

Application Note

Detection of Insertions and Deletions

in Resequencing Projects with Mutation Surveyor™ or Mutation Explorer®

Megan Kellander, ChangSheng Jonathan Liu

SoftGenetics, 200 Innovation Blvd. Suite 241, State College, PA 16803

Introduction

Insertions and deletions (indels) in a gene may cause severe genetic diseases, because translation of the gene often becomes frame-shifted. Indels frequently create a new stop codon which terminates the amino acid translation. Large Resequencing projects demand fully automated calls of the heterozygous indels in hunting for the disease cause mutations. Among the diseases caused by indels are Fragile X syndrome and many instances of breast cancer. Mutation Surveyor™ & Mutation Explorer® are the only software packages to detect the homozygous indels and heterozygous indels by deconvoluting its DNA sequence traces into two sets of the sequences¹. Mutation Surveyor & Mutation Explorer also provides an accuracy >99% in detecting homozygous and heterozygous point mutations², when both forward and reverse traces are of Phred 20 quality.

There are many software packages available to detect the point mutations including homozygous and heterozygous point mutations. These software packages include SeqScape® from Applied Biosystems Inc., Foster City, CA, PolyPhred³ University of Washington, Seattle, and PolyPhred derivative software, inSNP, novoSNP, seeSNP, spotSNP, Codon Code Aligner, PolyBayer, and Paracel Agent. The Mutation Surveyor and Mutation Explorer packages both feature ease of use, low learning curve, full automation, and an exclusive migration time comparison to detect Indels.

Procedure

Data Entry


1. Add 1-Directional or 2-Directional sample files.
2. Enter a Gen-Bank sequence or appropriate reference file sequences.

Reference File Selection

It is essential that the user input the correct reference trace in order for the software to detect Heterozygous Indels. It is recommended to use a parental sample as a reference. In the case of samples including tandem repeats, the most appropriate reference sample would have the same number of repeats as the sample(s).

Analysis

Locate indels in sample traces. They are indicated by a red bar as homozygous indel (Figure 1) and brown bar as heterozygous indel (Figure 2). The bar is shown at the top of the mutation electropherogram, 3rd panel in the figures. The first panel is the reference trace synthesized from the GenBank file. The second panel is the experimental trace.

Open the Heterozygous Deconvolution page, shown in Figure 3, by clicking on the  icon in the software's main toolbar or by double-clicking in the desired mutation cell in the various reports.

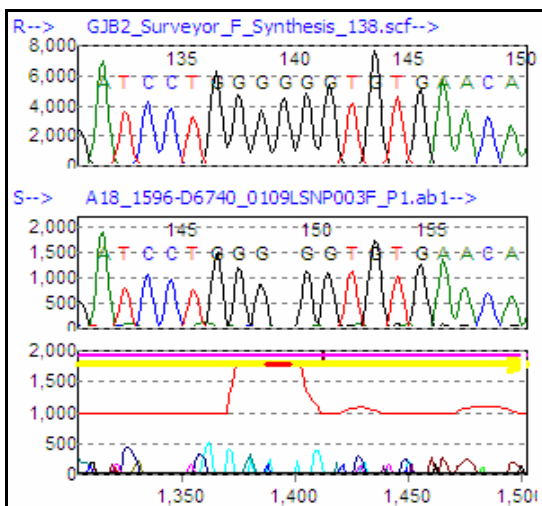


Figure 1. Example of Homozygous indel.

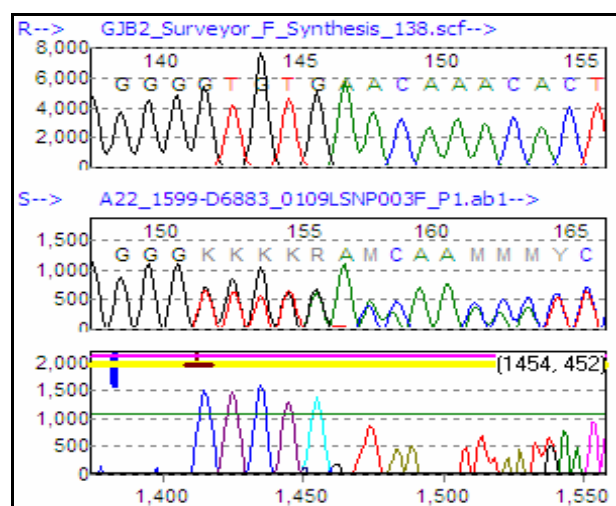


Figure 2. Example of heterozygous indel.

Results

The SoftGenetics software packages use the DNA migration time deviation plot before and after alignment to detect the indels. The local expansion of the sample trace aligning to the reference trace is a deletion, (Figure 1), and local shrinking of the sample is an insertion. The software's indel detection is not disrupted by basecalling error when DNA migration times of sample and reference traces are correlated.

Mutation Surveyor and Mutation Explorer can automatically deconvolute sample traces when a Heterozygous Indel has occurred, shown in Figure 3. **A** is the reference trace, and **B** is the sample trace. The two portions of the convoluted sample trace, the Conserved Sector, **C**, and the Mutation Sector, **D**, can be separated into two clean traces following the point of the insertion or deletion. Then the software is able to shift the Mutation Sector **D** to align with the reference **A**. The software accurately detects the 19 deleted bases, shown in **E** and **F**.

