Department of Science, Technology, Engineering, and Math (STEM), Qatar Science and Technology School

Bioinformatics Analysis of BRCA1 Founder Mutations for Breast Cancer Diagnosis

ABSTRACT

Aims: to investigate the potential use of founder mutations occurring to exons 2 and 20 of BRCA1 gene as genetic biomarkers in the diagnosis of BC.

Tools: Bioinformatics online tools and software for mutation analysis and validation including webcutter, primer3, and In silico PCR.

The results: Forward and reverse primers were successfully designed to detect 187delAG and 5385insC mutations occurring to exons 2 and 20, respectively in BRCA1. In silico PCR has validated the success of the designed primers. Analyzing exons 2 and 20 of the BRCA1 gene revealed that the reference sequence has *18* different restriction sites for a wide range of restriction enzymes such as Alw21I, ApaLI, BstXI and others. The restriction sites can be used in the detection of mutant BRCA1.

CONCLUSION

BRCA1 single nucleotide polymorphisms in exons 2 and 20 could be useful biomarkers of BC.

The objective of the current research is to evaluate the usage of exon 2 and 20 somatic and inherited mutations 187delAG and 5385insC of BRCA1 as biomarkers in the diagnosis and prognosis of Breast cancer.

INTRODUCTION

Breast cancer usually results from an accumulation of many mutations thus the incidence increases with age. At the DNA level, a cancerous cell is usually characterized by at least one active oncogene and the mutation of several tumor-suppressor genes. Breast cancer results from a series of dysregulated mechanisms involving the gain-of-function of oncogenes (OCGs) and the loss-of-function of tumour suppressor genes (TSGs) including BRCA1, BRCA2, P53, PTEN, STK11, CHEK2, ATM, BRIP1, and PALB2. Abnormal repression of tumour suppressor gene (BRCA1) results in deregulation of the cell cycle or fail to repair DNA damage mainly through the modification in the protein products types and quantities thereby leading to the mutations in other genes and thus developing the transformed phenotypes of the breast cancer cells.

Materials and Methods

1. NCBI

provide information about the intron / exon structure, the coding sequences, and references.

1. OMIM

provide all the allelic variants of the BRCA1/2 genes and the corresponding disorders.

1. Ensembl

Ensembl data base will be used to determine the domains of the gene in a graphical view.

1. PDB

The protein data bank (PDB) was used to download the amino acid sequence coded by the BRCA1/2 gene.

1. RasMol (RasWin)

The RasMol (RasWin) program will be used to view the protein sequence in 3D.

1. COSMO

This tool was used to obtain the complementary DNA sequence (cDNA) that contains the exon sequences only and to obtain the amino acid sequence coded by the BRCA1/2 gene as well as to determine the position of the mutations in the BRCA1/2 gene.

1. Webcutter

Webcutter was used to generate the restriction sites and to construct the restriction map.

1. Prime3

Primer3 was used to generate the primers for the given DNA sequences.

1. In silico PCR

In silico PCR was used to validate and confirm the success of the designed primers.

Table 1. Workflow of the project.

Results

The BRCA1 gene (gene ID: 1956) is located on the short arm of chromosome 7 with the cytogenetic location (7p11.2) as indicated by the arrow on figure 1. The size of the BRCA1 gene is 189,060 bp comprising 31 exons (table 2). The exact molecular location of the BRCA1 gene is between the base pairs 55,019,021 to 55,208,080 (Homo sapiens Annotation Release 109, GRCh38.p12).

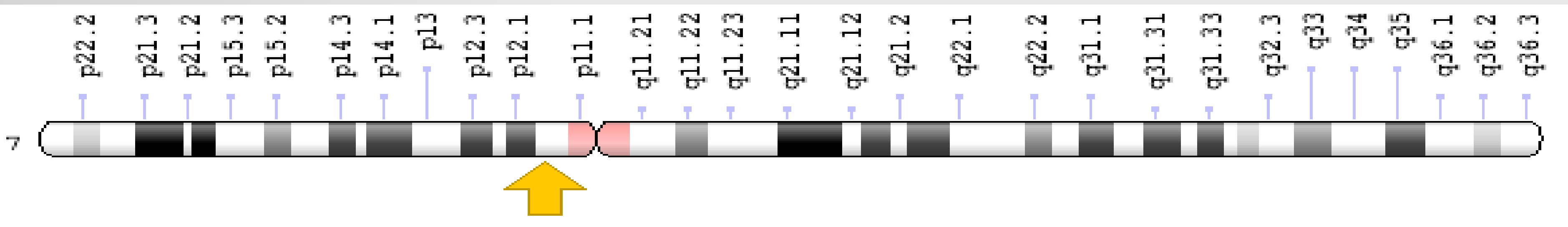


Fig. 1. Cytogenetic location of the BRCA1 gene as indicated by the arrow.

