f_3333_G02.ab1	chr17_F_Synthesis_510.scf	1-F		Noname-Gene1		1	82	944	863	36	0	
32	plate62col10_8_f_3333_H02.ab1	chr17_F_Synthesis_510.scf	1-F		Noname-Gene1		1	82	944	863	39	0	
33	plate64col1_1_f_3333_A03.ab1	chr17_F_Synthesis_510.scf	1-F		Noname-Gene1		1	77	939	863	30	0	
34	plate64col1_2_f_3333_B03.ab1	chr17_F_Synthesis_510.scf	1-F		Noname-Gene1		1	82	944	863	29	0	
35	plate64col1_3_f_3333_C03.ab1	chr17_F_Synthesis_510.scf	1-F		Noname-Gene1		1	82	939	858	34	0	
36	plate64col1_4_f_3333_D03.ab1	chr17_F_Synthesis_510.scf	1-F		Noname-Gene1		1	77	944	868	27	0	
37	plate64col1_5_f_3333_E03.ab1	chr17_F_Synthesis_510.scf	1-F		Noname-Gene1		1	82	939	958	36	0	
38	plate64col1_6_f_3333_F03.ab1	chr17_F_Synthesis_510.scf	1-F		Noname-Gene1		1	71	956	886	15	1	354C>T,11P>P/S\$9
39	plate64col1_7_f_3333_G03.ab1	chr17_F_Synthesis_510.scf	1-F		Noname-Gene1		1	82	944	863	24	0	
40	plate64col1_8_f_3333_H03.ab1	chr17_F_Synthesis_510.scf	1-F		Noname-Gene1		1	82	944	863	29	0	
41	plate69col8_1_f_3333_A06.ab1	chr17_F_Synthesis_510.scf	1-F		Noname-Gene1		1	82	944	863	26	0	
42	plate69col8_2_f_3333_B06.ab1	chr17_F_Synthesis_510.scf	1-F		Noname-Gene1		1	82	944	863	30	0	
43	plate69col8_3_f_3333_C06.ab1	chr17_F_Synthesis_510.scf	1-F		Noname-Gene1		1	73	945	873	18	0	
44	plate69col8_4_f_3333_D06.ab1	chr17_F_Synthesis_510.scf	1-F		Noname-Gene1		1	82	939	858	34	0	

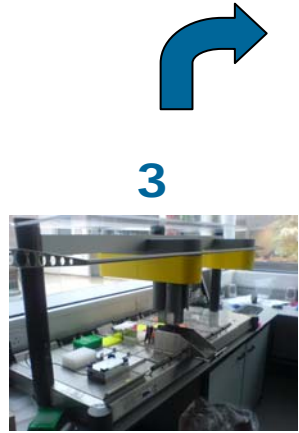
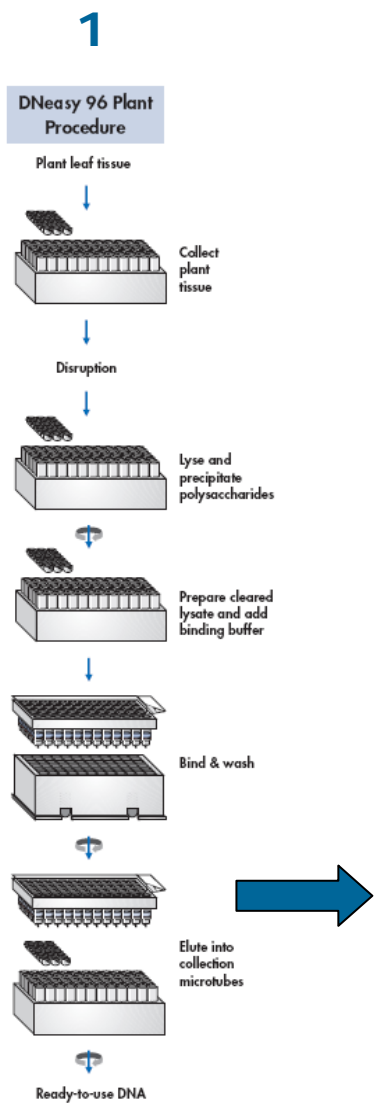


8 sequences representing candidate pool

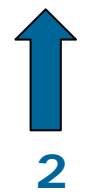
It aligns the sequences and then gives a table of all the mutations

# Preparation of Populations

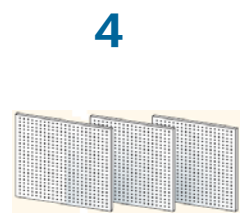
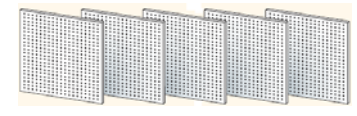
D  
N  
A  
  
e  
x  
t  
r  
a  
c  
t  
i  
o  
n



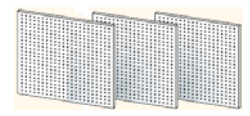
Normalisation



Picogreen assay



Pool



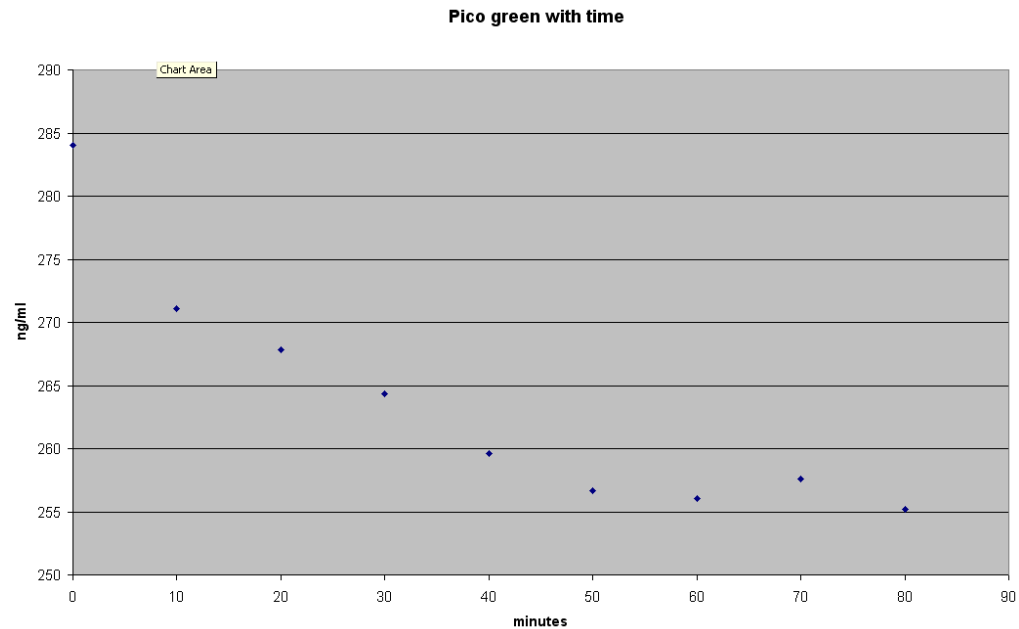
Aliquot into 5ul volume ready for stage 1 screen

It is critical to accurately normalise DNA so that each sample is equally represented within a pool

# Quantification

- normalise the individuals to 5ng/ul
- pool by hand
- dilute usually to 0.5ng/ul  
(With wheat we dilute to 1ng/ul)

10% reduction in signal over first 50 min

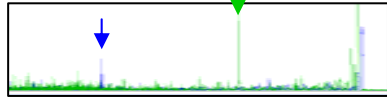


Reproducible PicoGreen  
quantification after 50 minutes

# Organisms

## TILLING using 3730 demonstrated on different organisms

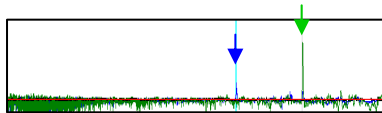
*Medicago*



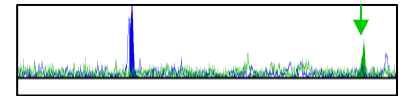
Wheat



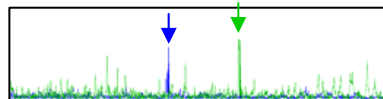
*Lotus*



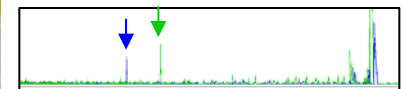
Barley



*Brassica rapa*



*Arabidopsis*  
(Adapted-TILLING)



# Data- *Medicago*



## *Medicago* (GLIP-JIC)

4300 M2 Plants EMS 2

42 Genes

485 Mutations

11.5 mutations per gene region

309 heterozygous (63.7%)

176 Homozygous (36.3 %)

21 induced stops (4.3%)

6 non C-T or  
G-A changes (1.2%)

## *Medicago* (GLIP-Dijon)

3072 M2 Plants EMS 2

20 Genes

133 Mutations

6.7 mutations per gene region  
(9.4 proportional to 4300 plants)

66\* heterozygous (69.5%)

29\* Homozygous (30.5 %)

5 induced stops (3.8%)

0 non C-T or  
G-A changes (0.0%)

\* Ignoring silent mutations

## *Arabidopsis* (ATP)

3072 M2 Plants

192 Genes

1890 Mutations

9.8 mutations per gene region  
(13.7 proportional to 4400 plants)

1276 heterozygous (67.5%)

614 Homozygous (32.5 %)

93 induced stops (4.9%)

16 non C-T or  
G-A changes (0.8%)

**ABI 3730 xl**

**Li-cor**

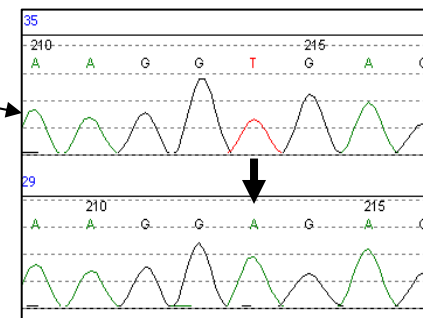
**Li-cor**

# Non C-T or G-A



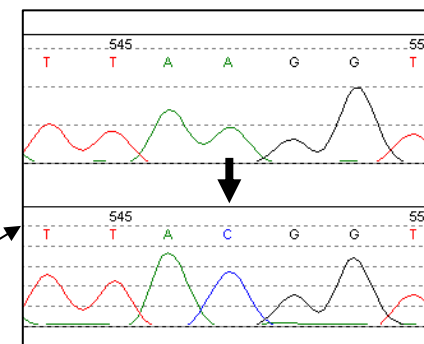
	DOSE	GENO M1	GENO M2	BoiteDNAStoc	Ligne	Colonne	PoolNor	Line	Column	Harvest M3 se	mutation and position
1	2 (0.15 %EMS)	953	5428	65	D	8	plate1	E	8		1223G>A,103G>E
2	3 (0.20 %EMS)	636	5112	62	B	4	plate1	B	4		1110T>A
3	2 (0.15 %EMS)	1572	6047	71	H	3	plate2	C	3		1243C>T,110L>L
4	2 (0.15 %EMS)	2004	6479	76	C	6	plate2	H	6		1498C>CT,195F>R/C
5	2 (0.15 %EMS)	1710	6185	73	C	7	plate2	E	7		1248G>GA,111E>E/E
6	2 (0.15 %EMS)	2576	7051	82	B	1	plate3	F	1		1402G>GA,163D>D/N
7	2 (0.15 %EMS)	2792	7267	84	B	5	plate3	H	5		1497C>CT,194P>P/P
8	2 (0.15 %EMS)	2552	7027	81	H	3	plate3	E	3		1108G>A
9	2 (0.15 %EMS)	2074	6549	77	A	3	plate3	A	3		1061G>GA
10	2 (0.15 %EMS)	2921	7396	85	C	7	plate4	A	7		1293C>CT,126L>L/L
11	2 (0.15 %EMS)	4517	8992	102	B	10	plate6	B	10		1527C>CT,204L>L/L

## 1110 T-A Silent



	DOSE	GENO M1	GENO M2	BoiteDNAStoc	Ligne	Colonne	PoolNor	Line	Column	Harvest M3 se	mutation and position
1	2 (0.15 %EMS)	861	5336	64	D	12	plate1	D	12		3798C>T,610T>I
2	2 (0.15 %EMS)	851	5326	64	D	2	plate1	D	2		3598G>A
3	2 (0.15 %EMS)	1765	6240	73	H	1	plate2	E	1		3191G>A,467K>K
4	2 (0.15 %EMS)	1959	6434	75	G	11	plate2	G	11		3739C>T,530P>R
5	2 (0.15 %EMS)	1362	5837	69	F	9	plate2	A	9		3317C>T
6	2 (0.15 %EMS)	1450	5925	70	F	1	plate2	B	1		3524C>CT
7	2 (0.15 %EMS)	2522	6997	81	E	9	plate3	E	9		3885C>CT,639T>T/I
8	2 (0.15 %EMS)	2197	6672	78	C	1	plate3	B	1		3719C>CT,584P>P/S
9	2 (0.15 %EMS)	2364	6839	80	A	3	plate3	D	3	stérile	3285C>CT,492S>S/F
10	2 (0.15 %EMS)	2076	6551	77	A	5	plate3	A	5		3496A>A,844K>T

