

RD-Connect Training:
Variant annotation and interpretation
with Alamut®

Séverine Lair - André Blavier
Interactive Biosoftware

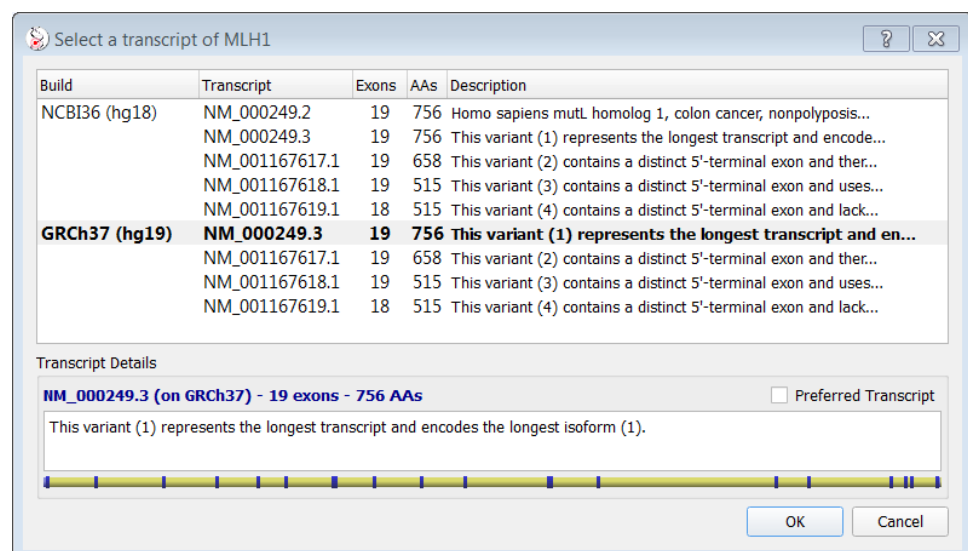
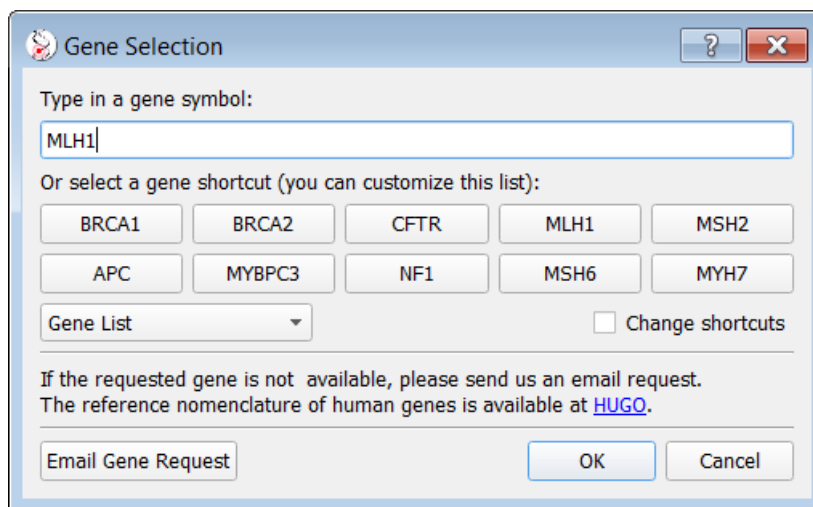
Alamut was designed to help you in this interpretation by entering a variant

- Which type of variant is it?
- Has this sequence variation been previously reported?
- Is there a functional impact for this variant?

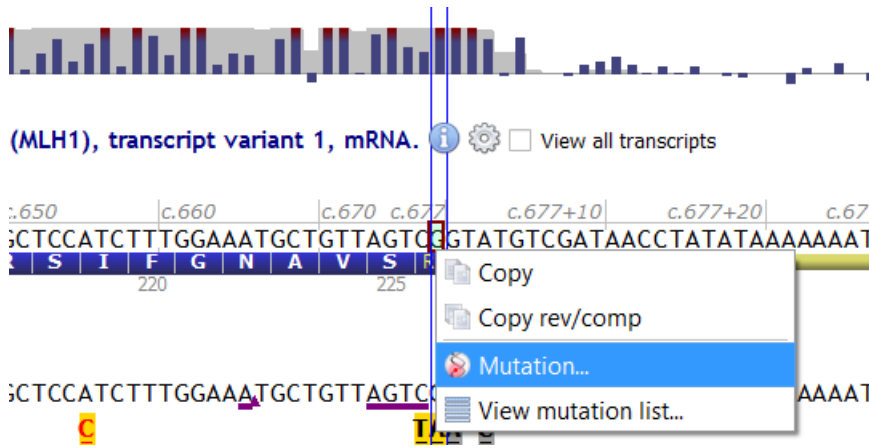
A documentation is available from Alamut by clicking on “Help” and on “Software documentation” from the menu bar or by clicking on the info icon ⓘ.

