NGS in diagnostics:

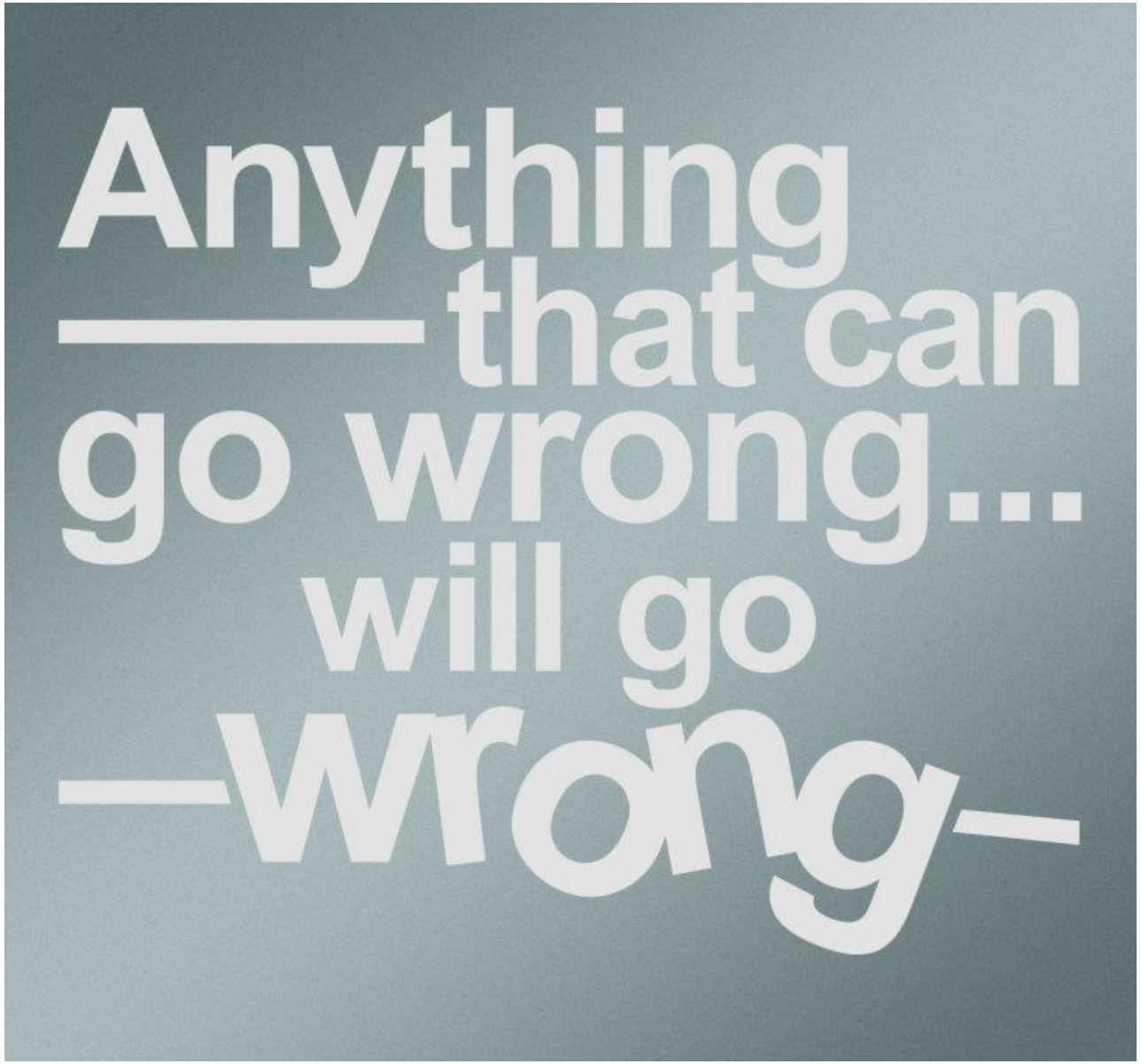
where things can go wrong

Johan den Dunnen and Anna Benet-Pagès

VEPTC – Johor 2018

NGS Workflow

- Sampling
- Sample preparation
- Sequencing
- Mapping
- Variant Calling
- Variant Annotation
- Checks
- Variant Interpretation



NGS Workflow

- Sampling
- Sample preparation
- Sequencing
- Mapping
- Variant Calling
- Variant Annotation
- Checks
- Variant Interpretation

Sample Swaps and contamination - frequency and most common reasons

Pre-analytical (collection, purchase)	Analytical (lab. Analysis)	Post-analytical (assessment – reporting)	
60%	**19%**	**15%**	**Molecular genetic testing 88.394 patients , 1 year**

Percentage of laboratory errors by processing phase; *Green, 2013*

- LIMS
- Barcoding
- Double check (2-man rule)
- Automated liquid handling
- Paralell sample identification analysis (second methodology)
- Programming of system interfaces

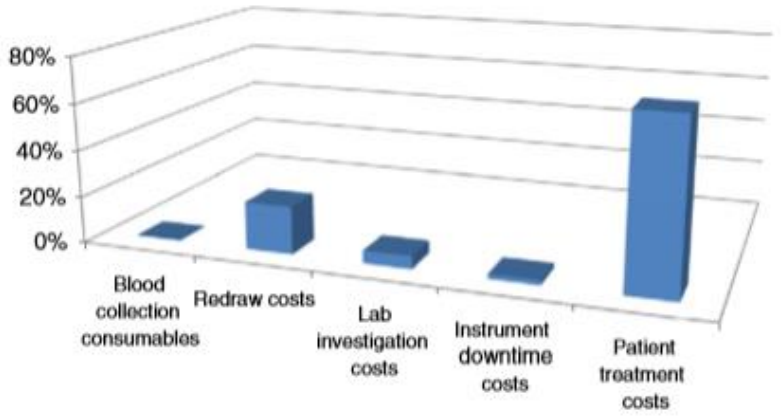
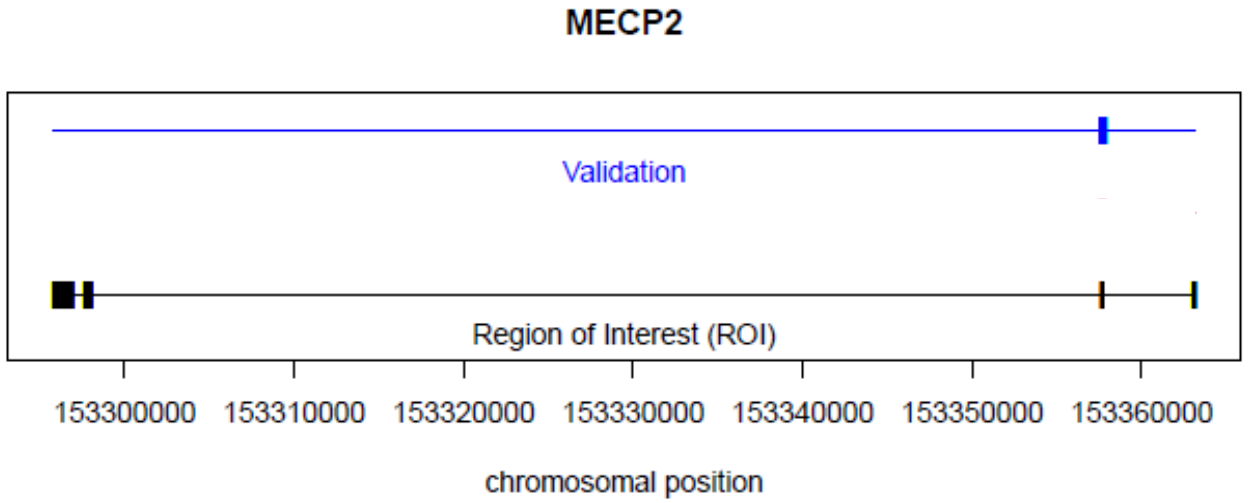
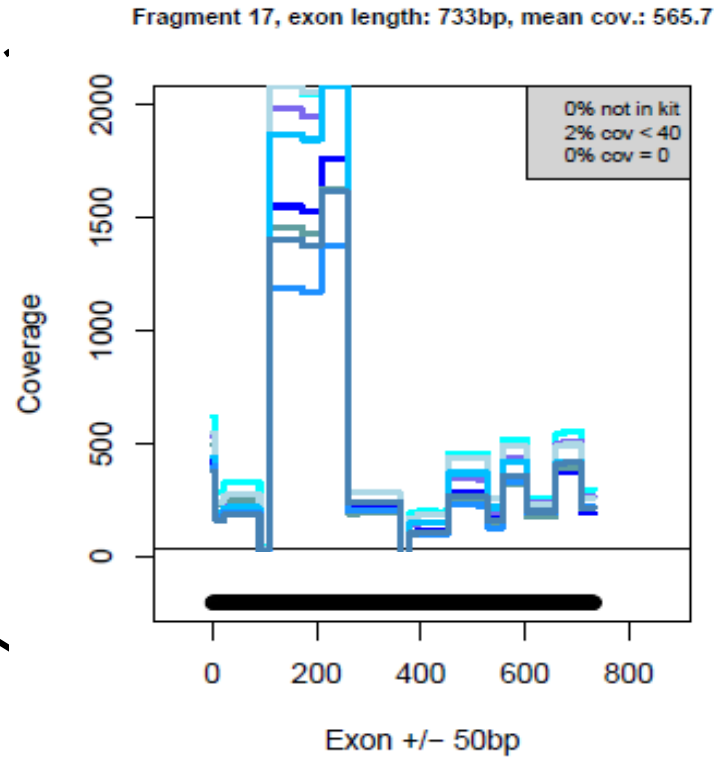


Fig. 2. The total cost of specimen rejection can be quantified by cost category.

Sample preparation – PCR-based – regions not amplified

Challenge: Primer design in GC, repeats, SNPs
Consequences: amplicon amplification reproducibility low
allelic drop out
uneven coverage





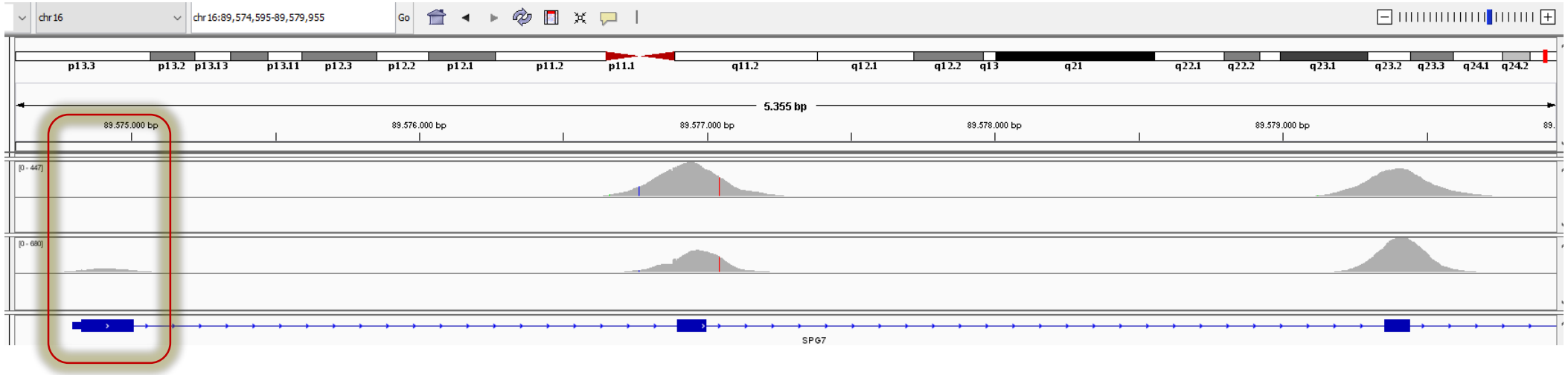
Region of interest: exons we want to analyze within the panel

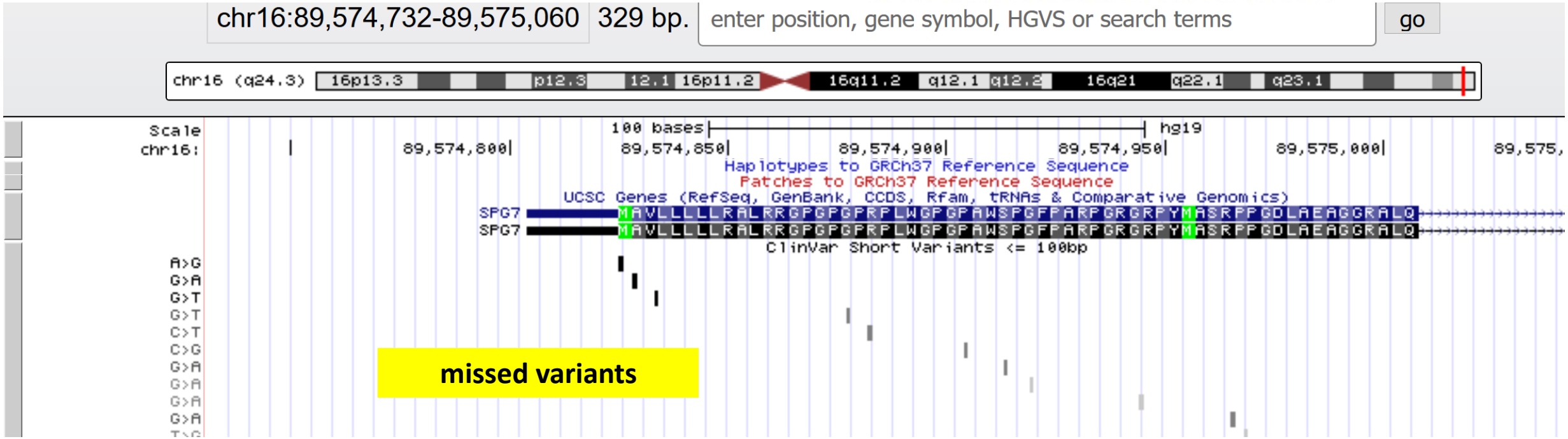
Validation: covered regions from the NGS results

coverage of a single amplicon

CNV detection not possible

Sample preparation – capture-based – uneven coverage



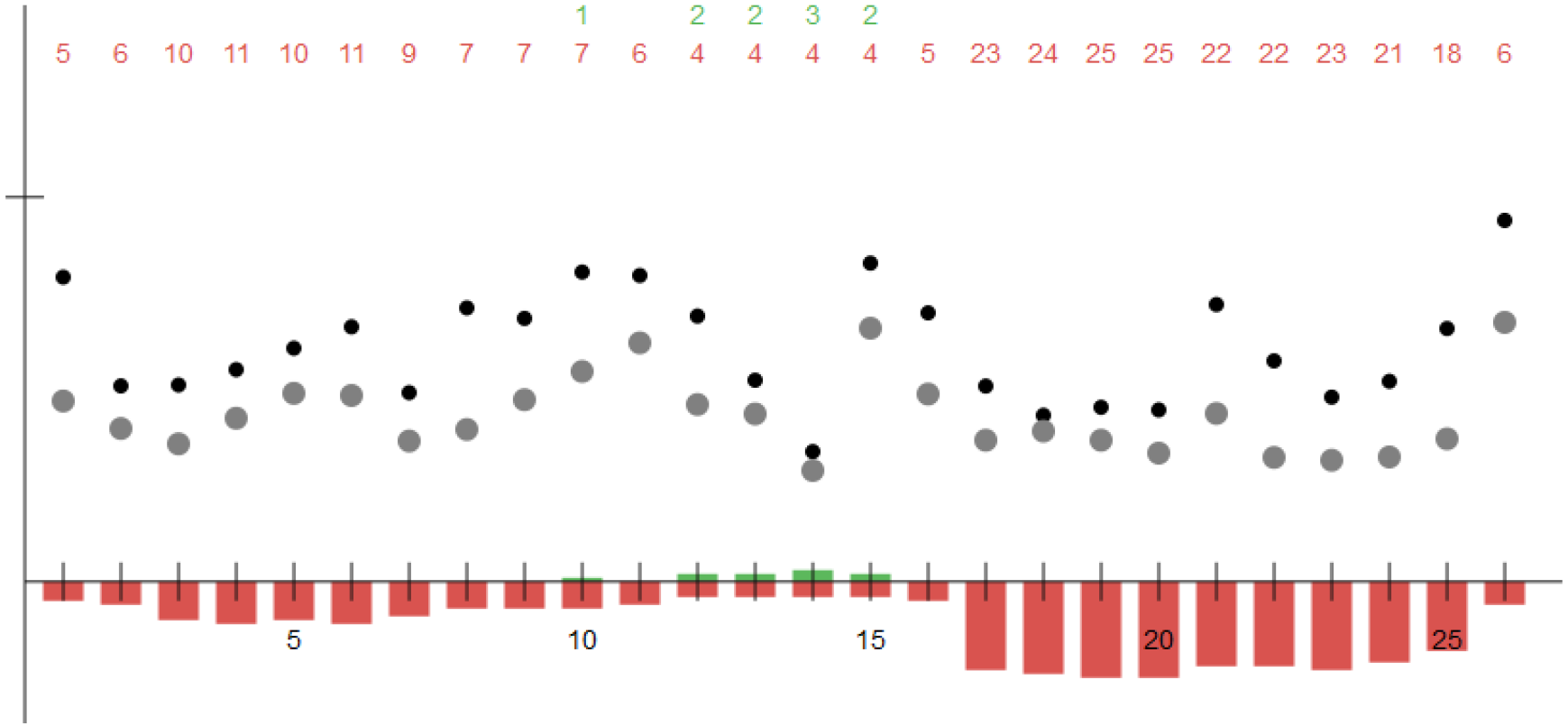


Sample preparation – uneven coverage – false CNV calls

2. SCN1A

Number of exons	26
Duplications in	5 exons and 4 calls (frequency = 0.00)
Deletions in	26 exons and 39 calls (frequency = 0.01)

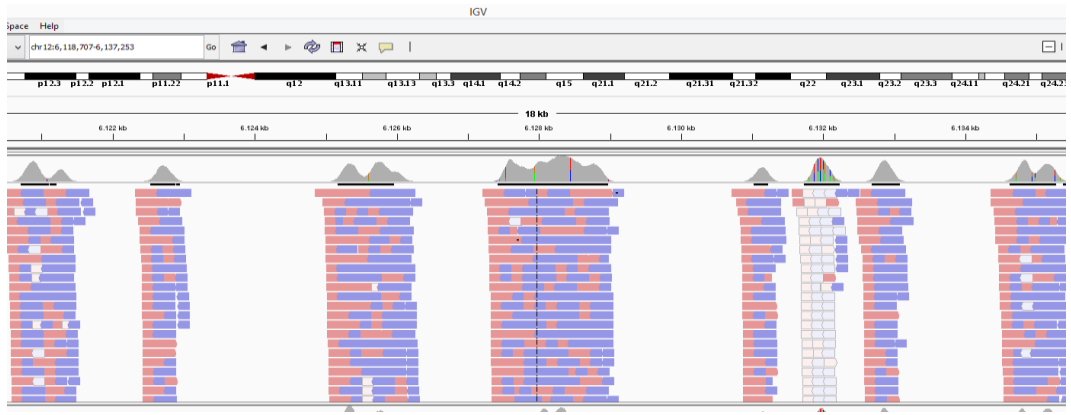
False positive deletion calls in the *SCN1A* gene



