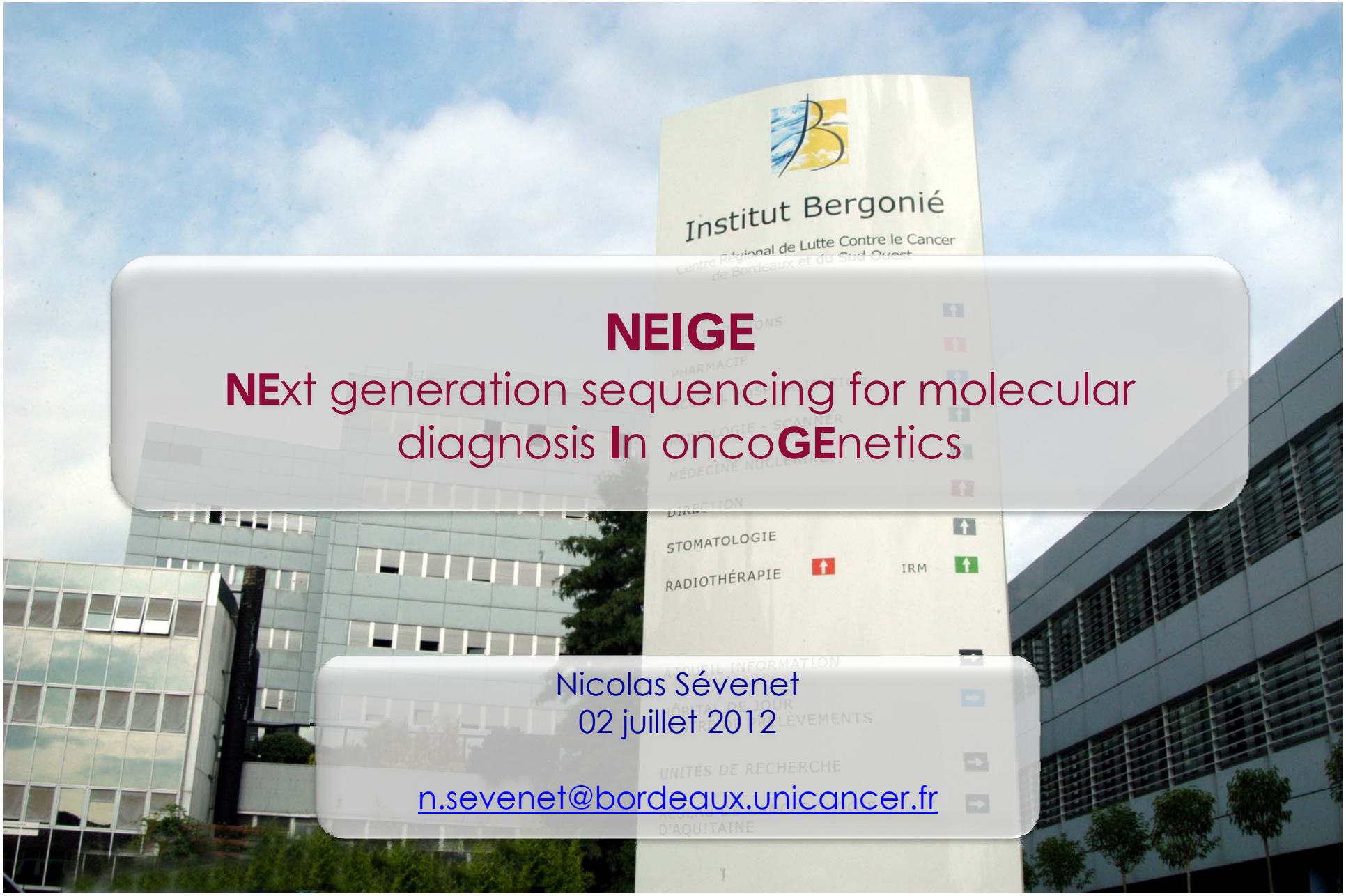


NEIGE

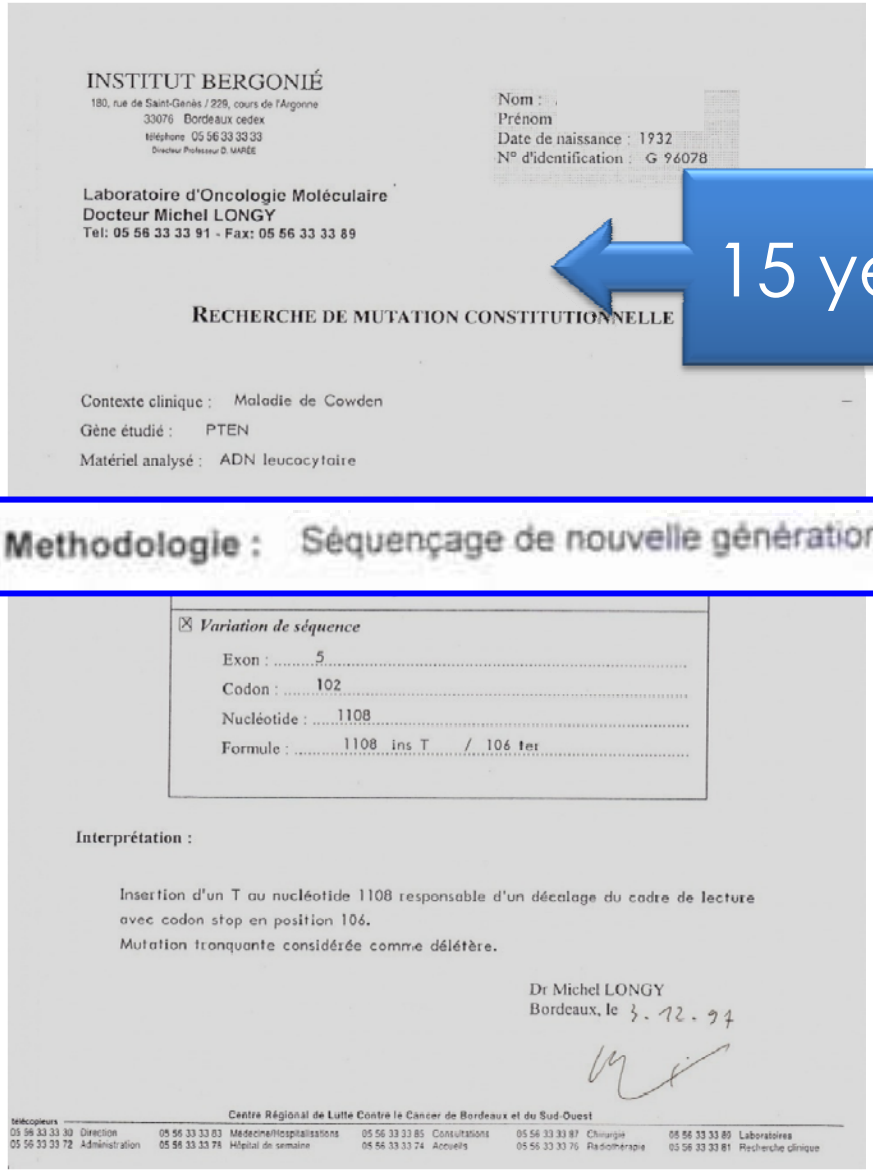
NExt generation sequencing for molecular diagnosis **I**n onco**GE**netics

Nicolas Sévenet

02 juillet 2012

n.sevenet@bordeaux.unicancer.fr

Reports



INSTITUT BERGONIÉ
180, rue de Saint-Genès / 229, cours de l'Argonne
33076 Bordeaux cedex
téléphone 05 56 33 33 33
Directeur Professeur D. MAREE

Nom :
Prénom
Date de naissance : 1932
N° d'identification : G 96078

Laboratoire d'Oncologie Moléculaire
Docteur Michel LONGY
Tel: 05 56 33 33 91 - Fax: 05 56 33 33 89