1. Interpretation of a sequence variation

- Load the gene “MLH1” and choose the transcript “NM_000249.3”.



- Go to the position g.37055942 or c.677 in the exon 8 of this transcript.
- Left click to select the nucleotide “G” in the transcript track.
- Right click and then select “Mutation...”.



(MLH1), transcript variant 1, mRNA. View all transcripts

c.650 | c.660 | c.670 | c.677 | c.677+10 | c.677+20 | c.677

CTCCATCTTTGGAAATGCTGTTAGTCGGTATGTCGATAACCTATATAAAAAAAT

S I F G N A V S

Copy

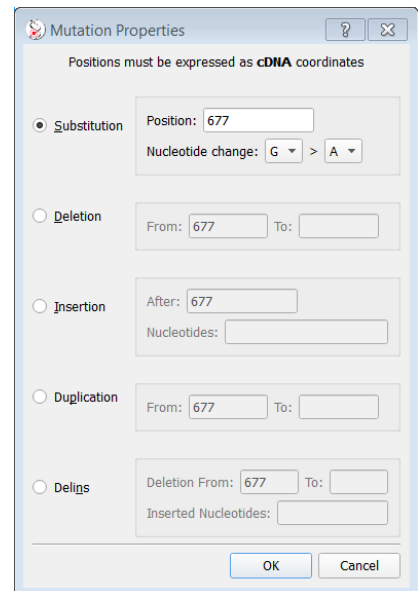
Copy rev/comp

Mutation...

View mutation list...

A first mutation window, named “Mutation Properties”, appears to define the property of the observed mutation

- Enter that the nucleotide “G” is changed to a “A”.



Mutation Properties

Positions must be expressed as cDNA coordinates

Substitution Position: 677 Nucleotide change: G > A

Deletion From: 677 To:

Insertion After: 677 Nucleotides:

Duplication From: 677 To:

Deletins Deletion From: 677 To: Inserted Nucleotides:

OK Cancel

- Click on “OK”.



The “mutation interpretation” window appears. This window allows you to have all current known relevant information for this variant.

Let us have a look at the different parts:

The “variant features” part

- What is the HGVS nomenclature for this variant?
- What is the type of this variant?
- What is its coding effect?

The “known variations” part

- Is it a known variant?
- If so in which databases is this variant known?
- Have you got information about its allele frequency?

External links to PubMed Abstracts, Locus specific databases... are available to check quickly whether it is previously described in other sources.



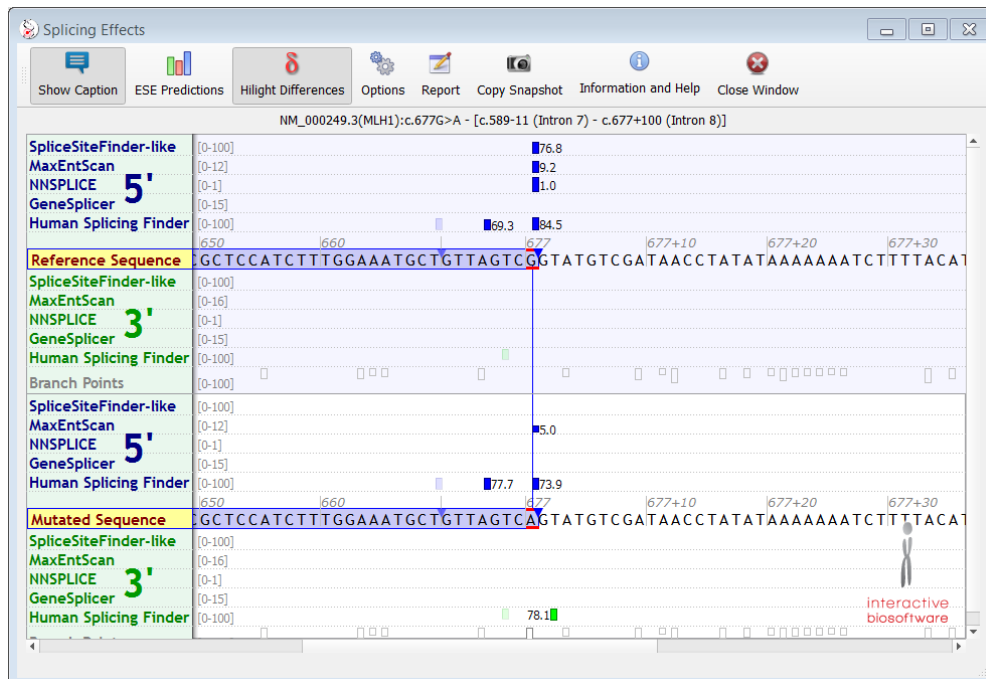
The prediction tool part can also be used to assist in variant classification.