A user may add a second heterozygous indel in addition to the one the software automatically detected. The user may modify the start position, and the software will do the calculation based on user constraints.

Discussion

SoftGenetics' Mutation detection software utilizes a unique physical comparison of the migration time of the sample traces and the reference traces, rather than relying on the text basecall. By monitoring the migration time of the sample and the reference, Mutation Surveyor and Mutation Explorer can accurately detect Het Indels while eliminating false calls caused by text based comparison and alignment.

Please note that the reporting number of heterozygous indel should always be referred to in the forward direction, because Het Indels may be called differently in the reverse orientation. This is due to the nature of PCR reaction reading from forward and reverse directions.

Mutation Surveyor and Mutation Explorer have been used in somatic mutation detection studies that resulted in the Human Cancer Genome Project^{4,5}.

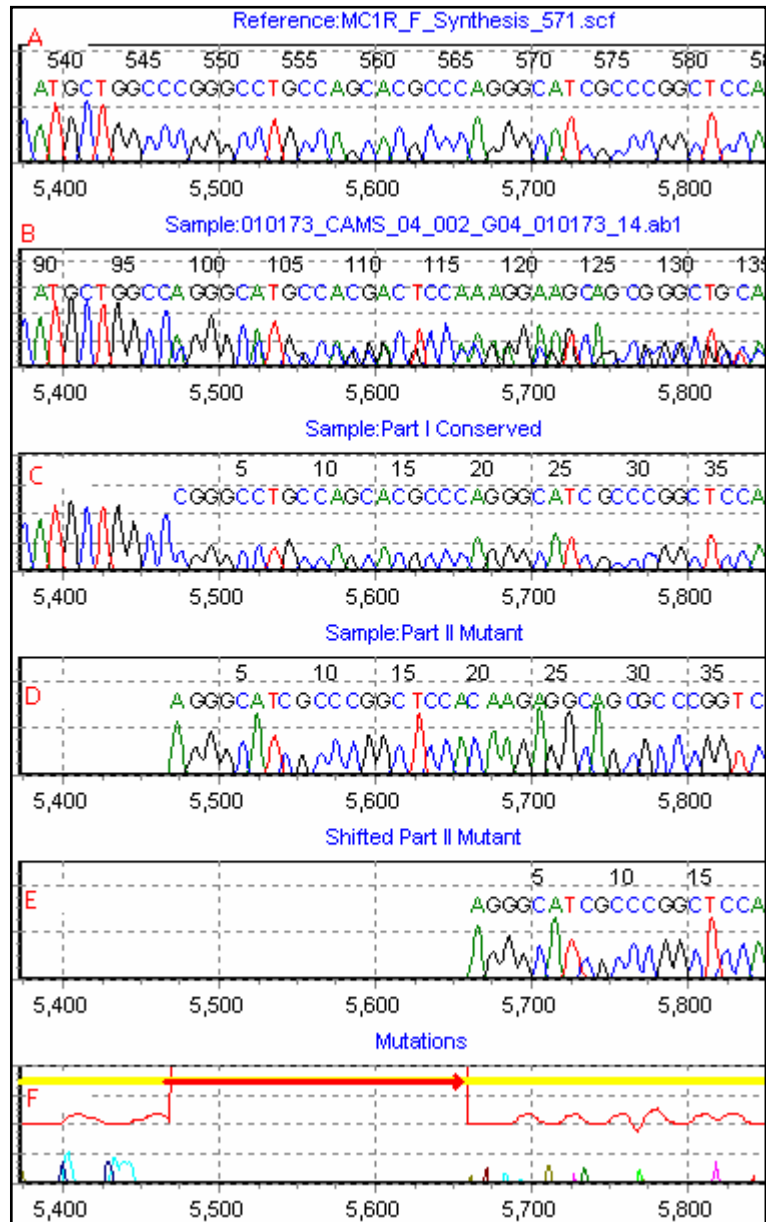


Figure 2: Heterozygous Indel Deconvolution page separates complex samples into two clean traces automatically.

References:

1. C.-S. J. Liu, S. Ni, "Sequence trace recovery of heterozygous indels using wide-type sequence trace as reference." July 2003, Mutation Detection Meeting, Australia. <http://www.genomic.unimelb.edu.au/PC%20pre%20proof.pdf>
2. Haines, et al, "Complement Factor H Variant Increases the Risk of Age Related Macular Degeneration", 2005. *Science*, **308**, 419-21.
3. Nickerson, D.A., Tobe, V.O., and Taylor, S.L. "PolyPhred: automating the detection and genotyping of single nucleotide substitutes using fluorescence-based resequencing." 1997. *Nucleic Acids Research*. **25**, 2745-2751.
4. Samuels et al. "High Frequency of Mutations of the PIK3CA Gene in Human Cancers." 2004. *Science*. **304**, 554.
5. Paez et al, "EGFR mutations in lung cancer..." 2004. *Science*. **304**, 1497-500.

SoftGenetics LLC □ 200 Innovation Blvd □ Suite 241 □ State College, PA 16803

Phone: 814/237-9340, Fax: 814/237-9343

www.softgenetics.com, tech_support@softgenetics.com