RECHERCHE DE MUTATION CONSTITUTIONNELLE

Contexte clinique : Maladie de Cowden
Gène étudié : PTEN
Matériel analysé : ADN leucocytaire

☒ *Variation de séquence*
Exon : 5
Codon : 102
Nucléotide : 1108
Formule : 1108 ins T / 106 ter

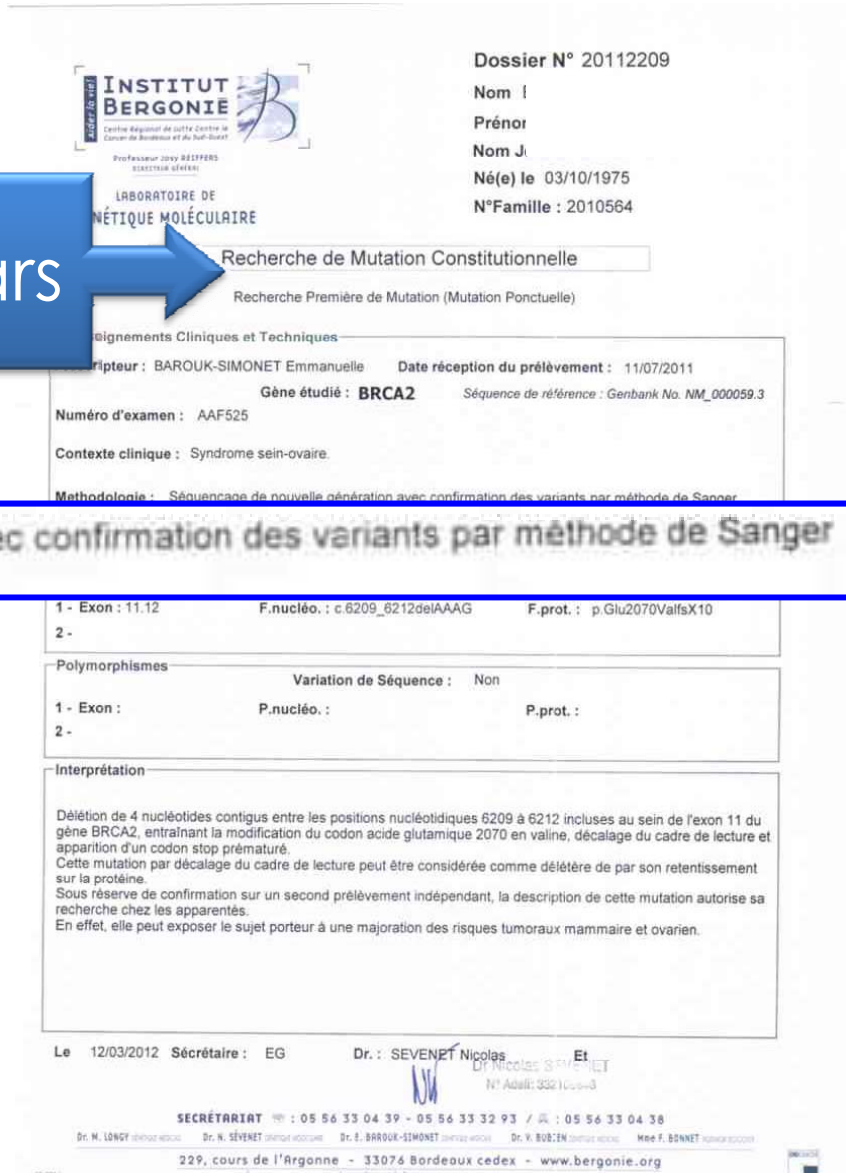
Interprétation :

Insertion d'un T au nucléotide 1108 responsable d'un décalage du cadre de lecture avec codon stop en position 106.
Mutation tronquante considérée comme délétère.

Dr Michel LONGY
Bordeaux, le 3.12.97

Centre Régional de Lutte Contre le Cancer de Bordeaux et du Sud-Ouest
télécopieurs
05 56 33 33 30 Direction
05 56 33 33 72 Administration
05 56 33 33 83 Médecine/Hospitalisations
05 56 33 33 78 Hôpital de semaine
05 56 33 33 85 Consultations
05 56 33 33 74 Accueils
05 56 33 33 87 Chirurgie
05 56 33 33 76 Radiothérapie
05 56 33 33 80 Laboratoires
05 56 33 33 81 Recherche clinique

15 years



INSTITUT BERGONIÉ
Centre Régional de Lutte Contre le Cancer de Bordeaux et du Sud-Ouest
Professeur Joxy REIFFERS
DIRECTEUR GÉNÉRAL

LABORATOIRE DE
GÉNÉTIQUE MOLÉCULAIRE

Dossier N° 20112209
Nom [illegible]
Prénom
Nom J
Né(e) le 03/10/1975
N°Famille : 2010564

Recherche de Mutation Constitutionnelle

Recherche Première de Mutation (Mutation Ponctuelle)

Renseignements Cliniques et Techniques

Prescripteur : BAROUK-SIMONET Emmanuelle **Date réception du prélèvement :** 11/07/2011
Gène étudié : **BRCA2** *Séquence de référence : Genbank No. NM_000059.3*
Numéro d'examen : AAF525
Contexte clinique : Syndrome sein-ovaire
Methodologie : Séquençage de nouvelle génération avec confirmation des variants par méthode de Sanger

Methodologie : Séquençage de nouvelle génération avec confirmation des variants par méthode de Sanger

1 - Exon : 11.12 **F.nucléo. :** c.6209_6212delAAAG **F.prot. :** p.Glu2070ValfsX10
2 -

Polymorphismes
Variation de Séquence : Non
1 - Exon : **P.nucléo. :** **P.prot. :**
2 -

Interprétation

Délétion de 4 nucléotides contigus entre les positions nucléotidiques 6209 à 6212 incluses au sein de l'exon 11 du gène BRCA2, entraînant la modification du codon acide glutamique 2070 en valine, décalage du cadre de lecture et apparition d'un codon stop prématuré.
Cette mutation par décalage du cadre de lecture peut être considérée comme délétère de par son retentissement sur la protéine.
Sous réserve de confirmation sur un second prélèvement indépendant, la description de cette mutation autorise sa recherche chez les apparentés.
En effet, elle peut exposer le sujet porteur à une majoration des risques tumoraux mammaire et ovarien.

Le 12/03/2012 **Sécrétaire :** EG **Dr. :** SEVENET Nicolas **Et**

SECRÉTARIAT : 05 56 33 04 39 - 05 56 33 32 93 / : 05 56 33 04 38
Dr. M. LONGY Dr. N. SÉVENET Dr. E. BAROUK-SIMONET Dr. V. BOBIEN Mme F. BONNET
229, cours de l'Argonne - 33076 Bordeaux cedex - www.bergonie.org

Next generation sequencing

- 06/2011
  - Earray (Agilent technologies) : targeted resequencing of 460 exons of 25 genes
    - Database : Ensembl 62, GRChp37
    - Repeat Masker UCSC
    - Oligos 120 mers, tiling 2X, +/- 50 bp around exons

    → 2666 oligos duplicated 19 or 38 X → 57750 oligos total)

    → 0,187 Mb

- 10/2011 → 02/2012

  2 experiments
  - NEIGE 1 = 12 samples
    - Positive controls
      - 9 substitutions
      - 5 delins
    - **Illumina GAIIX** → PGT-CGFB Bordeaux
    - **Illumina MiSeq** → GeT; INRA Toulouse
  - NEIGE2 = 30 samples
    - 16 positive controls with « complex » mutations
    - 14 samples double blind diagnostic series

Capture library selection

PNAS

Detection of inherited mutations for breast and ovarian cancer using genomic capture and massively parallel sequencing