1: * 131550. EPIDERMAL GROWTH FACTOR RECEPTOR; EGFR

EGFR/SEPT14 FUSION GENE, INCLUDED

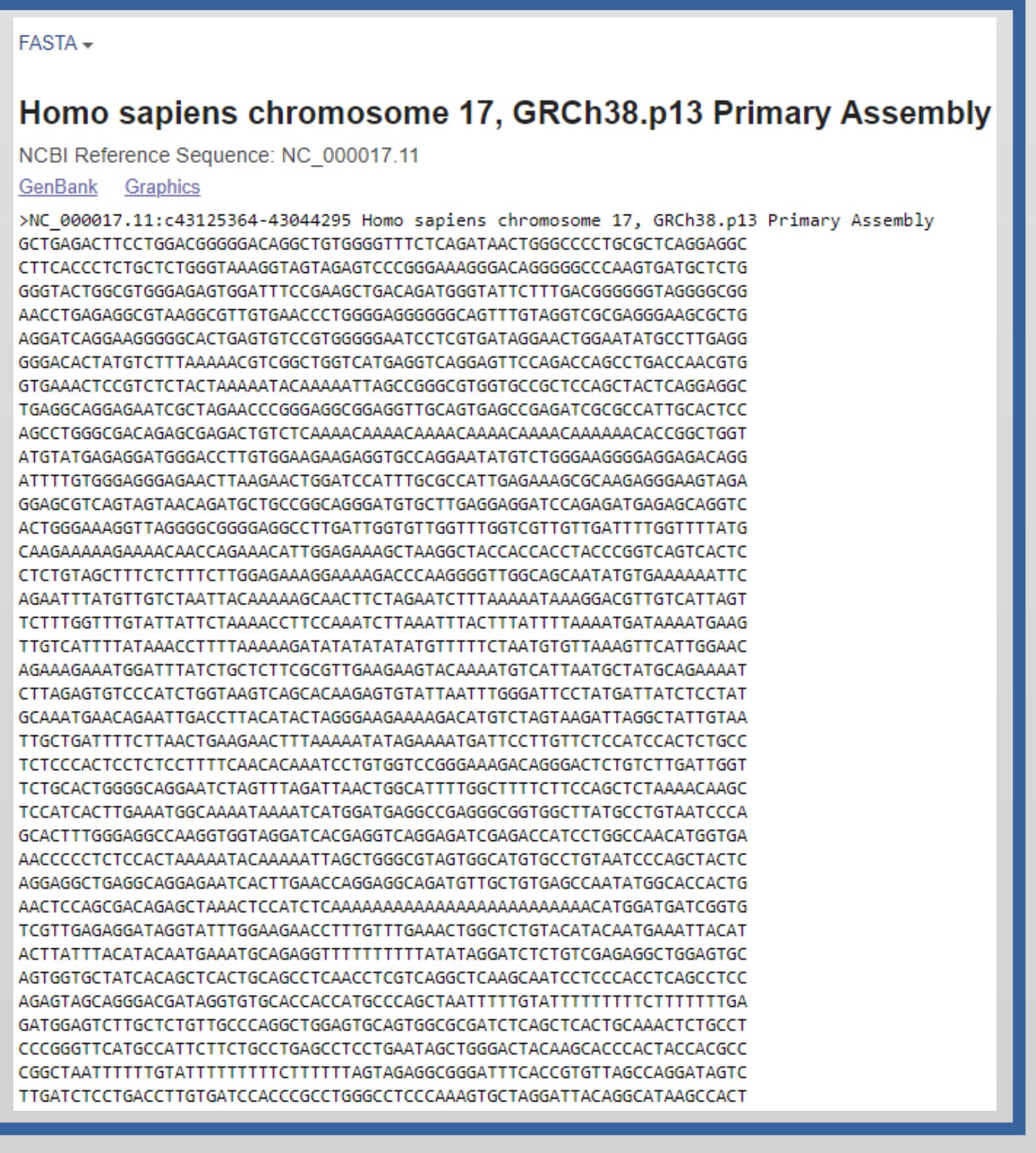
Cytogenetic location: 7p11.2, Genomic coordinates (GRCh38): 7:55,019,020-55,208,079