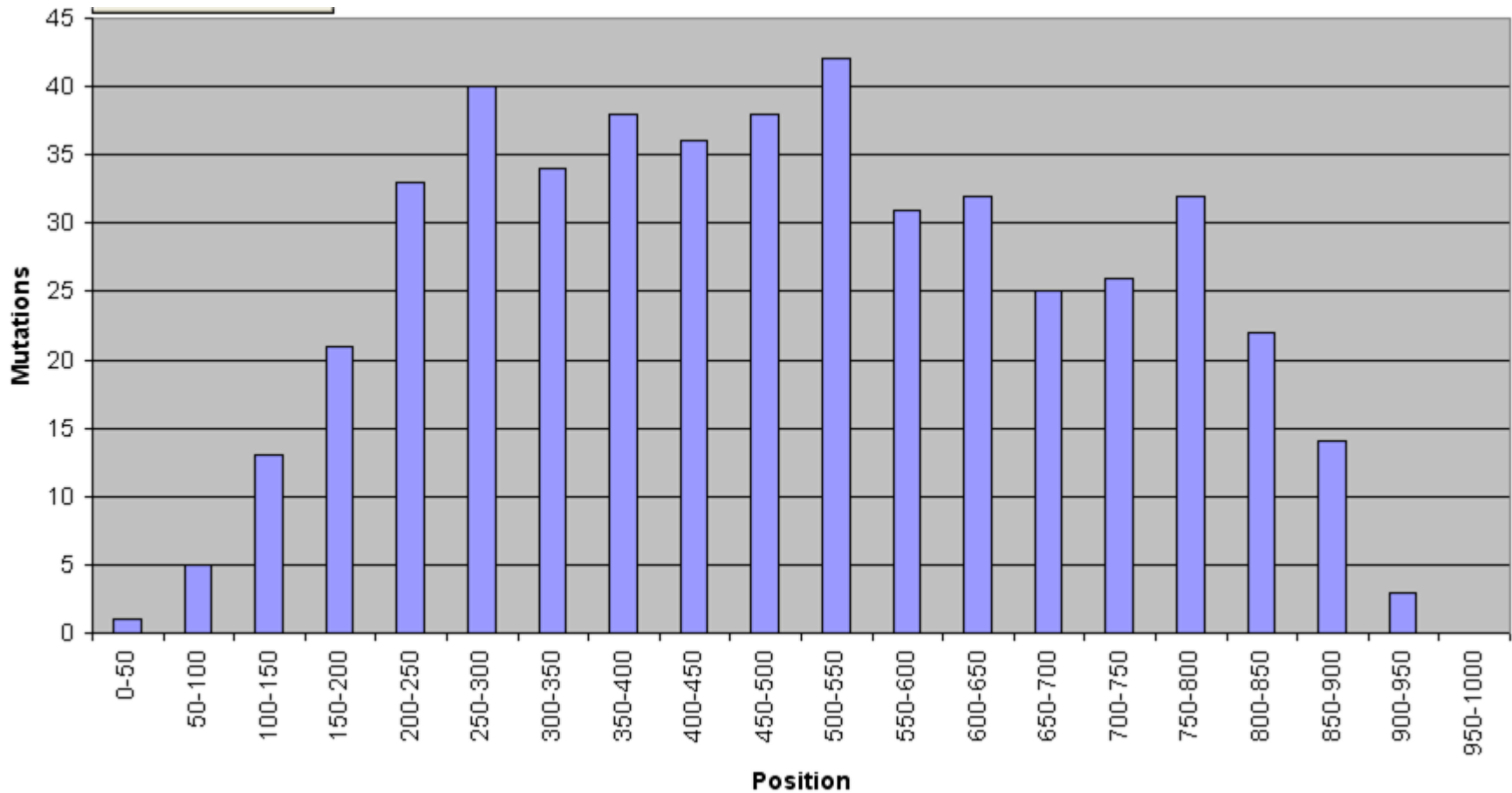
## 3496 A-C 844 K-T

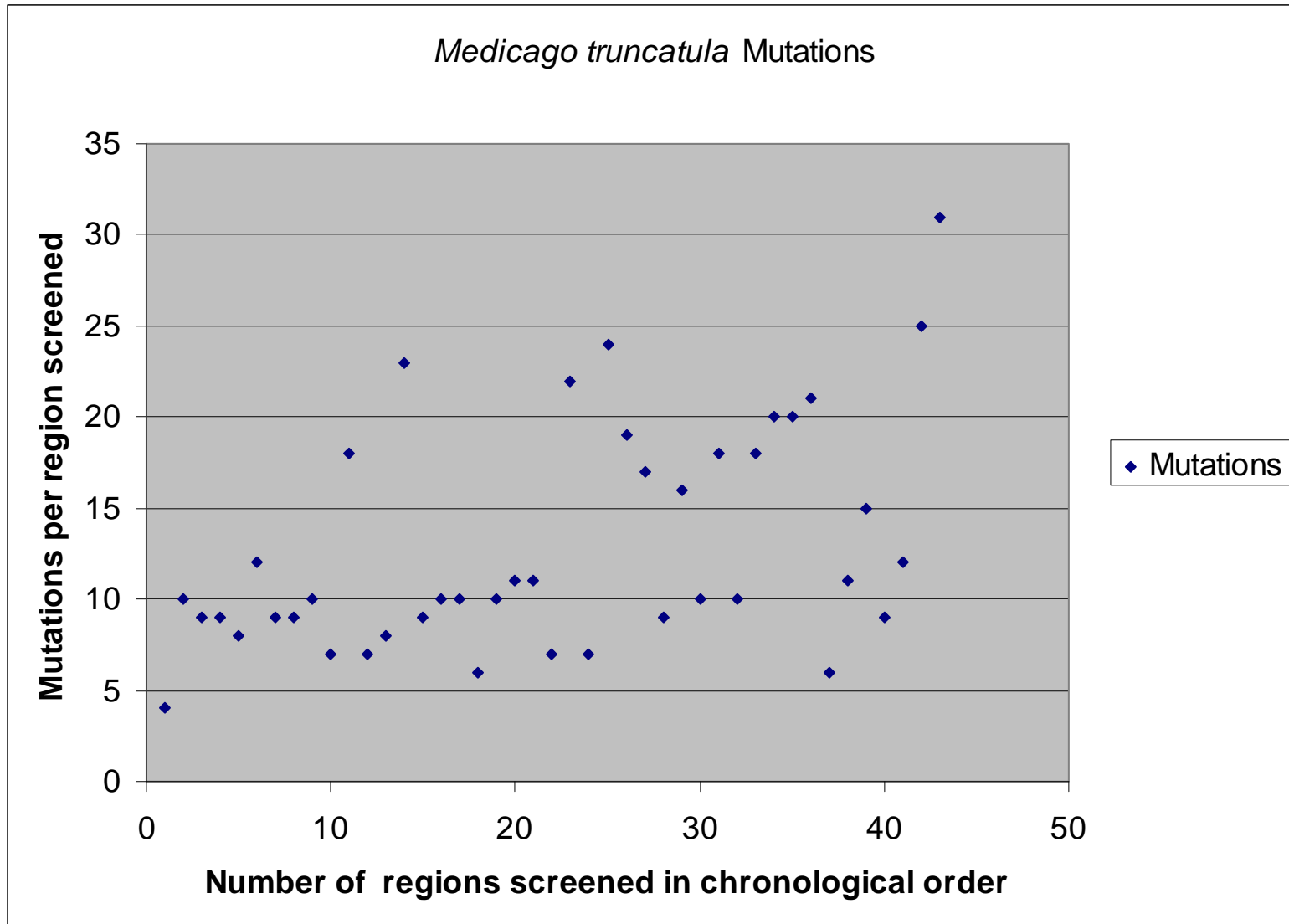


Stage 1 PCR where cleavage detected and PCR for confirmation sequencing done on independent DNA and using different taq polymerase



Normalised Distribution of Mutations



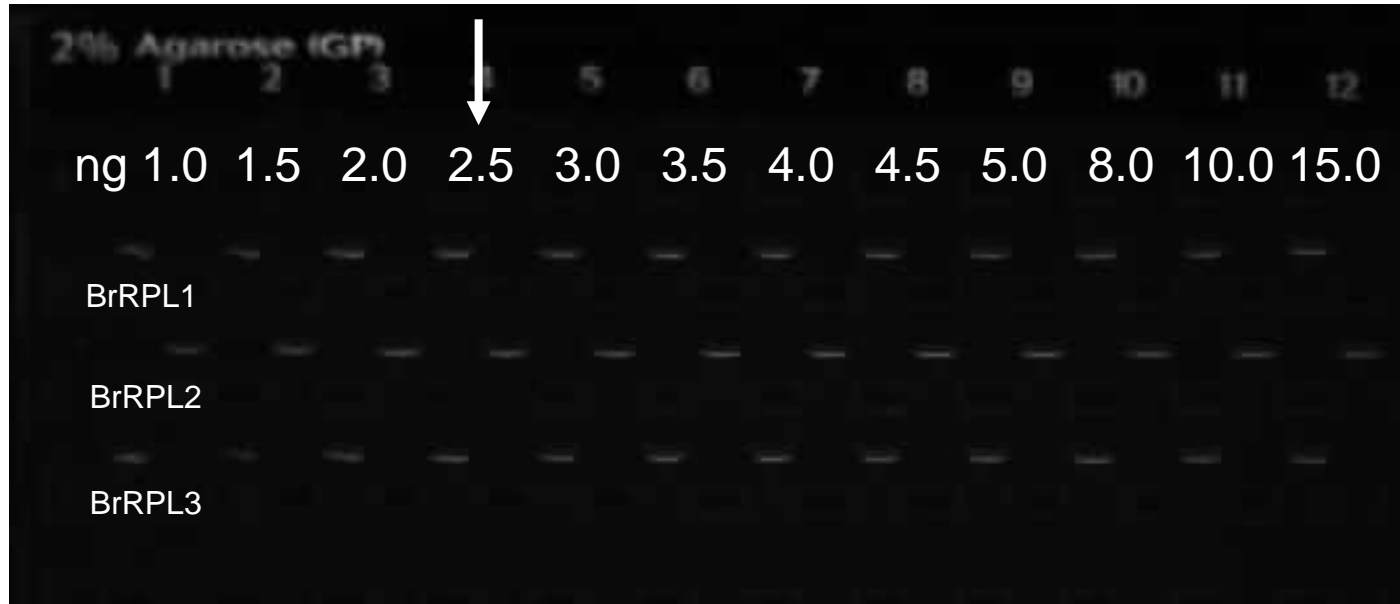


# Data- *Brassica rapa*



Due to its relative genome size compared to *Medicago* and *Arabidopsis* we tested the TIILING PCR at a range of genomic concentrations