The “missense predictions” with automatic prediction scores computed from several relevant tools can help you to predict the effect of amino acid changes.

- What are the pre-computed prediction given by the tools?
- What is the prediction given by PolyPhen2?
- What is the prediction given by KD4v?

The “splicing predictions” with a graphic view can help you to predict a splicing effect such as creation or loss of site.

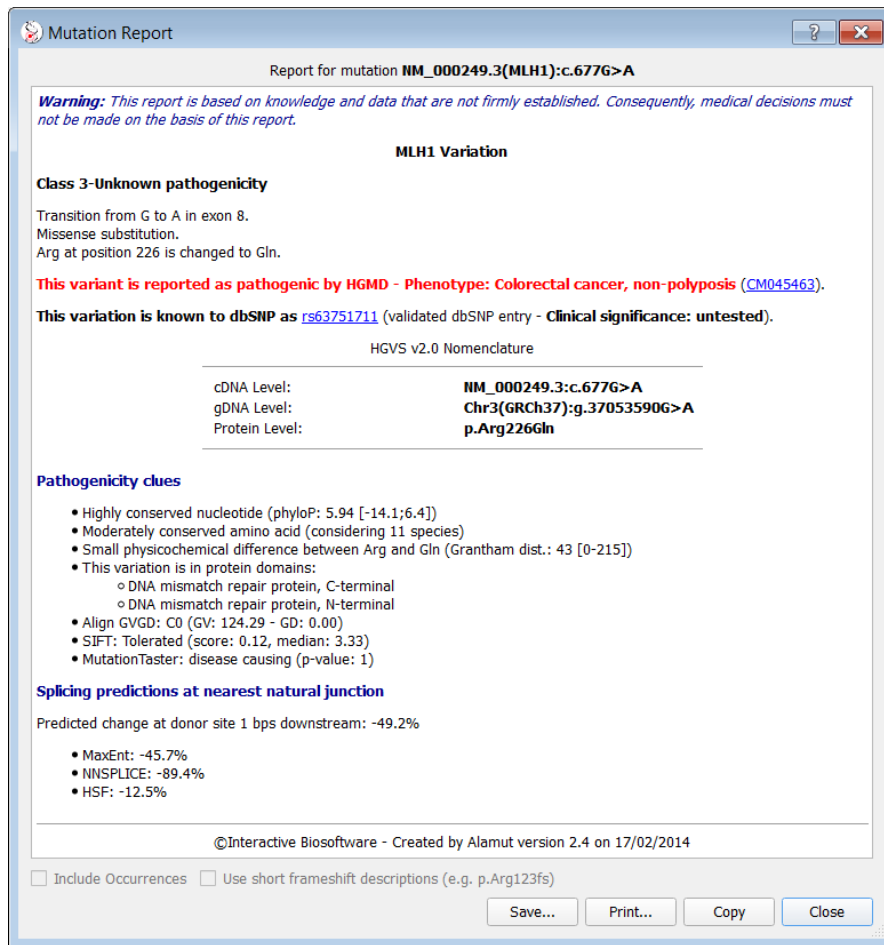
- What is the predicted splicing effect of the variant?
- Click on “Splicing window” and on “Highlight Differences” to show differences between the wild type sequence and the mutated sequence.
- What do you observe?



- Go back to the “Mutation interpretation” window.
- Select the appropriate pathogenicity class for the variant.

Alamut does not classify automatically a variant into a pathogenicity class. By default, the value is “unknown pathogenicity in the “5 classes”- classification” *ie* a VUS (variant of uncertain clinical significance).

- Edit a summary report.



Mutation Report

Report for mutation **NM_000249.3(MLH1):c.677G>A**

Warning: This report is based on knowledge and data that are not firmly established. Consequently, medical decisions must not be made on the basis of this report.

MLH1 Variation

Class 3-Unknown pathogenicity

Transition from G to A in exon 8.
Missense substitution.
Arg at position 226 is changed to Gln.

This variant is reported as pathogenic by HGMD - Phenotype: Colorectal cancer, non-polyposis (CM045463).

This variation is known to dbSNP as rs63751711 (validated dbSNP entry - Clinical significance: untested).