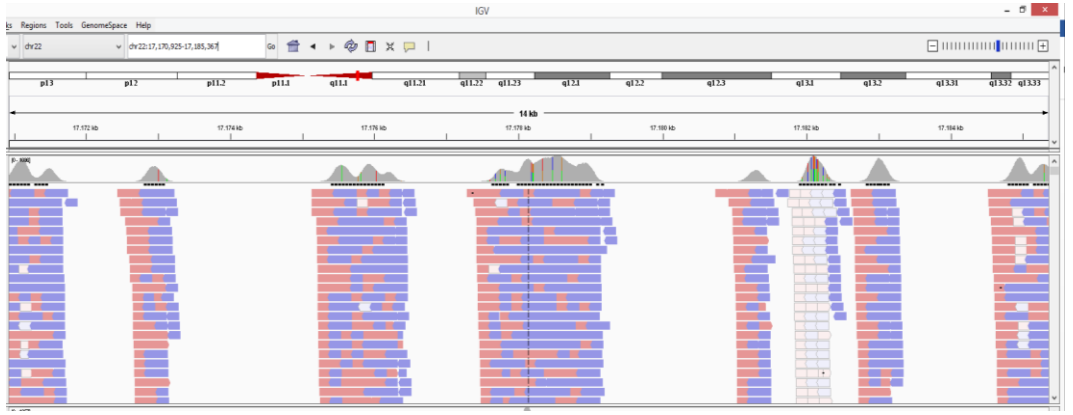
Sample preparation - capture – homologues sequences

VWF gene causes „*von Willebrand Syndrom*“:

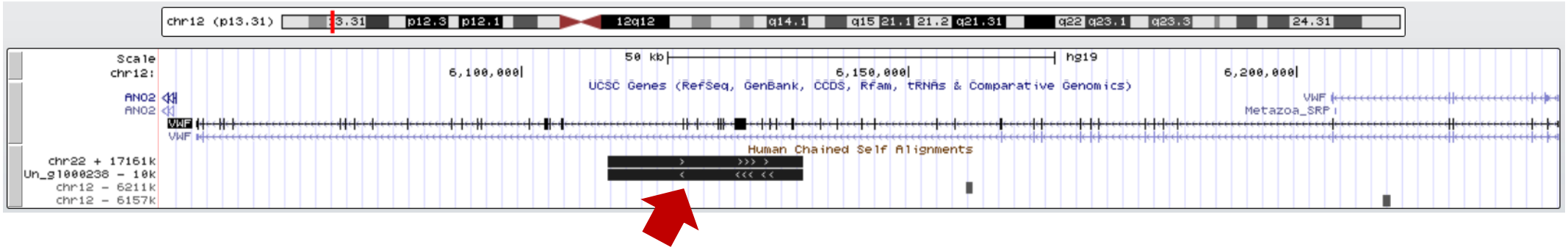
Pathogenic mechanismus: SNVs (~90%), CNVs (~10%), 6-335 bp geneconversions



VWF – von Willebrand factor gene
chr12:6,118,707-6,137,253

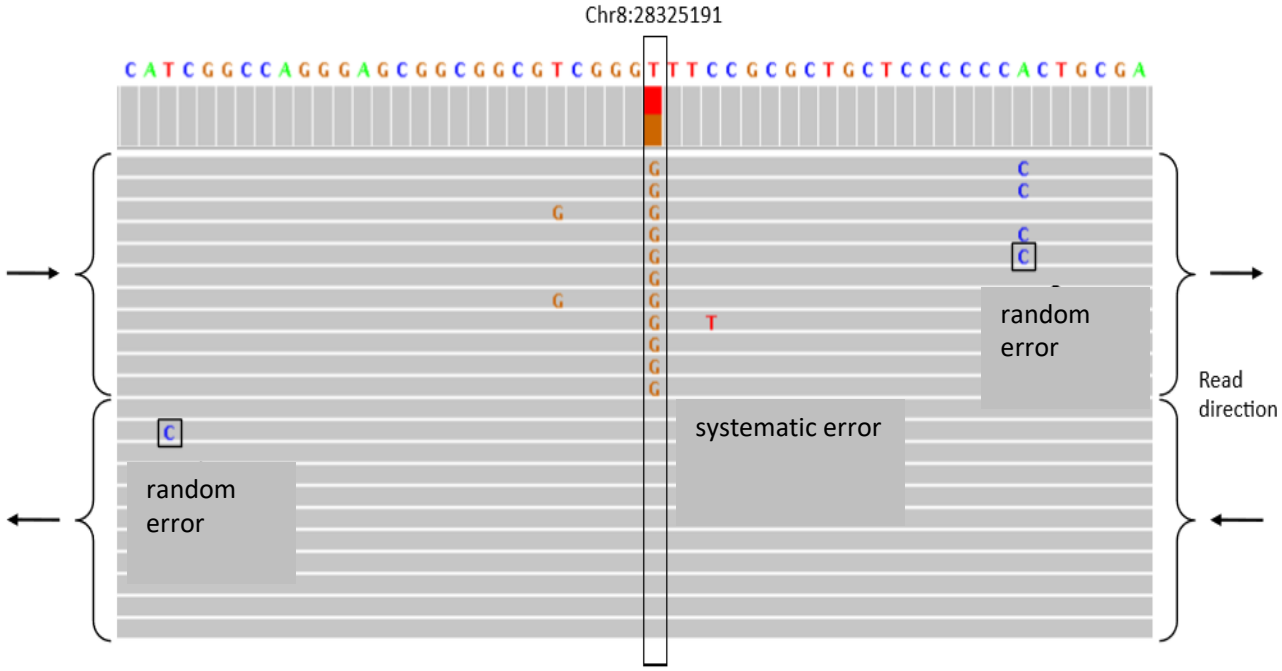


VWFP1 - von Willebrand factor pseudogene 1
chr22:17,170,925-17,185,367

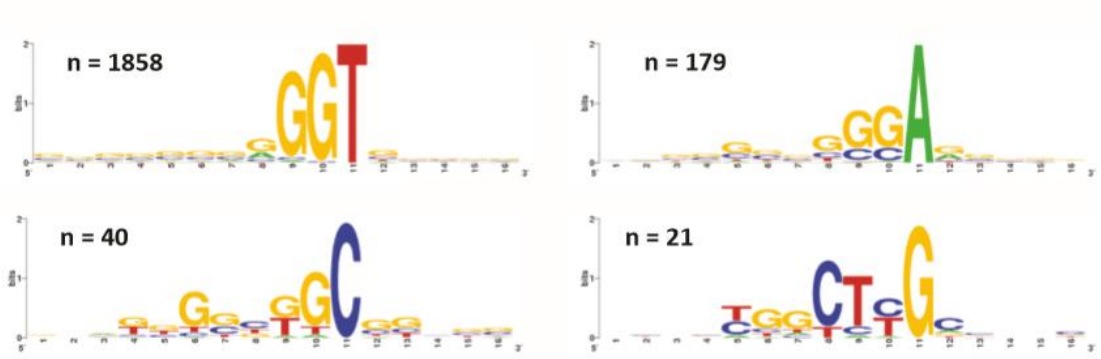


Sequencing – systematic errors – occur in 1 in 1000 base pairs

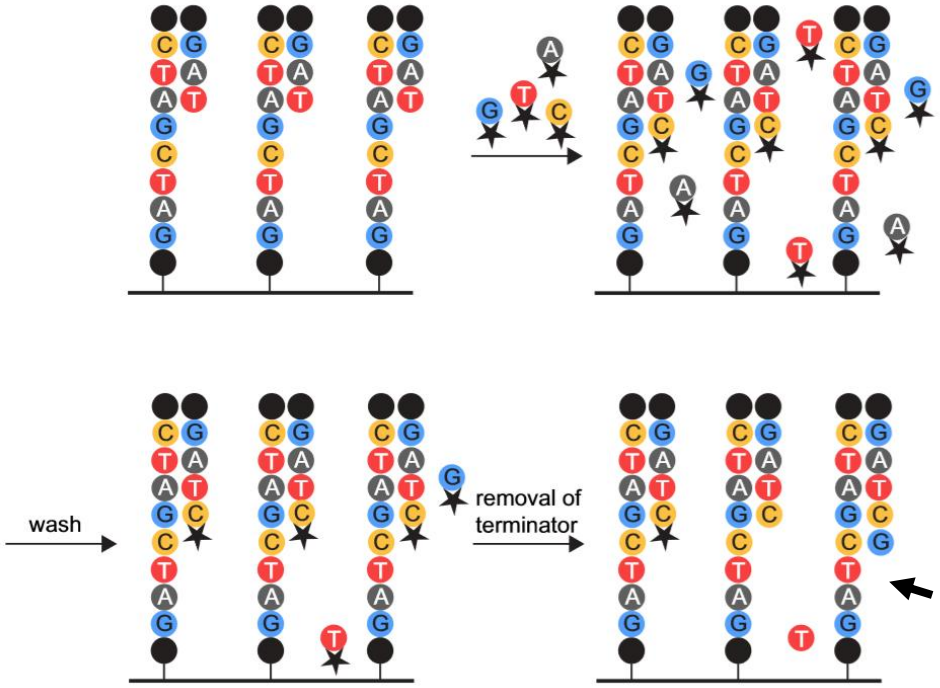
- base-call errors aggregate at specific genomic locations across multiple sequence reads



motiv around systematic errors



Origin phasing effect



Pfeifer et al.; 2018

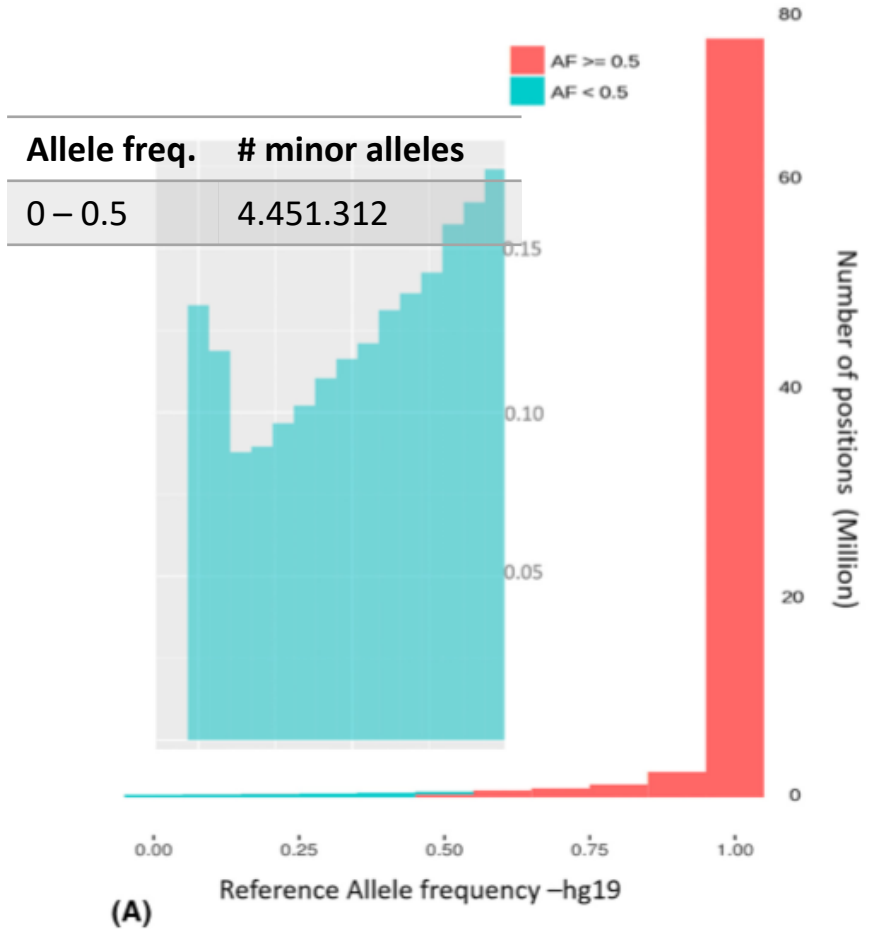
NGS Workflow

- Sampling
- Sample preparation
- Sequencing
- Mapping
- Variant Calling
- Variant Annotation
- Checks
- Variant Interpretation

reference genome – minor alelle - variants missed

- hg19 reference assembly do not represent major alleles at all three billion positions
- 7% of false negative calls and 30% of false positive calls

Allele freq.	# minor alleles
0 – 0.5	4.451.312



Hg19	A (minor allele)	A (minor allele)	A (minor allele)
sample	A	A/T	T
call	No call	Call T (major alelle)	Call T (major alelle)

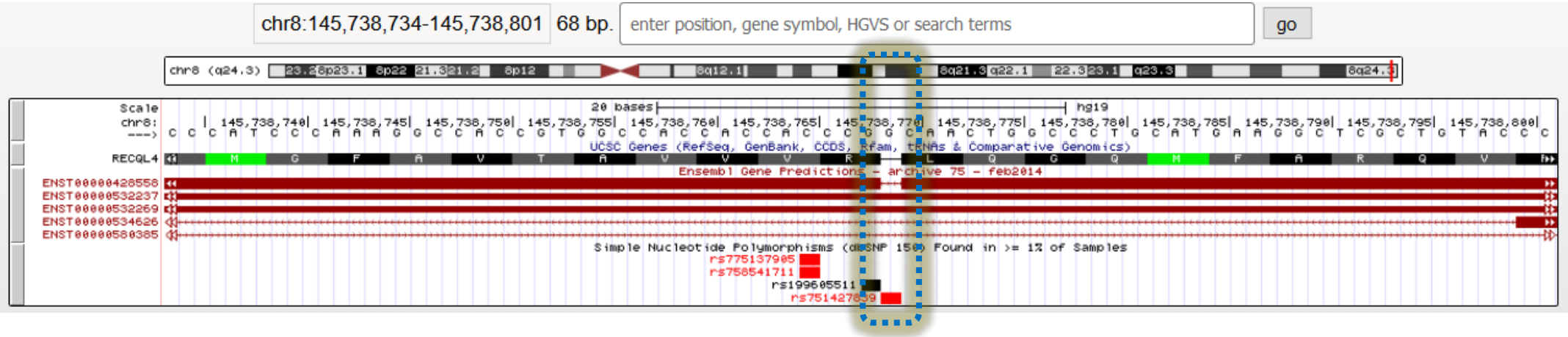
chr	Nucleotide substitution	Gene	Amino acid substitution	Alternate allele frequency
chr2	g.130949411T>G	*TUBA3E*	p.(Glu449Ala)	0.020567
chr6	g.29364787T>C	*OR12D2*	p.(Phe104Ser)	0.021166
chr6	g.33382288A>G	*PHF1*	p.(Lys304Arg)	0.022364
chr2	g.231149108A>G	*SP140*	p.(Lys516Glu)	0.026158
chr17	g.37101380C>T	*FBXO47*	p.(Arg209Gln)	0.027955
chr14	g.24901276T>G	*KHNYN*	p.(Leu270Trp)	0.040535
chr16	g.1389153A>C	*BAIAP3*	p.(Thr87Pro)	0.04373
chr19	g.29704010C>A	*UQCRFS1*	p.(Ala6Ser)	0.046126
chr22	g.22989256G>A	*GGTLC2*	p.(Gly70Glu)	0.046925
chr16	g.1370597C>G	*UBE2I*	p.(Ser164Arg)	0.049521
chr16	g.1370614G>C	*UBE2I*	p.(Arg170Thr)	0.049521
chr21	g.37617630G>T	*DOPEY2*	p.(Gly1118Cys)	0.050719
chr9	g.140130606T>A	*SLC34A3*	p.(Val513Glu)	0.053914
chr6	g.29141743G>A	*OR2J2*	p.(Ala111Thr)	0.058506
chr6	g.42666145T>C	*PRPH2*	p.(Lys310Arg)	0.058706
chr2	g.180810264T>A	*CWC22*	p.(Arg773Ser)	0.06869
chr17	g.80895933G>A	*TBCD*	p.(Gly1135Glu)	0.069089
chr2	g.96795857C>T	*ASTL*	p.(Arg222Gln)	0.072883
chr14	g.36789729G>T	*MBIP*	p.(Ser22Arg)	0.073083
chr3	g.12046364C>G	*SYN2*	p.(Pro37Ala)	0.073682
chr6	g.29911256G>T	*HLA-A*	p.(Glu185Asp)	0.074281
chr11	g.111749349T>A	*FDXACB1*	p.(Asn87Ile)	0.07528
chr2	g.44104925C>T	*ABCG8*	p.(Ala632Val)	0.077077
chrX	g.65382685C>T	*HEPH*	p.(Ala39Val)	0.079205
chr19	g.56047448G>A	*SBK2*	p.(Arg72Cys)	0.082069
chrX	g.88008423C>A	*CPXCR1*	p.(Ser3Try)	0.086358
chr19	g.36497358G>C	*SYNE4*	p.(His278Gln)	0.08726
chrX	g.84563135A>T	*POF1B*	p.(Leu349Met)	0.087947
chr17	g.72938100C>T	*OTOP3*	p.(Pro119Ser)	0.09385
chr5	g.741736T>G	*ZDHHC11B*	p.(Asp314Ala)	0.095248
chr5	g.140559596G>T	*PCDHB8*	p.(Val661Leu)	0.097644
chr1	g.155026942C>A	*ADAM15*	p.(Thr191Lys)	0.098043
chr1	g.34330067C>A	*HMGB4*	p.(Ala92Glu)	0.098842
chr7	g.150500729G>A	*TMEM176A*	p.(Ala122Thr)	0.09984

Lists 34 nonsynonymous mutations with allele frequencies <10%. Column 2 represents the nucleotide substitution as per hg19K reference. Column 3 represents the gene in which the mutation lies, and Column 4 gives the corresponding amino acid substitution. Column 5 is the allele frequency for the alleles in hg19.