Tom Walsh[a], Ming K. Lee[a], Silvia Casadei[a], Anne M. Thornton[a], Sunday M. Stray[a], Christopher Pennil[b], Alex S. Nord[a], Jessica B. Mandell[a], Elizabeth M. Swisher[b], and Mary-Claire King[a,1]

[a]Departments of Medicine and Genome Sciences and [b]Obstetrics and Gynecology, University of Washington, Seattle, WA 98195

Contributed by Mary-Claire King, June 8, 2010 (sent for review May 17, 2010)

Table 1. Genomic regions targeted for breast and ovarian cancer genes

		Captured genomic region	
Gene	Chromosome	Start	End
BRCA1	17	41,186,313	41,347,712
BRCA2	13	32,879,617	32,983,809
CHEK2	22	29,073,731	29,147,822
PALB2	16	23,604,483	23,662,678
BRIP1	17	59,759,985	59,940,755
p53	17	7,561,720	7,600,863
PTEN	10	89,613,195	89,738,532
STK11	19	1,195,798	1,238,434
CDH1	16	68,761,195	68,879,444
ATM	11	108,083,559	108,249,826
BARD1	2	215,583,275	215,684,428
MLH1	3	37,024,979	37,102,337
MRE11	11	94,140,467	94,237,040
MSH2	2	47,620,263	47,720,360
MSH6	2	48,000,221	48,044,092
MUTYH	1	45,784,914	45,816,142
NBN	8	90,935,565	91,006,899
PMS1	2	190,638,811	190,752,355
PMS2	7	6,002,870	6,058,737
RAD50	5	131,882,630	131,989,595
RAD51C	17	56,759,963	56,821,692

BRCA1	uc002ict.2
BRCA2	uc001uub.1
BRIP1	uc002izk.1
PALB2	uc002dlx.1
ATM	uc001pkb.1
BARD1	uc002veu.2
CHEK2	uc003adt.1
RAD51C	uc002iwu.2
PTCH1	uc004avk.3
PTCH2	uc010olf.1
SUFU	uc001kvy.1
PIK3CA	uc003fjk.2
PTEN	uc001kfb.2
MLH1	uc003cgl.2
MSH2	uc002rvy.1
MSH6	uc002rwd.3
MUTYH	uc001cnh.2
PMS2	uc003spl.2
APC	uc003kpy.3
EPCAM	uc002rvx.2
CDH1	uc002ewg.1
STK11	uc002lrl.1
MRE11	uc001peu.2
RAD50	uc003kxi.2
TP53	uc002gij.2