HGVS v2.0 Nomenclature

cDNA Level:	NM_000249.3:c.677G>A
gDNA Level:	Chr3(GRCh37):g.37053590G>A
Protein Level:	p.Arg226Gln

Pathogenicity clues

- Highly conserved nucleotide (phyloP: 5.94 [-14.1;6.4])
- Moderately conserved amino acid (considering 11 species)
- Small physicochemical difference between Arg and Gln (Grantham dist.: 43 [0-215])
- This variation is in protein domains:
 - DNA mismatch repair protein, C-terminal
 - DNA mismatch repair protein, N-terminal
- Align GVGD: C0 (GV: 124.29 - GD: 0.00)
- SIFT: Tolerated (score: 0.12, median: 3.33)
- MutationTaster: disease causing (p-value: 1)

Splicing predictions at nearest natural junction

Predicted change at donor site 1 bps downstream: -49.2%

- MaxEnt: -45.7%
- NNSPLICE: -89.4%
- HSF: -12.5%

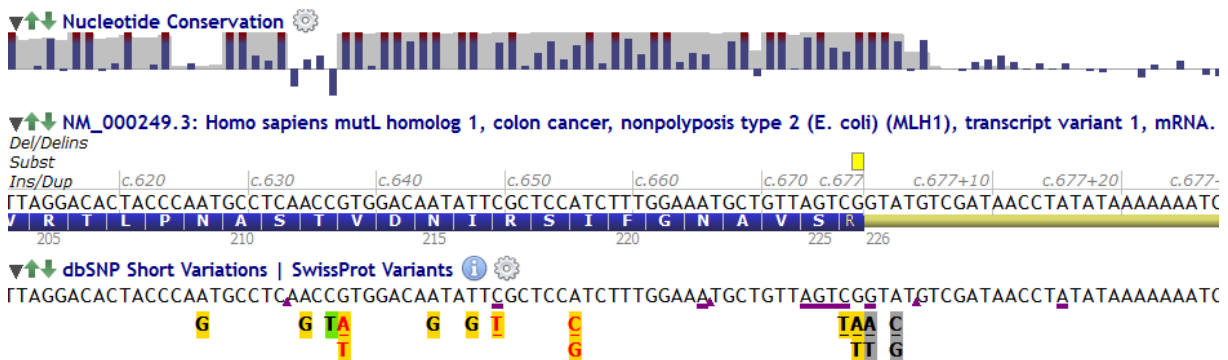
©Interactive Biosoftware - Created by Alamut version 2.4 on 17/02/2014

Include Occurrences Use short frameshift descriptions (e.g. p.Arg123fs)

Save... Print... Copy Close

- Click on “Close”.
- Go back to the “Mutation interpretation” window.
- Click on “Save” to save your annotated variant in Alamut.
- Where is the variant saved?
- Why is your variant yellow? Have you an idea on which information the color is associated to?

Your variant will appear on the transcript track for further analyses.



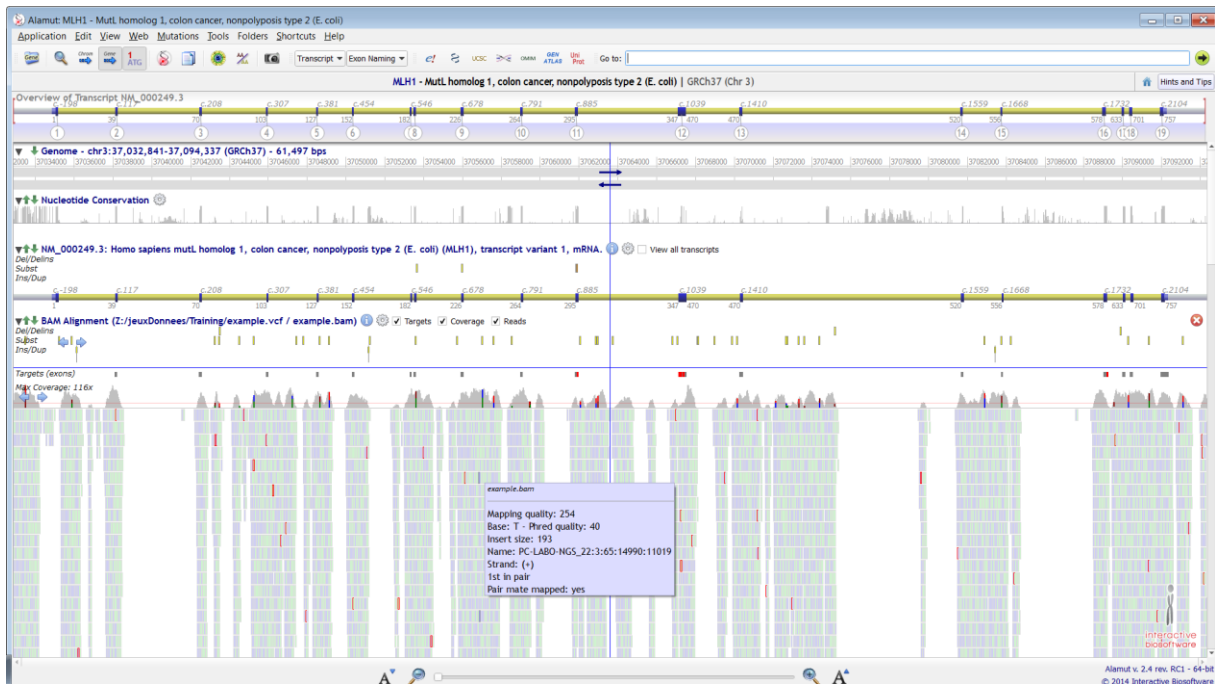


2. Working with NGS data

Alamut allows you to visualize BAM (read alignments), VCF (variant calls) and BED files from NGS analyses.