Karthikeyan et al. 2016

Mapping – reference genome – indel errors

Hg19 RECQL4 chr8:145738768-145738768



dbSNP build 150 rs751427839

dbSNP: rs751427839
Position: chr8:145738769-145738769
Band: 8q24.3
Genomic Size: 1
View DNA for this feature (hg19/Human)

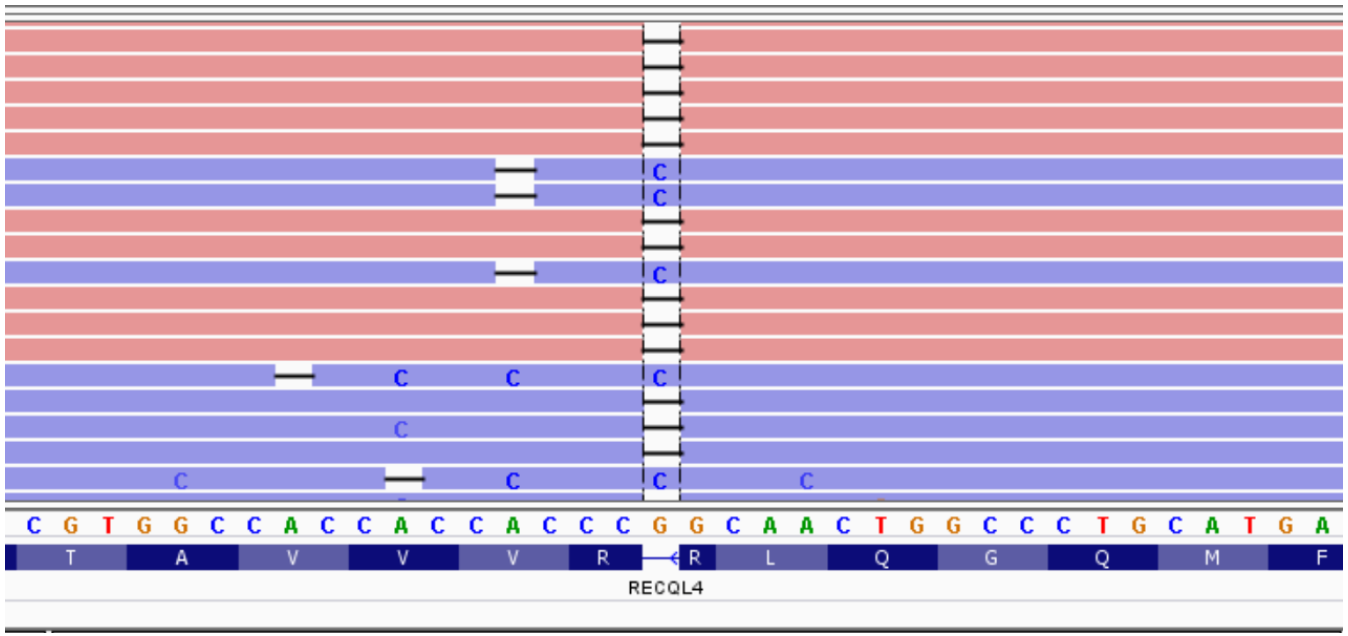
Strand: +
Observed: A/C/G/T
Reference allele: G

Class	single
Validation	by-frequency
Function	synonymous_variant, missense_variant
Molecule Type	genomic
Average Heterozygosity	0.328 +/- 0.293
Weight	1
Submitter Handles	EVA_EXAC, HUMAN_LONGEVITY, TOPMED
Allele Frequencies	A: 6.250% (1 / 16); C: 6.250% (1 / 16); G: 81.250% (13 / 16); T: 6.250% (1 / 16)

Coding annotations by dbSNP:
RECQL4 (NM_004260): synonymous_variant R (**C**GG) --> R (**A**GG)
RECQL4 (NM_004260): missense_variant R (**C**GG) --> G (**G**GG)
RECQL4 (NM_004260): missense_variant R (**C**GG) --> W (**T**GG)

UCSC Annotations:
There are other rsIds at this position with different variation.
This single-base substitution is quad-allelic.

UCSC's predicted function relative to selected gene tracks:
UCSC Genes RECQL4 (uc003zdj.3) splice_acceptor_variant



Mapping – reference genome – indel errors

- Reference genome (ABO O-type, del 1 nt) and reference transcript differ (ABO A-type)

Reference SNP (refSNP) Cluster Report: rs8176719 ****Clinical Channel****

RefSNP	
Organism:	human (*Homo sapiens*)
Molecule Type:	Genomic
Created/Updated in build:	117/151
Map to Genome Build:	108/Weight 1
Validation Status:	
Citation:	PubMed LitVar NEW
Association:	NHGRI GWAS

Allele	
Variation Class:	DIV: deletion/insertion variation
RefSNP Alleles:	-/G (REV)
Allele Origin:	
Ancestral Allele:	A
Variation Viewer:	VarView
Clinical Significance:	NA
MAF/MinorAlleleCount:	C=0.3764/45143 (ExAC) C=0.3438/1722 (1000 Genomes) C=0.3400/4154 (GO-ESP) C=0.3456/43394 (TOPMED)

HGVS Names
CM000671.2:g.133257521_133257522insC
NC_000009.11:g.136132908_136132909insC
NC_000009.12:g.133257521_133257522insC
NG_006669.1:g.20145_20147insG
NM_020469.2:c.260_262insG
NP_065202.2:p.Val87_Thr88=fs
XP_005276905.1:p.Thr87Aspfs
XP_005276906.1:p.Thr69Aspfs

Note the strange, non-HGVS, descriptions.
All are incorrect.

Variant: rs8176719

rs8176719 INSERTION

Most severe consequence frameshift variant | See all predicted consequences

Alleles -/C | MAF: 0.34 (C) | Highest population MAF: 0.48

Location Chromosome 9: between 133257521 and 133257522 (forward strand) | VCF: 9 133257521 rs8176719 T TC

Evidence status

HGVS names This variant has 9 HGVS names - Hide

Ensembl HGVS:

- NC_000009.12:g.133257521_133257522insC
- ENST00000453660.4:n.290_291insG
- ENST00000538324.2:c.259-1_259insG
- ENSP00000483018.1:p.Thr87AspfsTer107
- ENST00000611156.4:c.259-1_259insG
- ENSP00000483265.1:p.Thr87AspfsTer107
- ENST00000647353.1:n.54-6370_54-6369insG

dbSNP HGVS:

- NM_020469.2:c.260_262insG
- NP_065202.2:p.Val87_Thr88=fs

Synonyms PharmGKB PA166170086

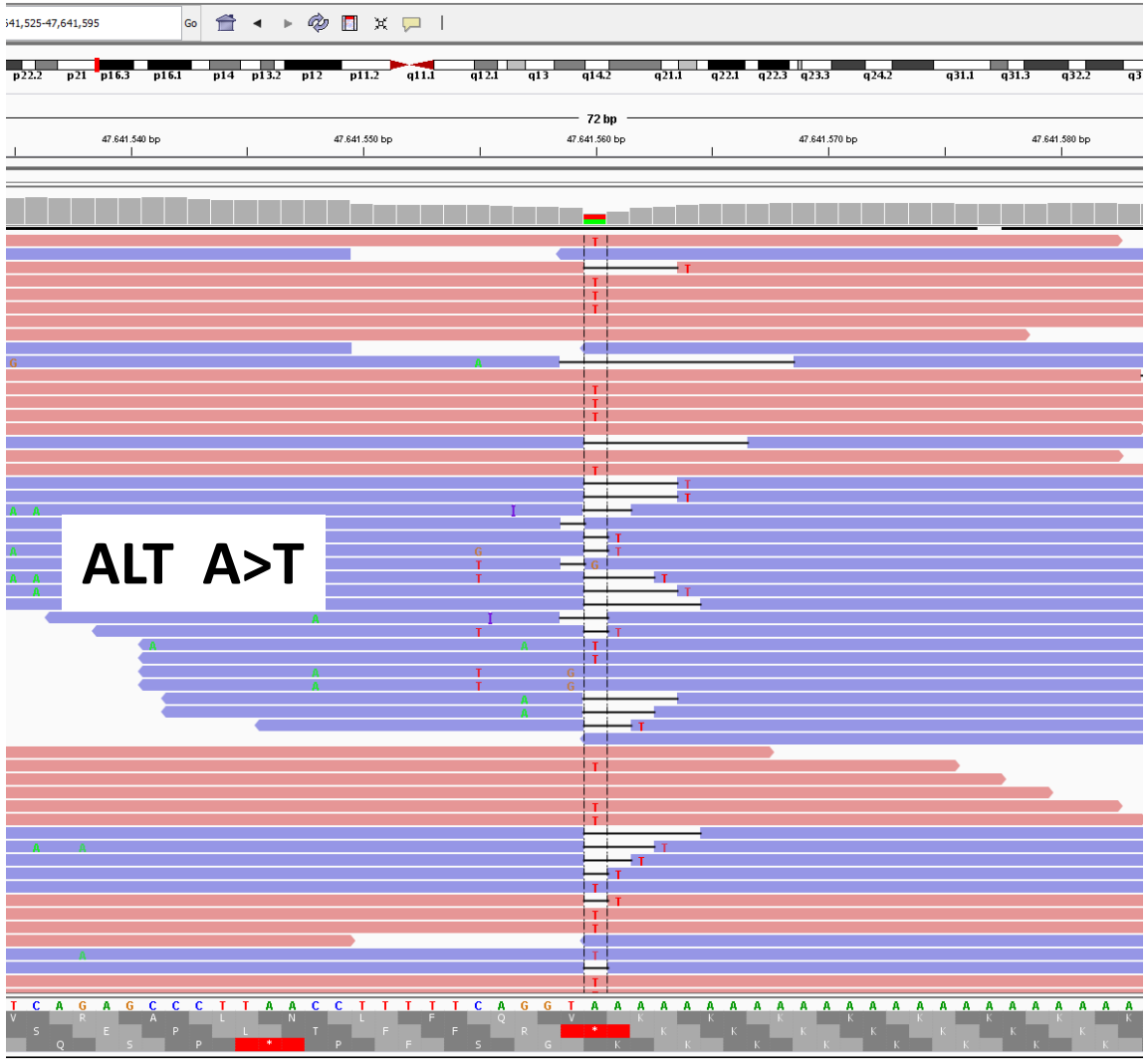
Original source Variants (including SNPs and indels) imported from dbSNP (release 150) | View in dbSNP

About this variant This variant overlaps 4 transcripts, has 2550 sample genotypes, is associated with 2 phenotypes and is mentioned in 54 citations.

Ensembl data copied from dbSNP

Mapping variants in simple sequences (homopolymers)

- MSH2 : NM_000251:c.942+3A>T Most frequent pathogenic MSH2 variant in HNPCC – Lynch Syndrome



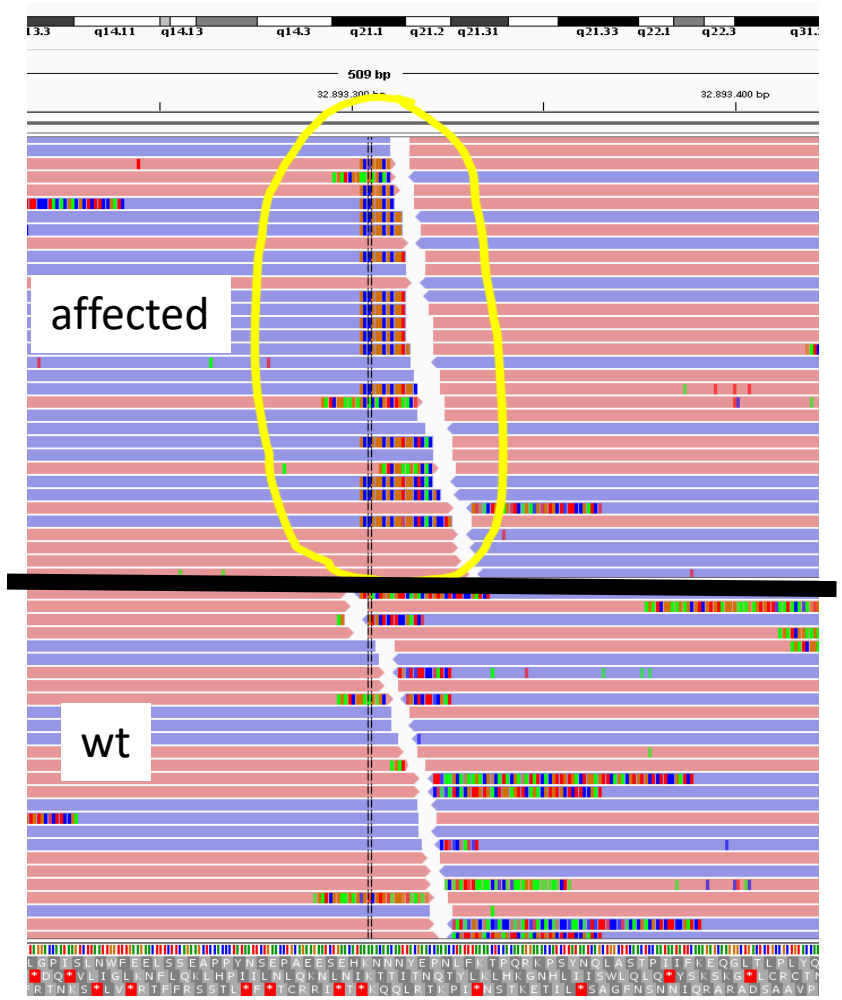
Mapping – missing insertions

- BRCA2

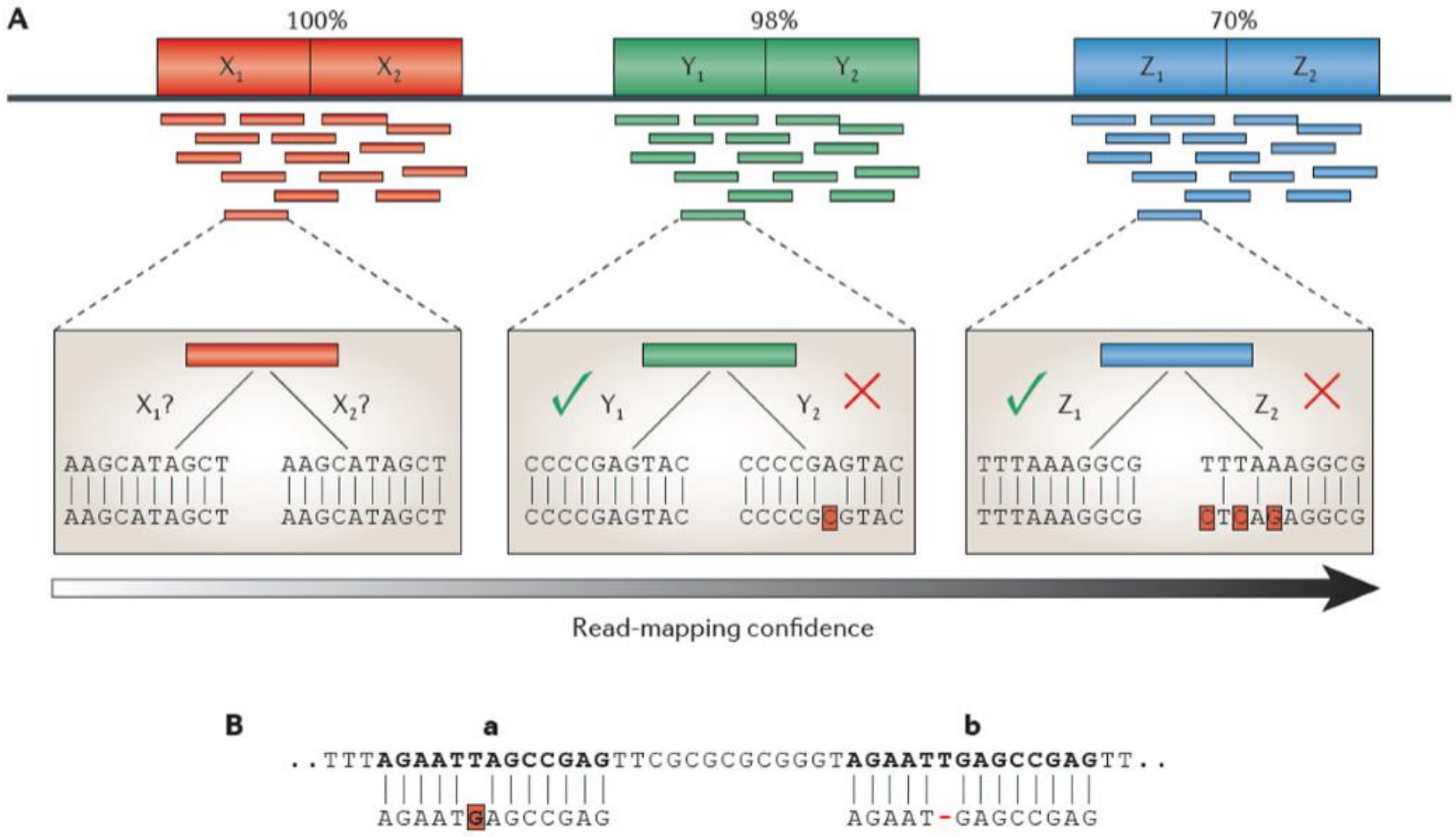
Breast Cancer Res Treat. 2009 Mar;114(1):31-8. doi: 10.1007/s10549-008-9978-4. Epub 2008 Mar 25.

The c.156_157insAlu BRCA2 rearrangement accounts for more than one-fourth of deleterious BRCA mutations in northern/central Portugal.

Peixoto A[1], Santos C, Rocha P, Pinheiro M, Príncipe S, Pereira D, Rodrigues H, Castro F, Abreu J, Gusmão L, Amorim A, Teixeira MR.