## Medicago

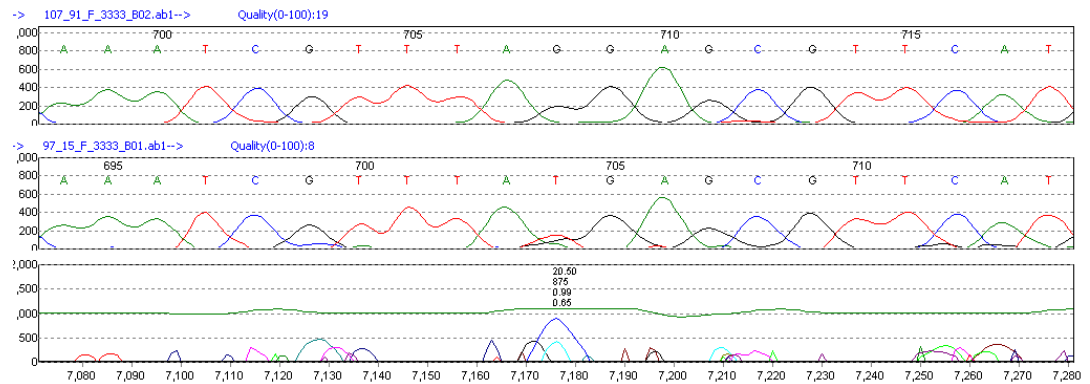
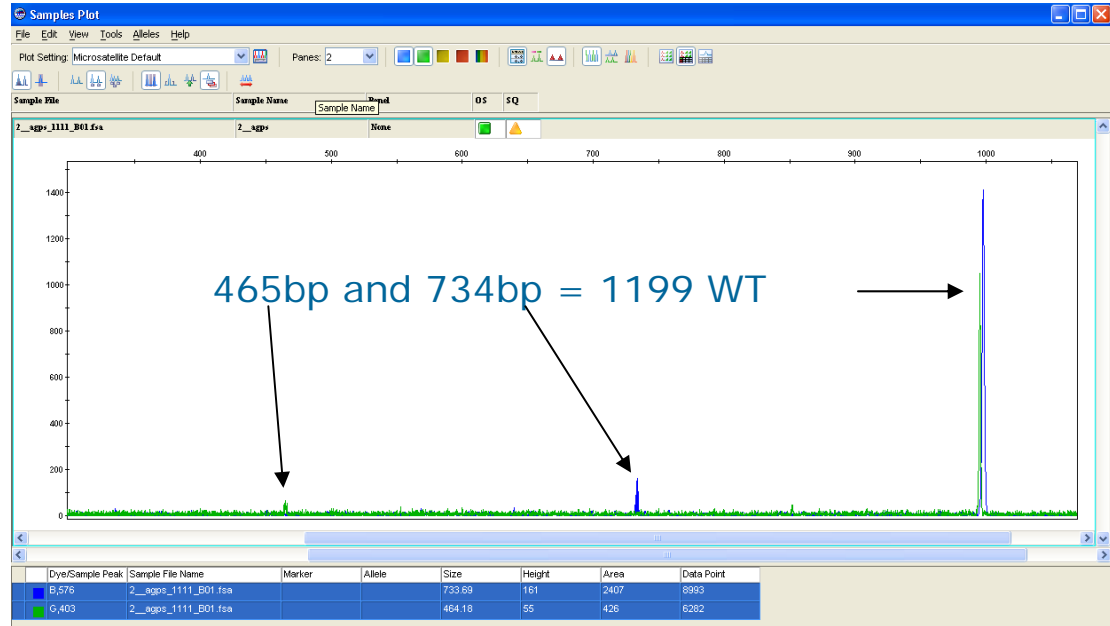


So decided to normalise and pool the *Brassica rapa* population the same as the *Medicago*



We tested to see if we could detect several known mutations in the Barley population at SCRI

We are in the process of testing the population of 4000 M2 DNA's kindly sent to us by SCRI

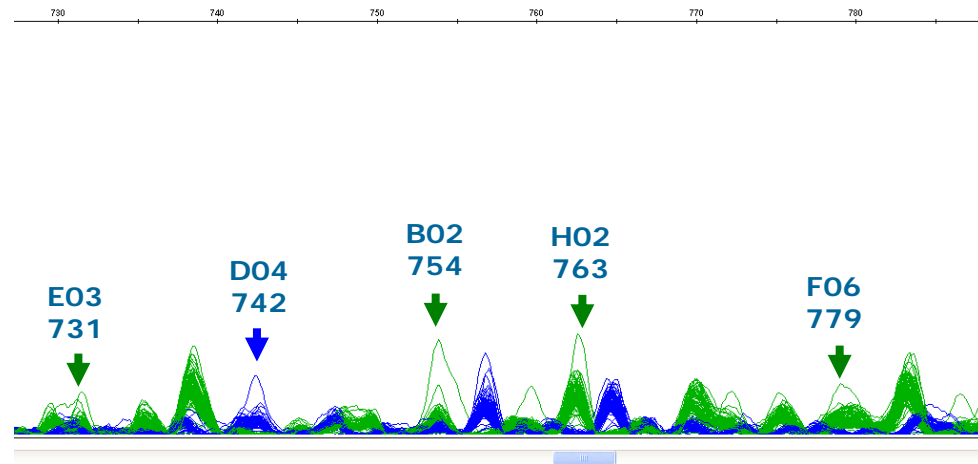
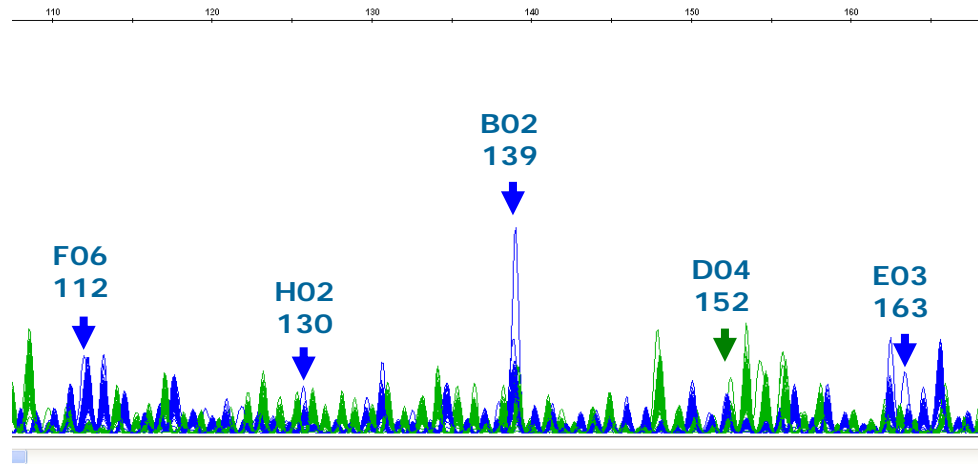




WT = 895bp

F06 112 + 779 = 891 bp  
H02 130 + 763 = 893 bp  
B02 139 + 754 = 893 bp  
D04 152 + 742 = 894 bp  
E03 163 + 731 = 894 bp

Analysis of the TILLING data showed that we had a high number of candidates with unique cleavage products adding up to the wild-type size.

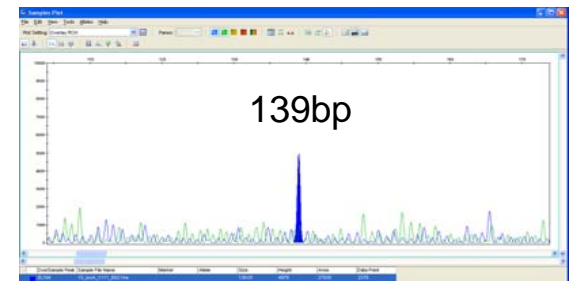
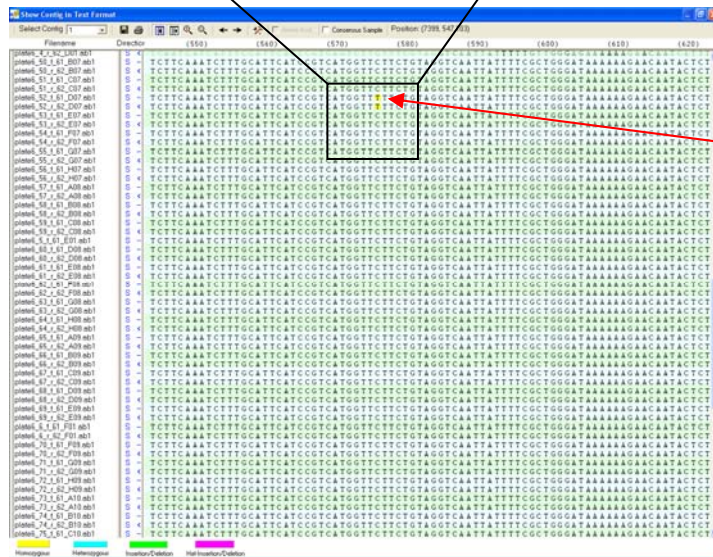


# Data- Wheat

To accurately deduce the number of missed positives, we sequenced the gene region for the entire test population (768 plants)

```
C A T G G T T C T T C T G
C A T G G T T T T C T G
C A T G G T T T T C T G
C A T G G T T C T T C T G
C A T G G T T C T T C T G
C A T G G T T C T T C T G
C A T G G T T C T T C T G
```

The results from the sequencing analysis revealed that the false-positive rate was very high and detected only one real mutation.



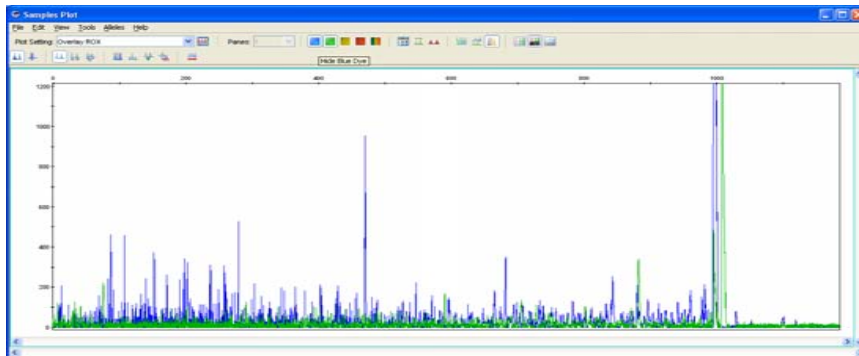
We will do more tests to determine why we are getting true cleavage at locations that do not contain SNP's

$$B02 \quad 139 + 754 = 893 \text{ bp}$$

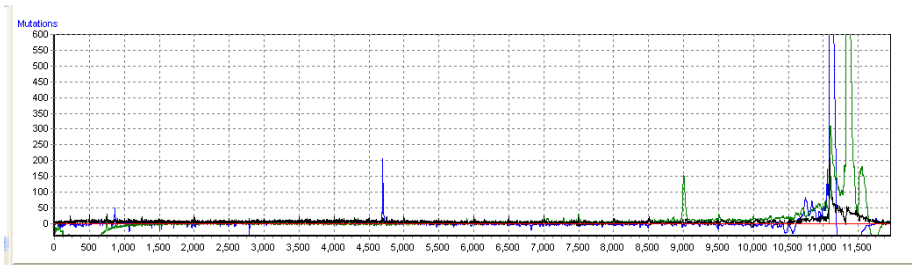


## Softgenetics Genemarker Module

Typical trace output in Genemapper



The same trace in Genemarker after average background subtraction from WT traces





## SoftGenetics Application Note

January 2008

### Automated TILLING® Analysis of Fluorescent Electrophoresis Data with GeneMarker® Kevin LeVan, David Hulce, Ni Shouyong, Wan Ning, CS Jonathan Liu

Can also analyse Li-cor gel and convert into ABI-like traces using JelMarker

#### Introduction

With the 2007 Science breakthrough of the year going to Human Genetic Variation (1), the importance of methods for studying variations are at an all-time high. The techniques of Targeted Induced Local Lesions in Genomes, or TILLING and EcoTILLING have been widely used since 2000 to detect Single Nucleotide Polymorphisms (SNPs) in organisms ranging from Arabidopsis (2) to zebrafish (3). The test samples may be experimentally mutagenized (ethylmethanesulfonate, radiation, etc) or from natural populations or derived from tumors or diseased tissues.

Briefly, the genes of interest are identified with gene-specific primers and PCR amplified. The amplicon's primers are labeled with two fluorescent dyes – the forward is often labeled with FAM-blue and the reverse primer is often labeled with HEX-green. The samples are mixed so heteroduplexes can be formed. The hybridized fragments are cleaved at the heteroduplex site by CEL I or Surveyor™ Nuclease, generating multiple pairs of fragments of complementary length and dye color (4). The denatured samples can be run through gel electrophoresis or mixed with an internal size standard and run through capillary electrophoresis. SNPs will yield two peaks of different color and the sum of the sizes will equal the amplicon length (see Figure 1).

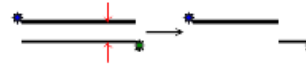


Figure 1: Double-stranded DNA with different colored dyes at each 5' end will be cleaved at the heteroduplex indicated by the red arrows, yielding two fragments of different color and having sizes that total the amplicon length.

Visual inspection, with the aid of GelBuddy® software, is a common method for determining all locations of point mutations that have been cleaved at the heteroduplex site (5). This leaves some ambiguity due to the efficiency of cleavage and level of noise. GeneMarker contains an application that has been optimized for detecting SNPs generated by the TILLING technique. By using the Internal size standard to align each capillary and generating a reference trace from all of the samples in the run, the reference can be subtracted from each individual sample trace yielding a plot highlighting the SNPs. Additionally, a table shows what the expected size of the complementary fragment should be to determine if each peak is a true variation.