- Select the MLH1 gene and then the transcript “NM_000249.3”.
- Click on “Application” then on “Load BAM file from URL”
- Paste this url <http://rd-connect.interactive-biosoftware.com/BAM/example.bam>

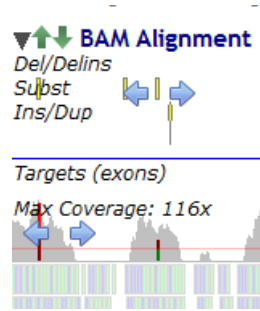
The BAM file is loaded in a new track.



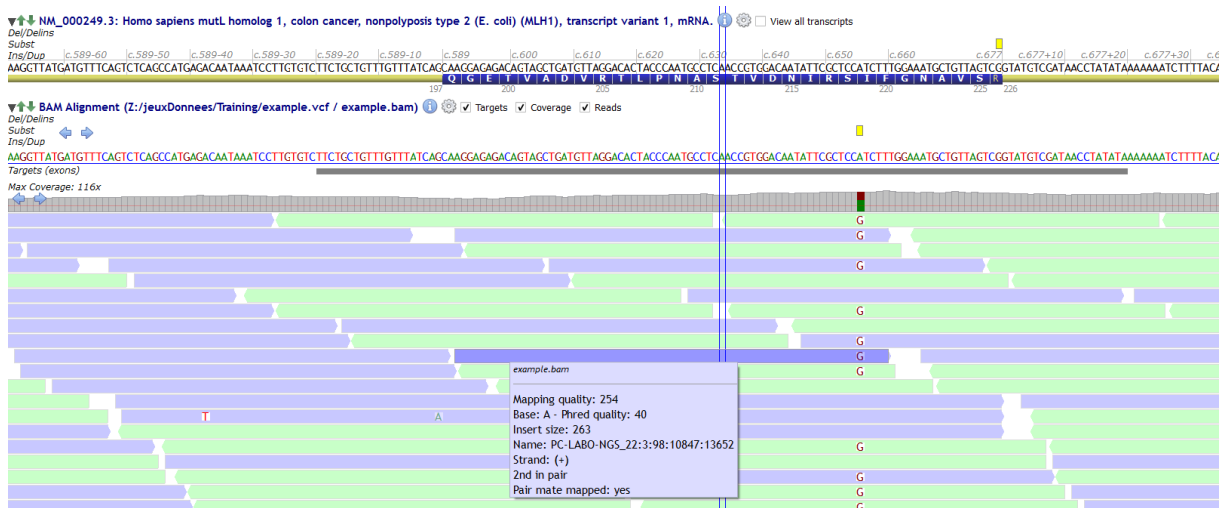
- What is the maximum depth of coverage for the MLH1 gene locus?
- What does the red line correspond to? What is the value marked by this red line?
- Which regions are tagged in the “Targets” sub-track by default?
- Click on the wheel icon from the BAM track then on “Load VCF file from URL”.
- Paste this url <http://rd-connect.interactive-biosoftware.com/BAM/example.vcf>

The called variants reported in the VCF file are loaded in the BAM track as a sub-track.

- Jump from one variant to the other by clicking the arrow on the left.



- Click on the exon 8 from the locus track to have a look at it.

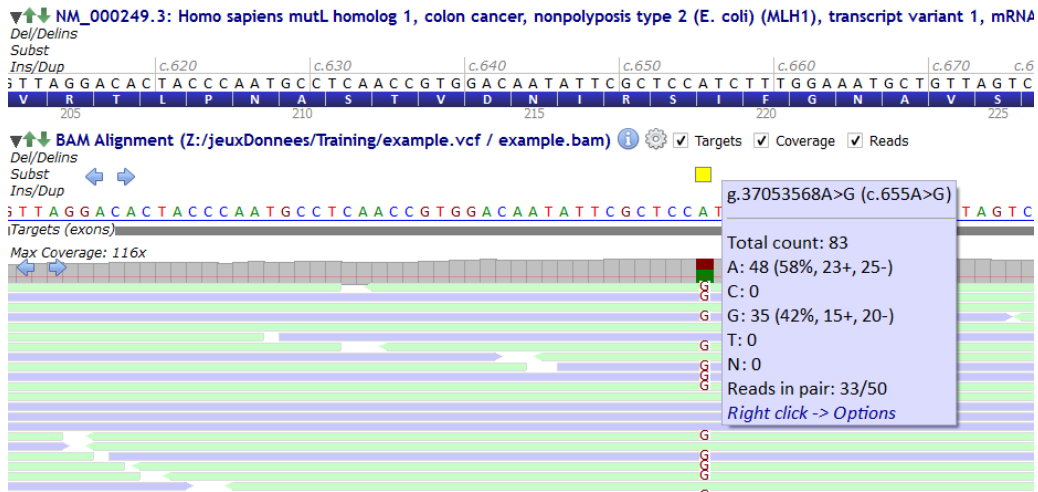


- Why are some reads blue or green?
- What do you observe at the position g.37053568?

