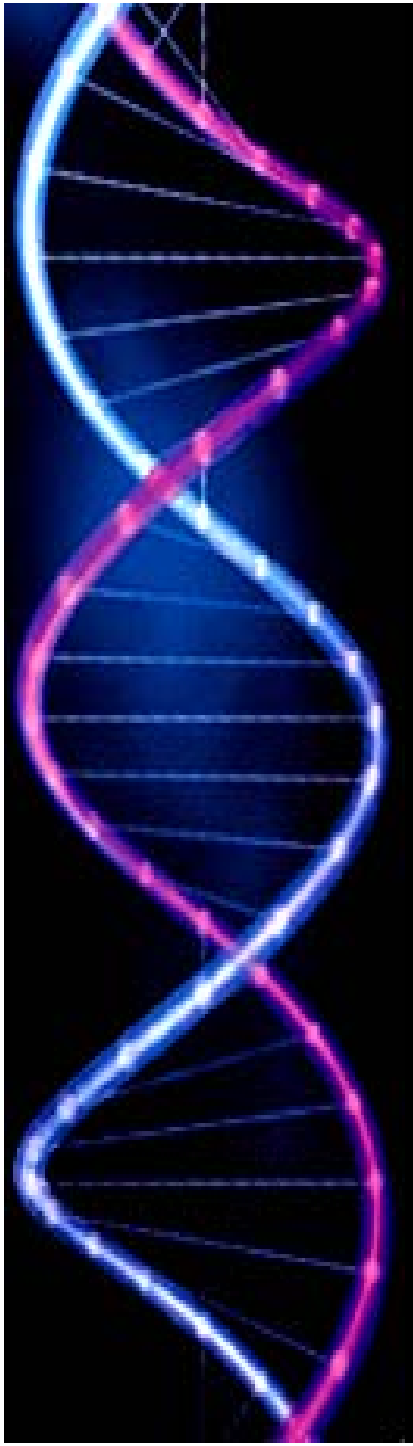
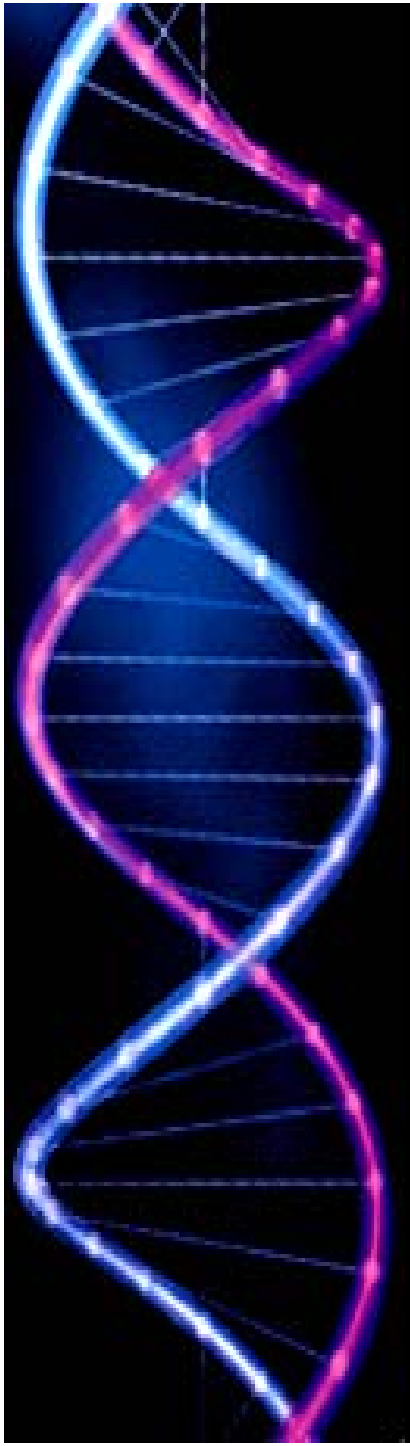


Evaluation of Four Automated Mutation Detection Programs for Clinical Re-sequencing

Dr. Stacey Bléoo

Assistant Professor, University of Alberta

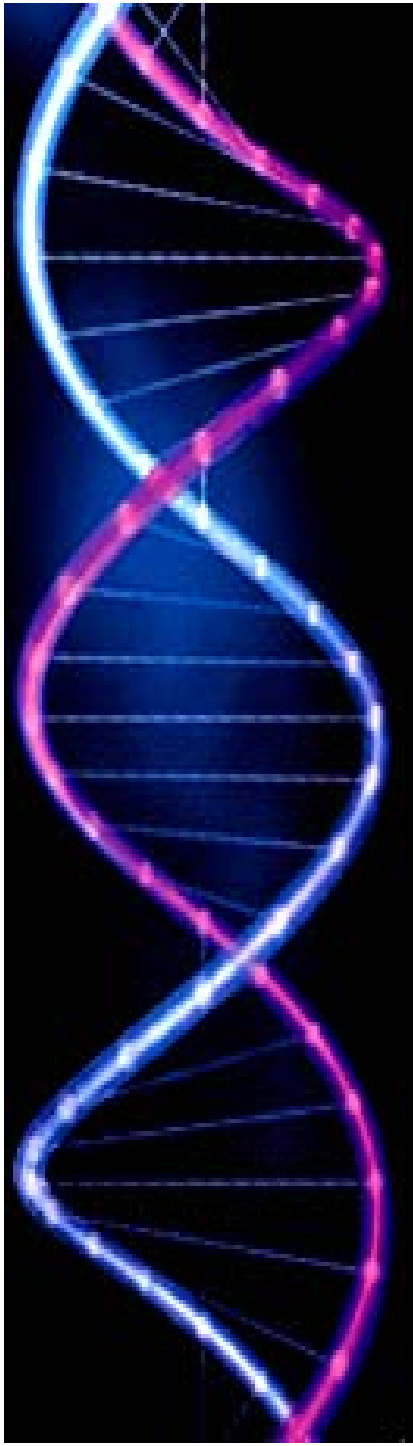
Director, Molecular Diagnostic Laboratory,
Stollery Children's Hospital



Objective: To evaluate all commercially available automated mutation detection software for use in the diagnostic testing of large genes.