Mapping – reference genome – repetitive regions



Mapping non-unique

GATTTGGGCAGAGCGATGG

GATTTGGGCAGAGCGATGG

GATTTGGGCAGAGCGATGG

GATTTGGGCAGAGCGATGG
+

GATTTGGGCAGAGCGATGG

GATTTGGGCAGAGCGATGG

GATTTGGGCAGAGCGATGG

GATTTGGGCAGAGCGATGG
+

GATTTGGGTAGAGCGATGG

GATTTGGGTAGAGCGATGG

Mapping non-unique

GATTTGGGCAGAGCGATGG

GATTTGGGCAGAGCGATGG

Non-unique

GATTTGGGCAGAGCGATGG

GATTTGGGCAGAGCGATGG

GATTTGGGCAGAGCGATGG
GATTTGGGCAGAGCGATGG

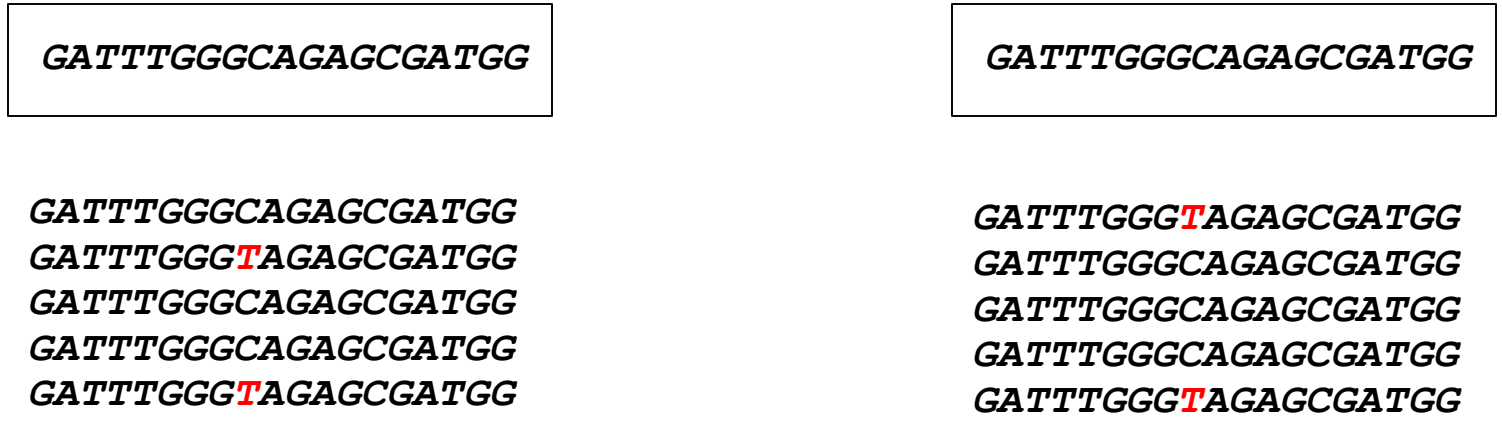
GATTTGGGCAGAGCGATGG

GATTTGGGCAGAGCGATGG
GATTTGGGCAGAGCGATGG
GATTTGGGCAGAGCGATGG

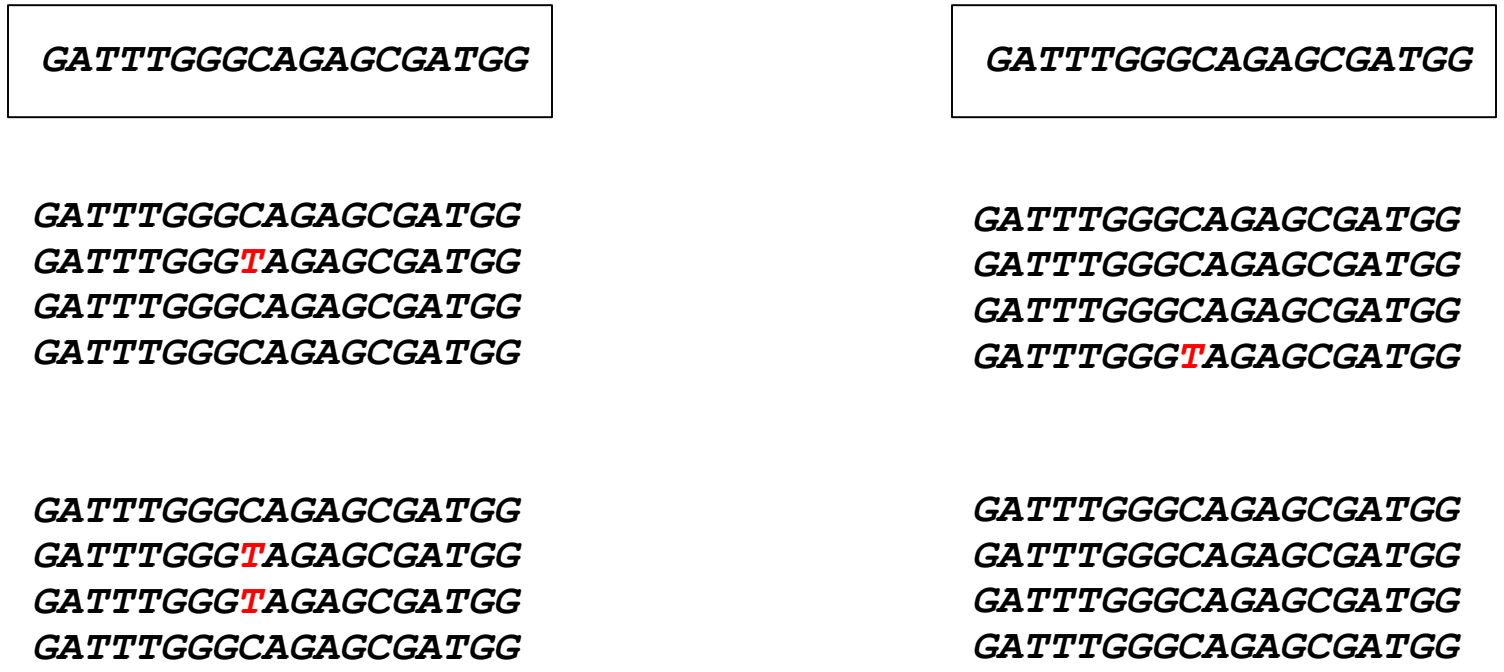
Non-unique

GATTTGGGCAGAGCGATGG
GATTTGGGCAGAGCGATGG
GATTTGGGTAGAGCGATGG
GATTTGGGCAGAGCGATGG
GATTTGGGCAGAGCGATGG
GATTTGGGTAGAGCGATGG
GATTTGGGCAGAGCGATGG
GATTTGGGCAGAGCGATGG

Non-unique

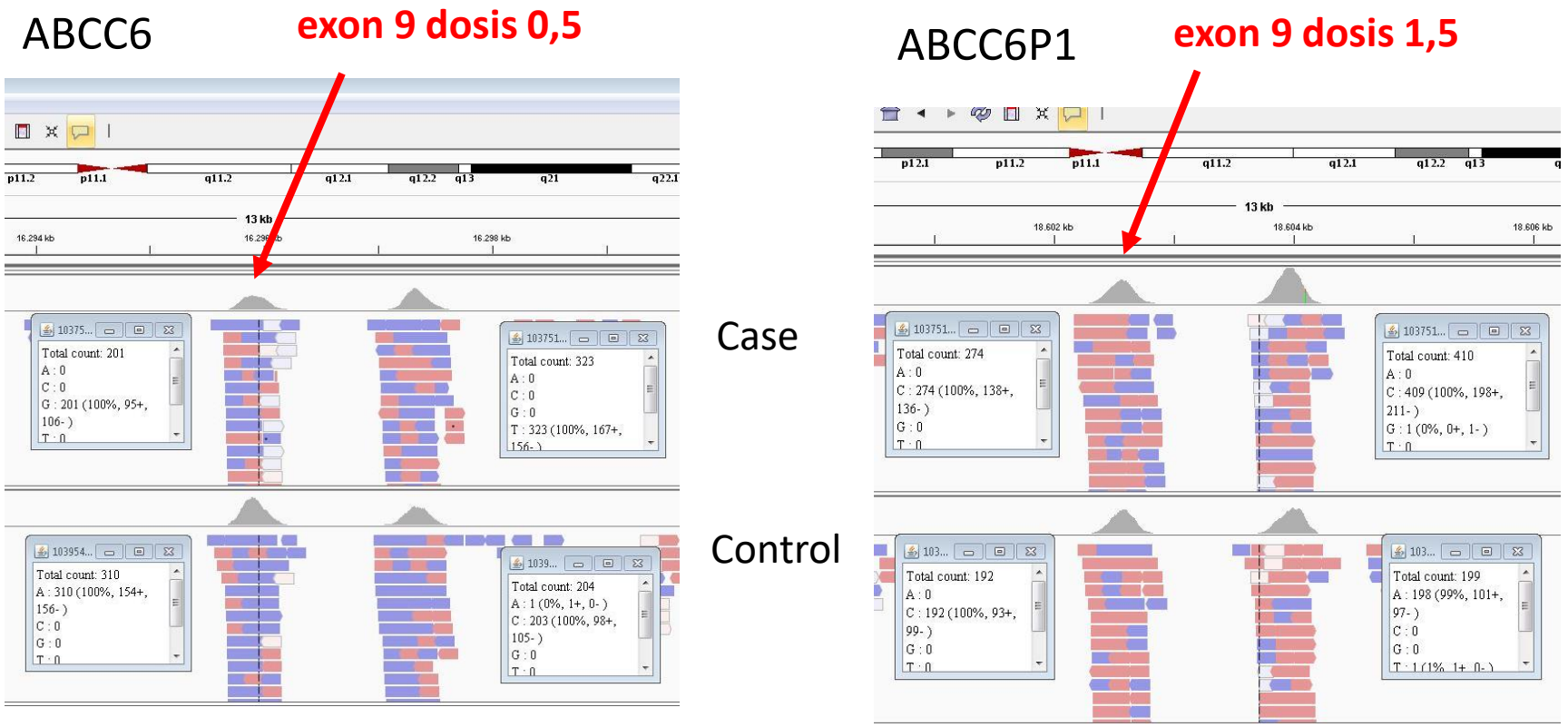


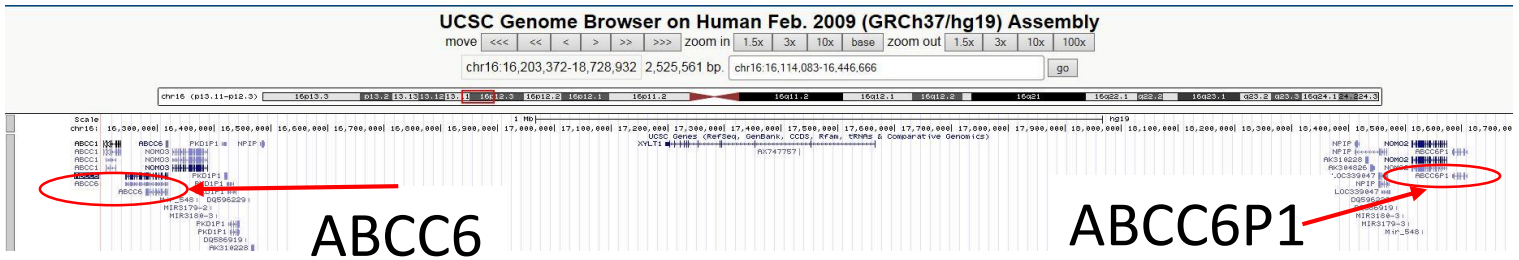
Map Non-unique



Mapping in pseudogenes – missing SNVs due to Gene conversions

- *Pseudoxanthoma elasticum;* c.4182del (p.Lys1394Asnfs*9) – class 5





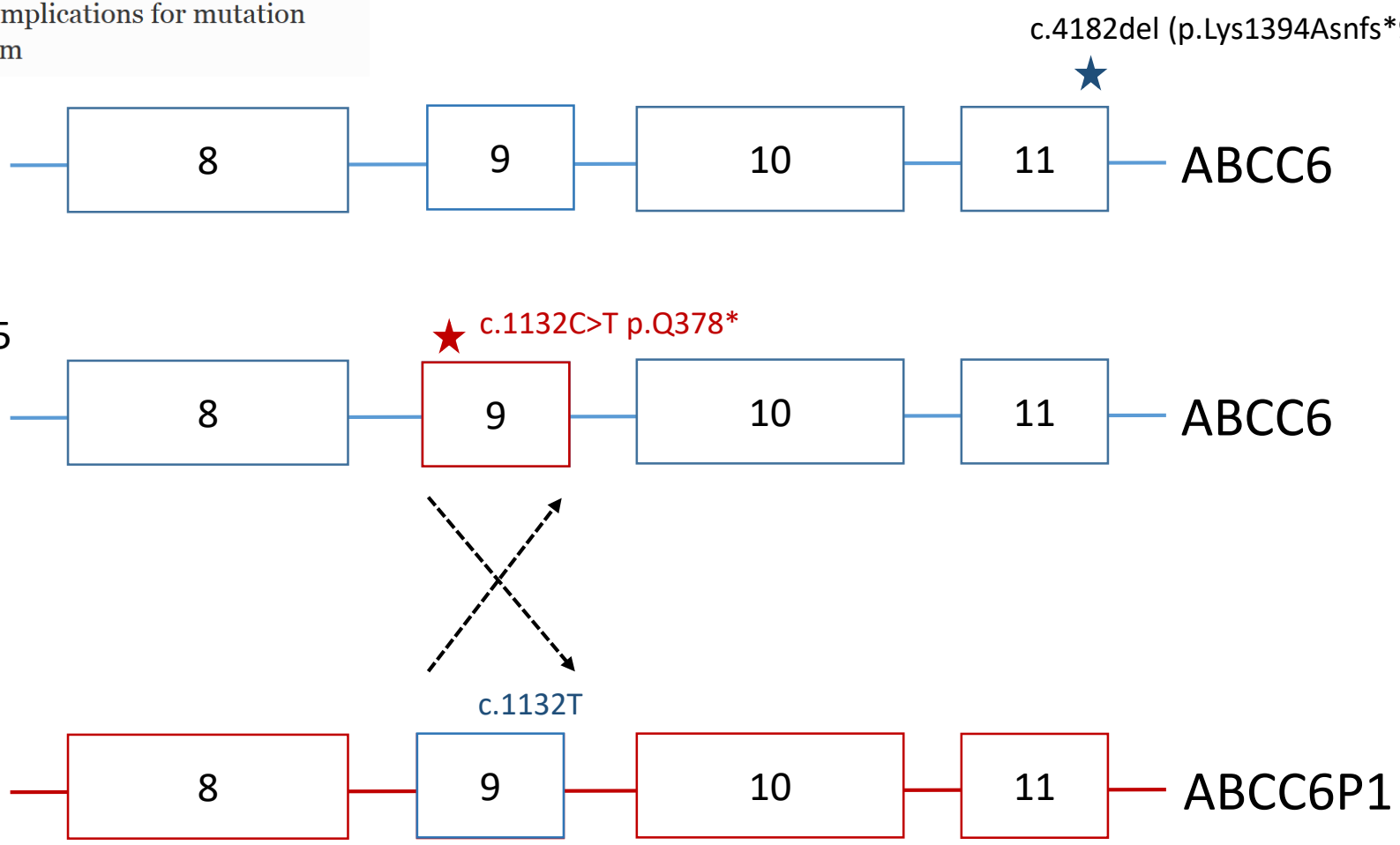
Variant Calling – missing SNVs due to Gene conversions

Journal of Molecular Medicine
September 2001, Volume 79, Issue 9, pp 536–546

A novel *Q378X* mutation exists in the transmembrane transporter protein *ABCC6* and its pseudogene: implications for mutation analysis in pseudoxanthoma elasticum

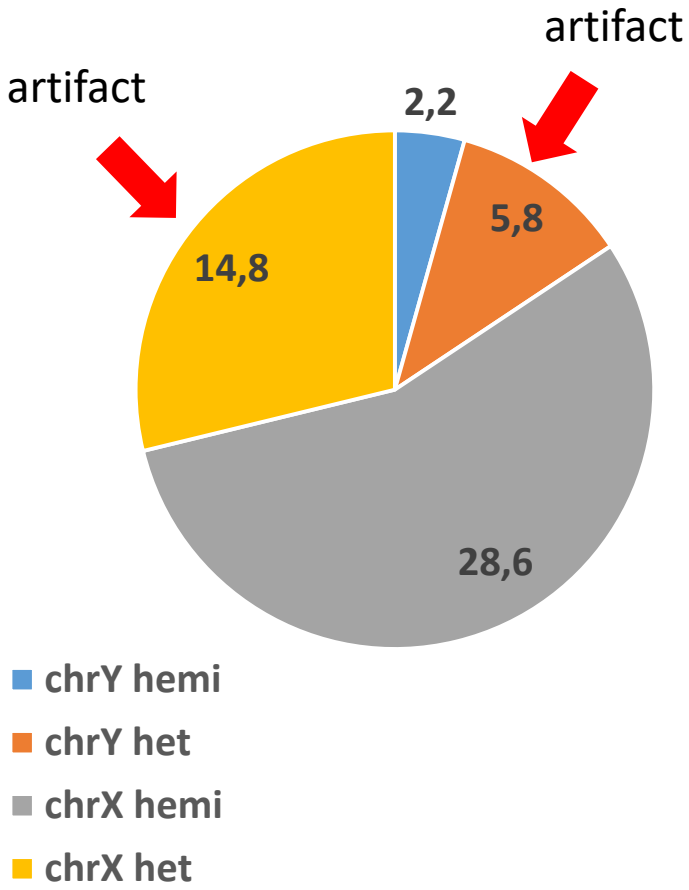
c.4182del (p.Lys1394Asnfs*9) - class 5

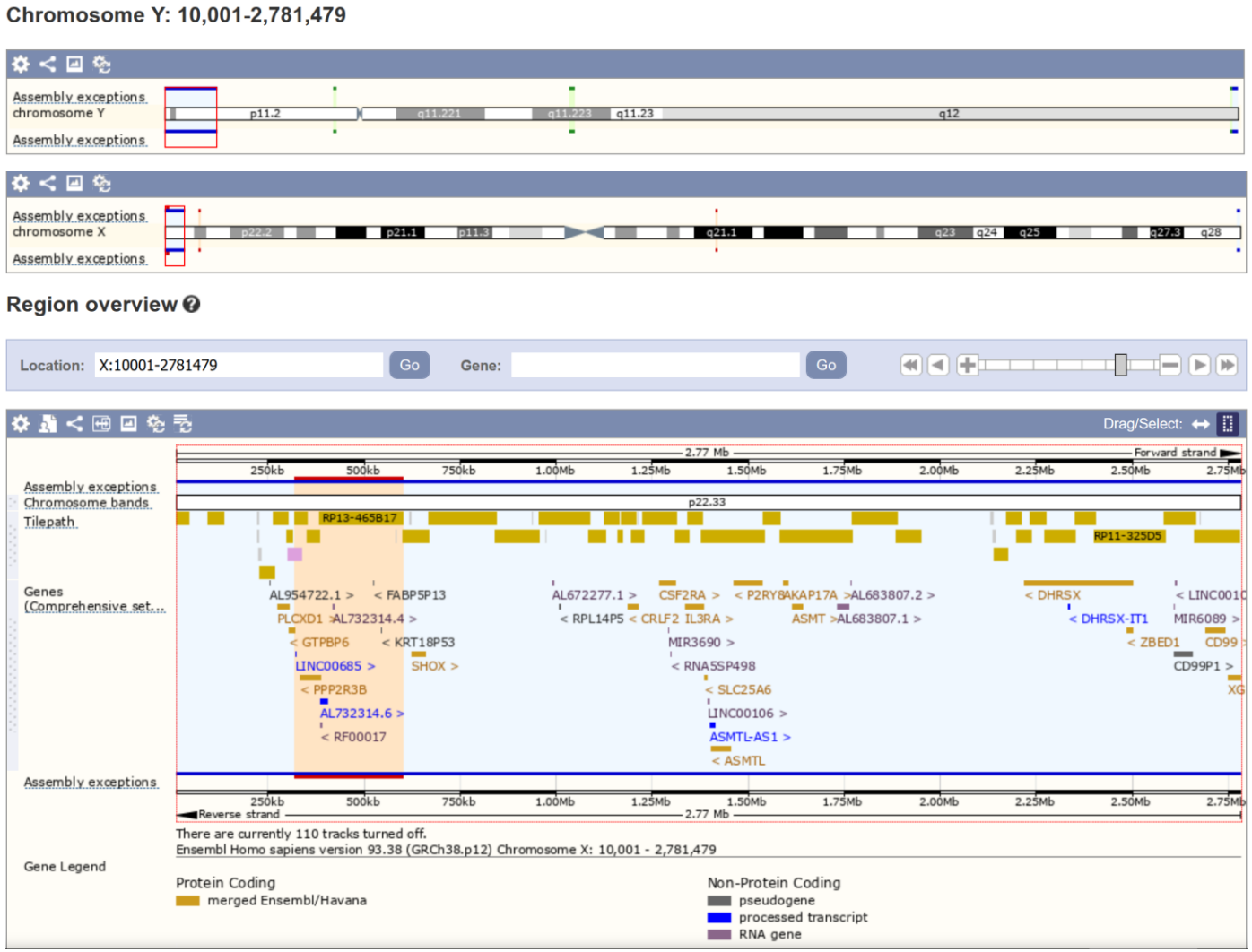
c.1132C>T p.Q378* - class 5



Variant Calling – male/female – X and Y chr pseudoautosomal region

- Number of SNVs called in X and Y chromosome in males





Variant Calling – male/female – X and Y chr pseudoautosomal region

For variant calling, where to draw the line (threshold)