Matching terms: egfr

► Gene-Phenotype Relationships ► Links

OMIM Allelic Variants - 113705 BREAST CANCER 1 GENE; BRCA1				
Number	Phenotype	Mutation	SNP	gnomAD SNP
.0001	BREAST-OVARIAN CANCER, FAMILIAL, SUSCEPTIBILITY TO, 1	BRCA1, CYS64GLY	rs80357064	rs80357064
.0002	BREAST-OVARIAN CANCER, FAMILIAL, SUSCEPTIBILITY TO, 1	BRCA1, CYS61GLY	rs28897672	rs28897672
.0003	BREAST-OVARIAN CANCER, FAMILIAL, SUSCEPTIBILITY TO, 1;;PANCREATIC CANCER, SUSCEPTIBILITY TO, 4, INCLUDED	BRCA1, 2-BP DEL, 185AG	rs80357914,rs1555600876	
.0004	BREAST-OVARIAN CANCER, FAMILIAL, SUSCEPTIBILITY TO, 1	BRCA1, 59-BP INS	rs67284603	
.0005	BREAST-OVARIAN CANCER, FAMILIAL, SUSCEPTIBILITY TO, 1	BRCA1, 1-BP INS	rs80357569	rs80357569

Table 2. Some of the allelic variants of BRCA1.



FASTA

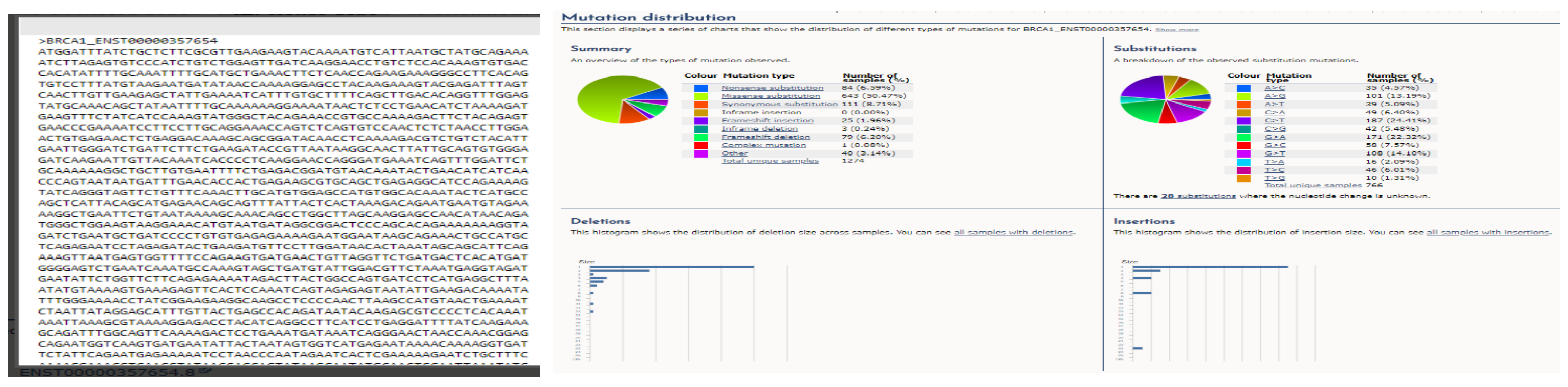
Homo sapiens chromosome 17, GRCh38.p13 Primary Assembly

NCBI Reference Sequence: NC_000017.11

GenBank Graphics

Results

Complementary DNA sequence of BRCA1 and mutation analysis..