Study Workflow

- 15 Qualitative Traits
- Ability to flag 114 variants
- Ability to name 114 variants
- Sensitivity & specificity
- Results from 125 BRCA patients



Method of Variant Detection

Mutation Surveyor v3.25

-Uses anti-correlation method (compares a reference electropherogram to a test – indicates discordance)

Variant v1.0/Seqscape v2.6

-uses base caller only

Seqpilot v3.2.1.2

-variant detection done using base caller however, all peak morphology including noise are stored and used for statistical comparisons of later peaks

Qualitative Aspects

	Variant Reporter	Seqscape	Seqpilot	Mutation Surveyor
Ease of template creation	X	X	√	√
Ability to lock template	X	√	X	X
Ease of sequence assembly	√	√	√	X
Able to auto-import directly from CE	X	√	√	√
Ability to create ROI	√	√	√	√
Audit trail capability	X	√	√	X
Visualization of 2X coverage of ROI	X	√	√	X
Measurement for seq. quality	√	√	X	√
Record variant type	X	√	√	√
Attach variant information (ie. PDF)	X	X	√	X
Program can run on a server	X	X	√	√
Variants named using HGVS*	W	W	√	√
Able to de-convolve frameshifts	X	X	√	√
Automatic variant detection on de-convolved fragments	X	X	X	X
Overall user friendliness^	3	2	1	4

* W indicates workaround possible

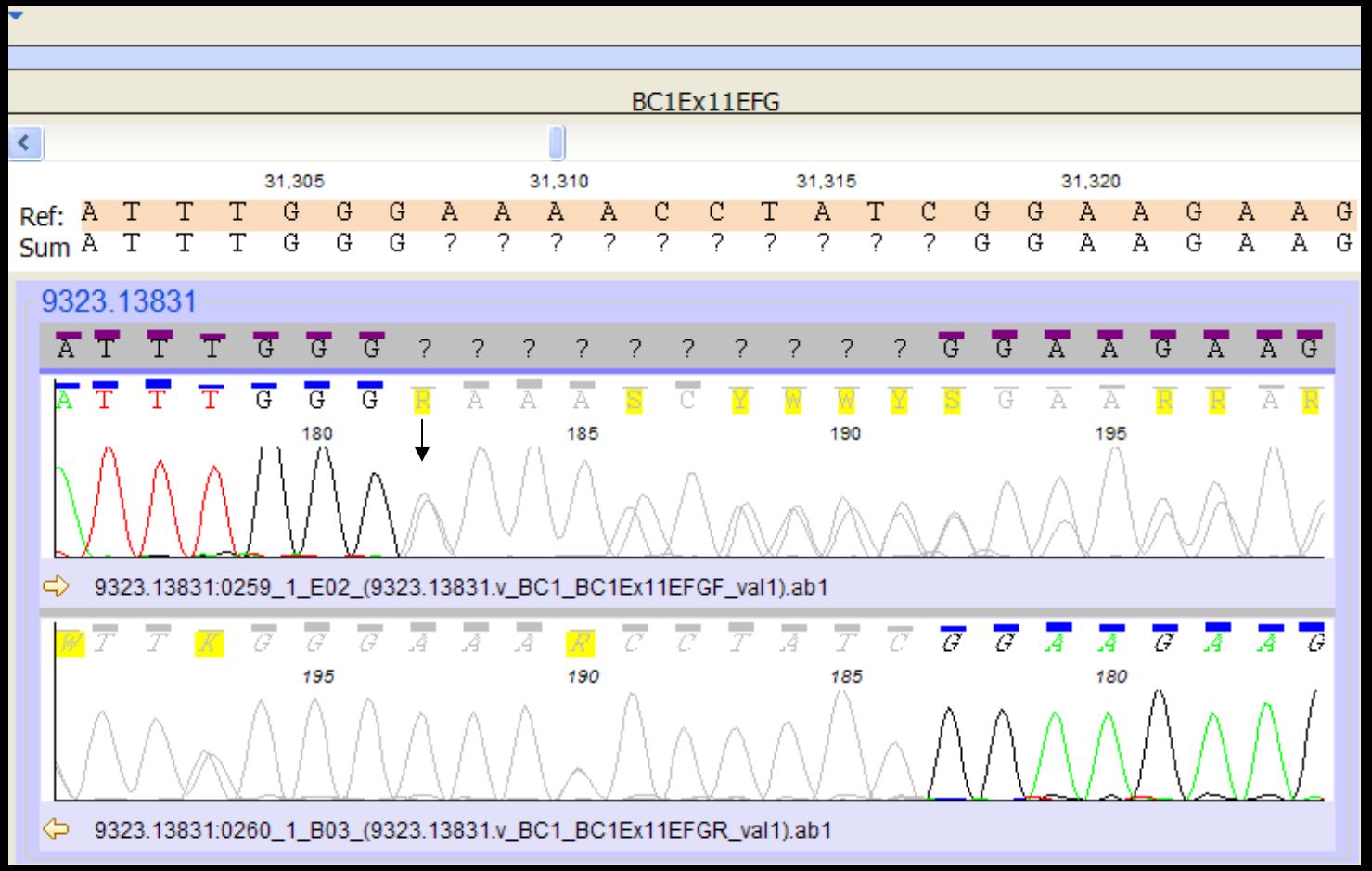
^ 1 being the most user friendly

Ability to 'Flag' 114 Variants

	Point Mutation (76)	Frameshift (36)	Complex Change (2)
Variant Reporter	99% (75)	85% (31)	100%(2)
Seqscape	99% (75)	94% (34)	100% (2)
Seqpilot	100% (76)	97% (35)	100% (2)
Mutation Surveyor	100% (76)	100% (36)	100% (2)