Breast & Ovary cancers

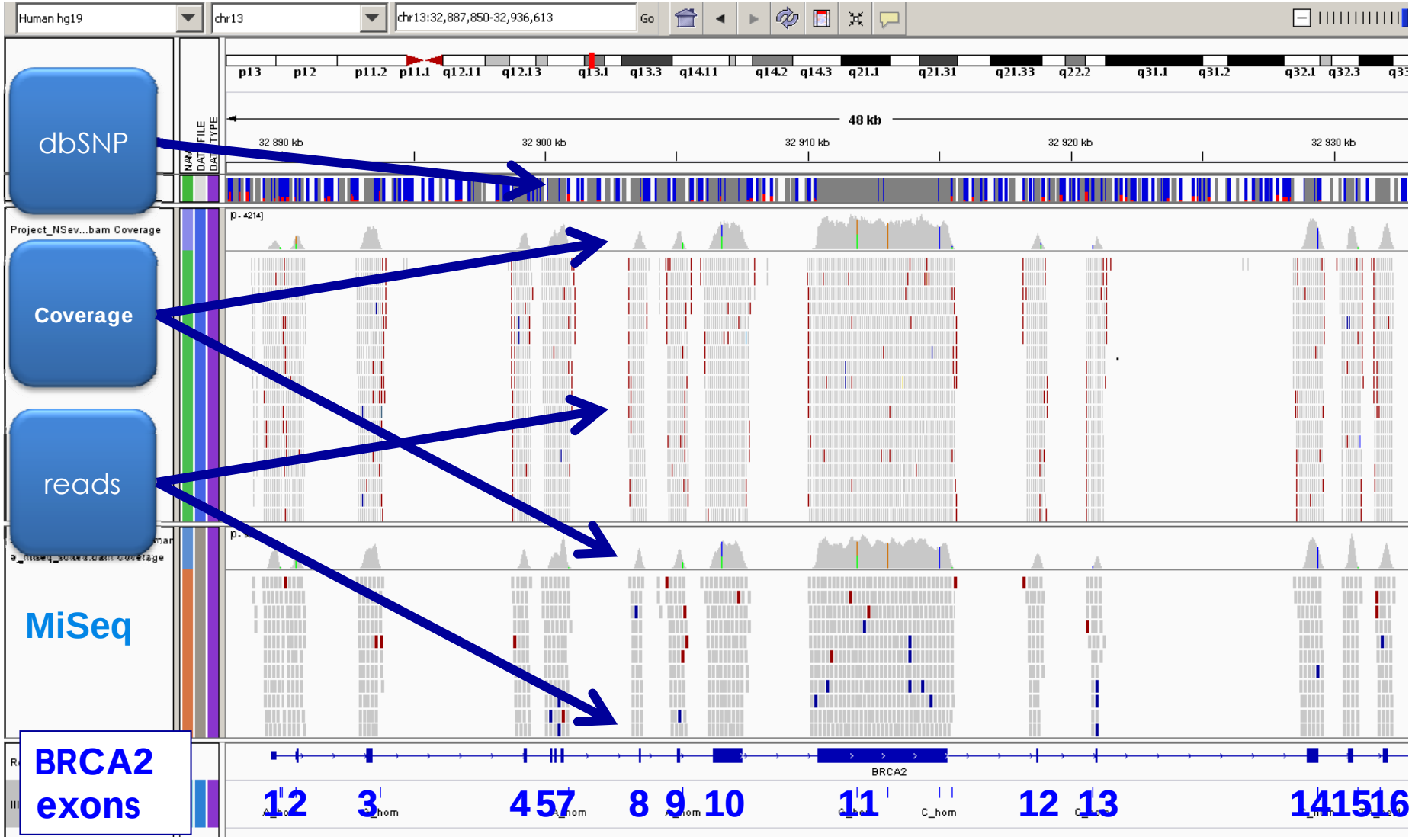
Transduction

Colon Cancer

Other

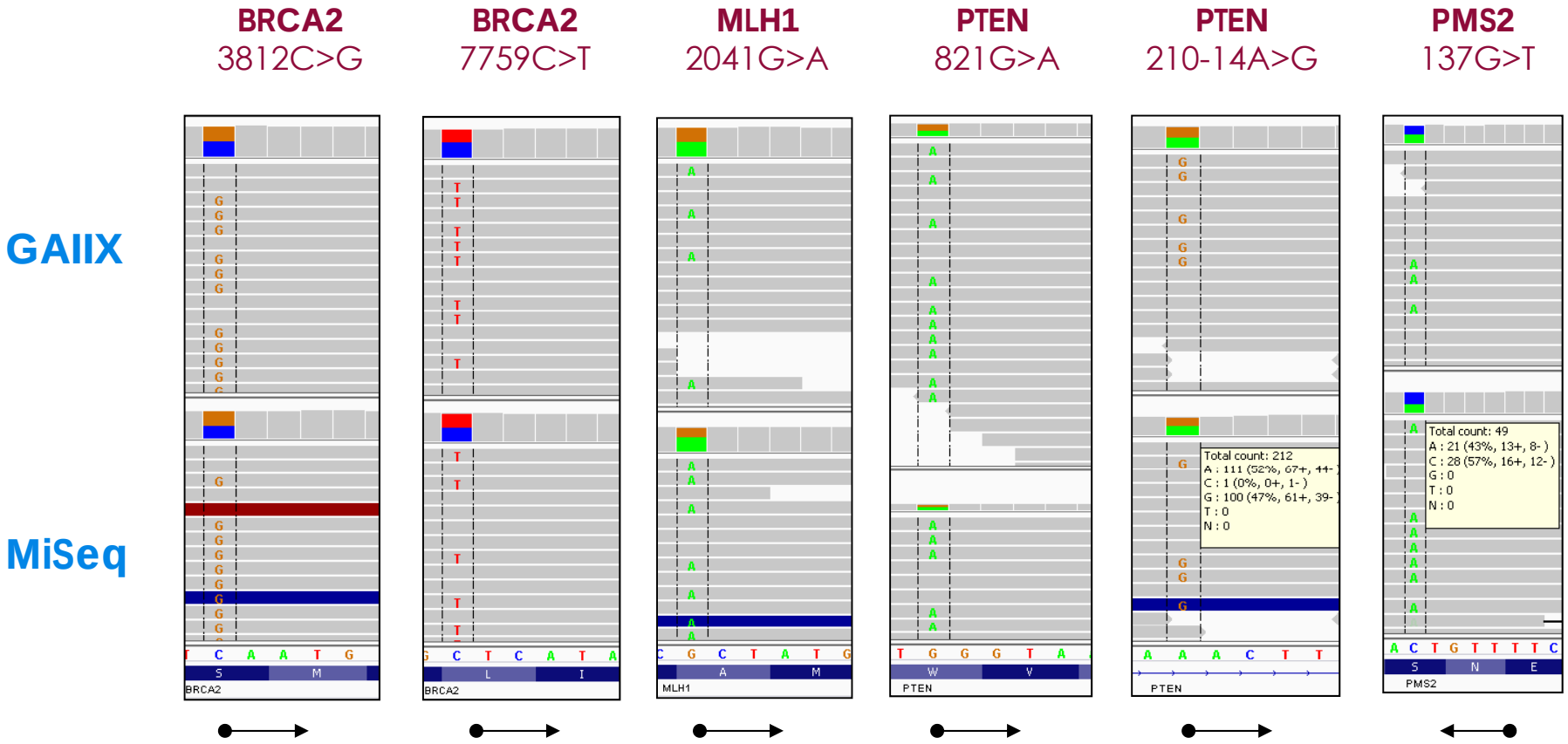
Integrative Genome Viewer (Broad Institute)

BRCA2 exons 1-16

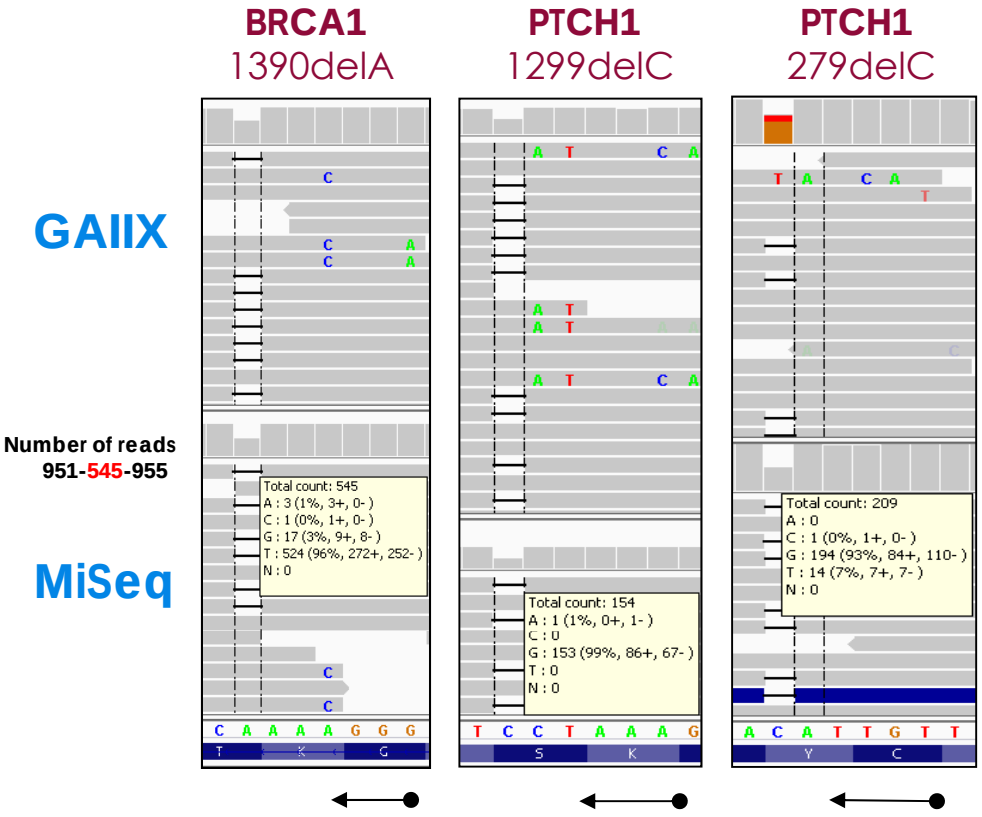


Integrative Genome Viewer (Broad Institute)

Substitutions



Integrative Genome Viewer (Broad Institute)
1-bp deletion



Results of the first set of experiment
6 genes in 12 samples

Variants attendus **(total = 72)**	GAIIX (CASAVA)	MiSeq (MiSeq Rep)
polymorphismes	**58**	**58**
substitutions	53	53
délétions	3	3
insertions/2	1*	2
mutations	**14**	**14**
faux sens	9	9
délétions	3	3
duplication	1	1
delétion insertion	1	1*

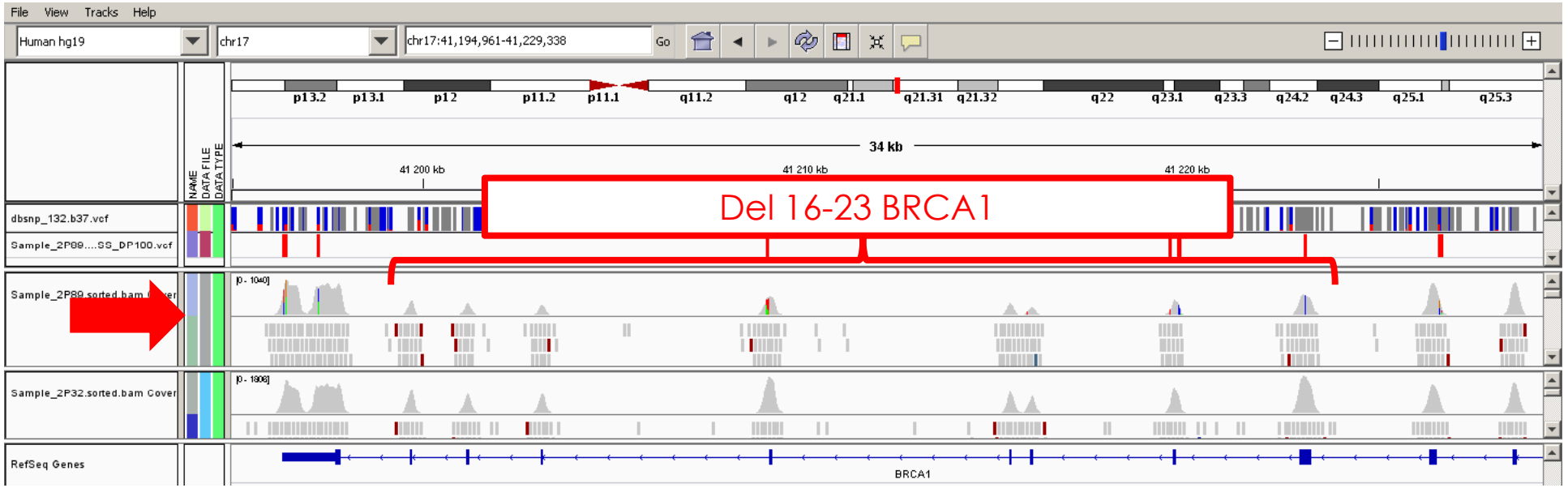
Variants additionnels **(total=151)**	GAIIX (CASAVA)	MiSeq (MiSeq Rep)
non vus par précriblage	3	3
nouveaux polymorphismes (hétérozygotes)	**90**	**90**
substitutions	80	80
délétions	6	6
insertions	4	4
nouveaux polymorphismes (homozygotes)	**61**	**61**
nouveaux polymorphismes (hétérozygotes) en +/-50	1	1
nouveaux polymorphismes (homozygotes) en +/-50	18	18
off Target (Nb de régions)	**9**	**9**