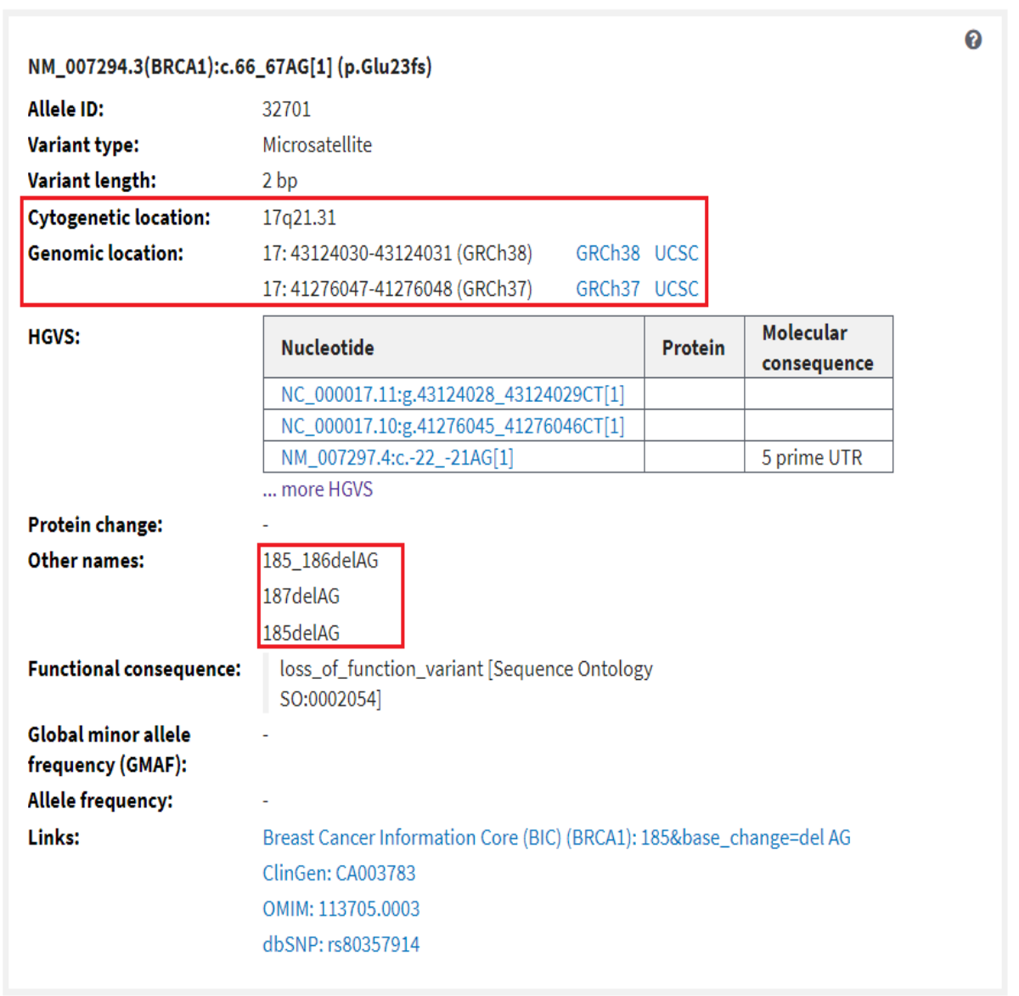
There are 28 substitutions where the nucleotide change is unknown.

This histogram shows the distribution of deletion size across samples. You can see all samples with deletions.

This histogram shows the distribution of insertion size. You can see all samples with insertions.

Founder mutation	Gene	Exon	Length of sequence	Primer (literature reported)
187delAG (c.68_69delAG, 185delAG)	BRCA1	2	340 bp	Forward: TTCGTATTCTGAGAGGCTGCT Reverse: ACGCCTCTCAGGTTCCG
5385insC (c.5266dupC, 5382insC)	BRCA1	20	271 bp	Forward: TGCAGATGCTGAGTTTGTGT Reverse: AGCTTATCTGAACAAAGTGATATT
6174delT (c.5946delT)	BRCA2	11	270 bp	Forward: ATATGTCTGGATTGGAGAAAGTT Reverse: AGCTGGTCTGAATGTTCGTT

Table 3. The studied founder mutations in BRCA1.