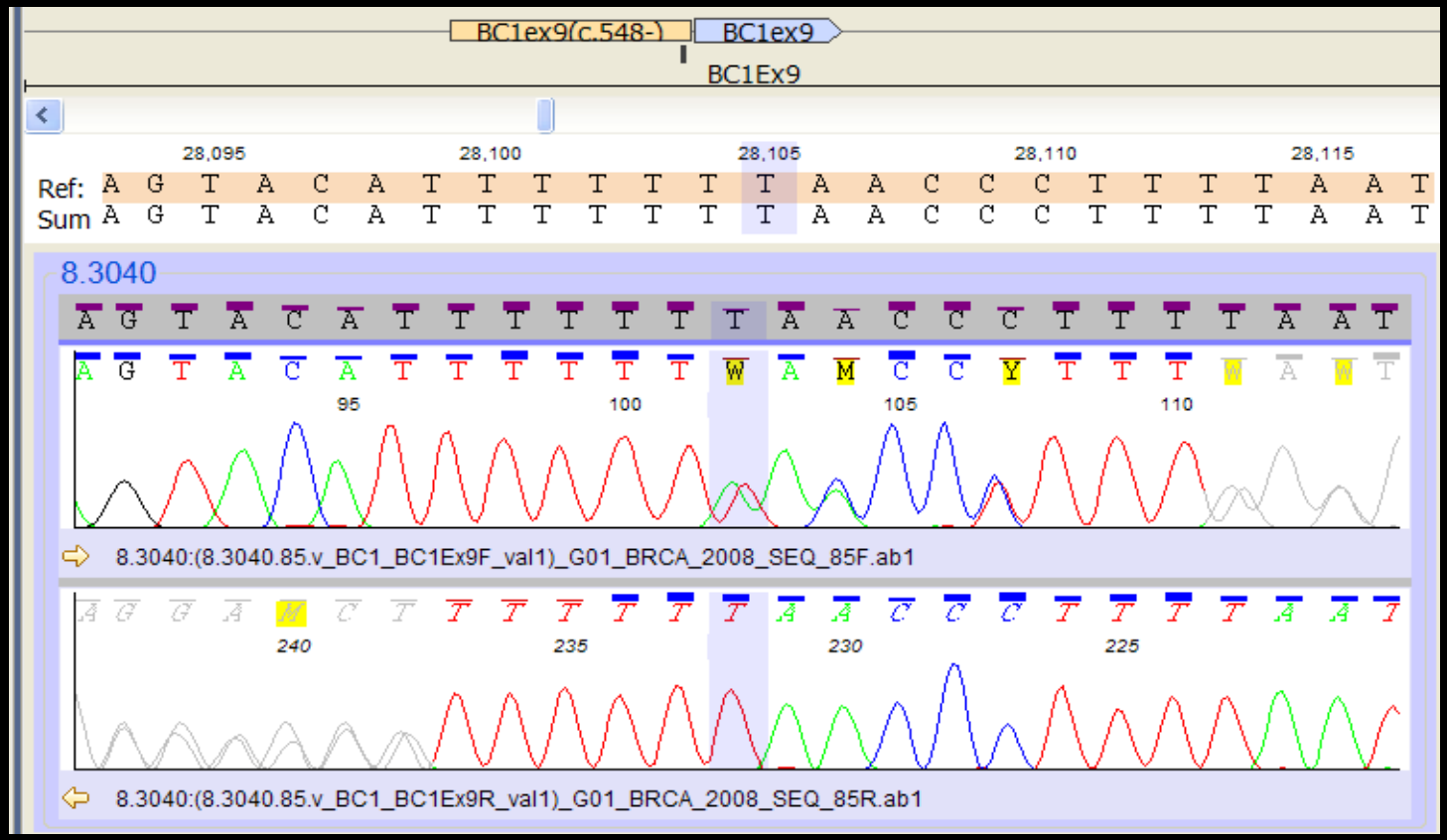
- variants must be listed in 'variant' table (or 'possible heterozygous indel mutation' table for seqscape)
- complex changes included: BRCA2 c.10095delCins11 & MSH6 c.866GC>AA
- Note: one frameshift mutation not detected by seqpilot, variant reporter and seqpilot was present only on one allele

Variant Reporter – Examples of missed point mutation & indel



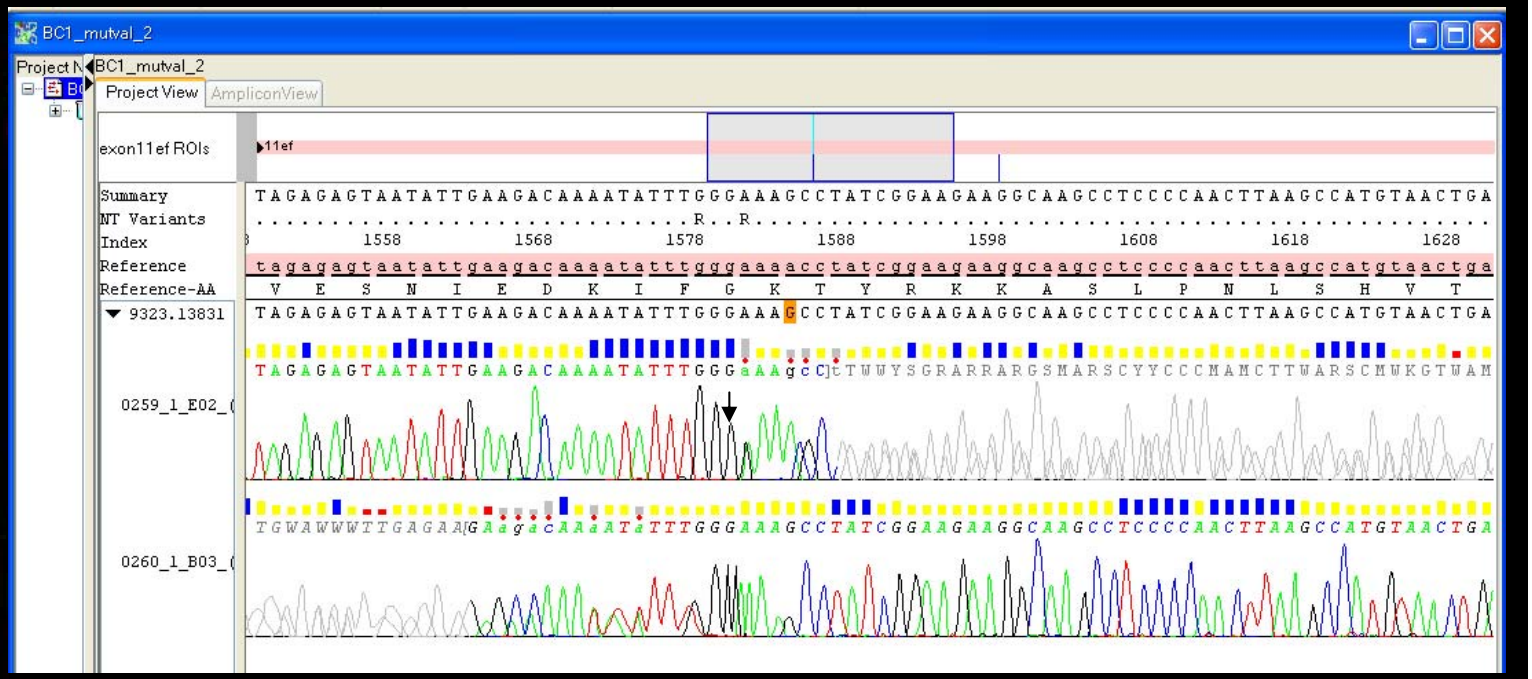
- BRCA1 c.1390_1391insG (1/2) and c.1387A>G (1/2) or c.1386dupG & c.1390 A>G
- VR trims off the “low quality” sequence, leaves a gap, but does not document it in the variant table, or a “position of interest”.
- This variant is completely missed.