#### Procedure:

1. Launch GeneMarker software.
2. Click Open Files icon, Click Add and select data (up to 1000 lanes maybe added)
3. Click open and then OK
4. Select Tilling Analysis from Applications menu
5. Select Size Standard, enter Data, Process options and select data size range including whole fragment of amplicon.
6. Click OK to process data

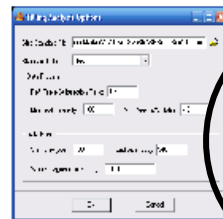


Figure 2: Tilling Options

#### Results:

When using GeneMarker's Tilling Application module, the size standard peaks are identified and all sample peaks are aligned. The peaks are smoothed, baseline is subtracted, and lane intensities are normalized. Low quality data is automatically rejected. A synthetic reference trace (Synthetic Control Sample) is constructed using median peak intensities from all of the high quality traces. This reference is subtracted from each sample trace generating a Mutation Chart that automatically identifies the sample's variations as shown in Figure 3, with the amplicon length of 1049bps.

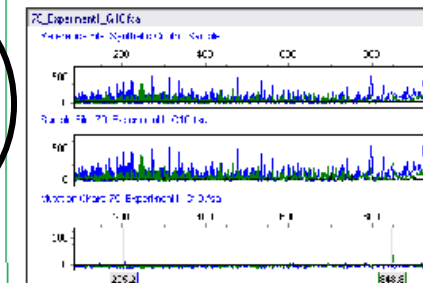
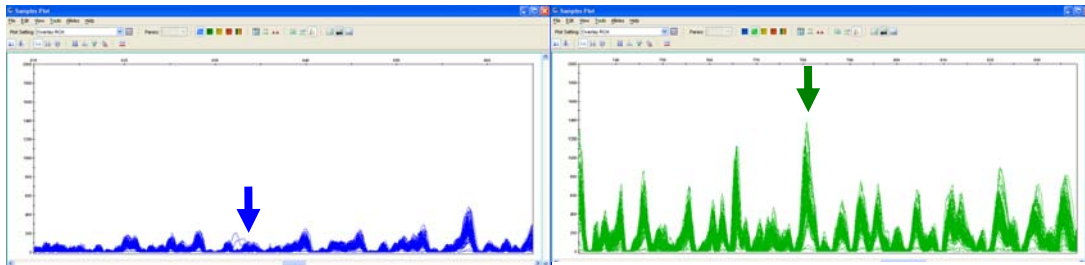


Figure 3: The top panel shows the Synthetic Control Sample obtained from the median intensity after peak alignment. The middle panel shows sample 79. The bottom panel shows the Mutation Chart, generated by subtracting the reference from the sample, identifying individual variations. A blue peak at 205.2bps and a green peak at 848.8bps have been automatically identified. The original amplicon size is 1049bps.

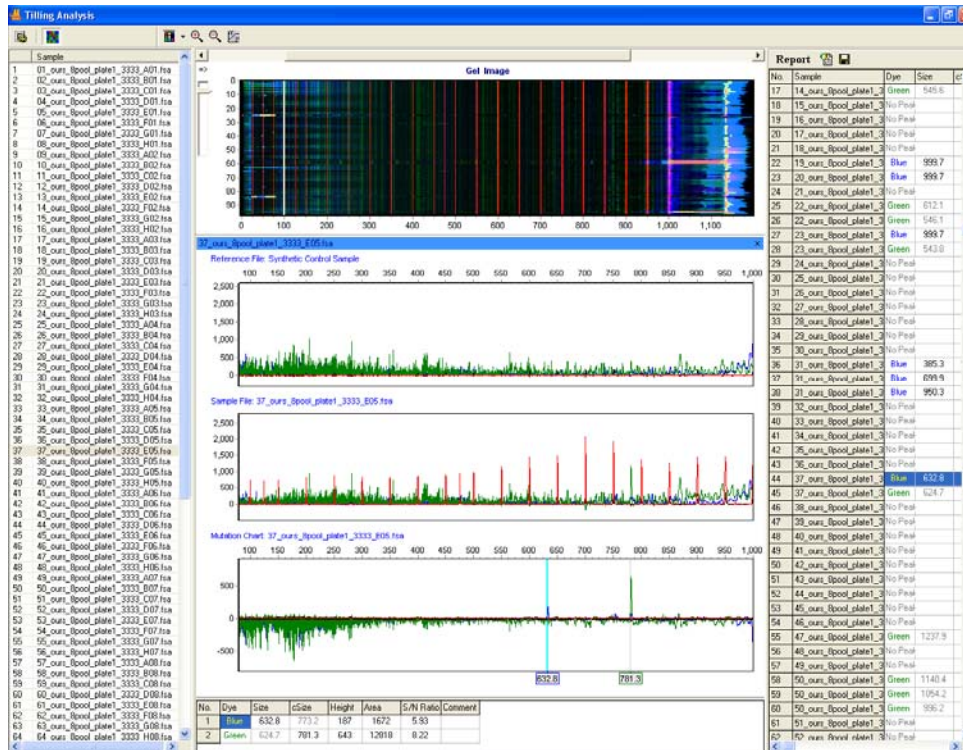
# Genemapper vs Genemarker



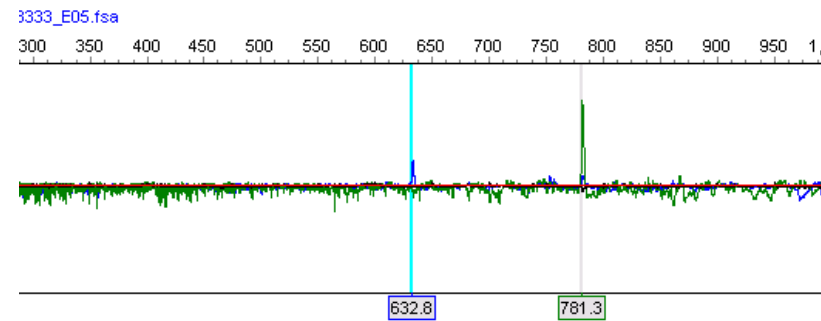
Subtracting the average background signal from wild-type traces emphasises the cleavage products in Genemarker 1.71

Lotus mutation not detected analysing 3730 traces in Genemapper 4.0

Report shows possible candidates

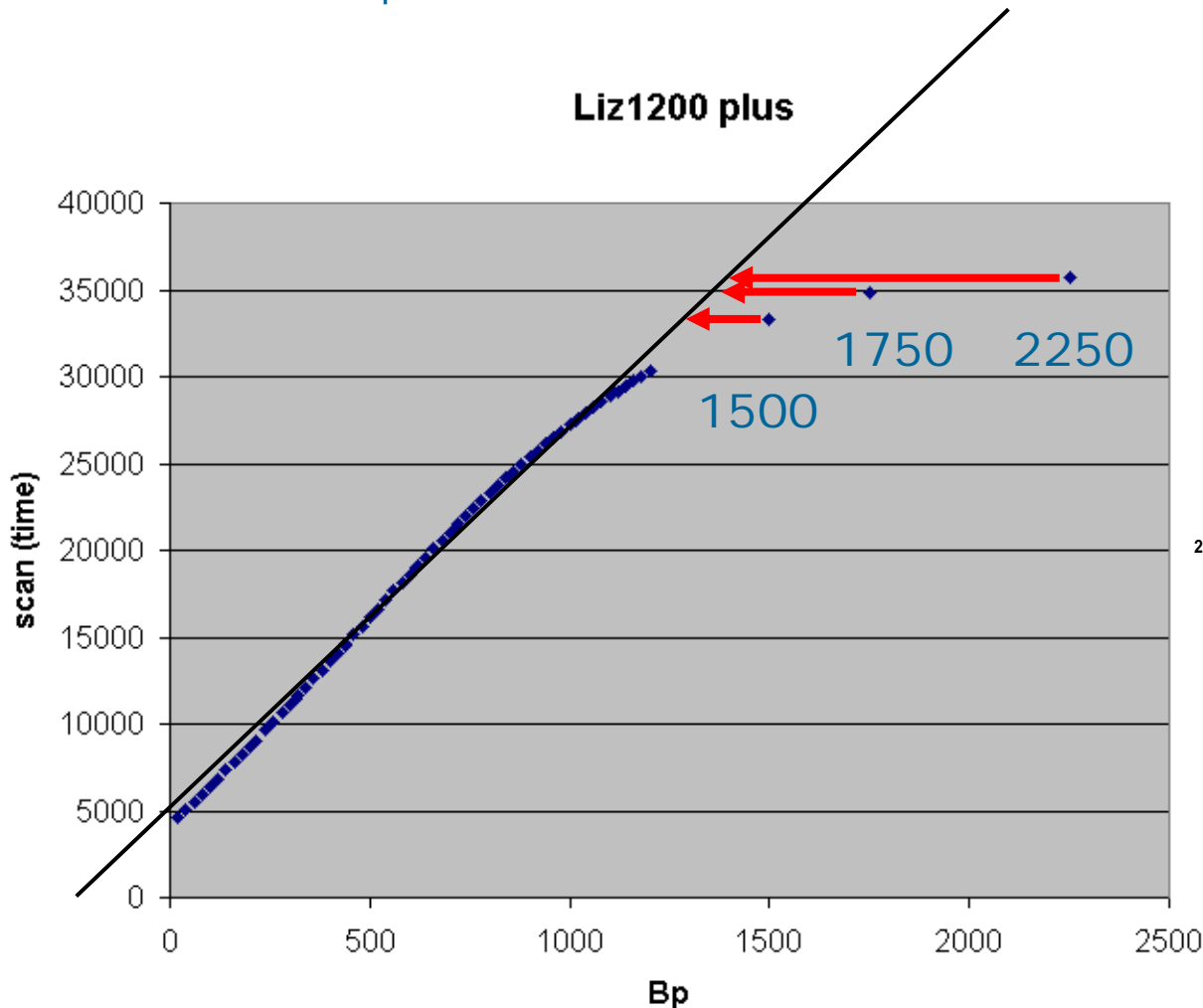


82	35_ours_8pool_plate1_3	Green	692.7	720.3	77	573
83	36_ours_8pool_plate1_3	Green	684.8	728.2	90	668
84	37_ours_8pool_plate1_3	Blue	632.8	780.2	185	1648
85	37_ours_8pool_plate1_3	Green	631.7	781.3	648	12956
86	37_ours_8pool_plate1_3	Green	546.0	867.0	104	1848
87	38_ours_8pool_plate1_3	No Peak				
88	39_ours_8pool_plate1_3	No Peak				

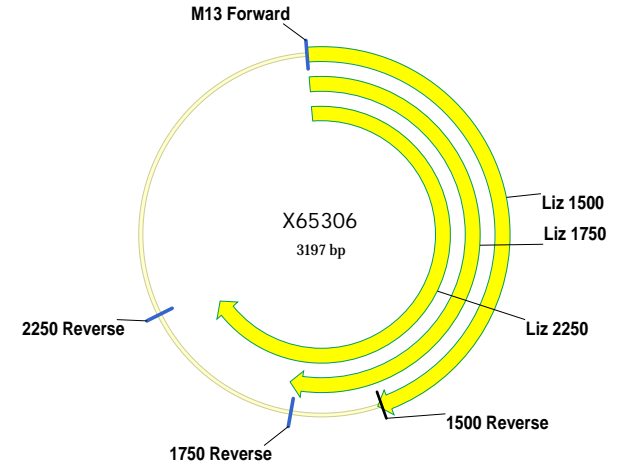


# New Developments – Extended Ladder

ABI have a 1200bp Liz ladder but we want to size up to 1.5kb  
We mixed the new products with the available ladder