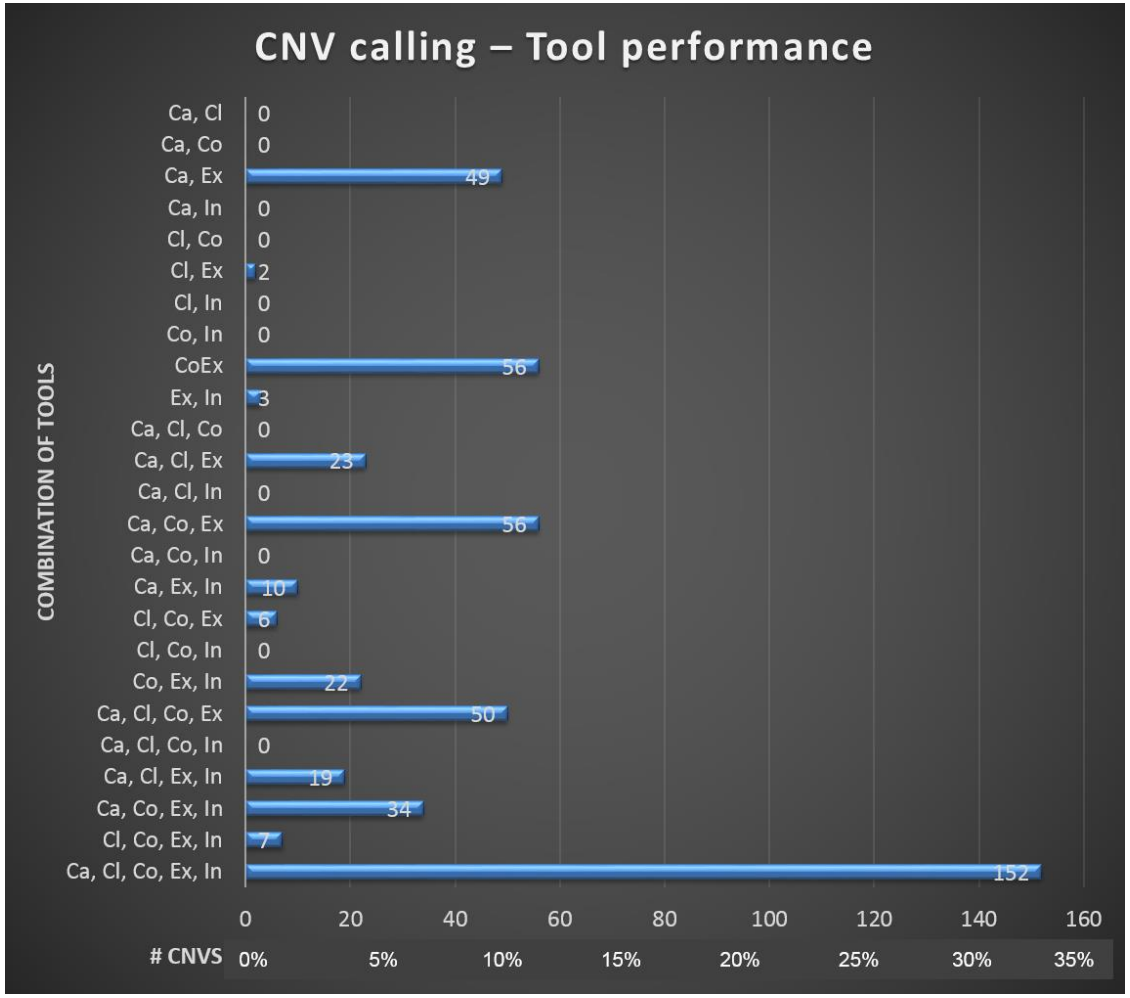
We set it initially at 10 reads or more, i.e. variants are not reported when we had 10 reads or less.

As a consequence we missed 9 reads, all with the same variant, on the X-chromosome in a male patient causing the disease !

A proper NGS pipelines treats the X-/Y- chromosomes different to call variants in males.

Variant Calling – SVs and CNVs – different tools deliver diferent calls

- Sample set ~1000 individuals



Tool	Mean Calls per Sample
Canoes	~ 1.4
Clamms	~ 5.1
Codex	~ 2.6
ExomeDepth	~ 7.7
Inhouse	~ 4.4

Variant Calling – limits of detection - mosaicism in PIK3CA

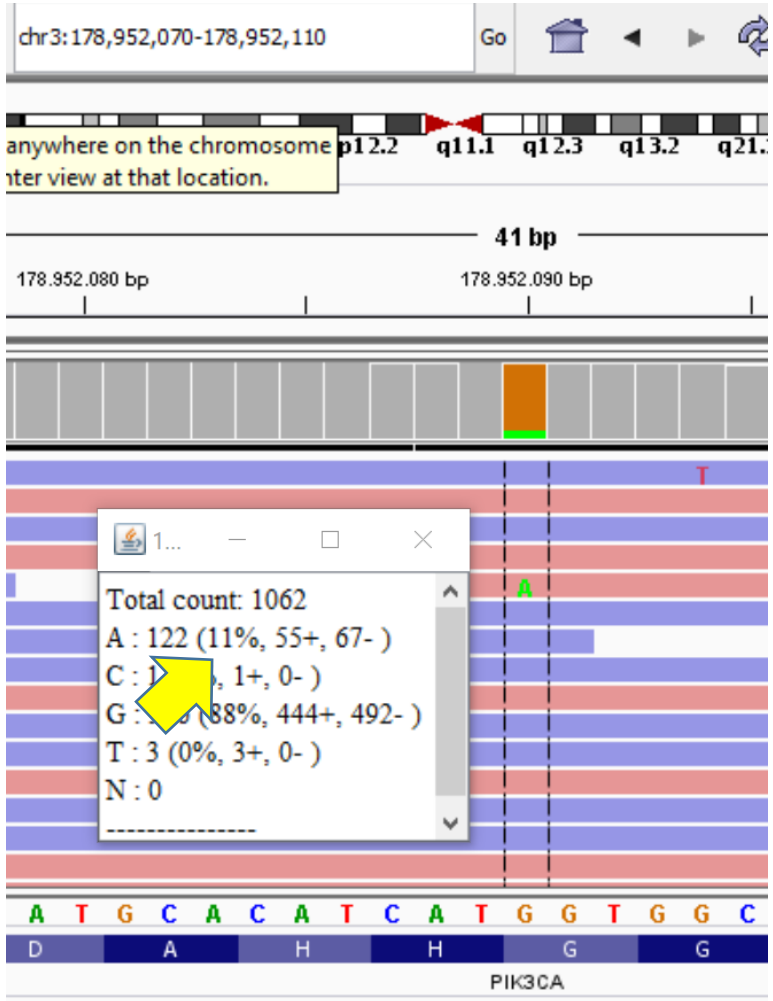
Somatic Mosaic Activating Mutations in *PIK3CA* Cause CLOVES Syndrome

Kyle C. Kurek,[1] Valerie L. Luks,[2] Ugur M. Ayturk,[2,9] Ahmad I. Alomari,[3,6] Steven J. Fishman,[4,6] Samantha A. Spencer,[2,6] John B. Mulliken,[5,6] Margot E. Bowen,[2,9] Guilherme L. Yamamoto,[7] Harry P.W. Kozakewich,[1,6] and Matthew L. Warman[2,6,8,9,*]

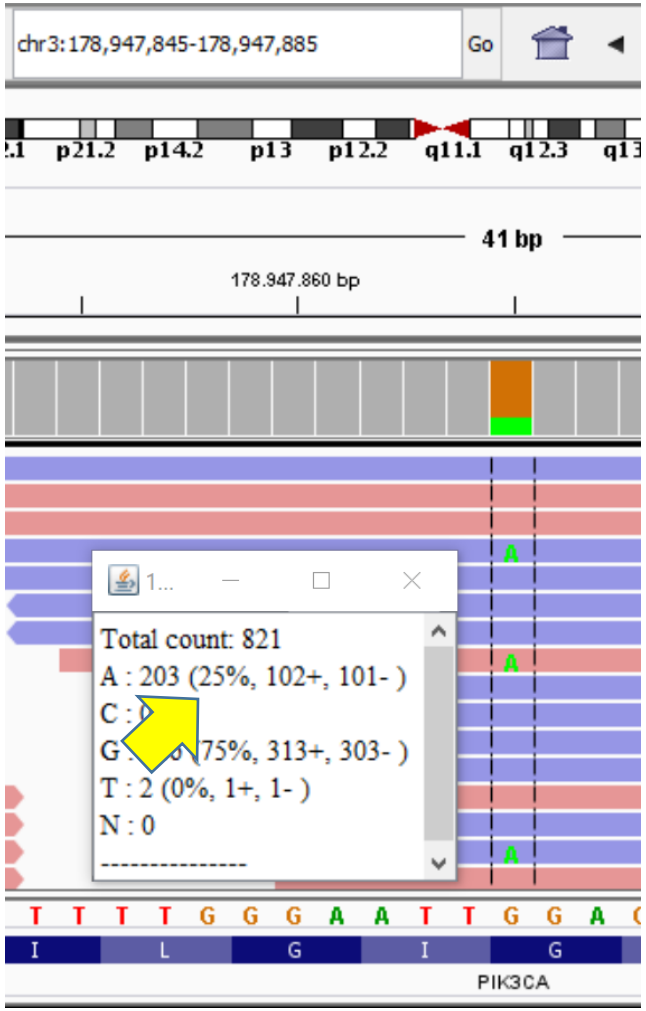
- Look for the right variant feature in the right gene/disease