Variant Reporter – Examples of missed indels



- BRCA1 c.548-58delT (1/2)
- VR needs filters removed in order to see this mutation, and it does not call it even if both sequences are assembled.
- It does show some of the mixed bases under the “position of interest” tab, but nothing is listed in the variant table

Seqscape – Missed base substitution



BRCA1 Variant:

c.1386dupG & c.1390 A>G

or

c.1387 A>G & c.1391insG

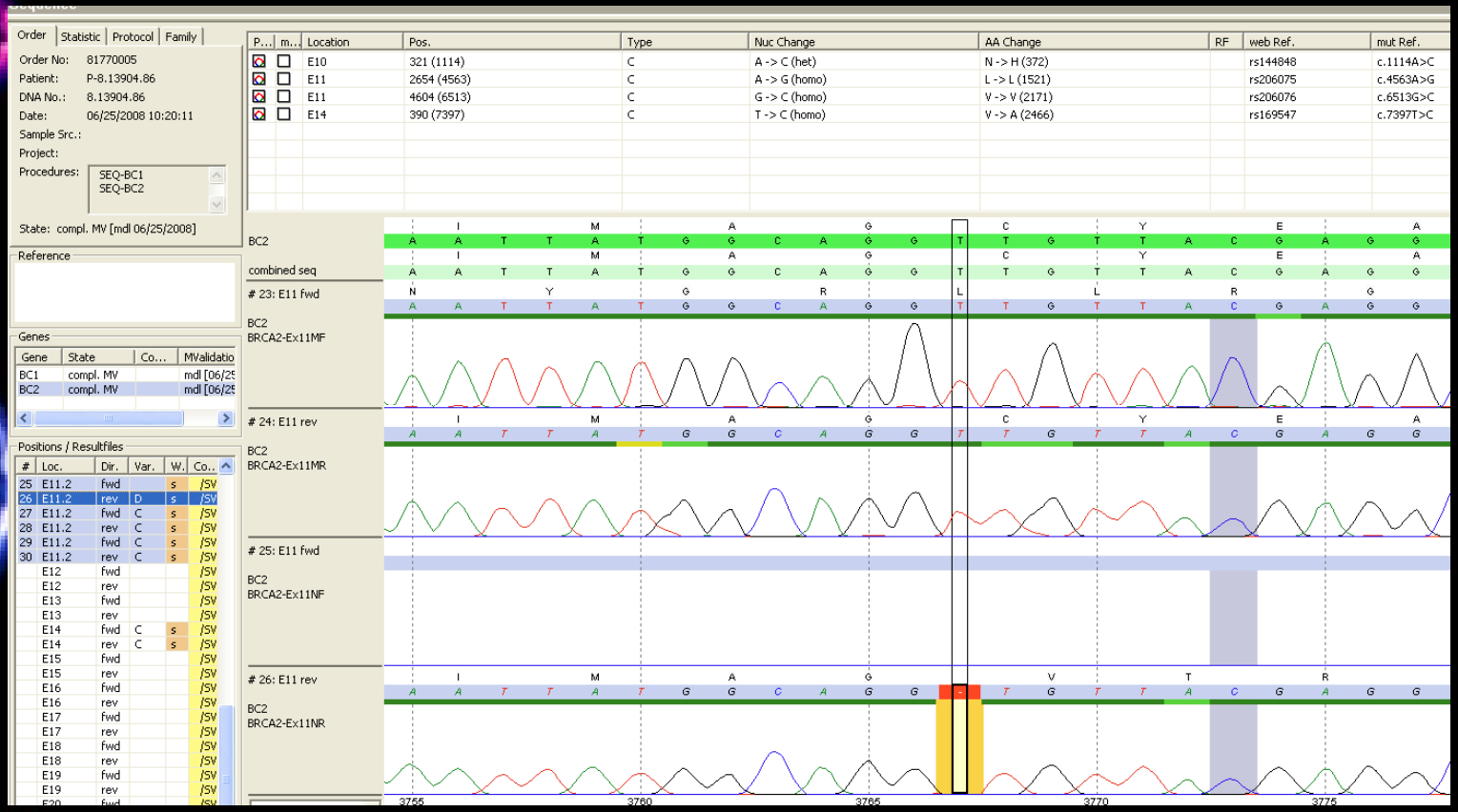
- seqscape picked up indel but missed base substitution

Seqscape – Missed indel

MSH6 homozygous c.3646+35_38delATCT

- not listed in either indel table or in mutation report; is manually searchable as a ‘discrepancy’

Seqpilot – Missed base substitution

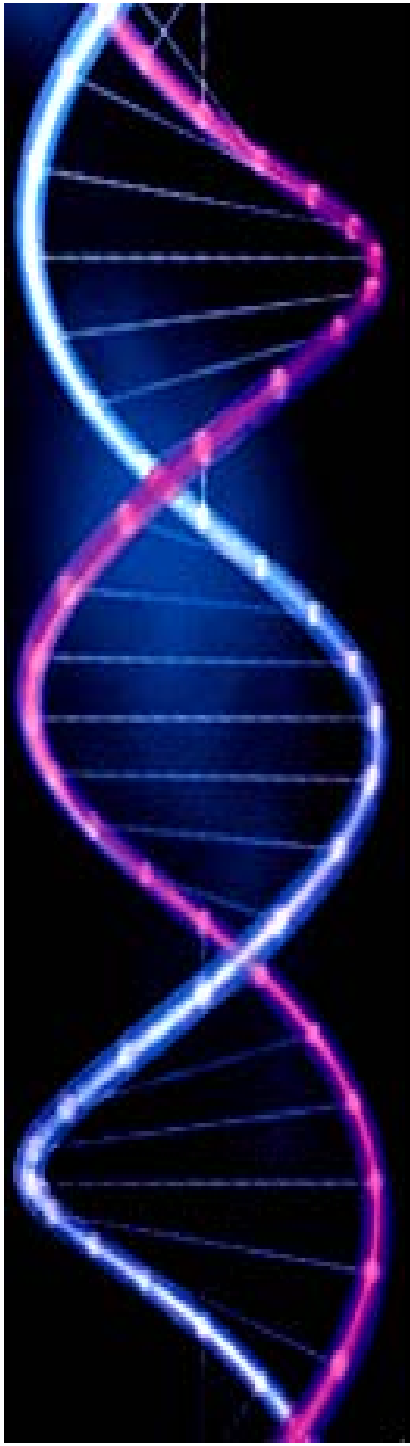


- This variant was due to an error in one primer and therefore was present on only one allele; although not found in the variant table, it is flagged with a ‘D’ in the position/resultfiles window (left)

Ability to Name (HGVS) 114 Variants

	HGVS	Point Mutation	Frameshift	Complex Change
Variant Reporter	cDNA	100%	63%	0%
	protein	100%	0%	0%
Seqscape	cDNA	100%	56%	0%
	protein	100%	0%	0%
Seqpilot	cDNA	86%*	46%	0%
	protein	100%	0%	0%
Mutation Surveyor	cDNA	100%	91%	0%
	protein	100%	3.6%	0%

- *Seqpilot point mutation calling incorrect for intronic mutations: “IVS” instead of referring to cDNA position



- decision not to use variant reporter:

  - variant reporter did not ‘catalog’ variants therefore, one would have to keep a reference list (ie. possibly excel?) of known polymorphisms when analyzing patients