*: mal annoté

NEIGE2

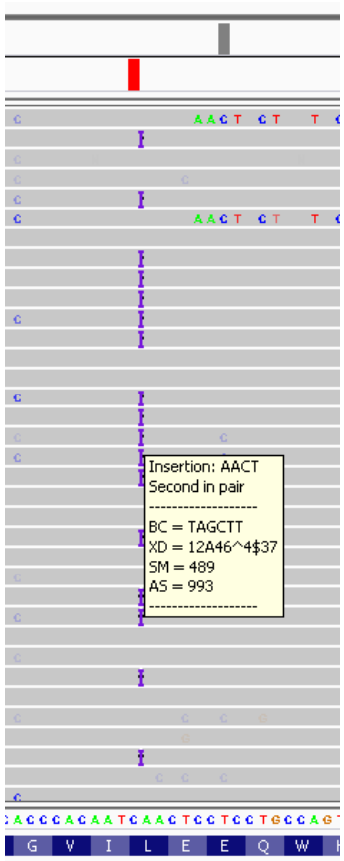
- 16 positive controls with « complex » mutations
  - 4 Gross gene rearrangement (GGR)
  - 8 delins
  - 4 point mutation
- 14 DNA sample from patients not previously screened
  - Consultations : 07/2011
  - Double-blind study
    - Gaëlle Geneste, Françoise Bonnet : EMMA-Sanger sequencing
    - Delfine Lafon, Nicolas Sévenet : Next-Generation sequencing
  - 100% of concordance
    - NGS seems to be more sensitive than our current screening method (dHPLC, EMMA, HRM) due to its detection and identification of homozygous variants
  - Résultats
    - BRCA1 : 1 GGR (del exons 21-24 included), 2 Unknown Variant
    - BRCA2 : 1 frameshift

Positive controls
Gross Gene Rearrangement

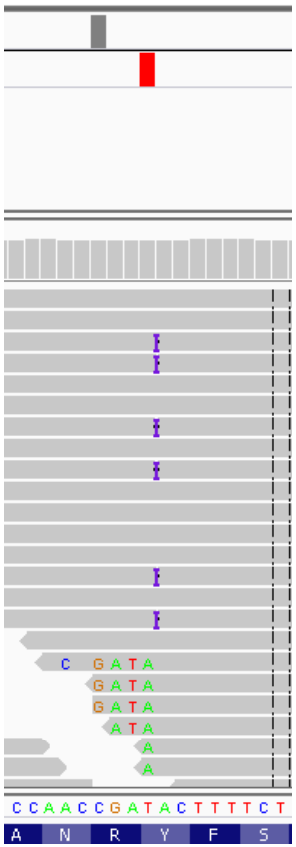


Positive controls
delins

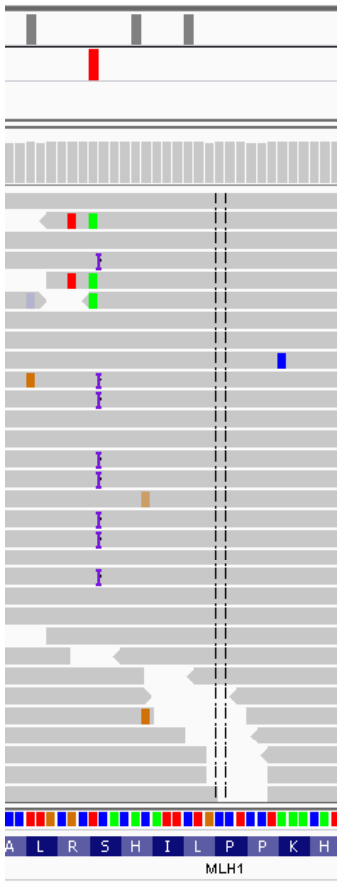
PTCH1
Exon 7
c.1019_1022dupAGTT
p.Ile342ValfsX96



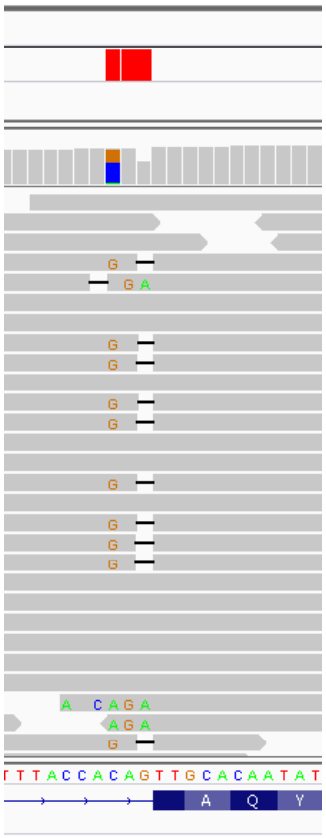
PTEN
Exon 8
c.1807dupA
p.Tyr336X



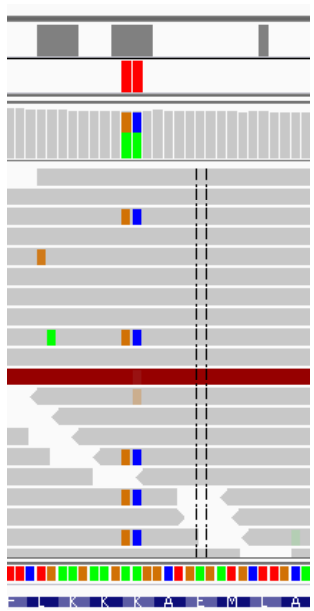
MLH1
Exon 19
c.2181_2182dupCA
p.Ile728ThrfsX56



PTEN
Intron 4
c.254-1delG

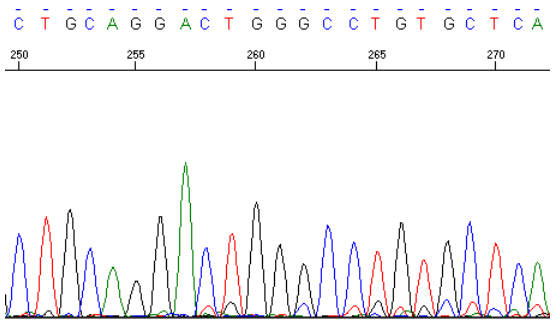


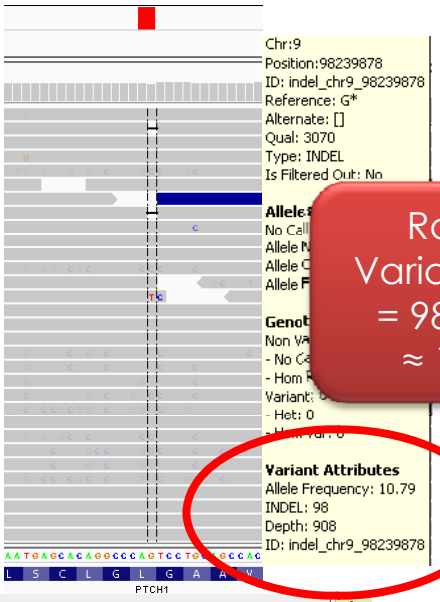
MLH1
Exon 16
c.1852_1853delinsGC
p.Lys618Ala