Hover over a read to learn more information about it.

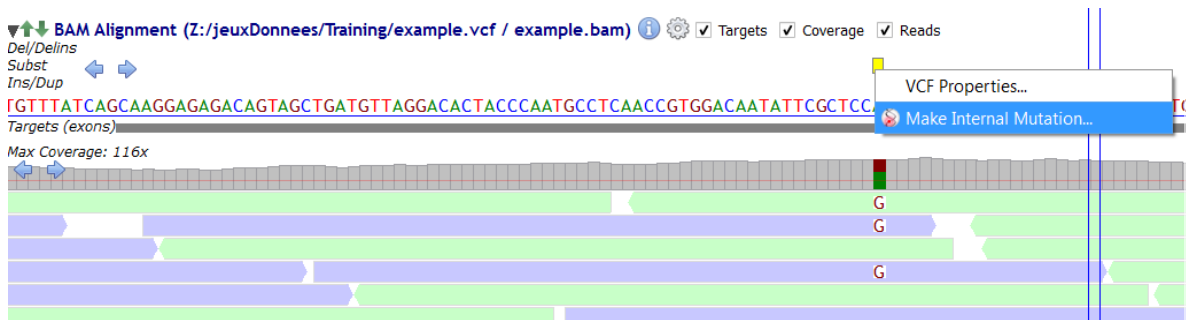
- What is the depth of coverage at the position g.37053568?
- What is the percentage of the reference allele observed at the position g.37053568?
- What is the percentage of the alternative allele observed at the position g.37053568?
- How many reads, aligned on the forward strand, hold the alternative allele?

Several display options about bases (show all or not) and reads (condense or squish views) are available when you right click on them.

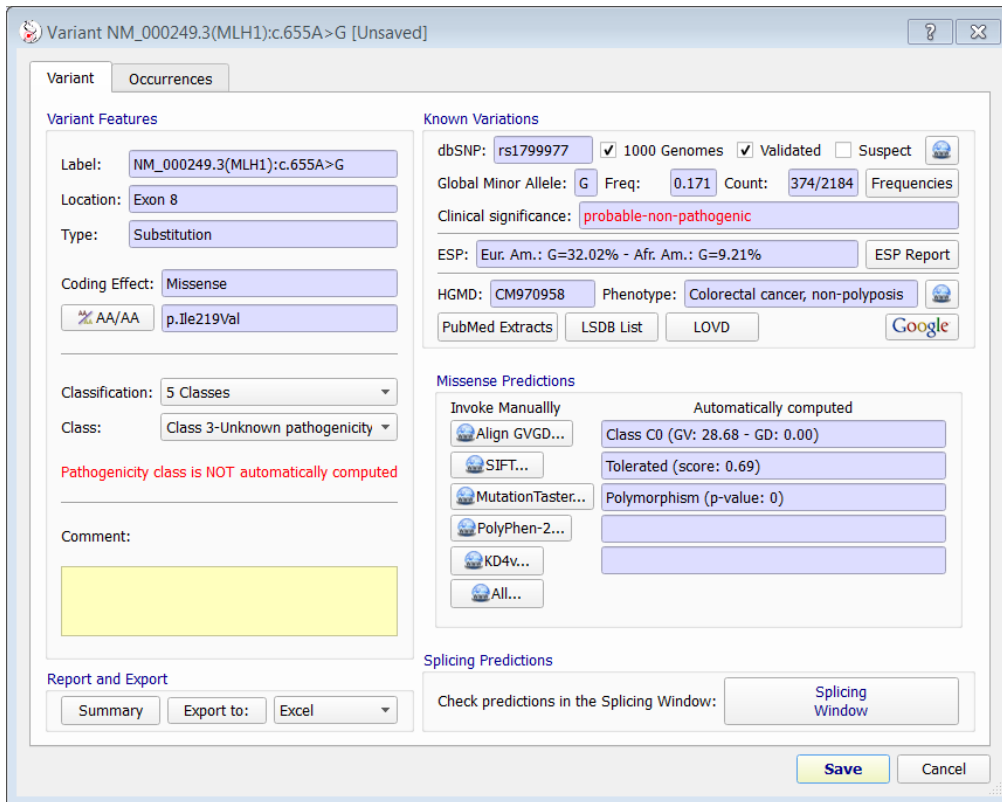


The next step is to interpret this observed variant.