Known Variants

Alignment	Type	Position	Reference	Variant
Yes	Substitution	68359	C	68359C>S

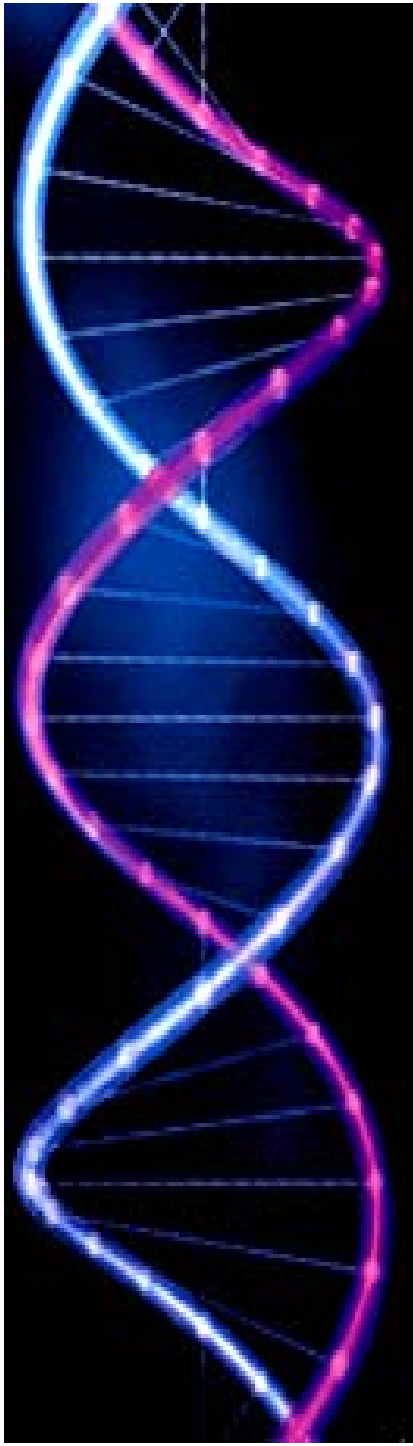
Genomic position only

This potential ‘comments’ field is locked (unusable)

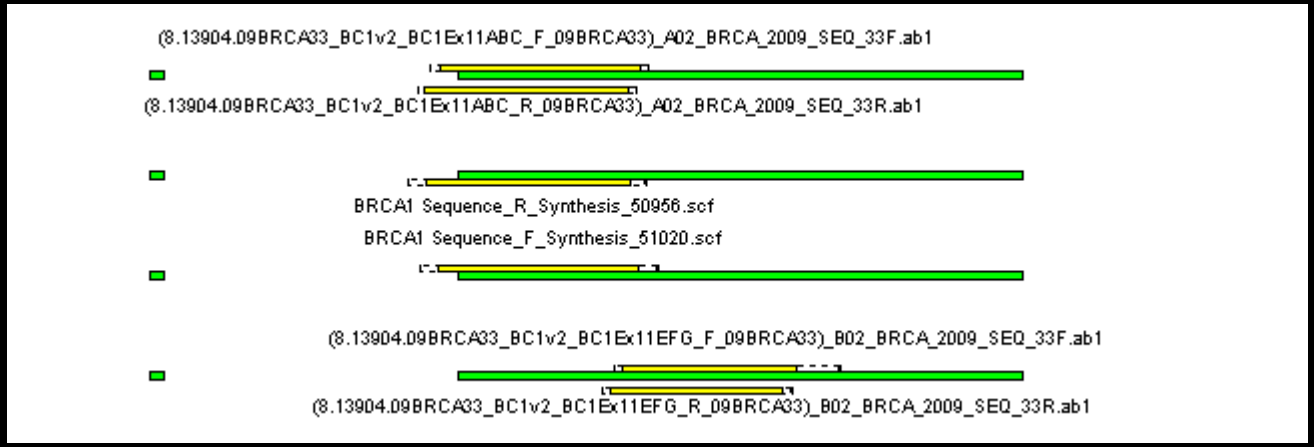
Sensitivity and Specificity

	Sensitivity	Specificity
Seqscape	**100%**	**36%**
Seqpilot (without Stats)	**100%**	**59%**
Seqpilot (with Stats)*	**100%**	**73%**
Mutation Surveyor	**100%**	**76%**

* Statistics included 80 analyzed patients



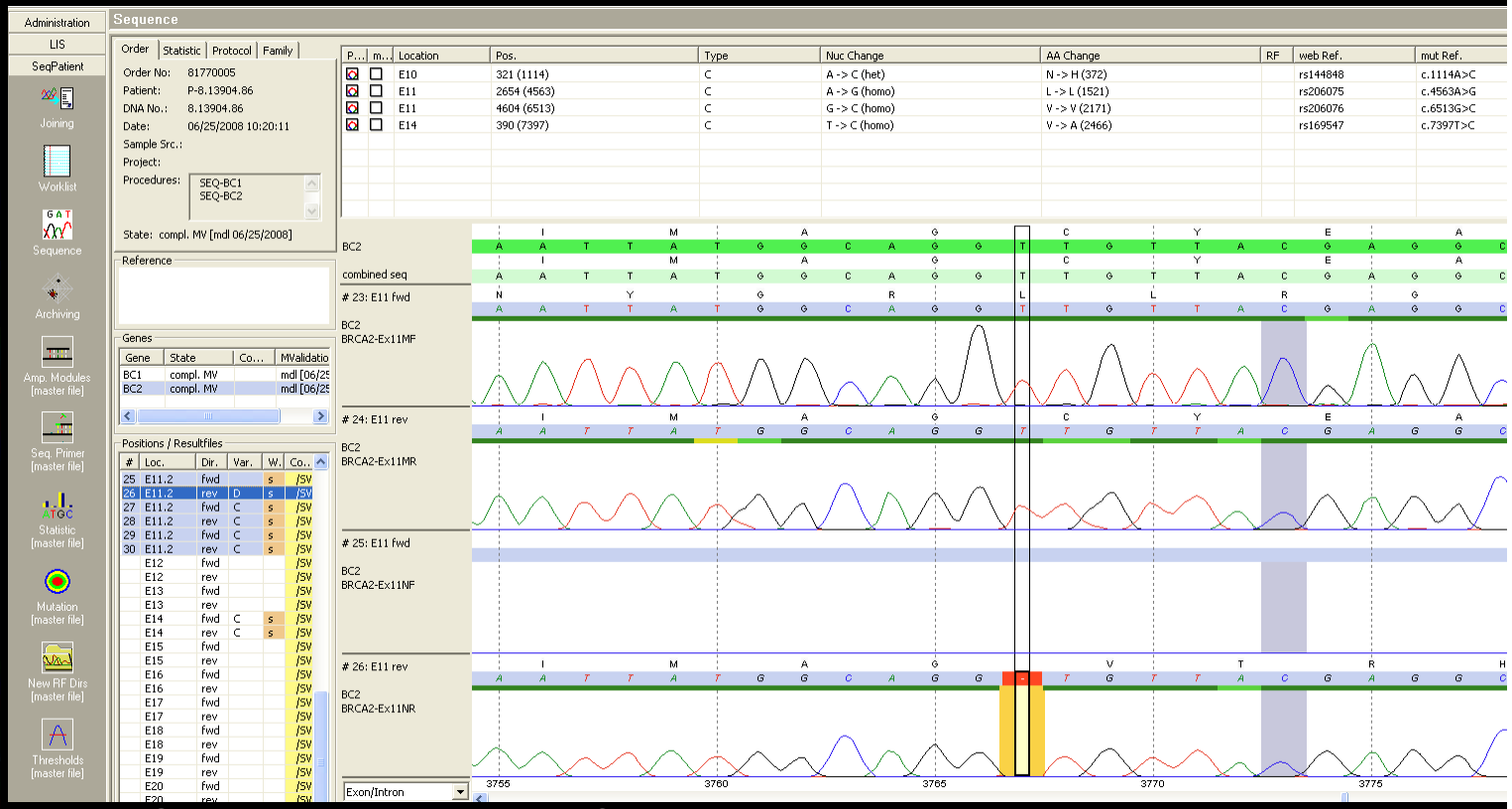
- decision not to use mutation surveyor for further validation studies
- mutation surveyor assembled BRCA1, but not BRCA2 into correct overlapping contigs therefore patients would have to be analyzed by amplicon (our workflow is patient centered not exon/amplicon centered)
- we found it difficult to determine graphically whether there was bi-directional overlap between contigs (however it can be done numerically)