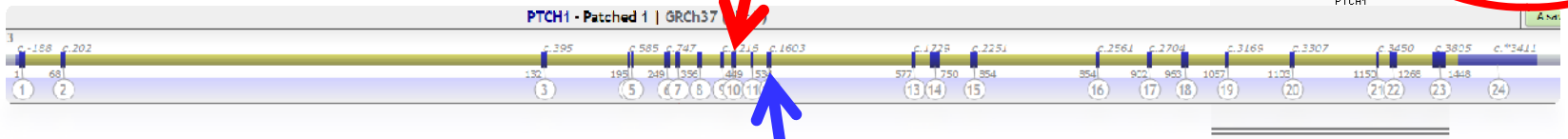
Positive controls
PTCH1 Mosaicism

Germline
Deleterious Mutation
Exon 10; c.1453delC;
p.Leu485TrpfsX6

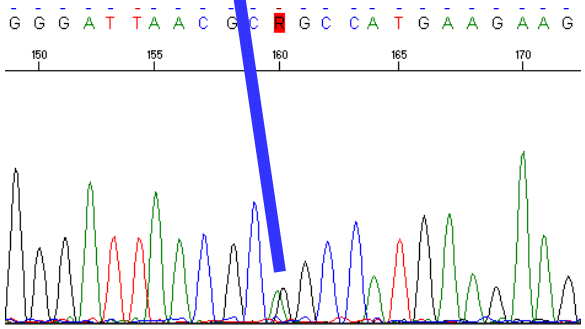


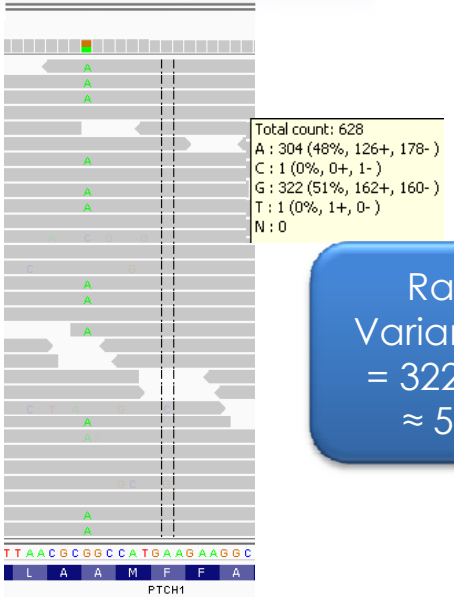


Ratio
Variant/WT
= 98/908
≈ 10%



Germline
polymorphism
Exon 12; c.1686C>T;
p.Ala562Ala
Reverse sequence



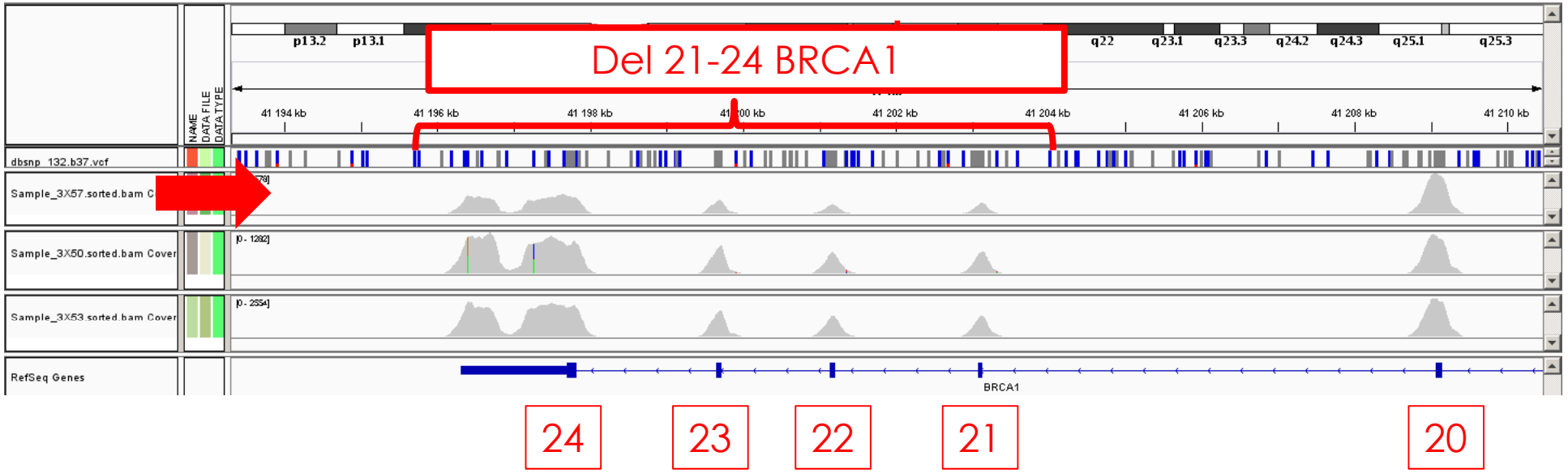


Ratio
Variant/WT
= 322/304
≈ 50%

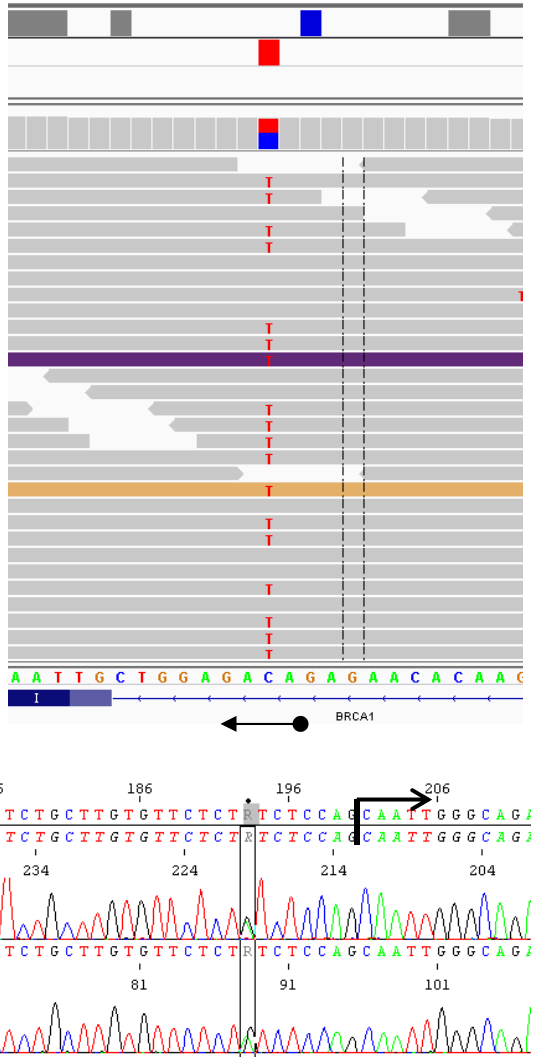