- Left click to select the called variant from the VCF in the BAM track.
- Right click and then select “Make Internal Mutation...”



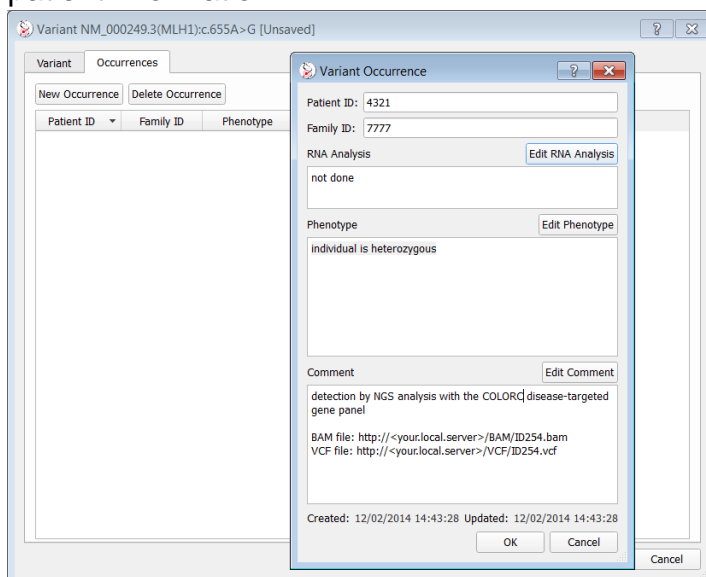
The “mutation interpretation” window appears.



- Interpret the variant for this patient.
- Annotate the variant for this patient.

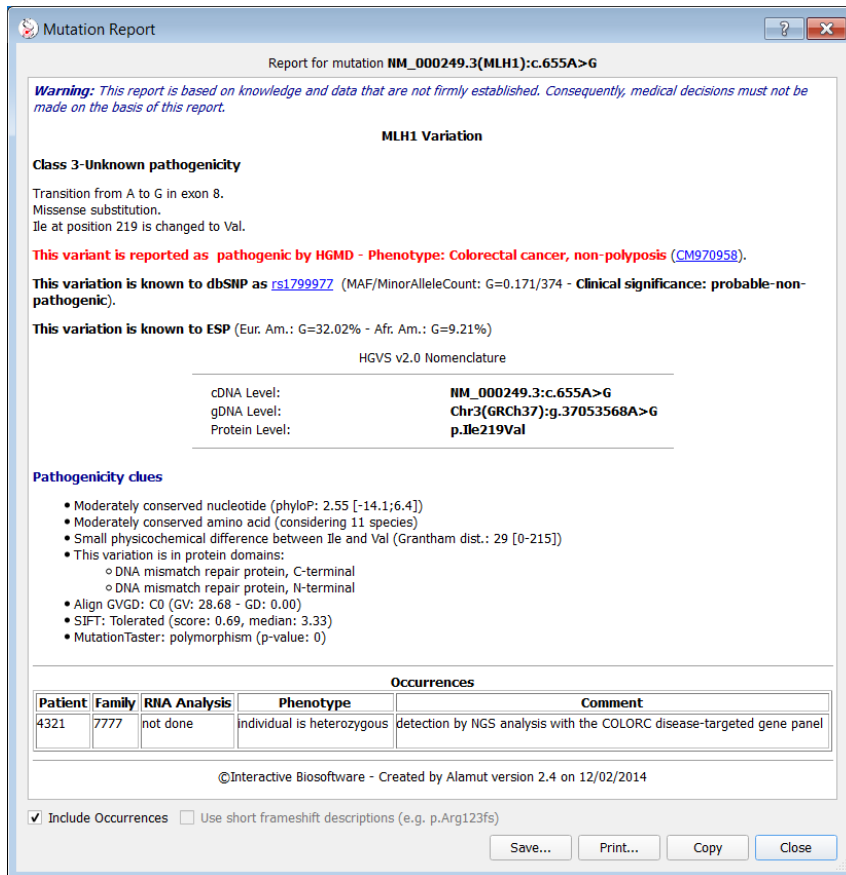
To do that:

- Click on the “Occurrences” then on “New Occurrence” and enter the following patient information:



- Click on “OK” to save the record.
- Click on “Variant” to go back to the “Mutation interpretation” window.

- Edit a summary report.



Mutation Report
Report for mutation **NM_000249.3(MLH1):c.655A>G**

Warning: This report is based on knowledge and data that are not firmly established. Consequently, medical decisions must not be made on the basis of this report.