- staff did not find this software to be user friendly (ie. familial mutations involve searching one cDNA position; difficult to determine with mutation surveyor)

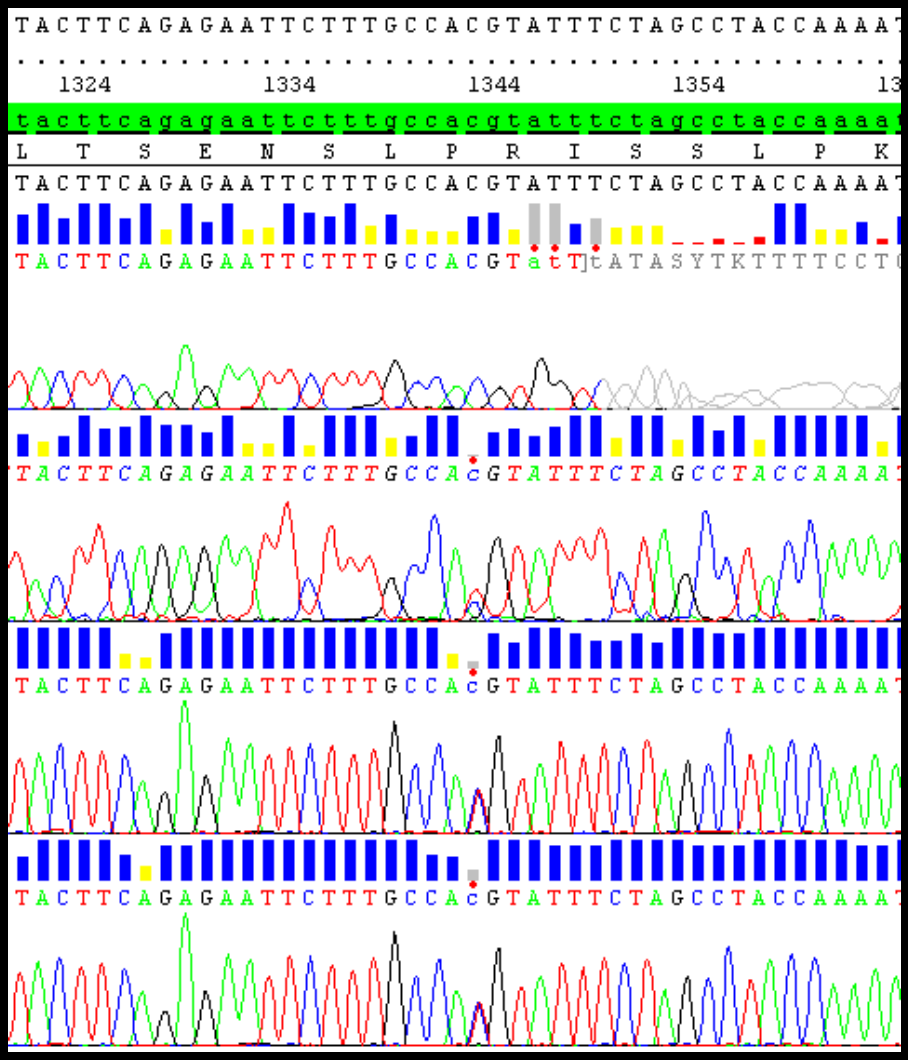
Final Phase of Evaluation

- 125 BRCA patients were analyzed in duplicate with seqscape and seqpilot: over 2 million bp sequenced in two directions



- BRCA2 primer error not found in variant table (noted as a ‘D’ in the positions/results files table); we questioned the potential for the software to miss homozygous deletions if spacing is adjusted differently on both strands