Double blind study

BRCA1

del 21-24 : c.5278-?_5592+?del

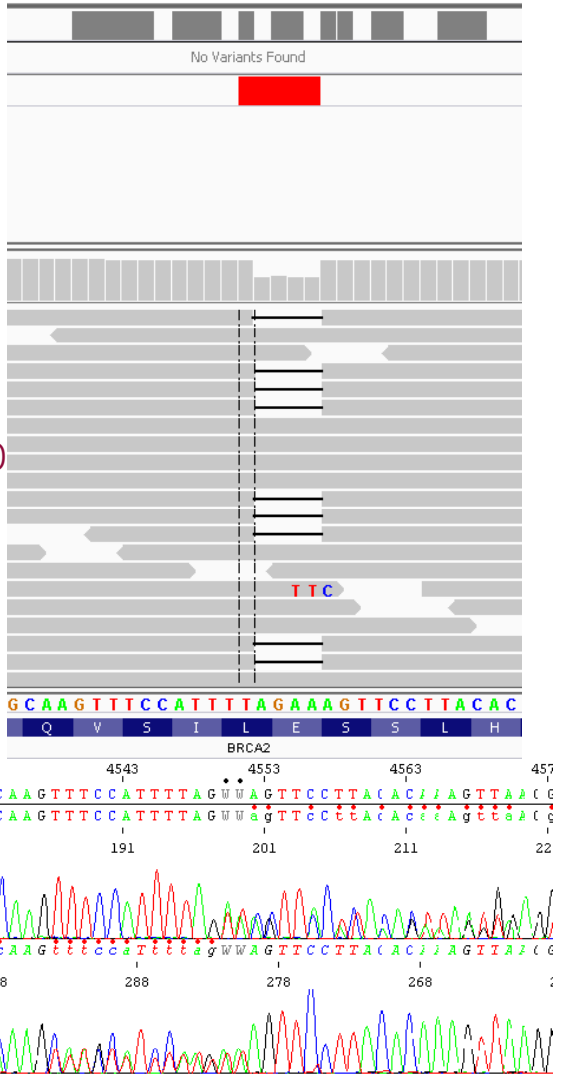


Double blind study
Point mutations

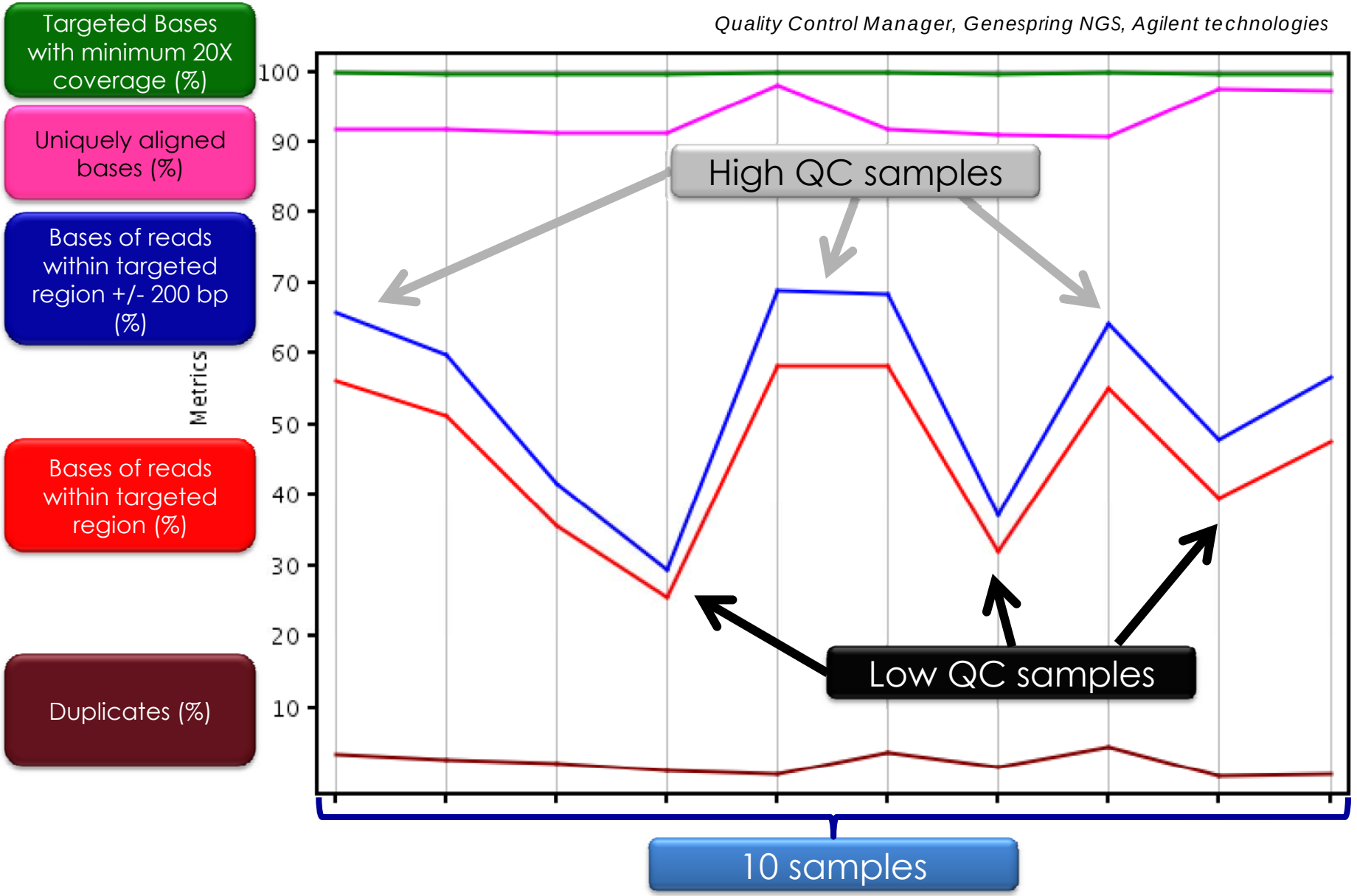


BRCA1
Exon 24
c.5468-8G>A

BRCA2
Exon 24
c.6209delAAAG
p.Glu2070ValfsX10



Capture statistics



Capture statistics

Uniquely aligned bases (%)

Bases of reads within targeted region (%)

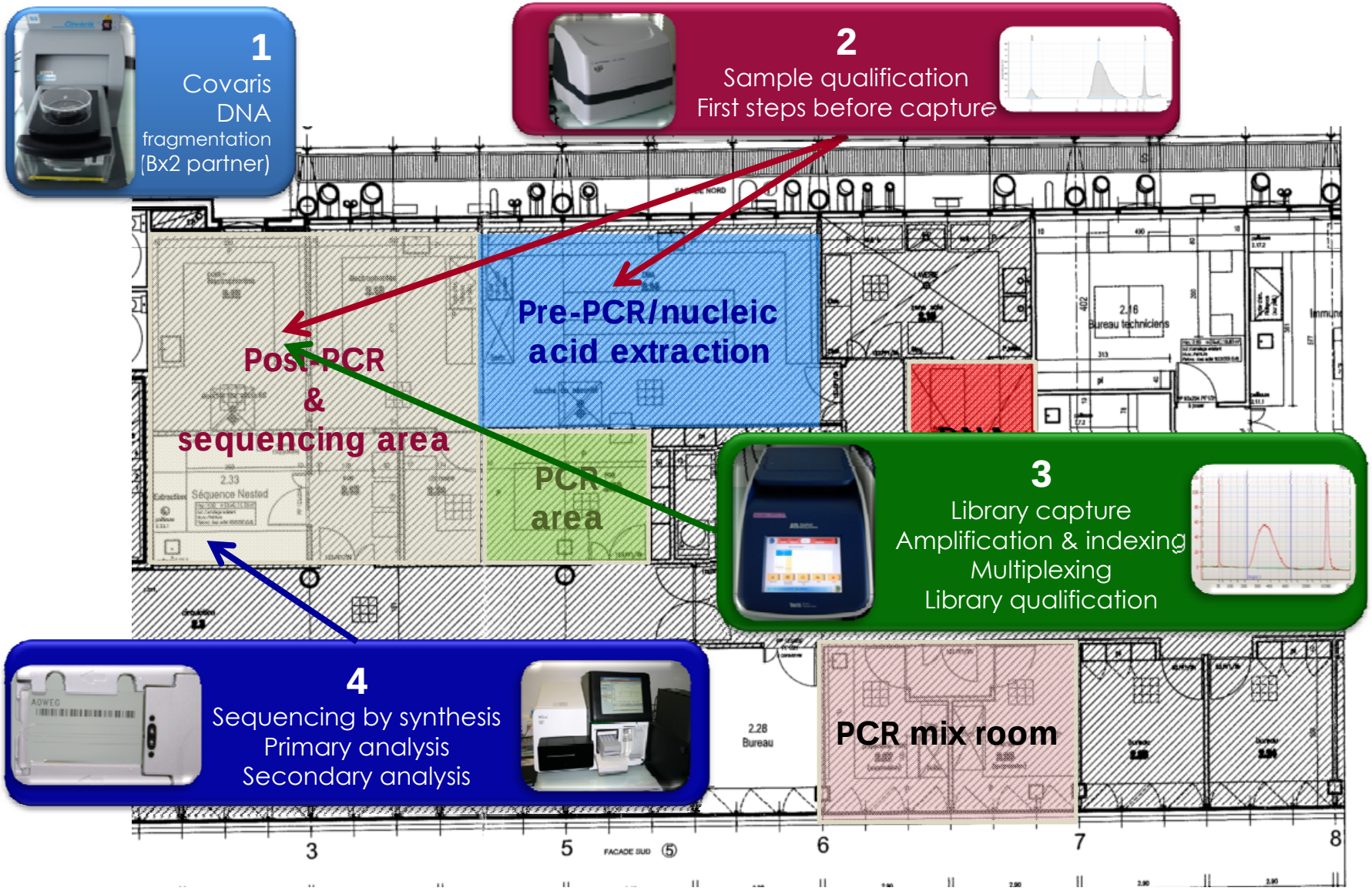
Duplicates (%)