MLH1 Variation

Class 3-Unknown pathogenicity
Transition from A to G in exon 8.
Missense substitution.
Ile at position 219 is changed to Val.

This variant is reported as pathogenic by HGMD - Phenotype: Colorectal cancer, non-polyposis (CM970958).

This variation is known to dbSNP as rs1799977 (MAF/MinorAlleleCount: G=0.171/374 - **Clinical significance: probable-non-pathogenic**).

This variation is known to ESP (Eur. Am.: G=32.02% - Afr. Am.: G=9.21%)

HGVS v2.0 Nomenclature

cDNA Level:	NM_000249.3:c.655A>G
gDNA Level:	Chr3(GRCh37):g.37053568A>G
Protein Level:	p.Ile219Val

Pathogenicity clues

- Moderately conserved nucleotide (phyloP: 2.55 [-14.1;6.4])
- Moderately conserved amino acid (considering 11 species)
- Small physicochemical difference between Ile and Val (Grantham dist.: 29 [0-215])
- This variation is in protein domains:
 - DNA mismatch repair protein, C-terminal
 - DNA mismatch repair protein, N-terminal
- Align GVGD: C0 (GV: 28.68 - GD: 0.00)
- SIFT: Tolerated (score: 0.69, median: 3.33)
- MutationTaster: polymorphism (p-value: 0)

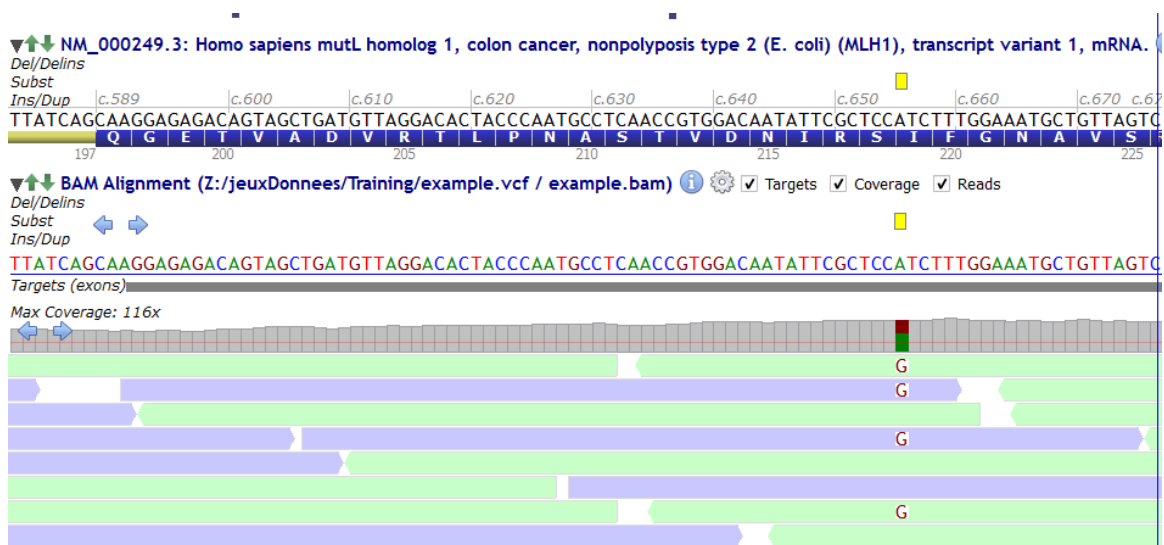
Occurrences				
Patient	Family	RIA Analysis	Phenotype	Comment
4321	7777	not done	individual is heterozygous	detection by NGS analysis with the COLORC disease-targeted gene panel

©Interactive Biosoftware - Created by Alamut version 2.4 on 12/02/2014

Include Occurrences Use short frameshift descriptions (e.g. p.Arg123fs)

Save... Print... Copy Close

- Save your annotated variant in Alamut.



▼▲ NM_000249.3: Homo sapiens mutL homolog 1, colon cancer, nonpolyposis type 2 (E. coli) (MLH1), transcript variant 1, mRNA.

Del/Delins
Subst
Ins/Dup

TTATCAGCAAGGAGAGACAGTAGCTGATGTTAGGACACTACCCAAATGCCTCAACCGTGGACAATATTCGCTCCATCTTTGAAATGCTGTTAGTC

197 200 205 210 215 220 225

▼▲ BAM Alignment (Z:/jeuxDonnees/Training/example.vcf / example.bam) [i] [g] [v] Targets [x] Coverage [x] Reads

Del/Delins
Subst
Ins/Dup

TTATCAGCAAGGAGAGACAGTAGCTGATGTTAGGACACTACCCAAATGCCTCAACCGTGGACAATATTCGCTCCATCTTTGAAATGCTGTTAGTC

Targets (exons)

Max Coverage: 116x

G
G
G
G