Seqscape – masking of variants

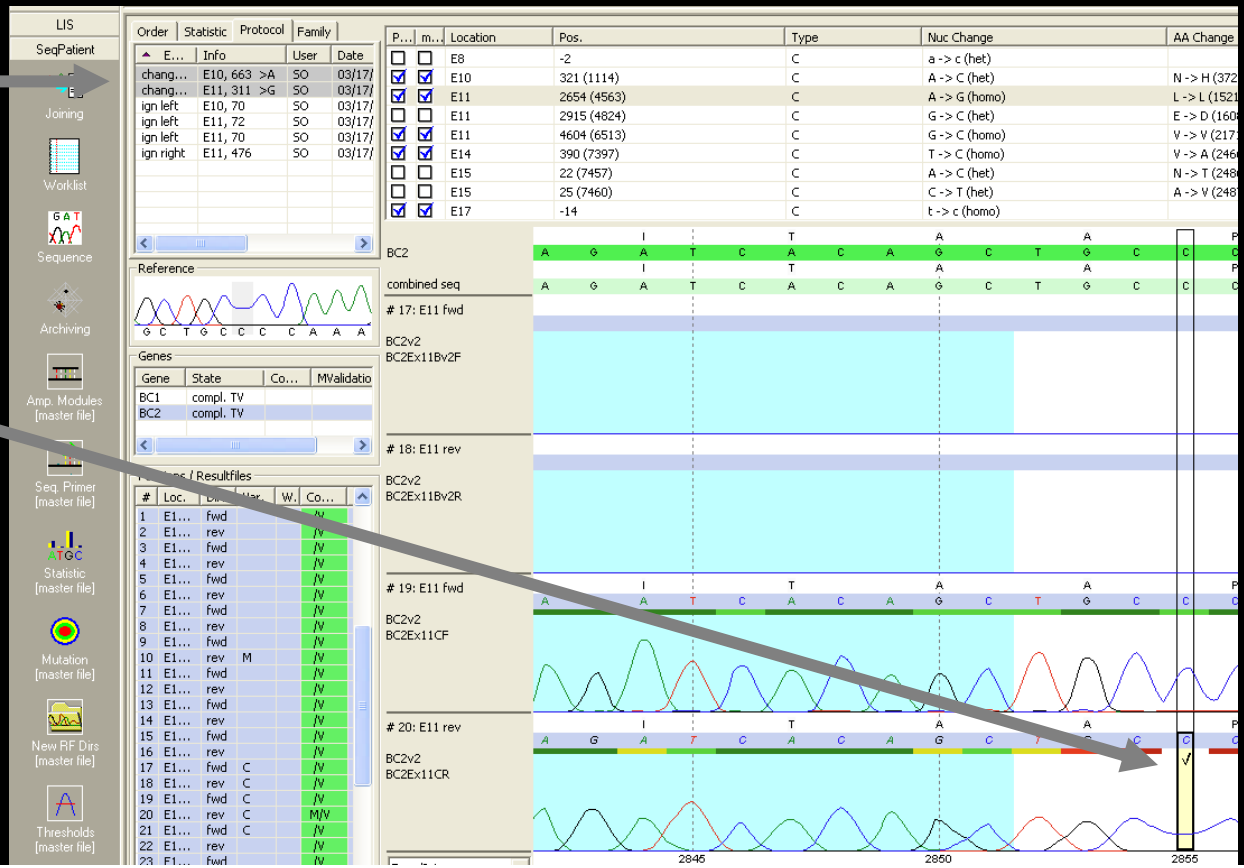


● Seqscape had a similar problem identifying changes that were not uniform across all alleles (ie. SNPs under primers). These changes are only identified by searching through all sequence discrepancies.

Final Decision

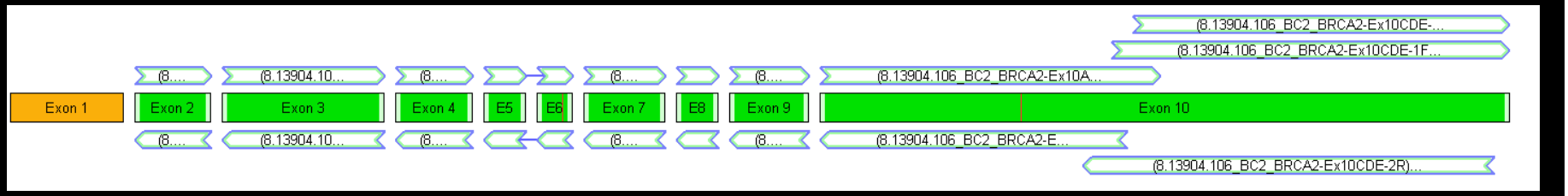
- final decision was to go with seqpilot – high sensitivity and specificity
- total time for either a technologist to analyze a BRCA1/2 patient or director to sign-off = 15 minutes
- has an audit trail important for clinical QA/QC:

Changes to base calls made by technologist are recorded and can be evaluated by lab director (using a single click to locate the position)

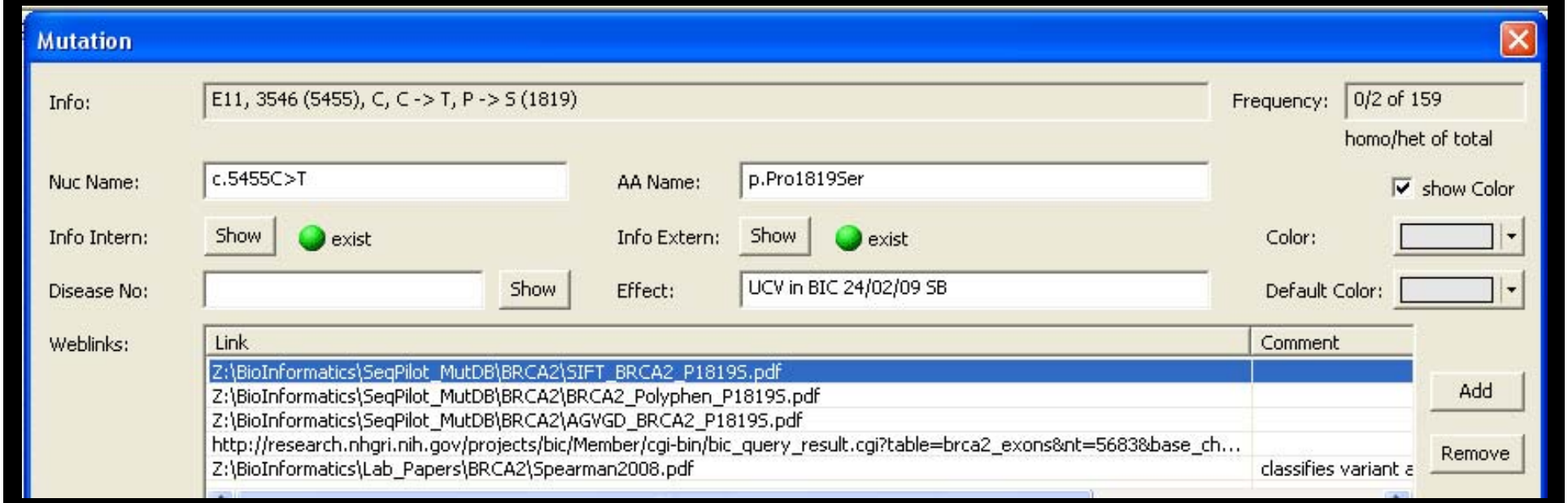


Final Decision

- fairly easy to determine if bi-directional coverage was achieved



- available as a server version
- able to hyper-link data on variants in mutation database