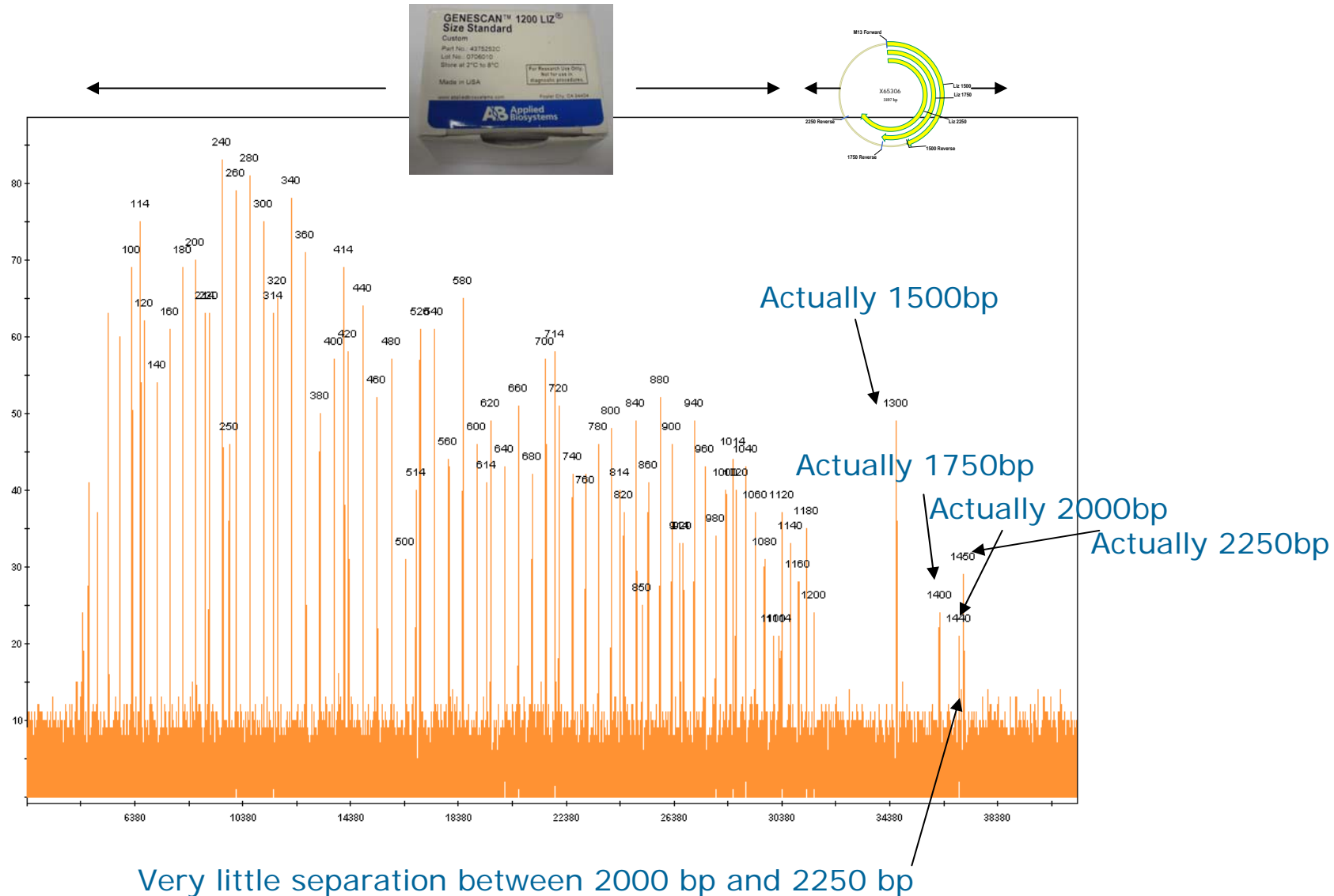
Using a specially produced 5' labelled Liz M13 primer PCR products were designed with various fragment lengths off pGEM-3Zf(+) of above 1.2kb



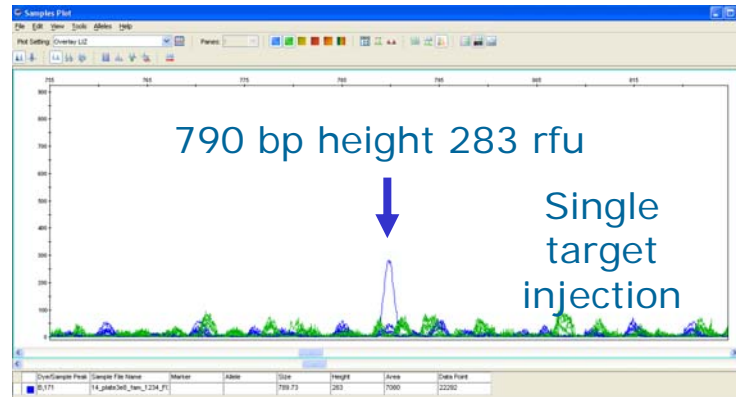
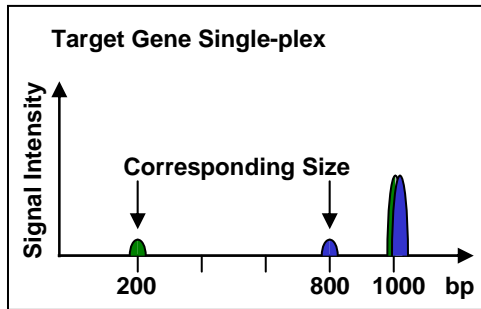
Sizing algorithm needs linearity to pass sizing within Genemapper 4.0

Developed also for the possibility of transferring de-TILLING onto 3730

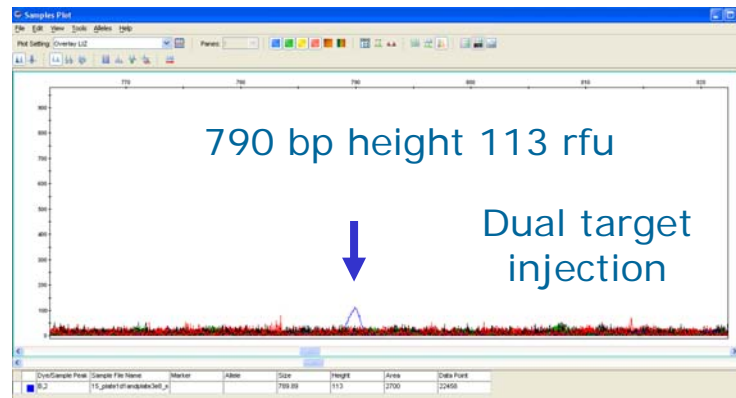
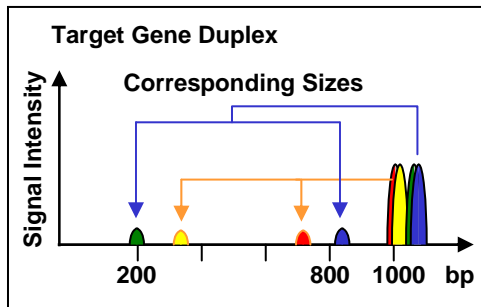
# New Developments - Ladder



# Multiple Dye Loading

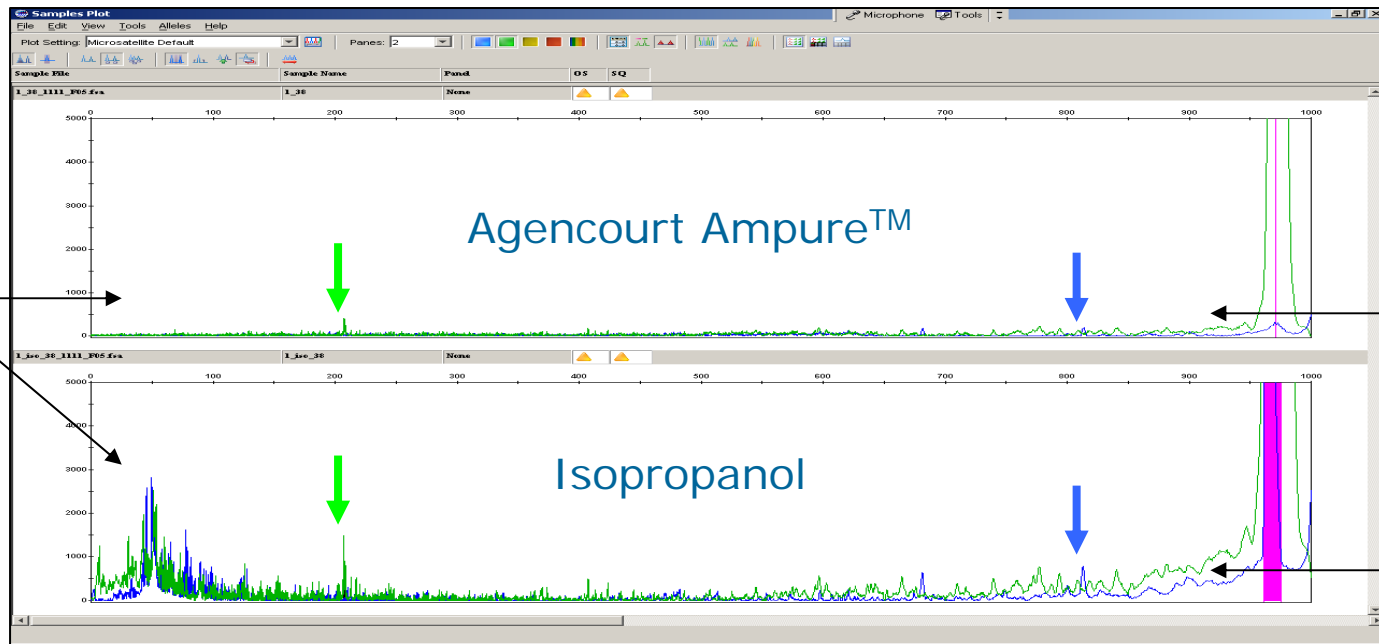


Injecting two targets in a single capillary causes less cleavage product to be injected



Although running costs would be reduced at present we only run a single target in a single capillary

# Post Cel1 Cleanup



Cleanup is an essential step to remove unwanted non-specific cleavage products that obscure specific cleavage

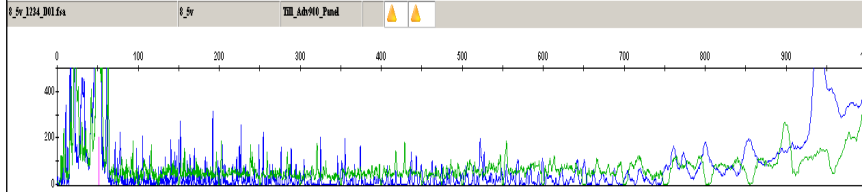
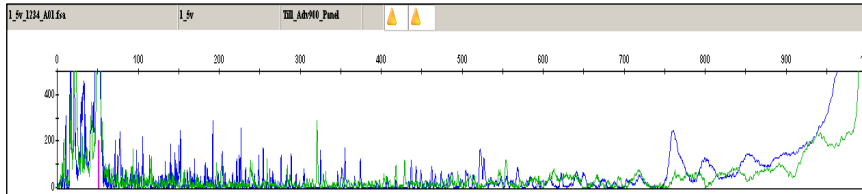
Both magnetic bead and isopropanol methods are effective

Magnetic bead method more expensive and less reproducible although the method is easily automatable on a liquid handling robot

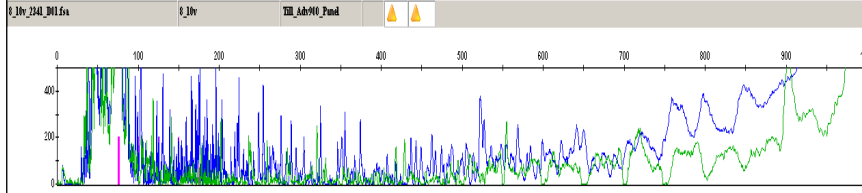
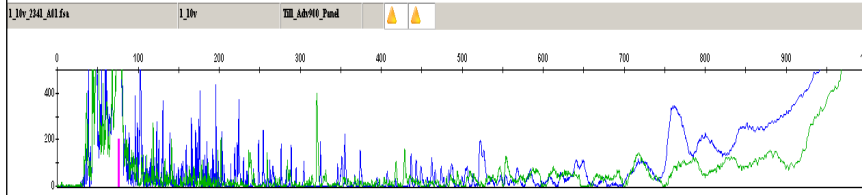
# Effect of Increased Injection Voltage



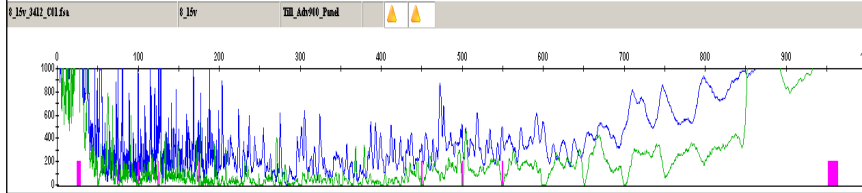
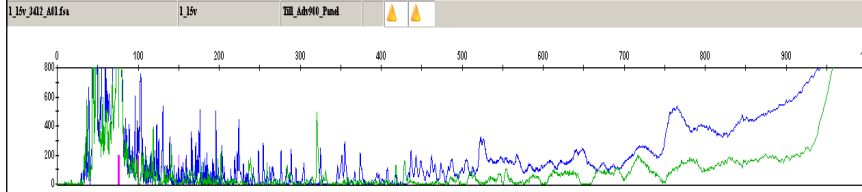
5.0 kV