NM_007294.3(BRCA1):c.66_67AG[1] (p.Glu23fs)

Allele ID: 32701

Variant type: Microsatellite

Variant length: 2 bp

Cytogenetic location: 17q21.31

Genomic location: 17: 43124030-43124031 (GRCh38) GRCh38 UCSC
17: 41276047-41276048 (GRCh37) GRCh37 UCSC

HGVS:

Nucleotide	Protein	Molecular consequence
NC_000017.11:g.43124028_43124029CT[1]		
NC_000017.10:g.41276045_41276046CT[1]		
NM_007297.4:c.-22_-21AG[1]		5 prime UTR

... more HGVS

Protein change: -

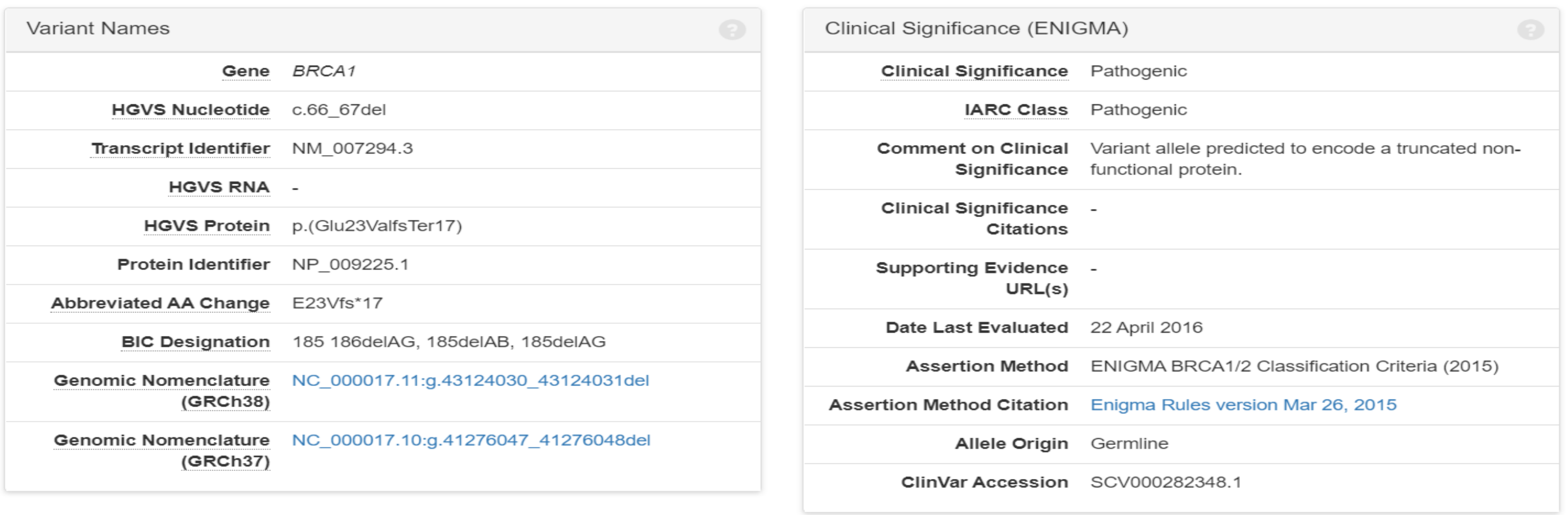
Other names: 185_186delAG
187delAG
185delAG

Functional consequence: loss_of_function_variant [Sequence Ontology SO:0002054]