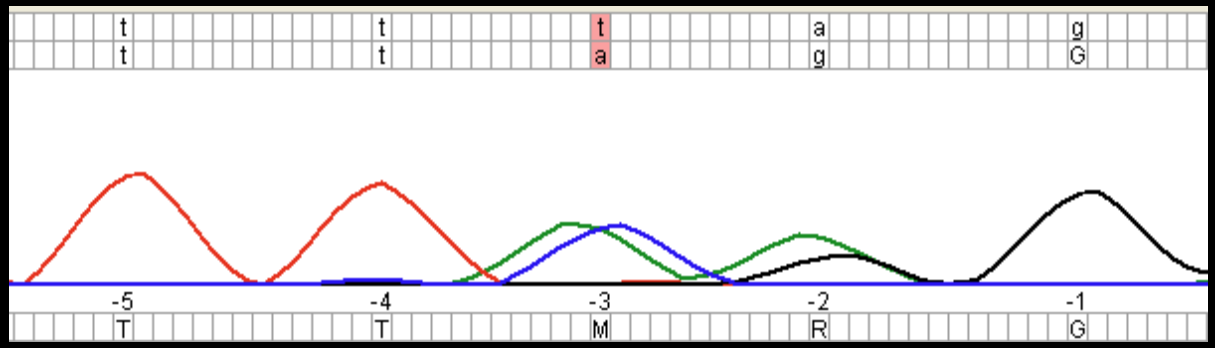


Final Decision

- able to deconvolve heterozygous indels; also provides a semi-automated method for sequence evaluation of each allele after deconvolution

Red 'flag' indicates that the allele call does not match the electropherogram

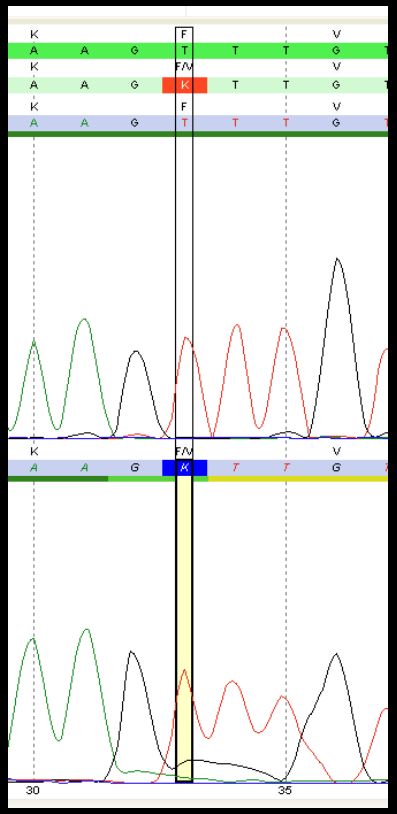
Expected call for each allele

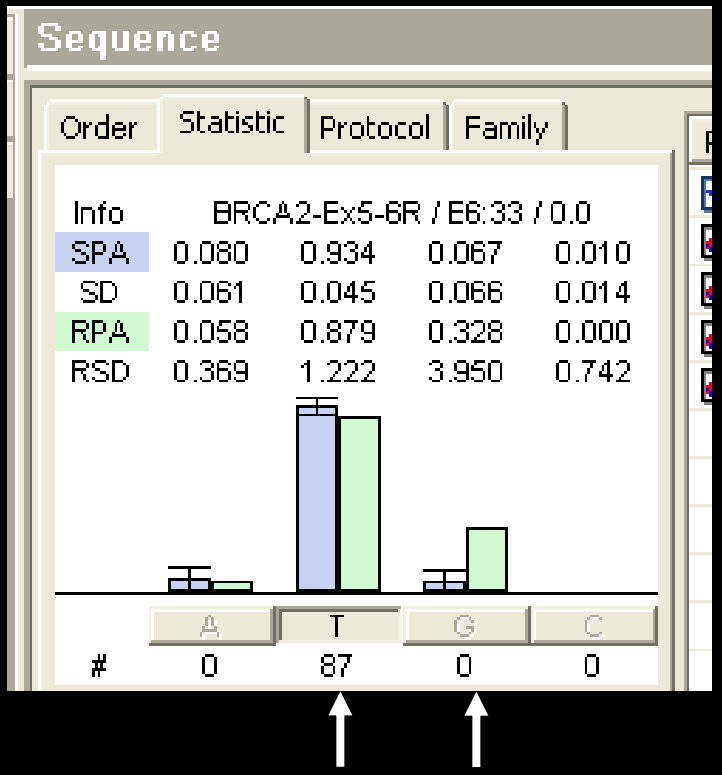


- able to mark primer locations; important for large 'multi-amplicon' exons whereby the presence of SNPs might result in a null allele in a neighbouring amplicon

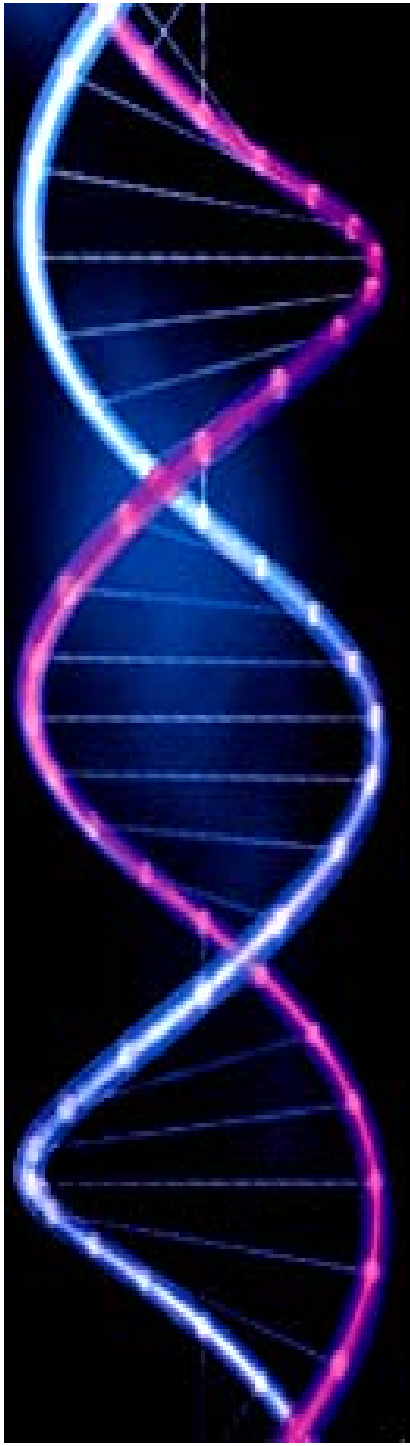
Final Decision

- does not require a similar quality reference trace: will use peak statistics to evaluate similarity of current peak to previously analyzed peaks





Bars indicate peak area; blue is average of all previous peaks and includes std dev.; green is current peak area



Acknowledgements

- My entire lab and notably the individuals who performed most of the evaluation:

  Robert Tomaszewski

  Christine Walker

  Mark Hicks

- JSI (Seqpilot), Softgenetics (Mutation Surveyor), ABI (Applied Biosystems)

- Dr. Martin Somerville – Divisonal Director, Medical Genetics Laboratories, Edmonton, Alberta

- Dr. Jillian Parboosingh – Associate Director Molecular Diagnostic Lab, Calgary, Alberta