Targeted Bases with minimum 20X coverage (%)

	Metrics	Low QC 1F18_low	High QC 1K24_high
Reads	Total reads	1217364	1614132
	Uniquely mapped reads (#)	1001698	1407682
	Unmatched reads (#)	149918	104719
	Avg read length(Reads in targeted regions only)	149,96	150,05
	Reads in targeted regions (#)	339327	990656
	Reads in targeted regions (%)	**31,79**	**65,63**
Bases	Total bases of reads (includes bases of unmatched reads)(#)	182332398	241968474
	Total bases of mapped reads(#)	**159694780**	**226155905**
	Uniquely Aligned Bases (#)	149768158	210796283
	Uniquely Aligned Bases (%)	**82,14**	**87,12**
	Bases of reads within targeted regions (#)	45140999	131991704
	Bases of reads within targeted regions (%)	**28,27**	**58,36**
Enrichment	Genome targeted (%)	0,01	0,01
	Enrichment in targeted regions (fold)	5234,16	10806,61
Coverage	Average coverage (fold)	240,07	701,97
	Median coverage (fold)	**239**	**693**
Targeted regions	Targeted regions (#)	460	460
	Targeted regions covered by at least 1 read (#)	460	460
	Targeted regions covered by at least 5 read (#)	460	460
	Targeted regions covered by at least 10 read (#)	460	460
	Targeted regions covered by at least 20 read (#)	459	460
	Targeted regions covered by at least 30 read (#)	459	460
	Targeted regions covered by at least 40 read (#)	458	460
	Zero Coverage Targeted Regions(#)	0	0
	Zero Coverage Targeted Regions(%)	0	0
Duplicates	Duplicates (#)	55597	147723
	Duplicates (%)	**10,32**	**19,44**
Bases in targeted regions	Bases in targeted region	**188010**	**188010**
	Targeted bases with minimum 1x coverage (%)	99,97	99,99
	Targeted bases with minimum 5x coverage (%)	99,87	99,95
	Targeted bases with minimum 10x coverage (%)	99,81	99,9
	Targeted bases with minimum 20x coverage (%)	**99,61**	**99,86**
	Targeted bases with minimum 30x coverage (%)	99,38	99,83
	Targeted bases with minimum 40x coverage (%)	99,17	99,8
	Targeted bases with minimum 1Kx coverage (%)	0	14,28

Molecular genetic lab, Institut Bergonié

200 m² technical platform



Molecular genetic lab, Institut Bergonié

NGS : bioinformatic analysis workflow

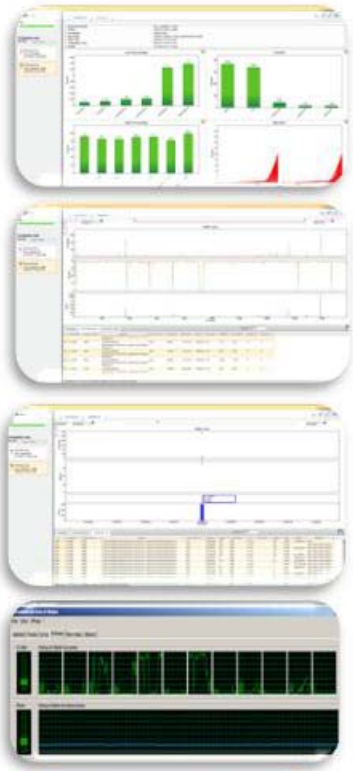
1
Sample sheet elaboration
Illumina Experiment Manager

2
Sequencing by synthesis (MiSeq)

Sequencing Analysis Viewer
Primary Analysis

3
Transfert on a server (HP)

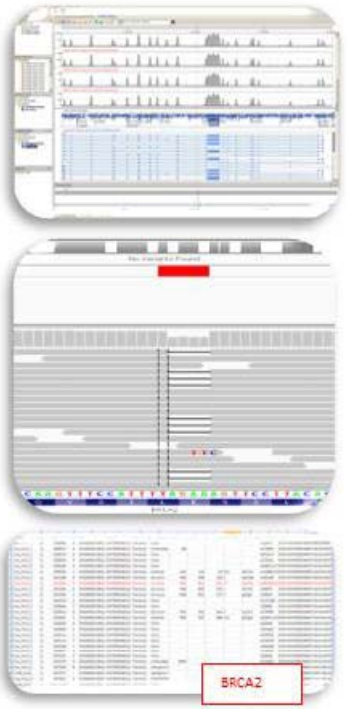
CASAVA 1.8



4
Secondary Analysis

MiSeq reporter
On MiSeq or PC

→ Bam files & vcf files



5
Third Analysis Qualiy Control
Genespring NGS
IGV
VEP (Ensembl)
Alamut HT
Specialized databases

Future

- Validation phase
  - Experiments :
    - 150 germline DNA samples double blind analyzed before 12/2012
  - Organization
    - 10 samples multiplexed each week
    - Analysis time < 2 weeks
- Developments
  - Haloplex design & test (09/2012)
  - V2 capture library design
    - Enrichment for BAP1 & RAD51 family involved in HBOC
    - Enrichment for DDB2 involved in Basal cell carcinoma syndrome
  - Multiplexing of 25 samples with the 7Gb flowcell
- Expected
  - NGS as the sole molecular diagnostic method from 02/2013

Acknowledgments

Institut Bergonié
Delfine Lafon
Françoise Bonnet
Équipe oncogénétique
Michel Longy
Nicolas Kihl
Équipe informatique hôpital

Integragen
Francis Rousseau
Bérengère Genin

Agilent technologies
Juliette Grégoire
Hervé Chaulet
Didier Goidin
Florent Brun

Plateforme Génome et transcriptome – Univ. Bordeaux Segalen
Christophe Hubert
Jennifer Dupiot

Plateforme Génome et transcriptome GeT – Inra Toulouse
Denis Milan
Cécile Donnadieu
Gérald Salin
Frédéric Escudié
Jérôme Lluch
Emeline Llhuillier

Illumina
Caroline Thureau
Amélie Castro
Tiphaine Godefroy

Bioinformatic infrastructure