10.0 kV



15.0 kV



An increase in the injection voltage increases the signal

However this causes a dramatic loss of resolution in the upper size ranges

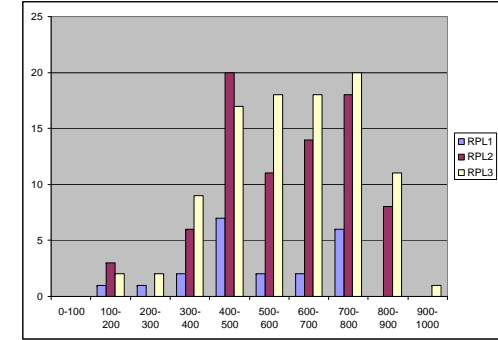
# Repetitive Regions



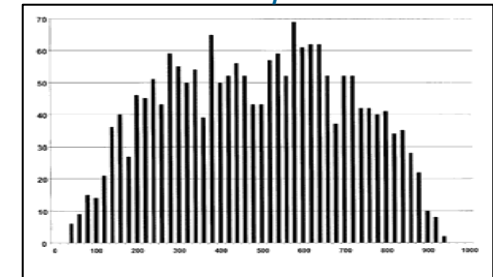
## RPL2 - 11,13bp region

1	CGGCACGTAA	AACTTCTTCA	TCATTTAAAT	GGATTAAGCT	TTAACTATAG	CTGGTAATAA	CATCAGATAA	TGAGCCTTGA	CCAAAACGTG	AATGCTGAAT
101	AAATTCTTCC	GCACAAACA	TTACAGGCTC	ATGAAGGCC	AGATATTTTA	GATGTTCTTC	TTTTTTCGAA	CTAGTAAATA	ATAGTACTAT	GTAGATTATA
201	TGATGGATAG	TGTTTCAAAA	AAAAGTGTTC	AAAAAAAAAA	AAGATTATAT	GATGGATAAT	AATATACACA	CTTGTCCGGC	CTTCAAAAAG	AAAGCAAAA
301	AAGCCTCTCT	CTCTCTCTCT	CTCTCTCTCT	CTTAACTCTC	CGTTTCACCC	GTCACCATTA	TCTCTTACAC	TGTTACTCAC	TTGTCTCTCT	CTCTCTCTCT
401	CCTCCTGCAT	CTCGAAATCA	TCCTTTGCCA	TGTCGATGC	ATACGAGCCA	TACCATGTTC	TCCAACATAG	CCGACGAGAT	AAACTTCGGA	TCCCTTCTCT
501	CGATTCCAC	TTTCACTTTC	ACCACCTTCC	TCCCTCTCTC	TCTTCTCCGC	CGCGCGCGCG	CGTGTTCCTC	ATCGCCGACT	CTGATTTTCT	CGCAGCCGGT
601	GGATTCCATT	CAAAACAATA	CAATACCCCT	AGTACAGTA	ACTTCATGGG	TTTTCTCGGT	GGGCAATCTT	CCTCTTCACT	CACCGCCGTC	CGCCGTCGCT
701	GAGATCATT	CTTCAACGCC	CTTCAACGCC	CTTCAACGCC	CTTCAACGCC	CTTCAACGCC	CTTCAACGCC	CTTCAACGCC	CTTCAACGCC	CTTCAACGCC
801	CGATCTCGTC	GTTCGCCGCG	TTGTTACTCT	CGGATTCGTG	AGATCCGCG	CTGAAGCCGC	CGCTGCCGCG	GTCACCGTAG	TTTCAAGTAG	CTCCGGTCTC
901	CTCGGCCGCT	TCACCTGGTA	CGCATCCATT	CTGAAGGAT	CAAAAGTCTT	GAAAACAGCA	CAGATGCTTC	TTGATGATT	TTGATGATT	AGTCCGCCGG
1001	TTTACACCGA	ACAAATCGTC	GACGACGAG	ATGACGATTC	CTCTTTGCTT	TTTGACCCAA	CGATTGATAA	TCTCTGCGGC	GTTTCTGAGC	CGCGCGTAGG
1101	AGAAAATGGC	AAG	TCTTTTACCG	TTC						

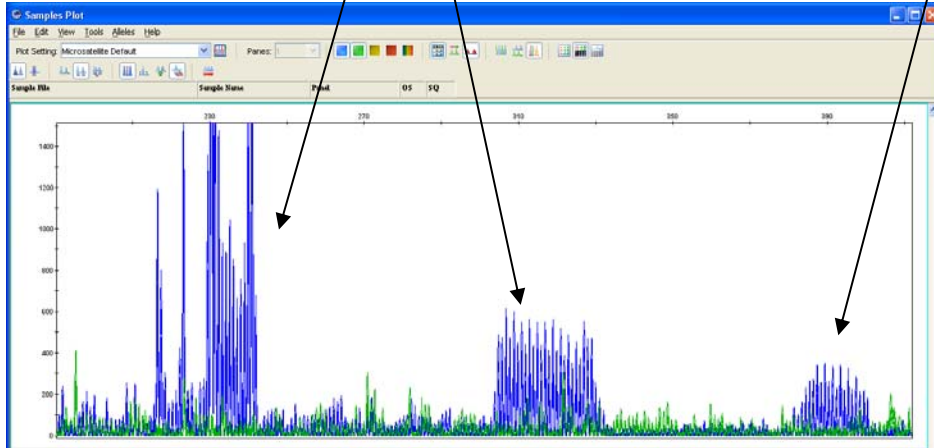
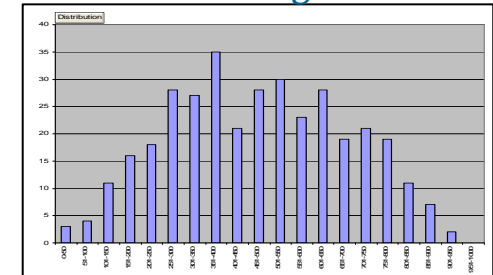
## Normalised distribution *Brassica rapa* BRL1,2, and 3



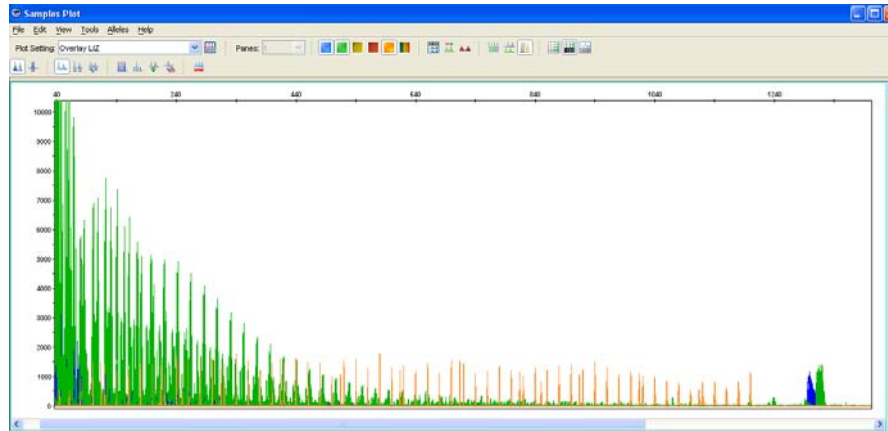
## *Arabidopsis*



## *Medicago*

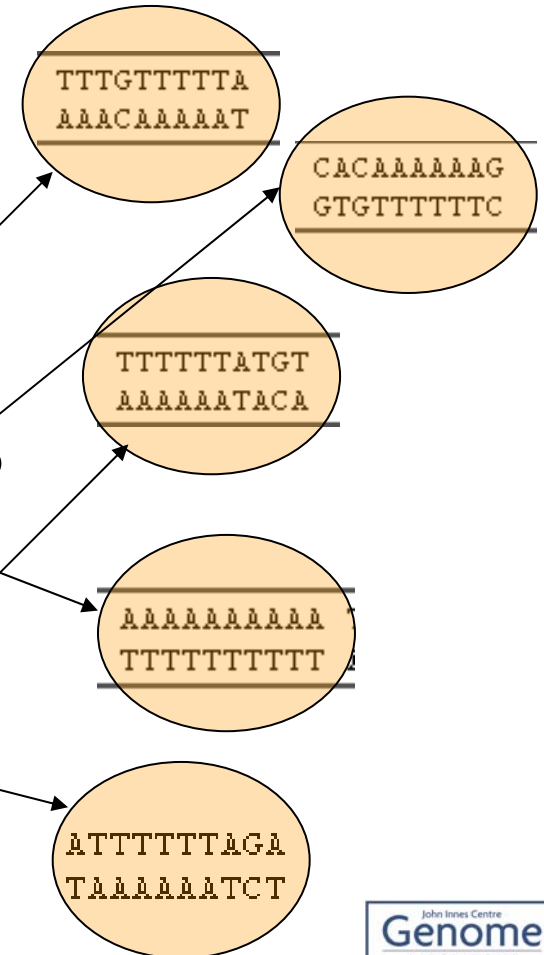


# Region Considerations



Non specific cleavage reduces the number of intact copies to cleave with genuine heteroduplex mismatch