Only detected with the

LOW-FREQUENCY pipeline



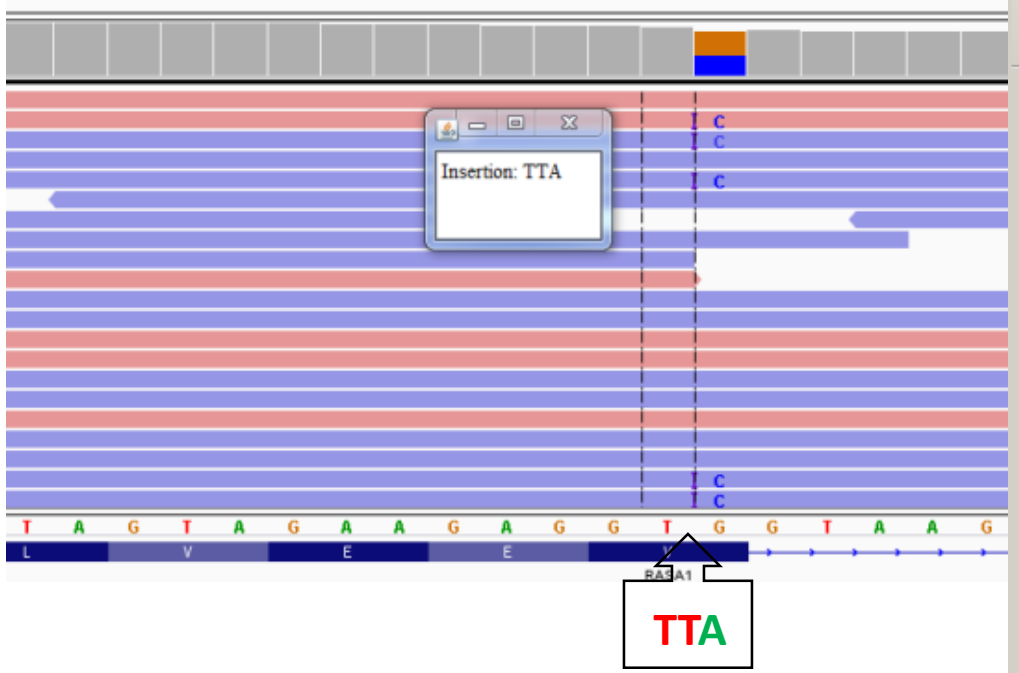
NM_006218
c.3145G>A:p.Gly1049Ser

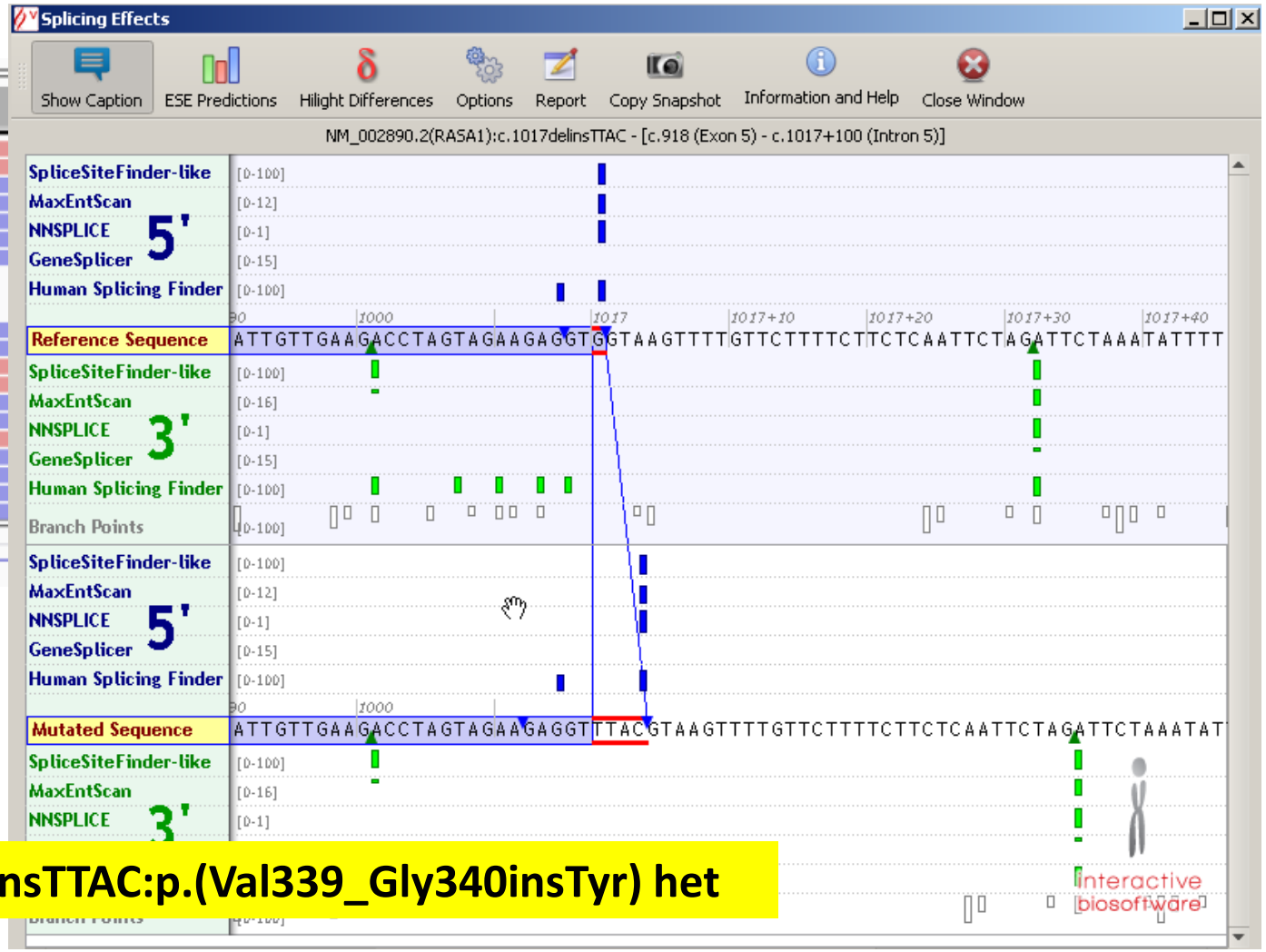


NM_006218
c.2740G>A:p.Gly914Arg

Annotation – delins in splice site regions

RASA1 : NM_002890:c.1016_1017insTTA:p.Gly340*; het; AD >>>> Stopp gain





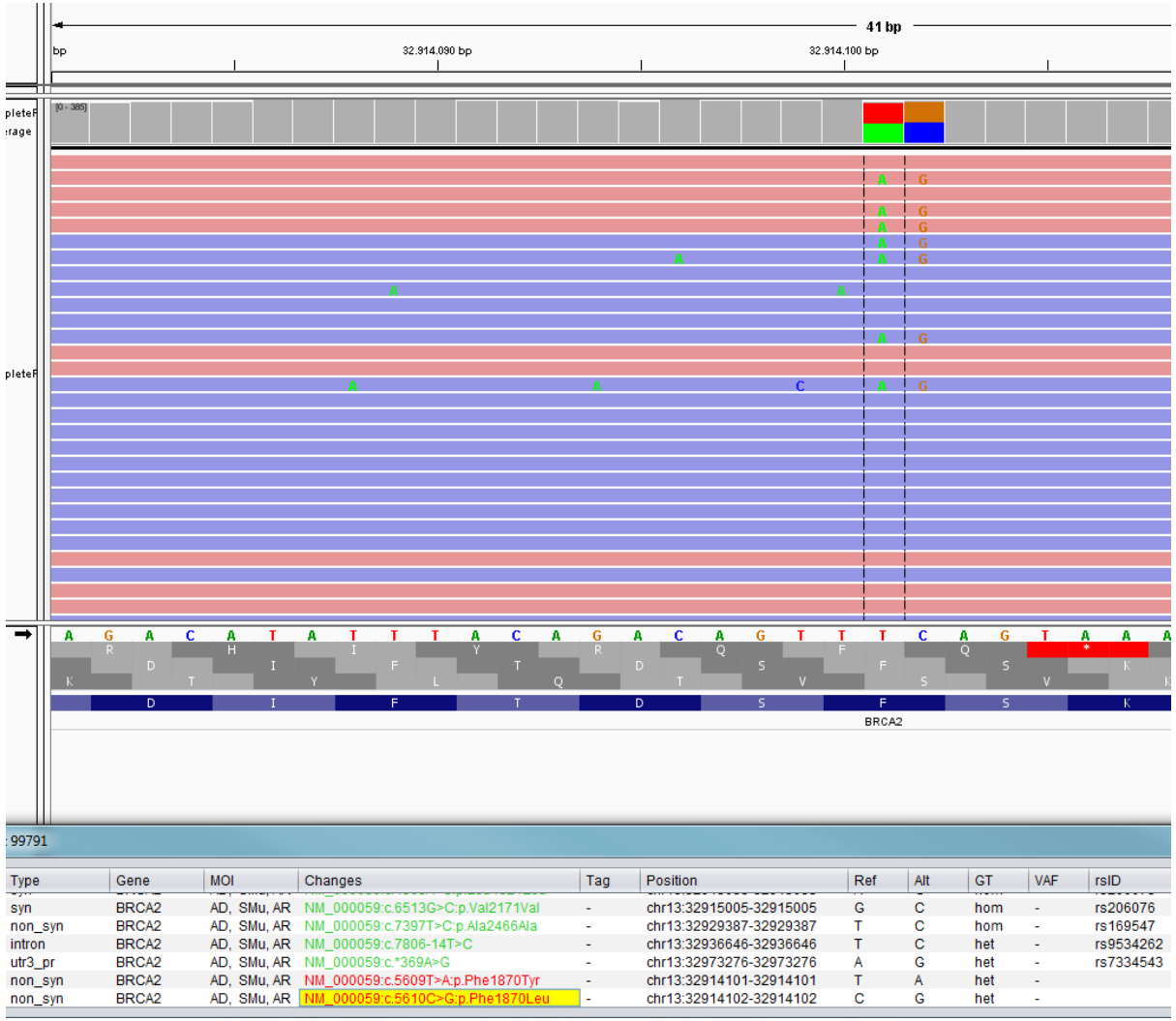
HGVS: RASA1: c.1017delinsTTAC:p.(Val339_Gly340insTyr) het

Addition of tyrosin > splice-site not affected (moves 3 positions)

Annotation – delins annotation as two single events

BRCA2 : NM_000059_c.5609T>A;p.Phe1870Tyr;het

BRCA2 : NM_000059_c.5609T>A;p.Phe1870Leu;het



Type	Gene	MOI	Changes	Tag	Position	Ref	Alt	GT	VAF	rsID
syn	BRCA2	AD, SMu, AR	NM_000059:c.6513G>C:p.Val2171Val	-	chr13:32915005-32915005	G	C	hom	-	rs206076
non_syn	BRCA2	AD, SMu, AR	NM_000059:c.7397T>C:p.Ala2466Ala	-	chr13:32929387-32929387	T	C	hom	-	rs169547
intron	BRCA2	AD, SMu, AR	NM_000059:c.7806-14T>C	-	chr13:32936646-32936646	T	C	het	-	rs9534262
utr3_pr	BRCA2	AD, SMu, AR	NM_000059:c.*369A>G	-	chr13:32973276-32973276	A	G	het	-	rs7334543
non_syn	BRCA2	AD, SMu, AR	NM_000059:c.5609T>A:p.Phe1870Tyr	-	chr13:32914101-32914101	T	A	het	-	
non_syn	BRCA2	AD, SMu, AR	NM_000059:c.5610C>G:p.Phe1870Leu	-	chr13:32914102-32914102	C	G	het	-	

NCBI Resources How To

ClinVar — ClinVar — Search ClinVar for gene symbols, HGVS — Advanced

Home | About | Access | Help | Submit | Statistics | FTP

NEW Click here to see the new Variation Report design!

NM_000059.3(BRCA2):c.5609T>A (p.Phe1870Tyr)

Variation ID: 431319

Review status: ☆☆☆☆ (0/4) no assertion criteria provided

Interpretation

Clinical significance: Uncertain significance

Last evaluated: Mar 5, 2014

Number of submission(s): 1

See supporting ClinVar records

NEW Click here to see the new Variation Report design!

NM_000059.3(BRCA2):c.5610C>G (p.Phe1870Leu)

Variation ID: 441502

Review status: ★☆☆☆ criteria provided, single submitter

Interpretation

Clinical significance: Uncertain significance

Last evaluated: Oct 11, 2016

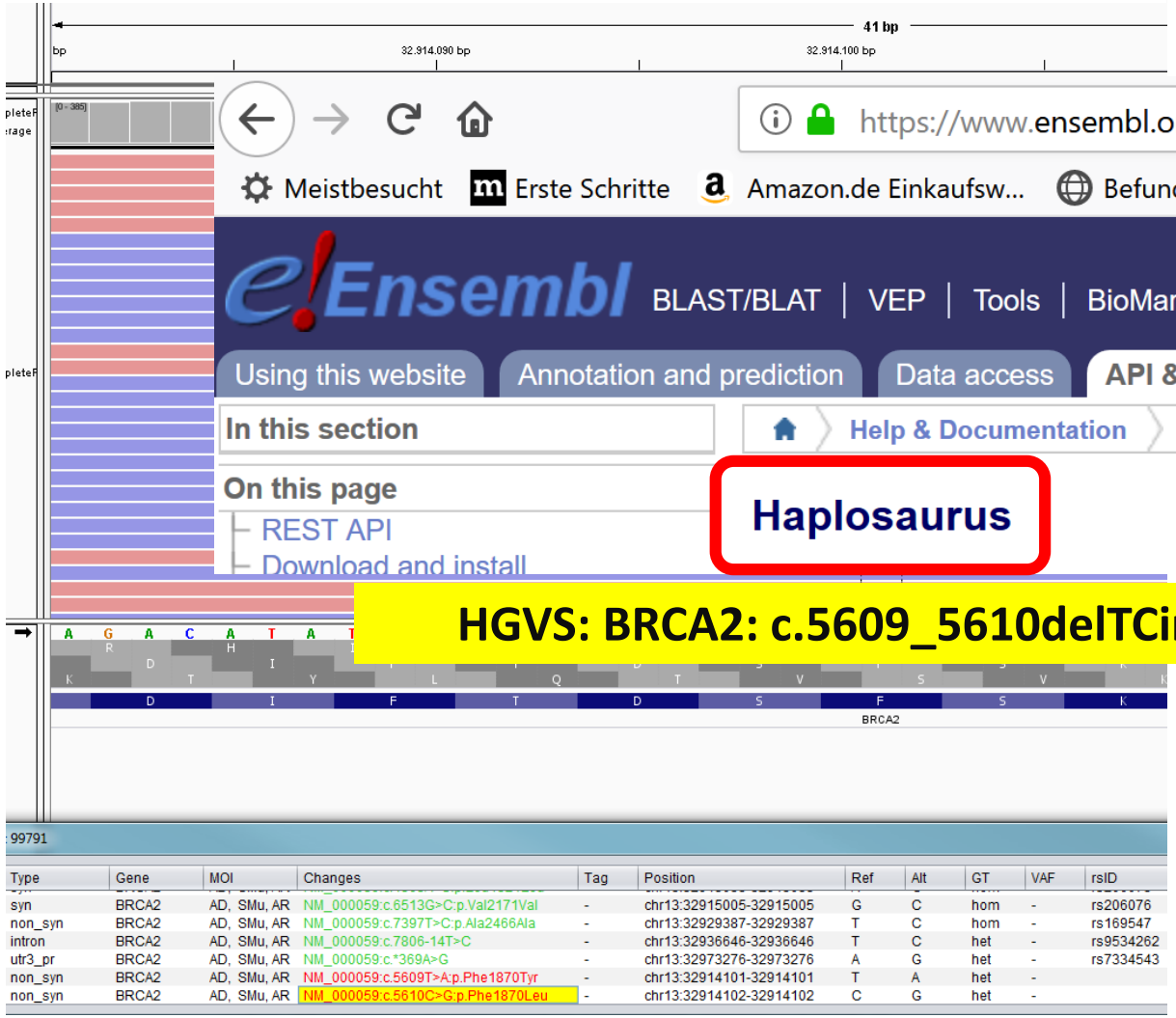
Number of submission(s): 1

Condition(s): Hereditary cancer-predisposing syndrome [MedGen]

Annotation – delins annotation as two single events

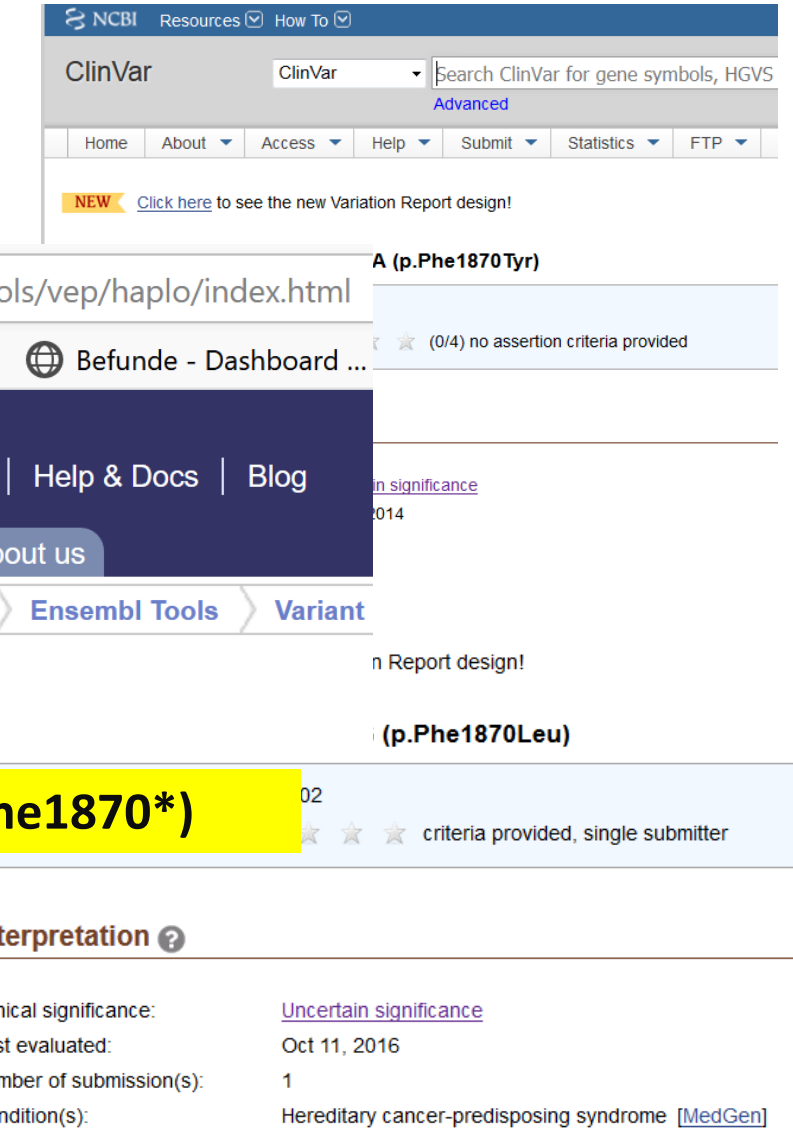
BRCA2 : NM_000059_c.5609T>A;p.Phe1870Tyr;het

BRCA2 : NM_000059_c.5609T>A;p.Phe1870Leu;het

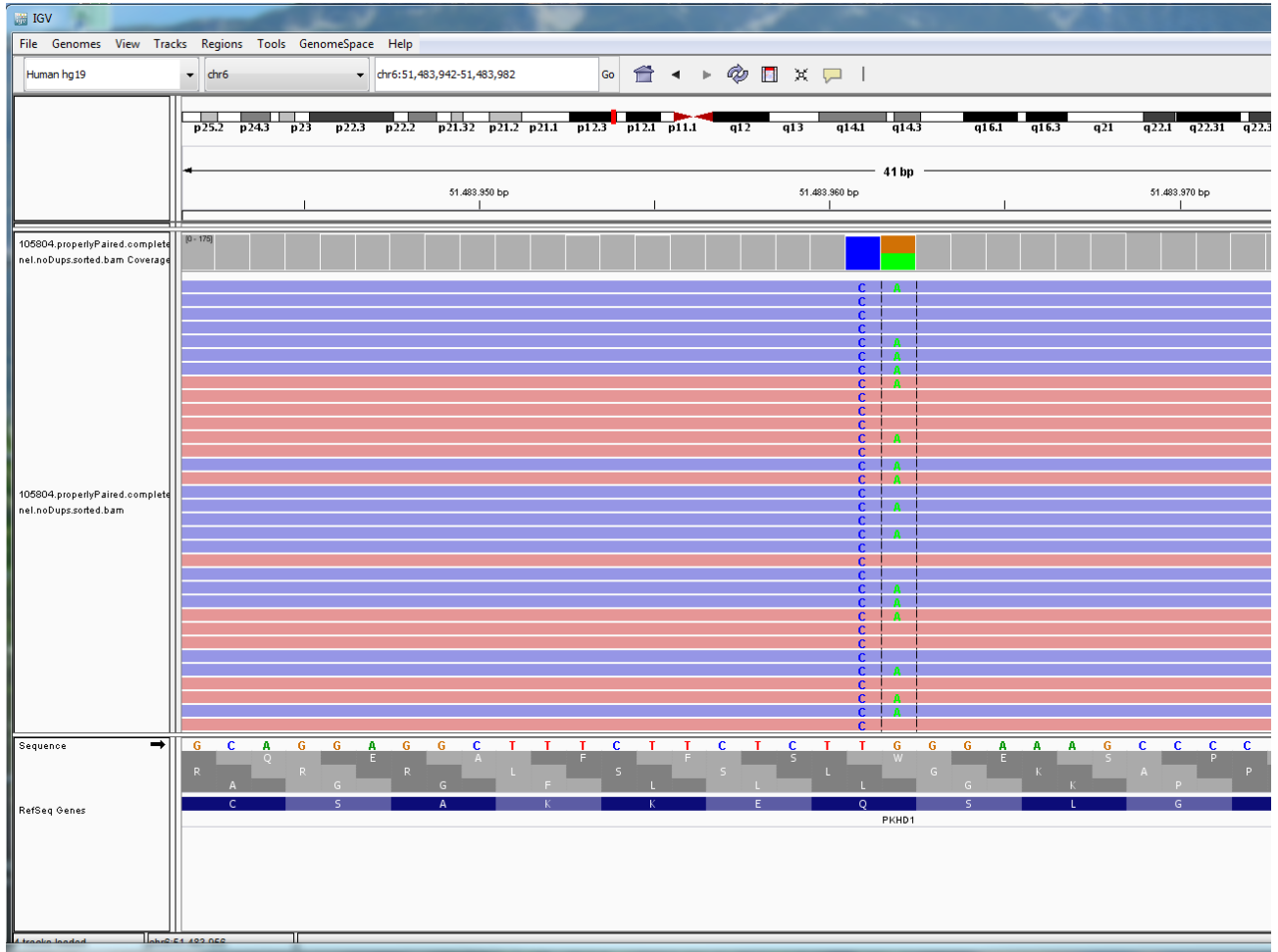


HGVS: BRCA2: c.5609_5610delTCinsAG (p.Phe1870*)

Type	Gene	MOI	Changes	Tag	Position	Ref	Alt	GT	VAF	rsID
syn	BRCA2	AD, SMu, AR	NM_000059:c.6513G>C:p.Val2171Val	-	chr13:32915005-32915005	G	C	hom	-	rs206076
non_syn	BRCA2	AD, SMu, AR	NM_000059:c.7397T>C:p.Ala2466Ala	-	chr13:32929387-32929387	T	C	hom	-	rs169547
intron	BRCA2	AD, SMu, AR	NM_000059:c.7806-14T>C	-	chr13:32936646-32936646	T	C	het	-	rs9534262
utr3_pr	BRCA2	AD, SMu, AR	NM_000059:c.*369A>G	-	chr13:32973276-32973276	A	G	het	-	rs7334543
non_syn	BRCA2	AD, SMu, AR	NM_000059:c.5609T>A:p.Phe1870Tyr	-	chr13:32914101-32914101	T	A	het	-	
non_syn	BRCA2	AD, SMu, AR	NM_000059:c.5610C>G:p.Phe1870Leu	-	chr13:32914102-32914102	C	G	het	-	



Filtering after population frequency = missed variant in complex call event



Val	Mat	Type	Gene	MOI	Changes	Tag	Position	Ref	Alt	GT	VAF	rsID	1000G	ESP
	G	intron	PKHD1	AR	NM_138694:c.5236+14A>G	-	chr6:51889358-51889358	T	C	het	-	rs12210725	0.0228	0.0232
	G	syn	PKHD1	AR	NM_138694:c.4035C>A:p.Gly1345Gly	-	chr6:51890573-51890573	G	T	het	-	rs140791735	0.0108	0.0085
	G	intron	PKHD1	AR	NM_138694:c.11507-11C>T	-	chr6:51497532-51497532	G	A	het	-	rs201478327	0.0004	
	G	syn	PKHD1	AR	NM_138694:c.3783C>T:p.Gly1261Gly	-	chr6:51890825-51890825	G	A	het	-	rs138388301	0.0004	0.0007
	G	stop_gain	PKHD1	AR	NM_138694:c.12142C>T:p.Gln4048*	-	chr6:51483962-51483962	G	A	het	-	rs201812542		0.0002