Global minor allele frequency (GMAF): -

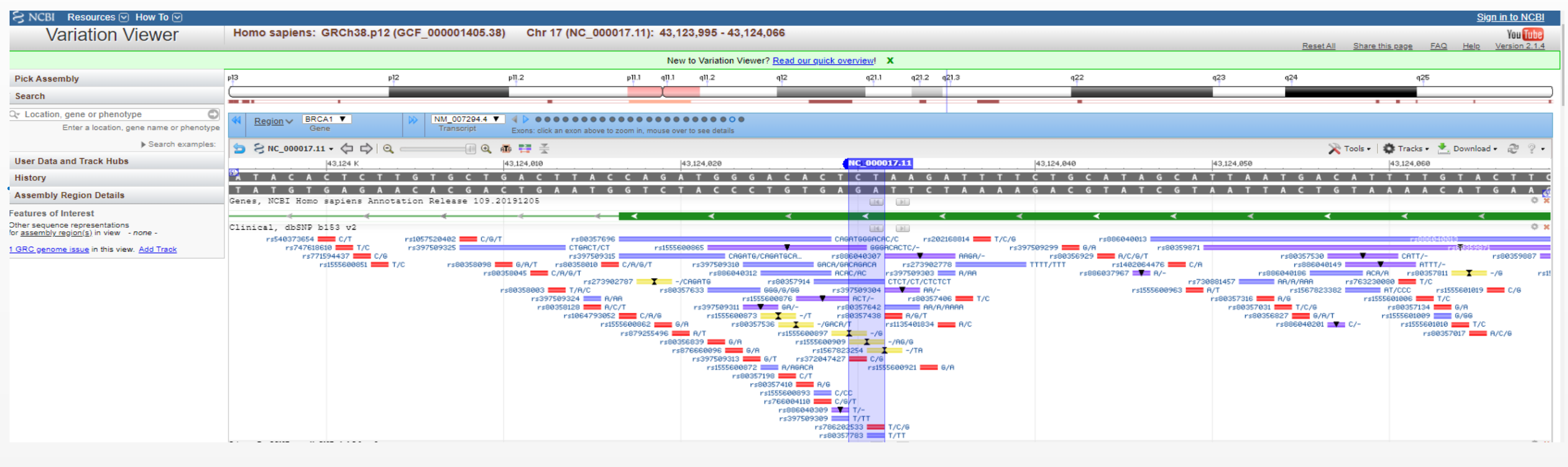
Allele frequency: -

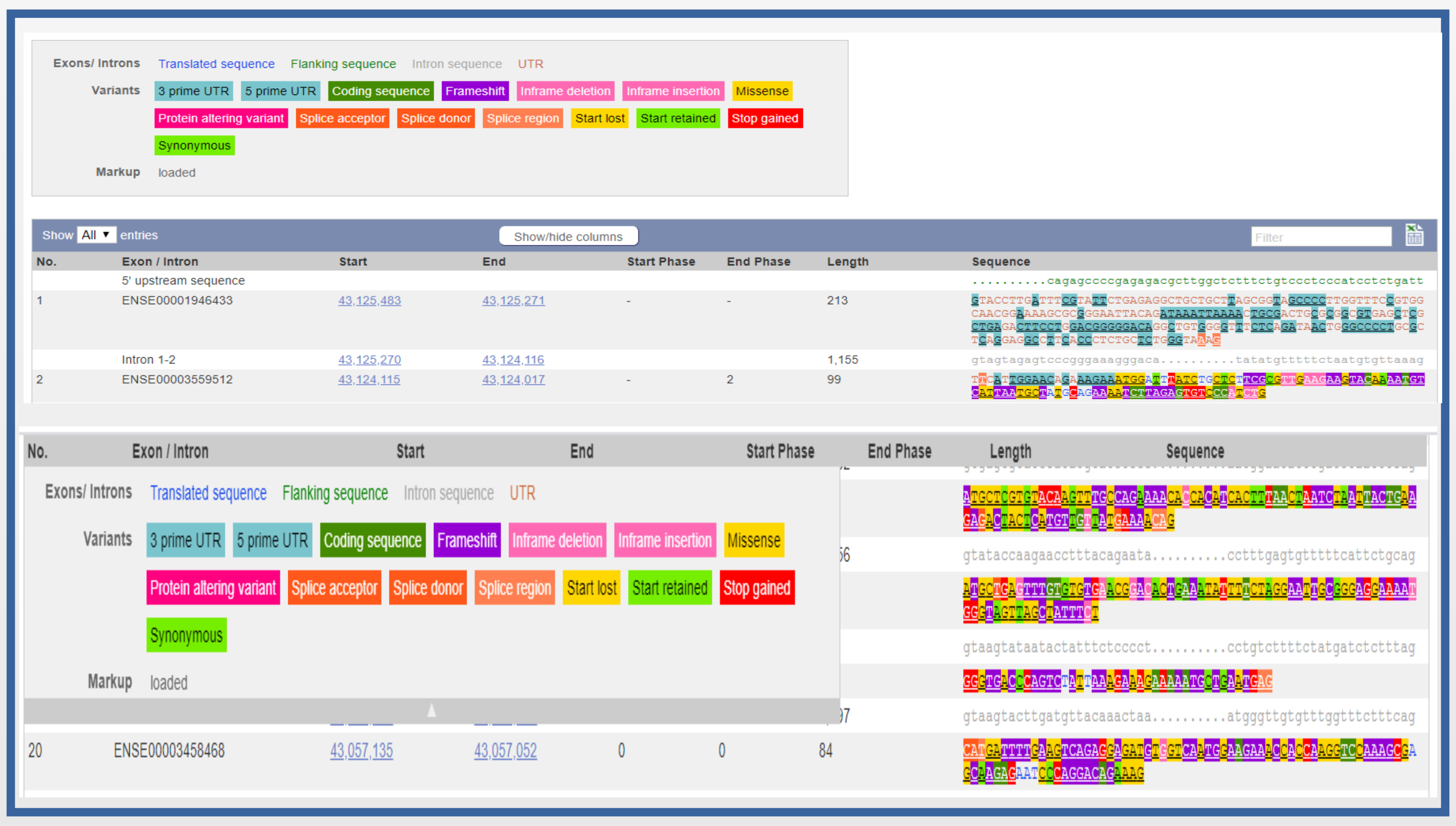
Links: Breast Cancer Information Core (BIC) (BRCA1): 185&base_change=del AG
ClinGen: CA003783
OMIM: 113705.0003
dbSNP: rs80357914