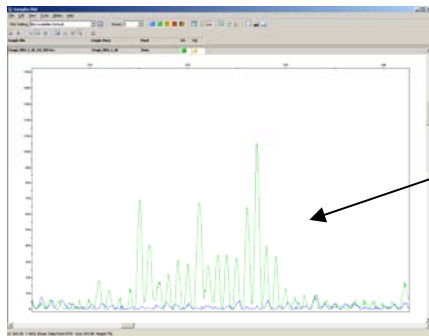
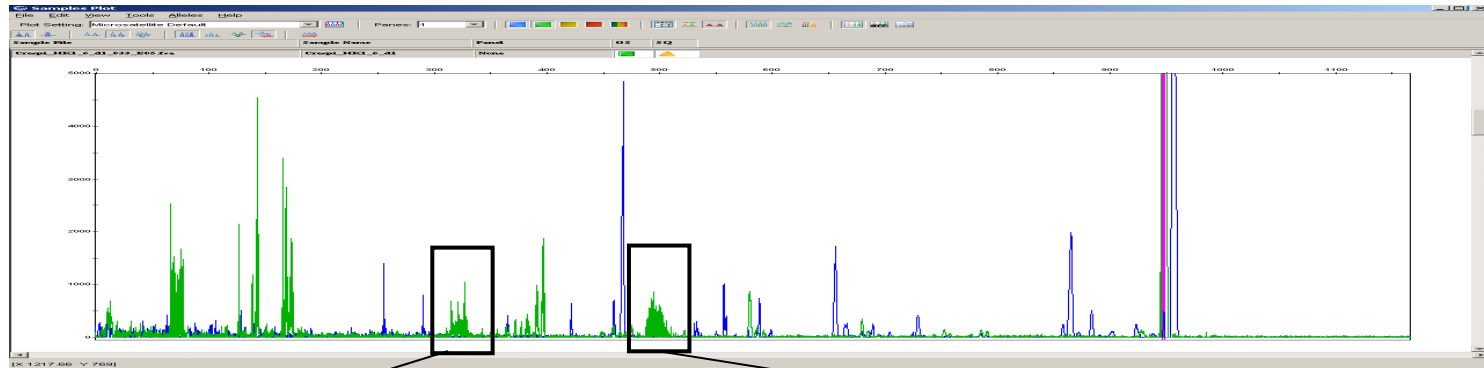
5201	GCTATTTTGG TAGGGTTTTT TTTTACTGG GAACAATFCA TTGCAAACT CTTCACATG CATCTGGATC CTCTATTGCA CCATGCGCG GTTTGCAAGG CGATAAAAAC ATCCCAAAG AAAAATGACC CTGTGTTAAGT AAACGTTTGA GAAGGTGTAC GTAGACCTAG GAGATAACGT GGTACGTCC CAACGGTCTT
5301	CGGTTGTATC TAACGGCCCT GCAGCCCTCT GTACTTTGCA GTAATTTATG GGGATCGAAA CTCAAATGAT TTAACAATCA TTTTGGCTC TTTGTTTTTA GGCAACATAG ATTGCCGGGA CCGCGGGAGA CATGAAACGT CATTAAATAT CCTAGCTTT GAGTTACTTA AATTGTTAGT AAAACGGCA AAAACAAAAAT
5401	CAATGAACAC AATATCGATA TGAAGCTATT TTCCTATAGT GTTTGTTTTT TACTGGGAAC AATTCATTTG CTAATAGTA TTCATCAATT ACCTTTATGT GTTACTTGTG TTATAGCTAT ACTTCGATAA AAGGATATCA CAAAACAAAA ATGACCCCTTG TTAAGTAAAC GATTTATCAT AAGTATGAA TGAATATAC
5501	CTAGATGTTT CAAAATCAAT AATATTATGG AAAAAAATAA TAACTATAT ACCTTATATC ATCTTCAATA TGGAACTTT GTGGAAATG ATCTGGACCC GATCTACAAA TTATAGTTTA TTATAATAC TTTTTTTTTT AATGATATA TGGAAATATAG TGAAGTATAG TACCTGTAAA CACCCCTTAC TAGACCTGGG
5601	TTTGAAGTGG AAAGAGCTGC AGAGTTTGGG GCAACAGCTT GACACTTGGC TAAAGCGAAT CCGAACAGA AAGGTAAGAA TGTTTATTCT TTAACAAGAA AAACTCACAC TTCTCGACG TCTCAAACCT CGTTGTGCAA CTGTGAAGCG ATTCGGCTTA GGCTTGTCTT TTCCATTCTT ACAATAAGA AATTGTTCTT
5701	AAAAGAAGAG AATAAATGAG TTAACATTTG ATTTTTCAGC TGAGGATCAA TTAACGTGAT AATATGATTT TTGCAGAAC AAGTTATGAA TCAATCCATC TTTTCTCTC TTATTTACTC AATTGTAAC TAAAAAGTCG ACTCCTAGTT AATTGACTAG TTTACTTAAA AACGCTTGG TTCAATACTT AGTTAGCTAG
5801	TCAGAGCTGC ATAAAAGGTT AAGATATTTA TGGTATAGAT TTGATTTAAT TTTTCTAAA TTGAATACTT ACAATAGTA CATTGAATG CACAAAAAAG AGCTCGACG TATTTCCCA TTCTATAAAT ACAATATCTA AACTAAAATTA AAAAGATTTT AACCTATGAA TGTATTTCAT GTAACCTAC GTGTTTTTTC
5901	TGACCGATC CAGAAATCT TTGATATACT TTTTGGAAAT GAAATTCATA TGCAAATGCG AATGACATG TAAAGCTAAG TGATATATAT AACAAAAATG ACTGGTCATA GTCTTATAGA AACTATATGA AAAACCTTTA CTTAAAGTAT ACGTTTAAAG TTACTGTACA ATTTCCGATC ACTATATATA TTGTTTTAAC
6001	TTTGAATCT CACAAGCAAT TAAATCATAT GCCATTTGAG TATTATCTTA GATATCATT GTGAAGTGT ACTAGTTTTT ATATAAATGG AAGCAAGATT AAACTATGAA GCTTGGCTTA ATTTAGTATA CGGTAAACCT ATAATAGAAT CTATAGTAAA CACTTCACAT TGATCAAAG TATATTTACC TTGTTCTTAA
6101	TTTACATAA ATTTTITAGA AATATAAATA TAAATATAAA ACTTTGCTAT GTATATGACA GCGAAGGACA TTGCAGGAG AAAACAACAA GCTAGCAAAG AAAGTGTAT TAAAAATCTT ATATATTTAT ATTTATATTT TGAACCGATA CATATACTGT CCGTTCCTGT AACGCTCCTG TTTTGTGTT CGATGGTCTG
6201	GTAGAATCCA CATTTCTGFC TAACAATFCA GCATACAAAT AGGGTTTCAA ATAAAGTCT GCGATCGCAA COTGAOCTGC ACACAACAAC TTTTTTATGT CATCTTAAAT GTAAGAGAG ATTGTAAGT GGTATTTTTA TCCAAAGTT TATTTCAGA CGCTAGCGTT GCACCTGACG TTGTTGTTG AAAAATAACA
6301	TGTCAAAACC GCAATGCGAC TGCAATTGAG GCCCATTGG CCGATTTTTA TTGTGATCT CCGTAATATC AAGAATCGCA ACATAGCTGC AACTGCTACG ACAGTTTGGG CGTTACGCTG ACGTTAACTC CGGGGTAAAC GGCATAAAAT AACACTAAGA GGCATTATAG TTCTTAGCGT TGATCGACG TTGACGATGG
6401	ACGACTGTTA TTTAAAACCT TCTAGCCAAT AAAGGGAAAT AAAATAAATA GACTAATGTT TTATTACATT TATGTTTAGA CAAAGGAGAA AGAGAAGACA TGCTGACAAT AAATTTTTGA AGATCGGTTA TTCCCTTCA TTTTATTTT CTGATTACAA AATAATGAAA ATACAAATCT GTTTCCTTCT TCTCTCTGT
6501	GTGAGTGAAC ATCCCAAAG ATGCCTAGAA ACCATAGGCA TAGGGCAATG TTGATCCACC CTCAACTTAA TTTGTGACC AGAGGTGCTA CCRCCACCAC CACTCACCTG TAGGTGTTTC TACGGATCTT TGGTATCCGT ATCCCGTTAC AAGTAGTGG GAGTTGAATT AAACAGTCGG TCTCCAGCAT GGTGGTGGTG
6601	AAAGACTGGT TCCTTCTCTA AATCTCAGGT ATTTCTTTTC CCTTCTAAGC ATAAATGAGA TCCAAATATG ATTTAGGAAC TACATATGA TGTGTTTAAA TTTCTGACCA AGGAAGAGAT TTAGAGTCCA TAAAGAAAAG GGAAGATTGG TATTACATCT AGGATTATAC TAAATCCCTG ATGTATAACT ACAACAATTT
6701	TAGCCACACT ACAGTGCTAT AGTGTAGCAG AATTTGGACA AATCGTTAAT GTTCCACACT ACGCTATATA GTACAAAATG TTGCAAAATA ACGCATATAG ATCGGTGTA TGCAAGATA TCACATCGTC TTAACCTGT TTAGCAATTA CAAGGTGTTA TGCGATATAT CATGTTTTAC AACAGTTTTAT TGCTGATATC
6801	TGGCCCTATG GGTAGCCAA AATTTAACAA ACCGCTATAT TCCAGATTT GCATTTGTCA TCATTACATA CAAATATAAC TCACCTGAGT TGGTCCGGTG ACCGGGATC CCAATCGCTT TAAATTTGTT TGCGCATATA AGGTACTAAA CGTAAACAGT AGTAAATGAT GTTTATATTT AGTGAACCTA ACCAGGCCAC
6901	ATCTTGGTTT AAGACCTTAG AATGTGCACC TTTTAAAGTT CTCAGGTTGC AATCTCCCTG ATGTCAATTT CCGTGGACTA GTTTGACTTC TTAAAAAAAA TAGAACAAA TTCTGGAATC TTACACGCTGG AAAAATTTCA GAGTCCAAAG TTAGAGGAC TACAGTTTAA GCCACCTGAT CAAACTGAAG AATTTTTTTT



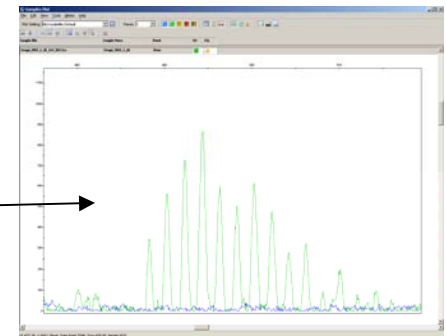
# Inaccessible regions – Taq Slippage



Repetitive DNA Sequences lead to Taq Slippage causing multiple cleavage. This example is in rice.



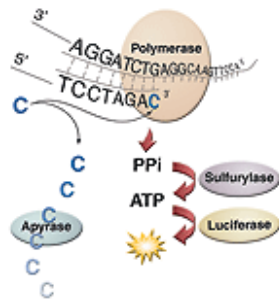
```
ggcaaccaat acgaagaaggtgatgttctcttgcctatgaaagt aagcttgattttgga  
gatccatcacggaacaacatcaaatgatgtaggtaggttotctctggctactgtaacattac  
atcttctggaactgacttccggcatttgcacgagaagatgcttatgatcaaaccttt  
gtacaggtatcaccagaagcccgcaaatggacagcacttctactgcaatcctatt  
ctttgtgatctctctctctaatcggttacatttatatgggtctgggaatcatat  
tgtcaagttagggatgatctcaccgaatgcaggagctaaaagttcgagcagaggcagc  
cgatgttgcaagtcacaggtacttctcttgacatgctgctggcacttggcatatttgg  
atattctcttgatcgcacaataactaaactgaaattttgttttggctcagttct  
Agctactgtctctcatgaaatagaacacctatgaatggcatttaggtaactgcaaaa  
tctctctctctttgcctaacatgtaacatattgcttctatagcttcaacttgctc  
atataaatgttatcacaacacaaactctgtatgattgtaggaatgcttggctgcttc  
tacgcacggaattgaactcaactcaacgggactatgctcagactgctcaagcatgggga  
aggctctgatagcattaataaatgaagtgcctgacogagcaaaaatgaaagctggcaat  
tagagctggaagcgggtccatttgacctcgctcactactcgatgatgctctctcttt  
ttcagagaagctcagacacaaaggttagaggatgctcaactcgttactaacctattg  
tgaaggccaatcagagttcataataatattcctaatctgggttctctcttatttggatg  
atcactcattatccatttacccttcatagctggcagtggttctctgataaagt  
tccagatattggtatgggagactctggg
```



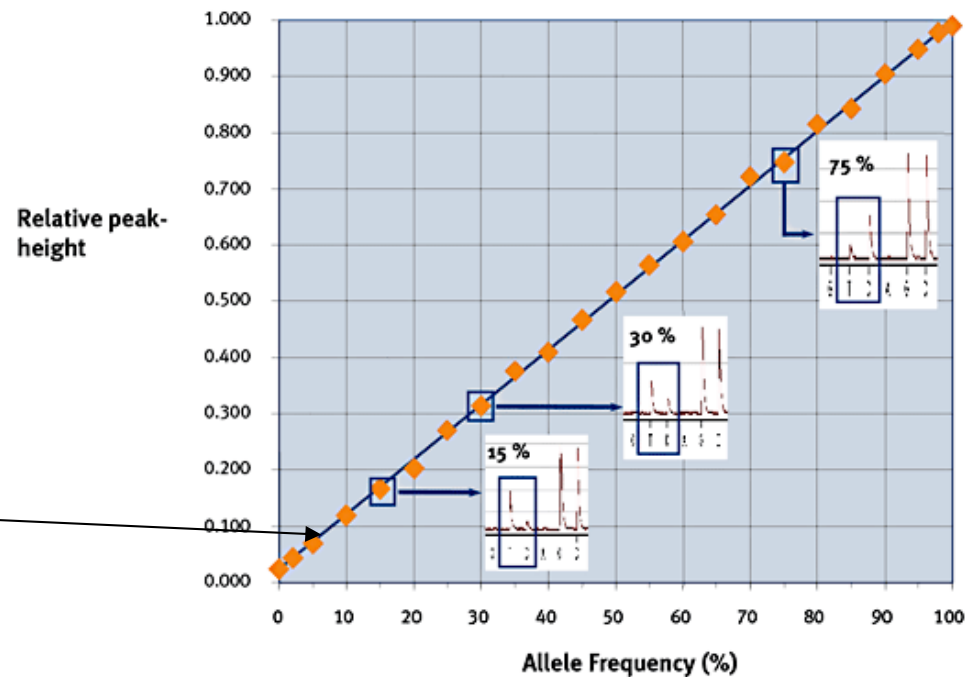
# Strategies

- Could these repetitive regions be overcome when using a method that does not use Cel1 cleavage as the primary detection method?
- If there are EMS populations with DNA's that contain a number of M1 plants can we also utilise these with a non cleavage method?
- If mutations rates are high as with more complex organisms that can survive with higher EMS dosages, can we just target the codons that could be mutated to a stop amino acid?

Could we use Pyrosequencing?