PKHD1

NM_138694:c.12142C>T:p.Gln4048*

- truncating strong

NM_138694:c.12141A>G:p.Gln4048Arg hom

- >5% frequency (automatically filtered out)

NGS Workflow

- Sampling
- Sample preparation
- Sequencing
- Mapping
- Variant Calling
- Variant Annotation
- Checks
- Variant Interpretation

Checks – % mapped reads/ average coverage / target coverage

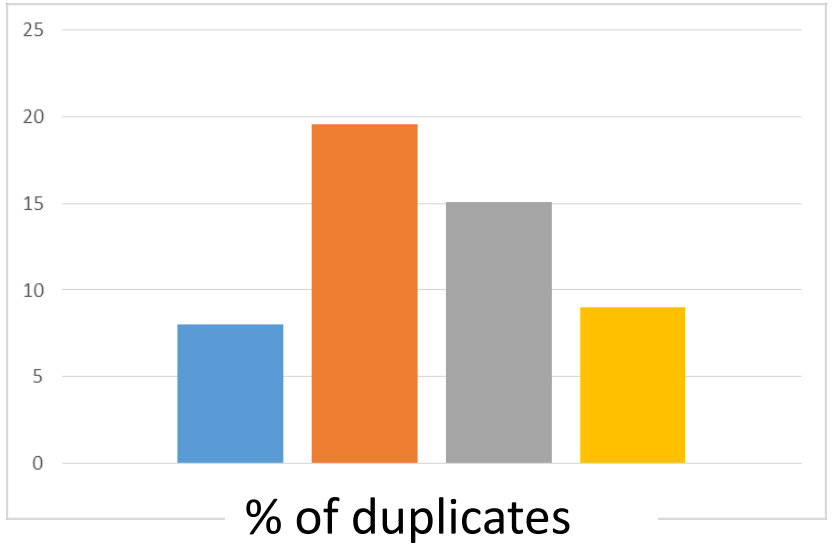
- Duplicates decrease the efficiency of sequencing

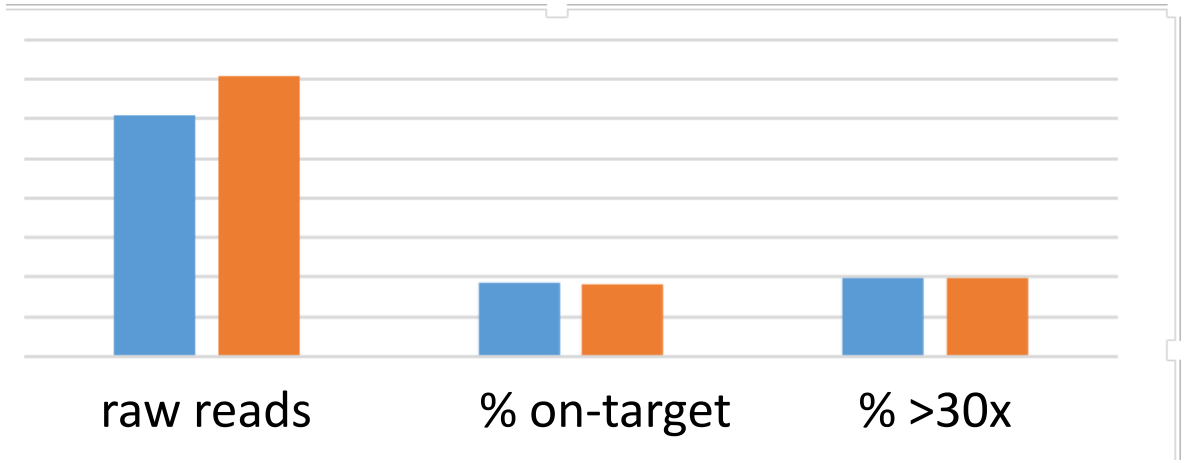
SSXT 1µg DNA / 390bp frag.size

SSXT-low input 100ng/ 390bp DNA fragment

SSXT-low input 100ng/ 430bp DNA fragment

SSXT-low input 200ng/ 455bp DNA fragment

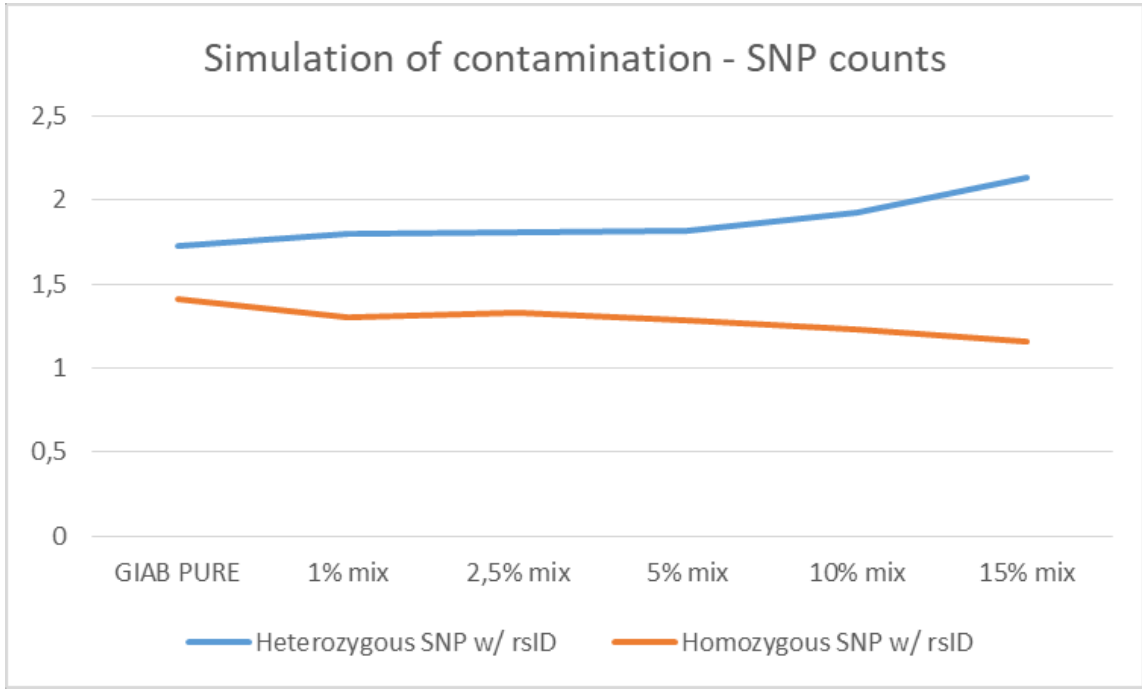


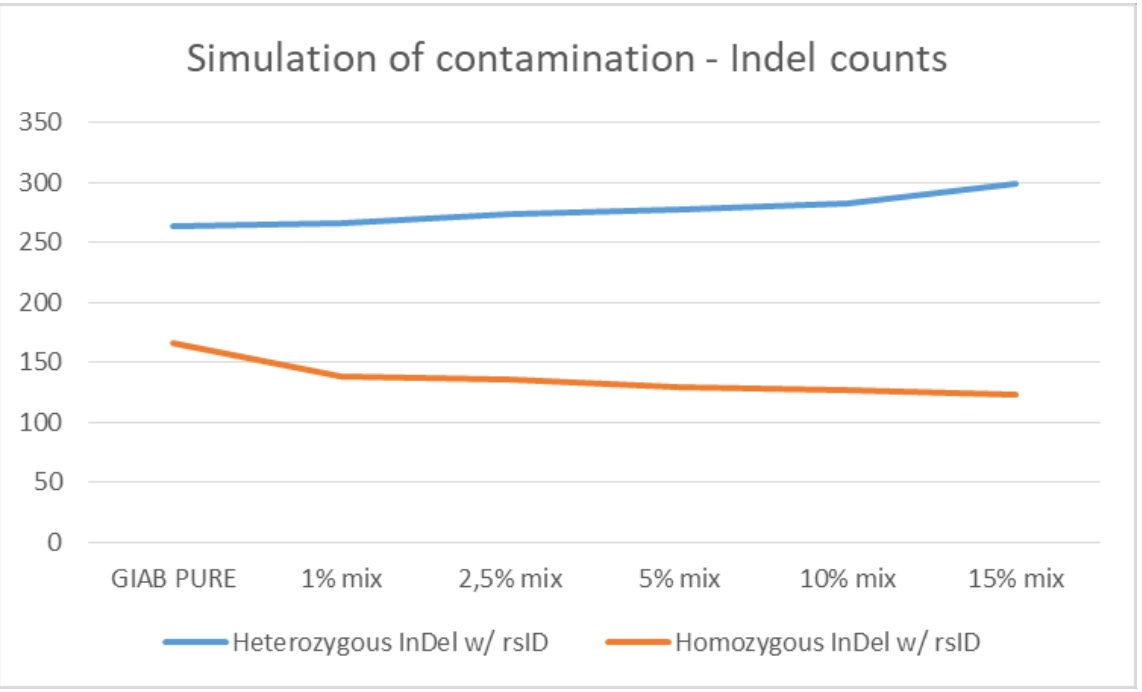


- Define a minimal coverage treshold in addition to mean coverage

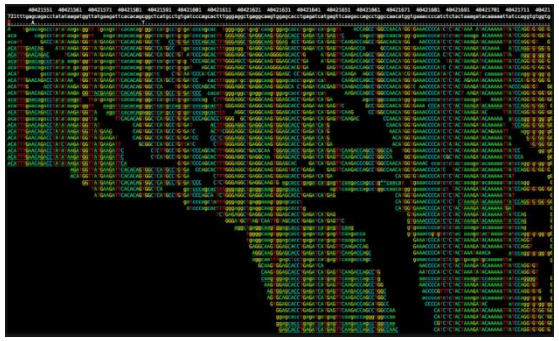
Check - sample contamination

- Check calls: Homozygote vs. heterozygote SNPs/Indels or Transversions vs. Transitions
- Check mapping : reads from other species

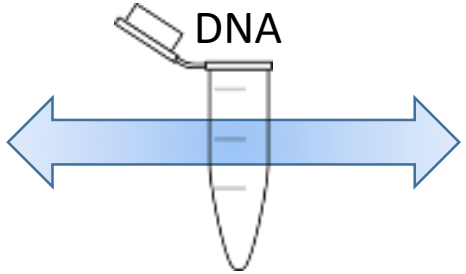




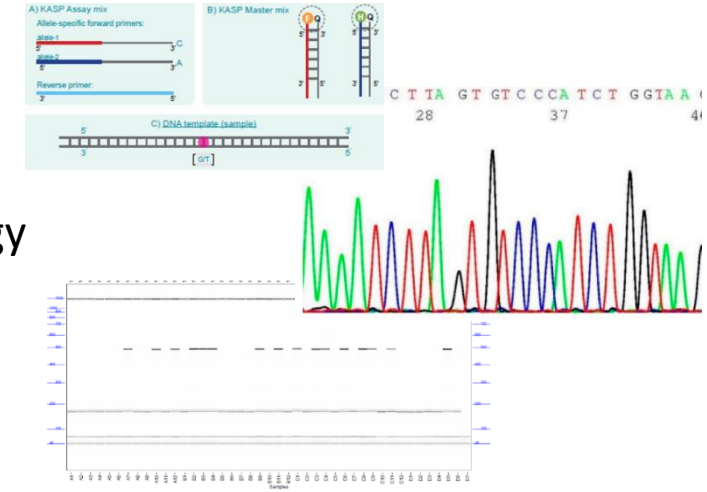
Checks – Exclusion of sample swap

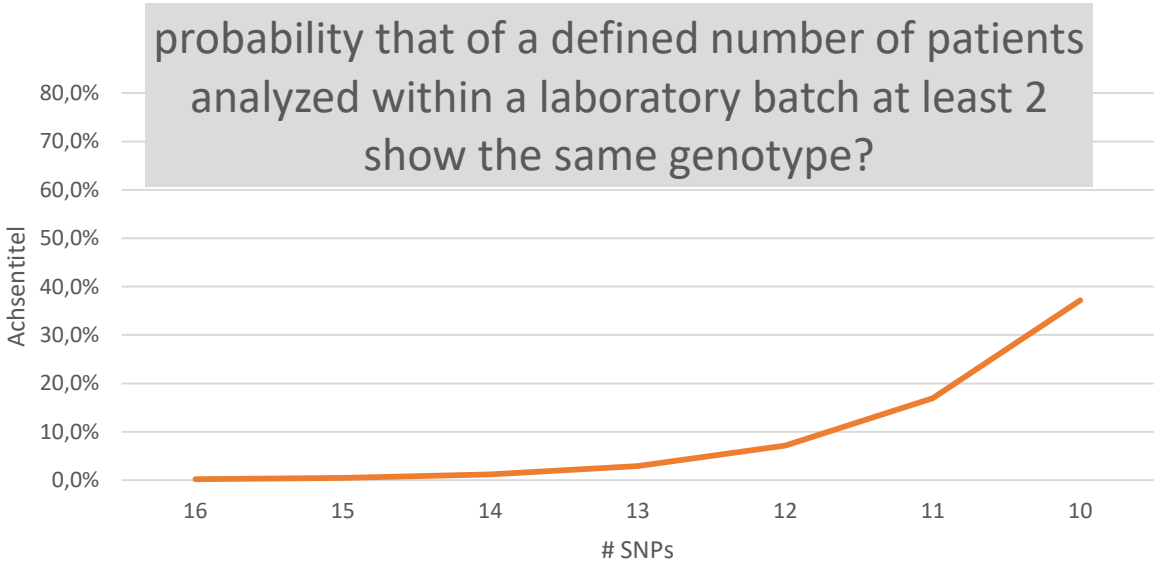


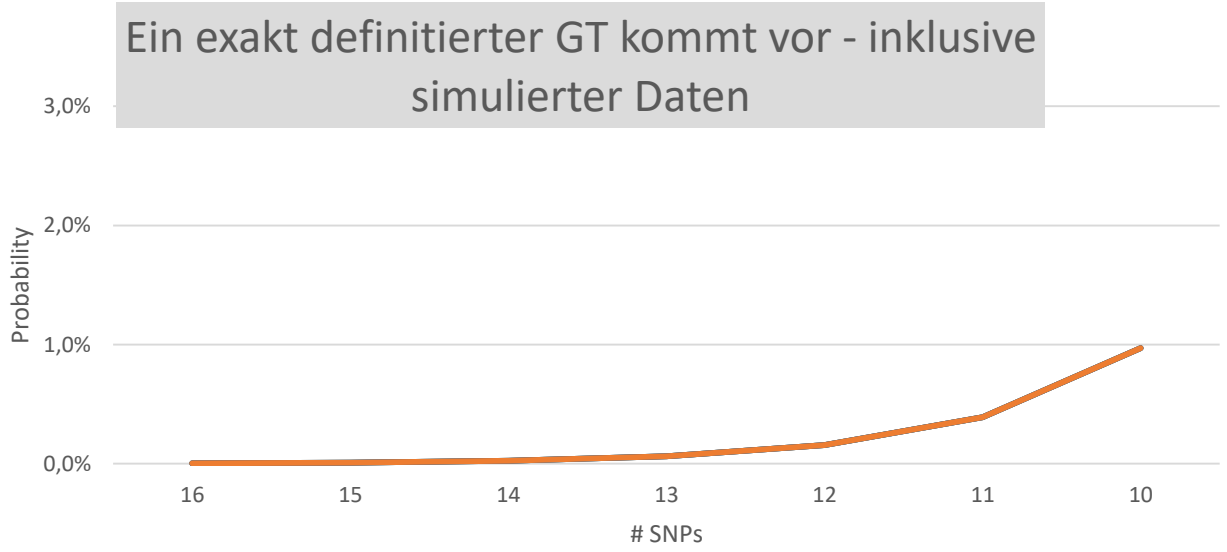
NGS raw data
SNPs



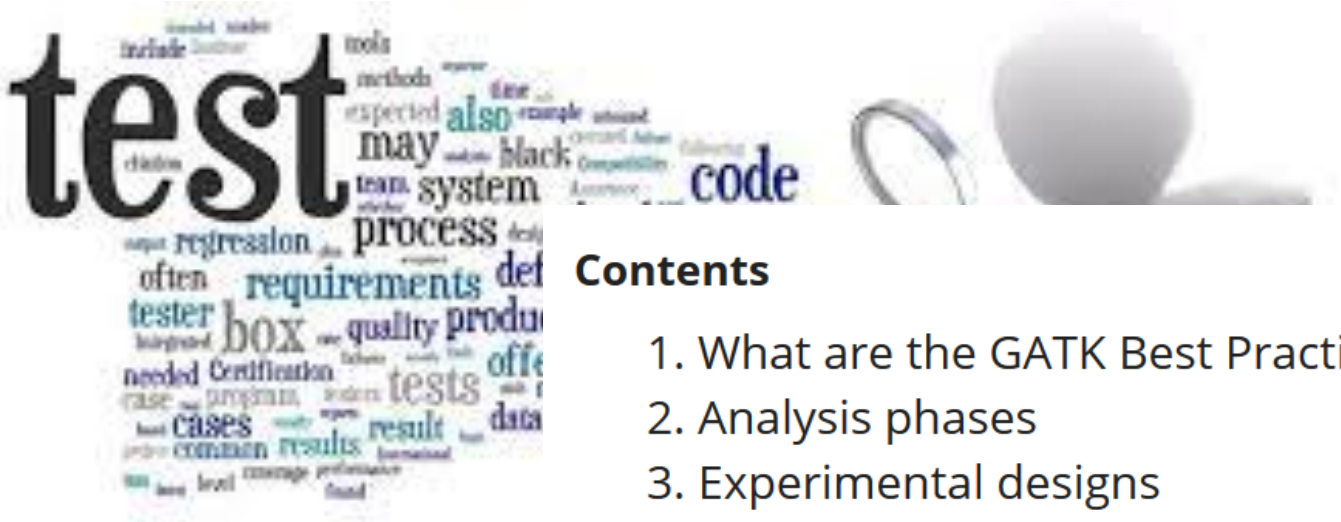
Other methodology
SNPs







Checks – new pipeline / component : re-run standard controls



Contents

1. What are the GATK Best Practices?
2. Analysis phases
3. Experimental designs
4. Workflow scripts provided as reference implementations
5. Scope and limitations
6. What is *not* GATK Best Practices?
7. Beware legacy scripts

https://software.broadinstitute

Erste Schritte | Amazon.de Einkaufsw... | Befunde - Bef

gatk | Best Practices | User Guide | Blog | Forum | Events | Download

Best Practices

Introduction to the GATK Best Practices

Best Practices Workflows | Created 2018-01-09 | Last updated 2018-01-09

NGS Workflow

- Sampling
- Sample preparation
- Sequencing
- Mapping
- Variant Calling
- Variant Annotation
- Checks
- Variant Interpretation