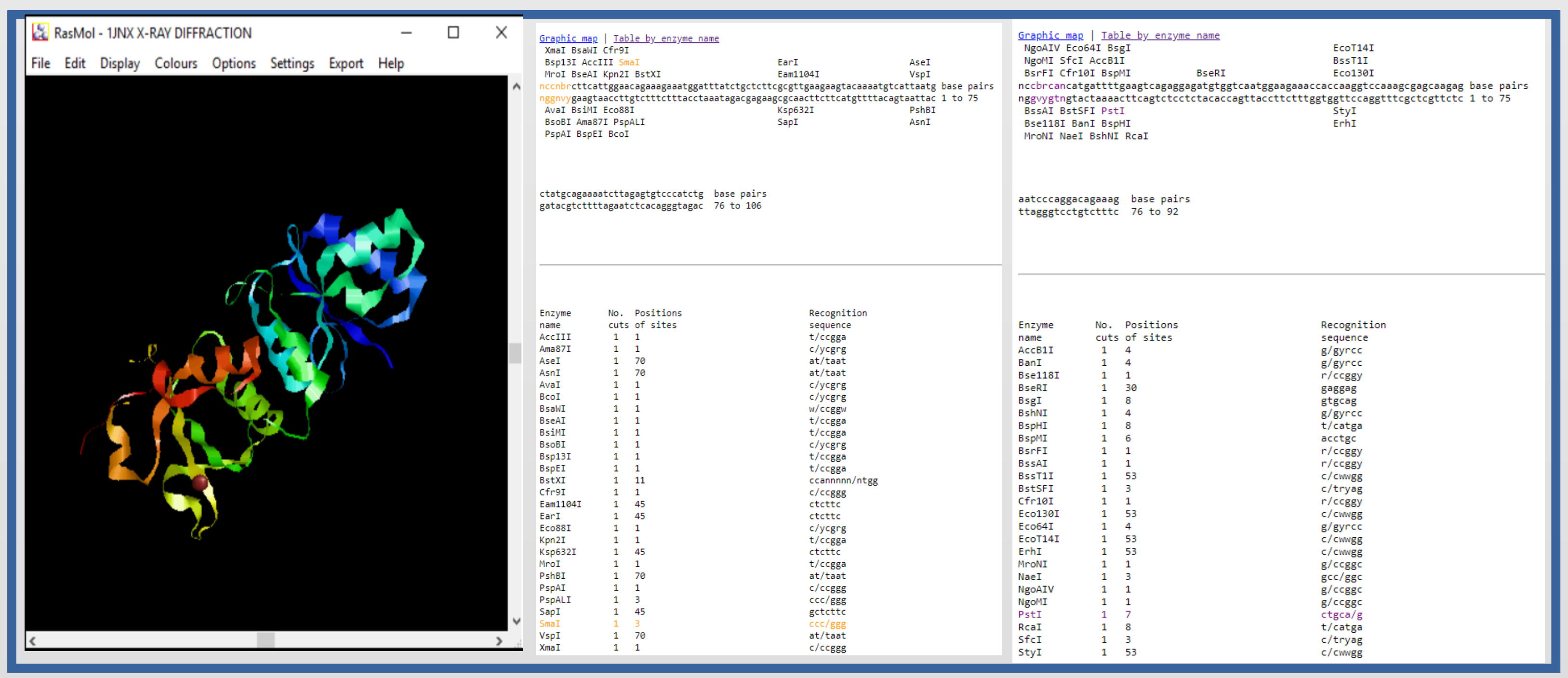


Variant Names	
Gene	*BRCA1*
HGVS Nucleotide	c.66_67del
Transcript Identifier	NM_007294.3
HGVS RNA	-
HGVS Protein	p.(Glu23ValfsTer17)
Protein Identifier	NP_009225.1
Abbreviated AA Change	E23Vfs*17
BIC Designation	185 186delAG, 185delAB, 185delAG
Genomic Nomenclature (GRCh38)	NC_000017.11:g.43124030_43124031del
Genomic Nomenclature (GRCh37)	NC_000017.10:g.41276047_41276048del

Clinical Significance (ENIGMA)	
Clinical Significance	Pathogenic
IARC Class	Pathogenic
Comment on Clinical Significance	Variant allele predicted to encode a truncated non-functional protein.
Clinical Significance Citations	-
Supporting Evidence URL(s)	-
Date Last Evaluated	22 April 2016
Assertion Method	ENIGMA BRCA1/2 Classification Criteria (2015)
Assertion Method Citation	Enigma Rules version Mar 26, 2015
Allele Origin	Germline
ClinVar Accession	SCV000282348.1







Primer3 Output

```
PRIMER PICKING RESULTS FOR NC_000017.11:43123992-43124146 Homo sapiens chromosome 17, GRCh38.p13 Primary Assembly

No mispriming library specified                                         Exon2 / BRCA1
Using 1-based sequence positions
OLIGO            start  len     tm     gc%    any    3'  seq
LEFT PRIMER         16   21  60.13   52.38   4.00  2.00 GCTGACTTACCAGATGGGACA
RIGHT PRIMER       122   22  59.06   36.36   4.00  0.00 CATTGGAACAGAAAGAAATGGA
SEQUENCE SIZE: 155
INCLUDED REGION SIZE: 155

PRODUCT SIZE: 107, PAIR ANY COMPL: 4.00, PAIR 3' COMPL: 0.00
```

Primer3 Output