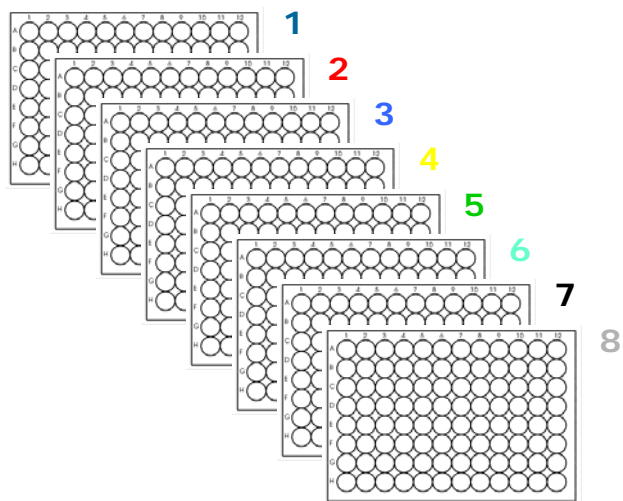


...and detect 1 mutant in 16 (6.25 %) i.e. heterozygous SNP in 8-pool





# Brassica- 2-D Pooling



Individual Plates

It becomes cost effective to screen in 2 dimensions when the mutation rate is above 10 mutations per 768 plants in a 1kb region

8-pool

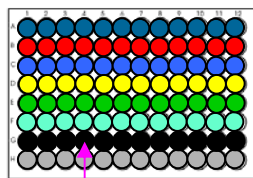
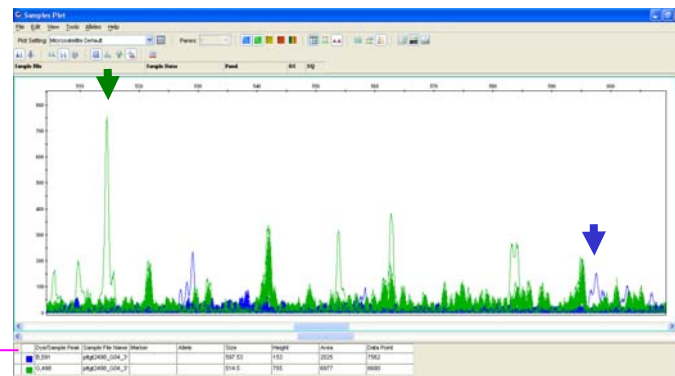
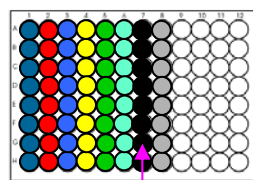


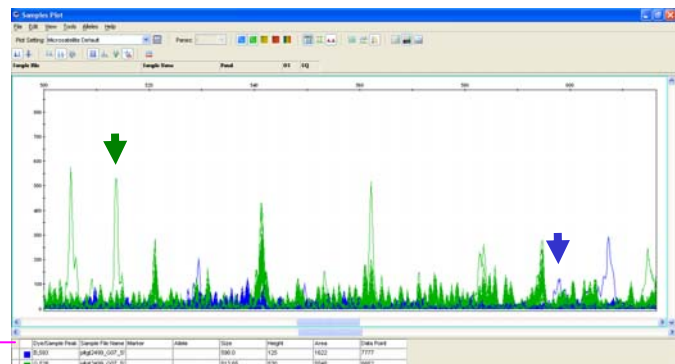
Plate 7 column 4



12-pool



Row G

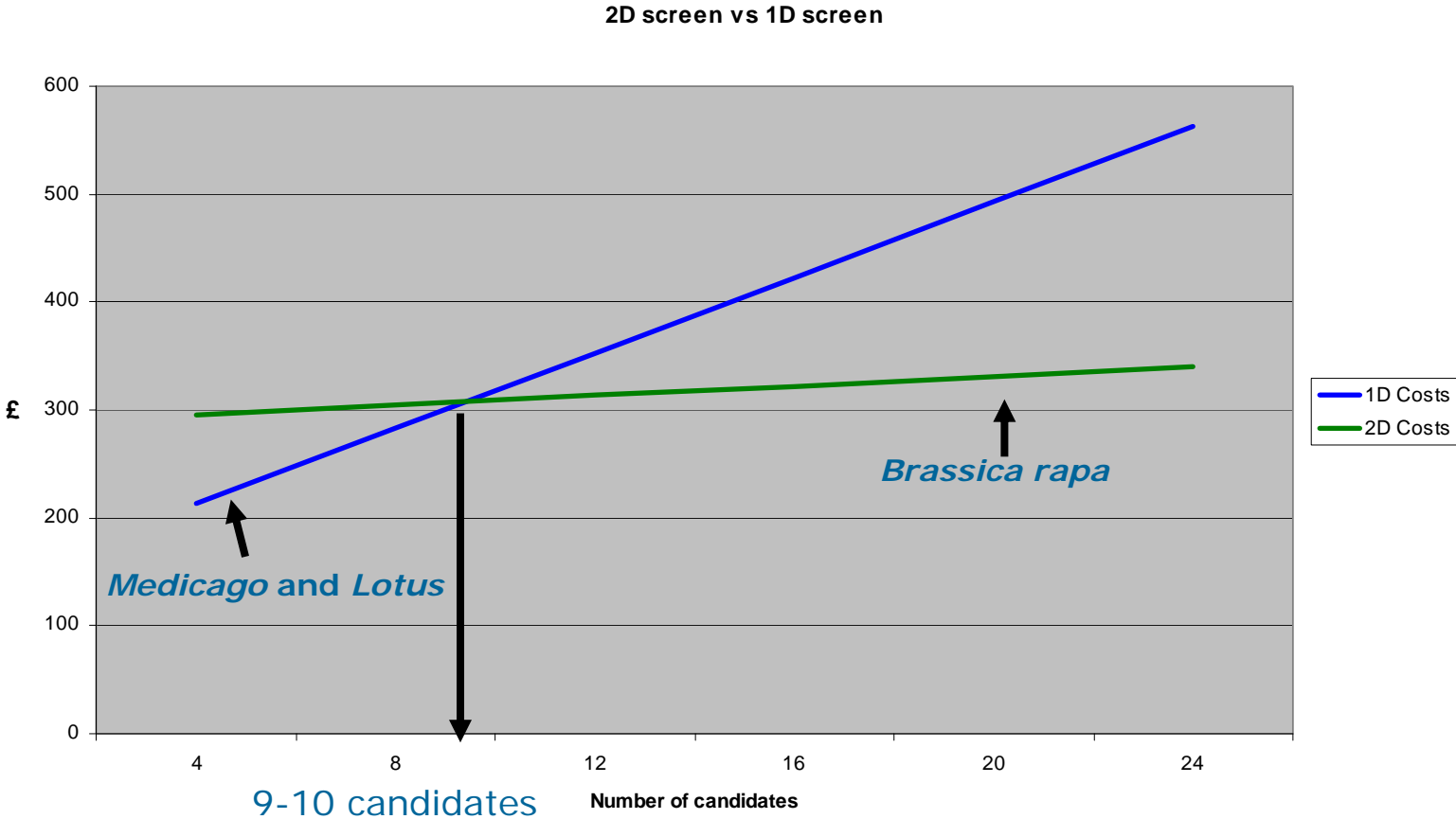


Individual Plate 7 G04 identified

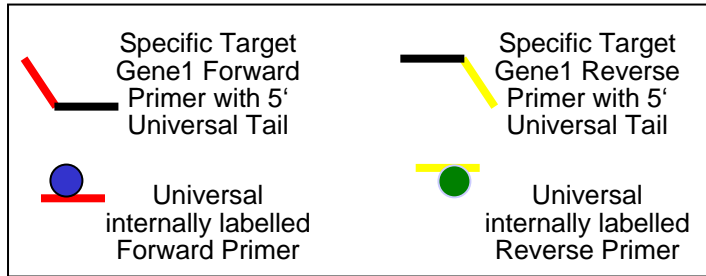
# 2D Pooling - Costing



The cost of screening a single 8-pooled plate and sequencing confirmation (768 plants)



# Arabidopsis – adapted TILLING



Due to the high number of regions a tailed primer approach was adopted for maximum cost effectiveness

After screening M2 seedlings corresponding to 7500 individually harvested M1 plants 78 mutants were identified in having a particular phenotype.

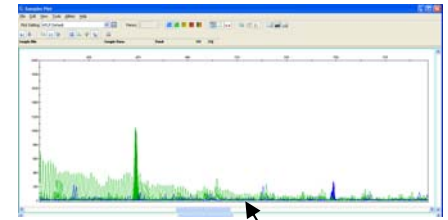
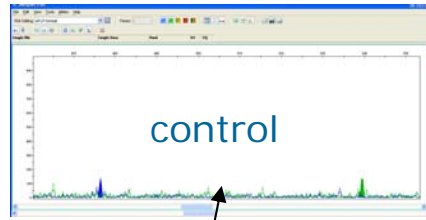
Mutations were then mapped to a chromosome.

To identify whether mutations lie in predicted candidate genes, a modified TILLING approach was used to screen for point mutations within approx. 400 *Arabidopsis* genes.

Mutant lines were mixed 50/50 with a WT line to allow for the detection of homozygous mutations.

# Adapted-TILLING

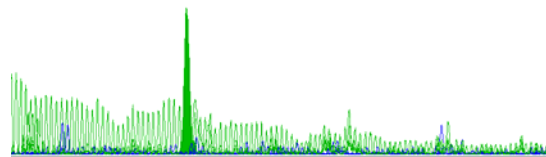
Each plate contained 23 candidate regions and 1 control



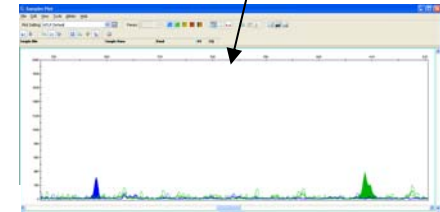
	1	2	3	4	5	6	7	8	9	10	11	12
A	control_WT (1021)	control_WT	control_mut (499+5)	control_mut	cnq7-1 A (1035)	cnq7-1 B	cnq7-1 C	cnq7-1 D	cnq7-2 A (1048)	cnq7-2 B	cnq7-2 C	cnq7-2 D
B	cnq7-3 A (991)	cnq7-3 B	cnq7-3 C	cnq7-3 D	cnq7-4 A (956)	cnq7-4 B	cnq7-4 C	cnq7-4 D	cnq7-5 A (951)	cnq7-5 B	cnq7-5 C	cnq7-5 D
C	ARF2-1 A (1049)	ARF2-1 B	ARF2-1 C	ARF2-1 D	ARF2-2 A (1069)	ARF2-2 B	ARF2-2 C	ARF2-2 D	RGAP2-1 A (1039)	RGAP2-1 B	RGAP2-1 C	RGAP2-1 D (green 469/1000 570)
D	RGAP2-2 A (994)	RGAP2-2 B	RGAP2-2 C	RGAP2-2 D	RGAP2-3 A (1119)	RGAP2-3 B	RGAP2-3 C	RGAP2-3 D	RGAP3-1 A (1167)	RGAP3-1 B	RGAP3-1 C (blue 552/998)	RGAP3-1 D
E	RGAP3-2 A (1027)	RGAP3-2 B	RGAP3-2 C	RGAP3-2 D	RGAP3-3 A (1089)	RGAP3-3 B	RGAP3-3 C	RGAP3-3 D	AGAP2-1 A (966)	AGAP2-1 B	AGAP2-1 C	AGAP2-1 D
F	AGAP2-2 A (1133)	AGAP2-2 B	AGAP2-2 C	AGAP2-2 D	AGAP2-3 A (1003)	AGAP2-3 B	AGAP2-3 C	AGAP2-3 D	AGAP2-4 A (933)	AGAP2-4 B	AGAP2-4 C	AGAP2-4 D
G	AGAP2-5 A (1149)	AGAP2-5 B	AGAP2-5 C	AGAP2-5 D	RGAP4-1 A (1006)	RGAP4-1 B	RGAP4-1 C	RGAP4-1 D	RGAP4-2 A (1047)	RGAP4-2 B	RGAP4-2 C	RGAP4-2 D
H	RGAP4-3 A (1044)	RGAP4-3 B	RGAP4-3 C	RGAP4-3 D	RGAP4-4 A (1100)	RGAP4-4 B	RGAP4-4 C	RGAP4-4 D	AP180'-1 A (1065)	AP180'-1 B	AP180'-1 C	AP180'-1 D

  poor tracer/difficult to score  
  Good but no cleavage products  
  Good with cleavage products

To aid detection using Genemapper, 3 WT reactions were performed to 1 mutant so that all 4 traces could be overlaid



A number of regions performed badly due to using the tailed adapter system





**[1] Capillary Electrophoresis on the ABI 3730 provides a cost effective, high-throughput platform for TILLING using Cel1 in a number of organisms**

**[2] ABI 3730 provides a single platform for TILLING and sequence confirmation**

**[3] GeneMapper v4.0 and Genemarker v1.71 are effective software applications for 3730 TILLING data analysis, Mutation Surveyor effective for sequence confirmation**

**[4] The JGL can offer a fast and reliable pipeline to prepare a population for TILLING and offer a cost effective approach to pooling**

**[5] Development of strategies to deal with new and difficult populations is ongoing in the JGL**

# Acknowledgments

## JIC Genome Laboratory

Bethany McCullagh  
Kilian Hann  
Richard Goram  
Gawain Bennett  
Darren Heavens  
Jonathan Clarke

## JIC

Lars Ostergaard (*Brassica rapa*)  
Pauline Stephenson (*Brassica rapa*)  
Kay Denyer (Barley)  
Sylviane Comparot-Moss (Barley)  
Simon Griffiths (Wheat)  
Simon Orford (Wheat)  
Trevor Wang (*Lotus*)  
Jillian Perry (*Lotus*)

## INRA (*Medicago*)

Richard Thompson  
Christine Le Signor

## SCRI (Barley)

Robbie Waugh  
Nicola McCallum

## University of Oxford (*Arabidopsis*)

Ian Moore  
Jose Perez-Gomez