Imprinting – maternal vs. paternal inheritance

- Paraganglioma and pheochromocytoma upon maternal transmission of *SDHD* mutations
- Only individuals who inherit the mutation from their father are affected
- If the mutation was inherited from the mother, they are not affected

Paraganglioma and pheochromocytoma upon maternal transmission of *SDHD* mutations

Jean-Pierre Bayley, Rogier A Oldenburg, Jennifer Nuk, Attje S Hoekstra, Conny A van der Meer, Esther Korpershoek, Barbara McGillivray, Eleonora PM Corssmit, Winand NM Dinjens, Ronald R de Krijger, Peter Devilee, Jeroen C Jansen, and Frederik J Hes

Author information ► Article notes ► Copyright and License information ► Disclaimer

Abstract

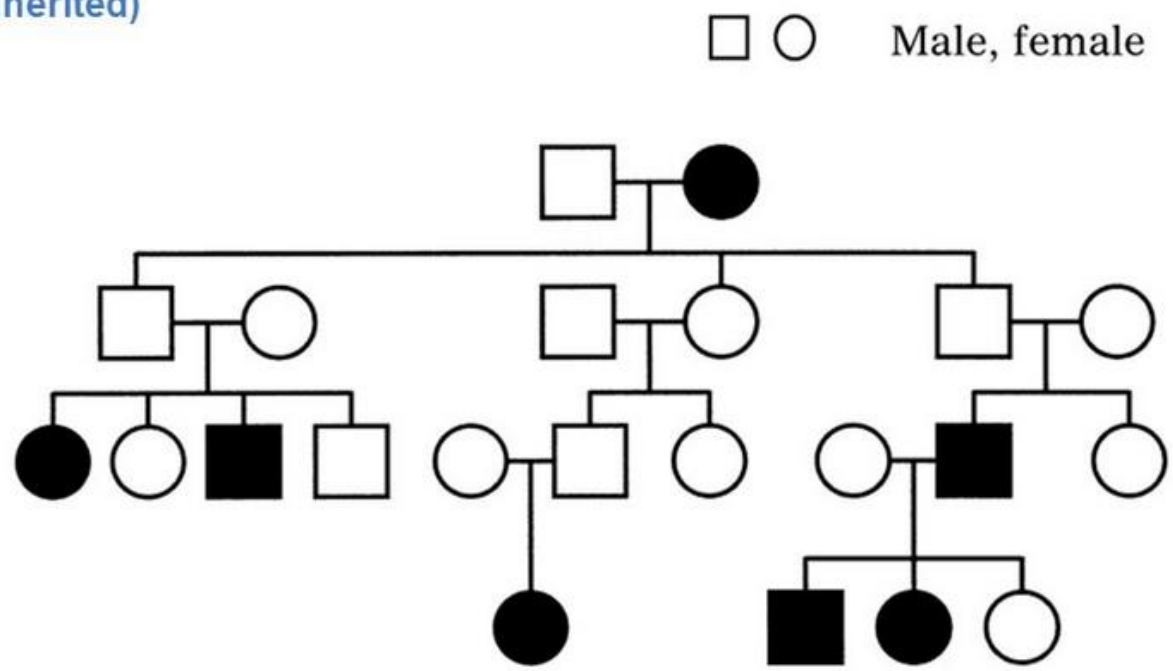
Go to:

Background

Go to:

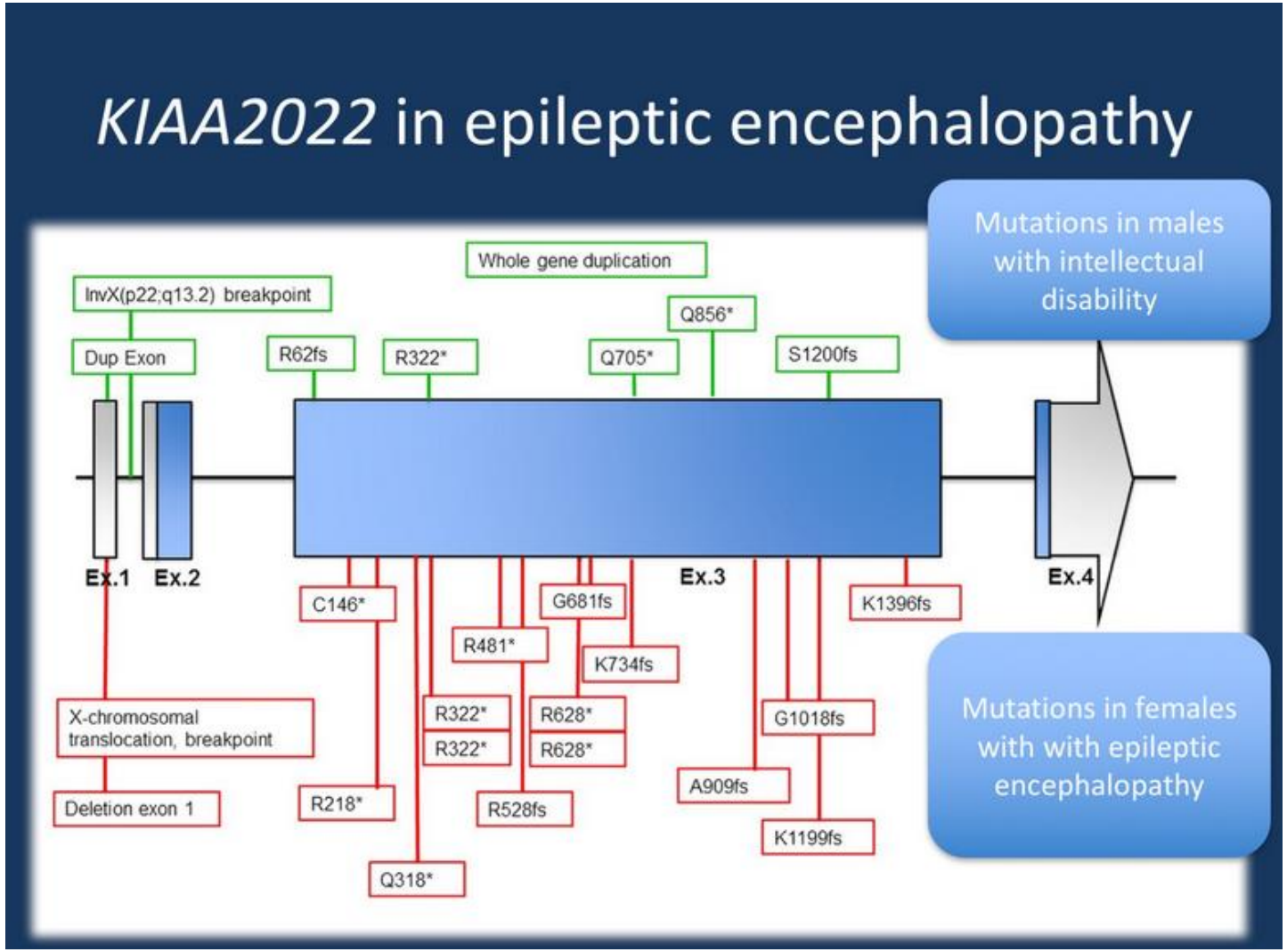
The *SDHD* gene encodes a subunit of the mitochondrial tricarboxylic acid cycle enzyme and tumor suppressor, succinate dehydrogenase. Mutations in this gene show a remarkable pattern of parent-of-origin related tumorigenesis, with almost all *SDHD*-related cases of head and neck paragangliomas and pheochromocytomas attributable to paternally-transmitted mutations.

Succinate dehydrogenase D (maternally imprinted, paternally inherited)



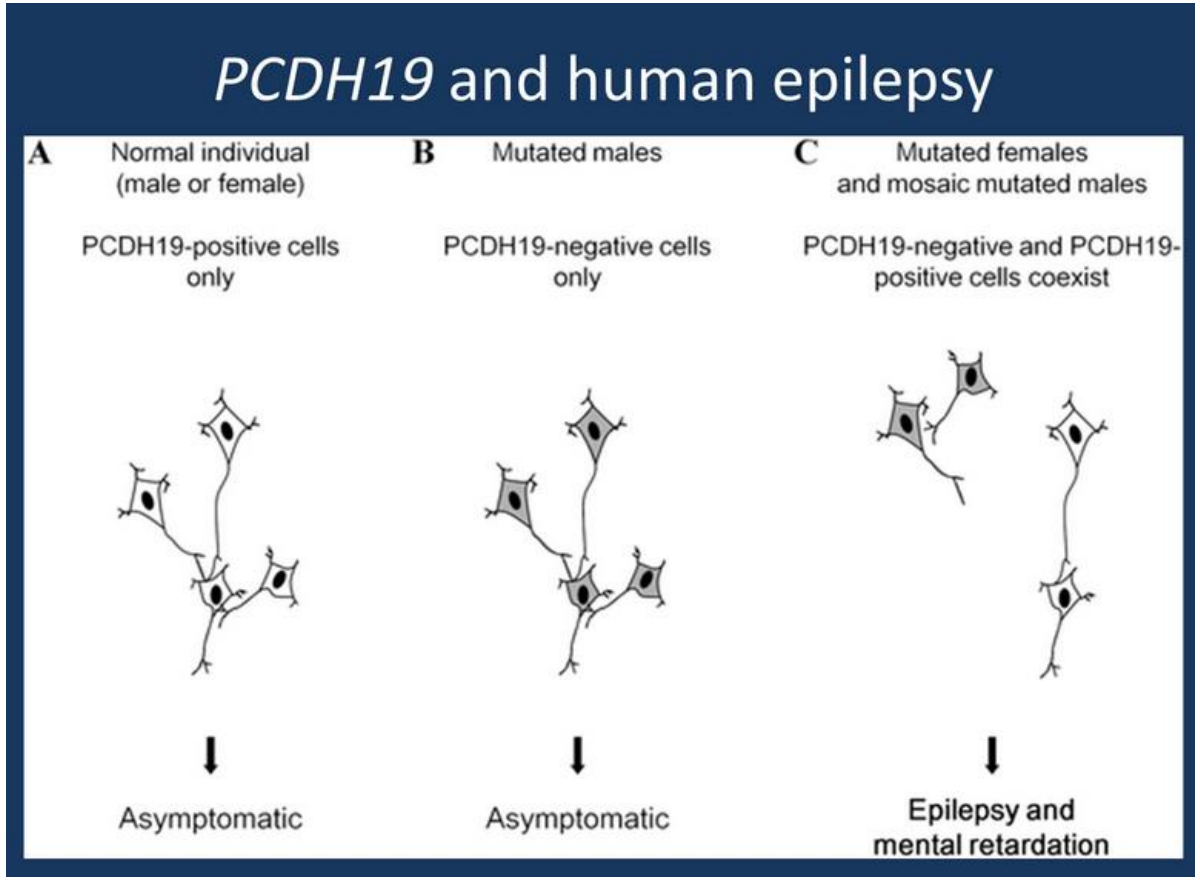
Interpretation – male / female different phenotypes

KIAA2022 – X-linked
Females (epileptic encephalopathy) / males (intelectual disability)



Interpretation – unexpected pathomechanismus

PCDH19: X-linked … but ONLY heterozygous females and mosaic males are affected



* 300460

PROTOCADHERIN 19; PCDH19

Alternative titles; symbols

KIAA1313

HGNC Approved Gene Symbol: PCDH19

Cytogenetic location: Xq22.1 ***Genomic coordinates (GRCh38): X:100,291,643-100,410,272*** (from NCBI)

Gene-Phenotype Relationships

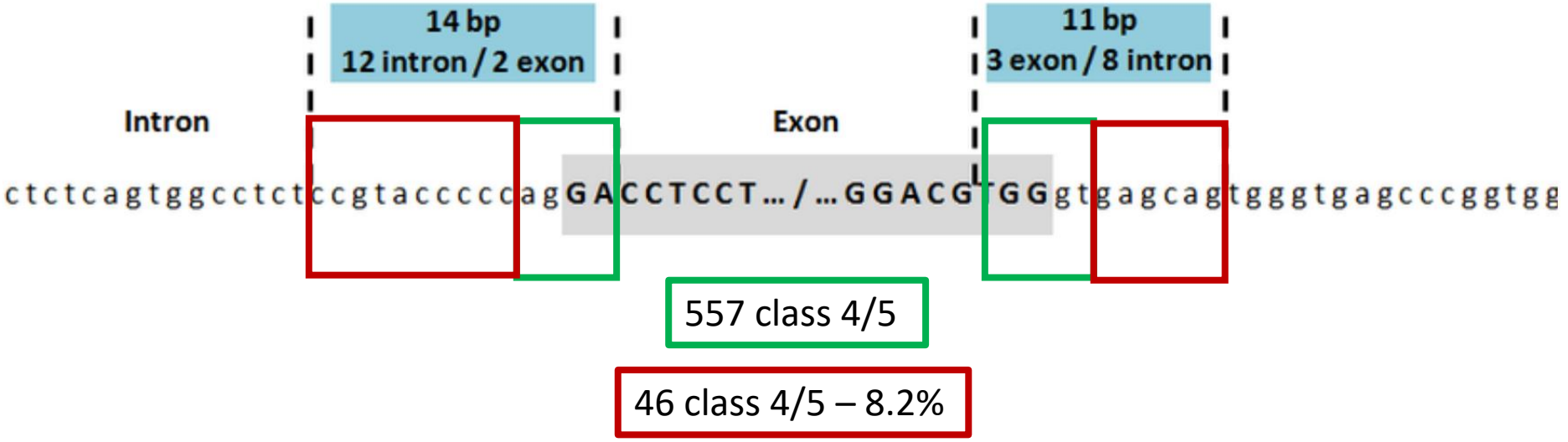
Location	Phenotype	Phenotype MIM number	Inheritance	Phenotype mapping key
Xq22.1	Epileptic encephalopathy, early infantile, 9	300088	XL	3

Interpretation softwares - Filtering intronic variants nearby exon sites???

- check for default filtering parameters, usually +/-2 !!

>10.000 Samples

6346 variants iat teh Cartegni site



„Guidelines for Splicing Analysis in Molecular Diagnosis Derived from a Set of 327 Combined *In Silico/In Vitro* Studies on *BRCA1* and *BRCA2* Variants" Tosi, Houdayer, Stoppa-Lyonnet et al. *Hum Mutat* 2013.

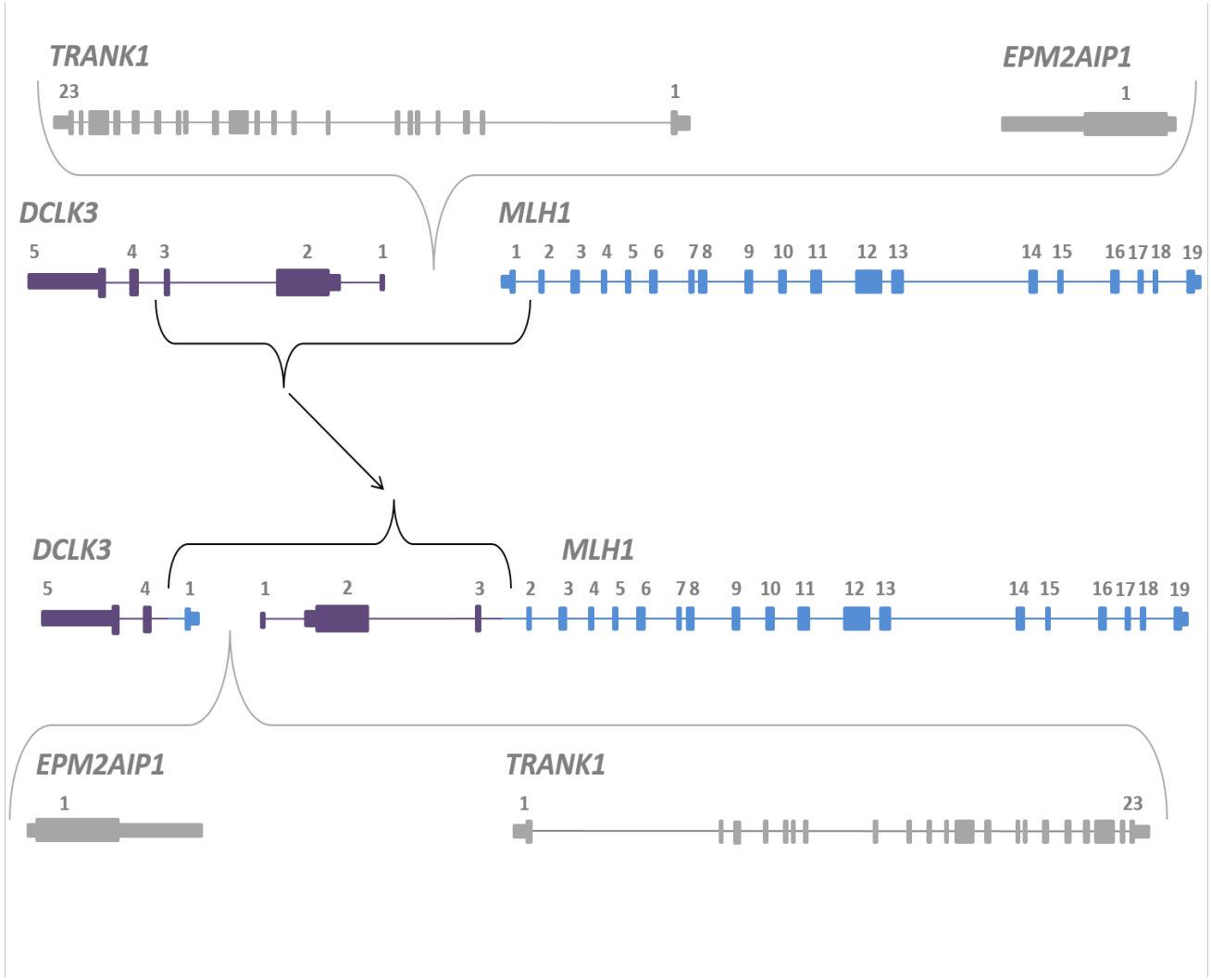
Interpretation – keep looking for the second hit!!

- IHC loss of MLH1/PMS2 in the tumour
- SNV analysis:

  heterozygous SNV in MLH1 exon 8
  NM_000249.3:c.799G>A+c.800T>G:p.Val267Arg

- CNV analysis: inconspicuous
- Deep-intronic sequencing:

Complex rearrangement on chromosome 2 that causes a paracentric inversion between the *DCLK3* gene and *MLH1*



Murphys law also apply for diagnostics!!!