```
PRIMER PICKING RESULTS FOR NC_000017.11:43057027-43057166 Homo sapiens chromosome 17, GRCh38.p13 Primary Assembly

No mispriming library specified
Using 1-based sequence positions
OLIGO            start  len     tm     gc%    any    3'  seq
LEFT PRIMER         39   20  60.10   50.00   3.00  0.00 GATTCTCTTGCTCGCTTTGG
RIGHT PRIMER       138   20  60.67   50.00   2.00  0.00 CCTGATGGGTTGTGTTTGGT
SEQUENCE SIZE: 140
INCLUDED REGION SIZE: 140

PRODUCT SIZE: 100, PAIR ANY COMPL: 2.00, PAIR 3' COMPL: 0.00
```

UCSC In-Silico PCR

```
>chr17:41276024+41276130 107bp GCTGACTTACCAGATGGGACA CATTGGAACAGAAAGAAATGGA
GCTGACTTACCAGATGGGACActctaagattttctgcatagcattaatga
cattttgtacttcttcaacgcgaagagcagataaaTCCATTTCTTTCTGT
TCCAATG
```

BRCA1 / Exon2

Primer Melting Temperatures

Forward: 60.1 C gctgacttaccagatgggaca
Reverse: 59.1 C cattggaacagaaagaaatgga
The temperature calculations are done assuming 50 mM salt and 50 nM annealing oligo concentration. The code to calculate the melting temp comes from Primer3.

UCSC In-Silico PCR

```
>chr17:41209082+41209181 100bp GATTCTCTTGCTCGCTTTGG CCTGATGGGTTGTGTTTGGT
GATTCTCTTGCTCGCTTTGGaccttggtggtttcttccattgaccacatc
tcctctgacttcaaaatcatgctgaaagaaACCAAACACAACCCATCAGG
```

BRCA1 / Exon20

Primer Melting Temperatures

Forward: 60.1 C gattctcttgctcgctttgg
Reverse: 60.7 C cctgatgggttgtgtttggt
The temperature calculations are done assuming 50 mM salt and 50 nM annealing oligo concentration. The code to calculate the melting temp comes from Primer3.