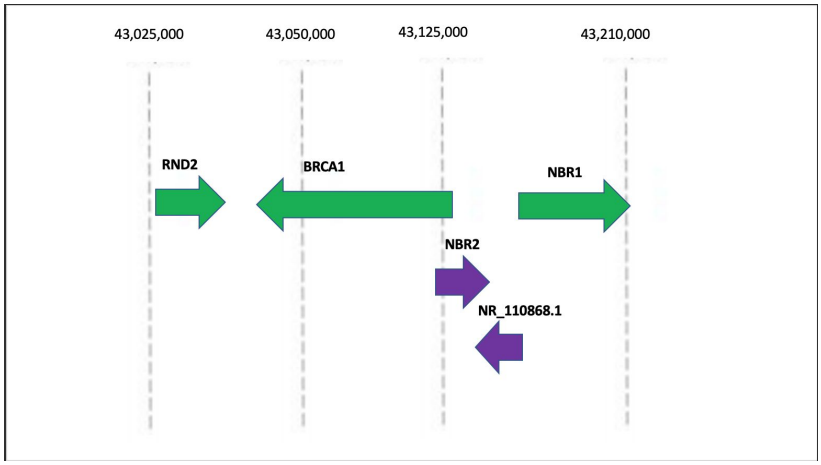
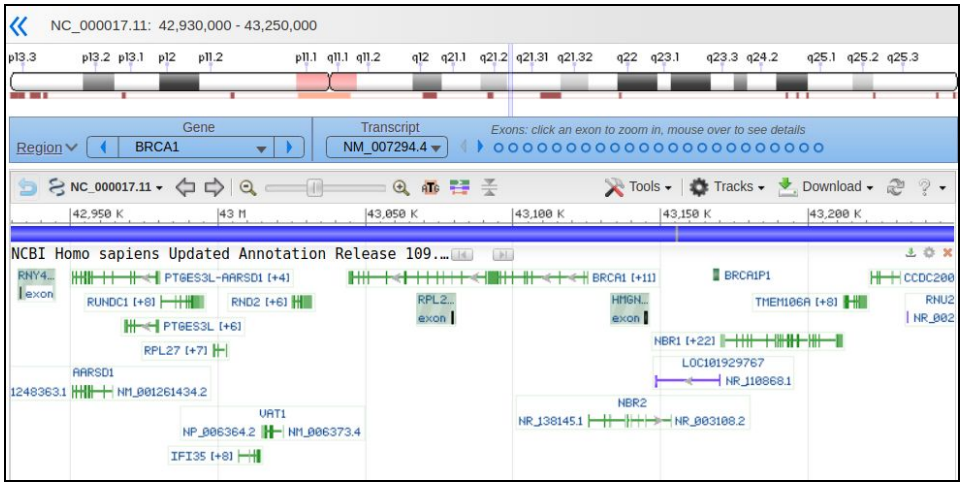


Figure 1: BRCA1 is located on Chromosome 17q21.31: 43,044,295-43,170,245 on the reverse strand, as indicated by the arrows (1). Green is labeled as genes (coding genes) and purple is labeled as a non-coding gene. The location within this figure is between 43,025,000 and 43,210,000 and shows the closest genes around BRCA1. A more comprehensive view of all the genes located near BRCA1 can be viewed in figure 2.

A

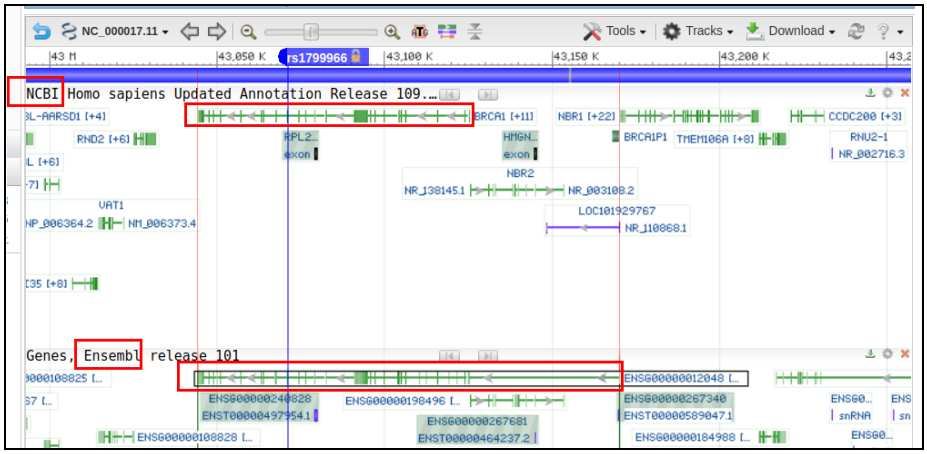


B

Feature Type	Color	Visual Examples
Miscellaneous features	purple	conserved nucle…
Regulatory	teal	minus_35_signal
Protein binding site	red	XerD
Recombination	golden brown	resIII subsite…
Mobile genetic element	blue	insertion seque…
Repeat region	blue	[repeat_region]
Mature peptide	golden brown	Clusterin
Gene	Green	CRYAA
RNA	Purple	XR_9377711
Coding Region	Red	CCDS13695.1
All other features	Black	STS_CEB240

Figure 2: Comprehensive view of NCBI RefSeq genes around BRCA1 (Chromosome 17q21.31 range: 42.93M-43.25M) using the NCBI Genome Data Viewer (1). A figure legend of the colors shown in A.) are listed in B.) (2).

A



B

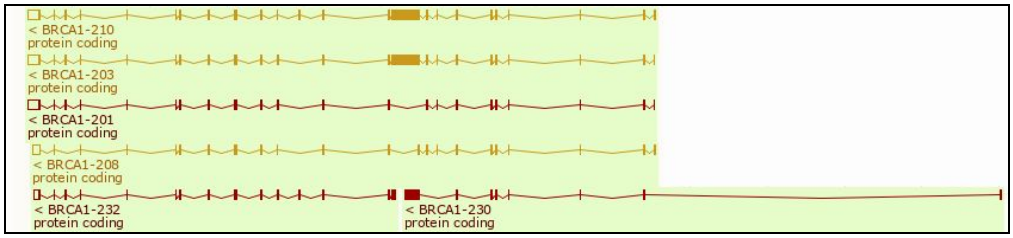


Figure 3. In A.), note the Ensembl gene is longer than the RefSeq gene in the NCBI genome viewer (1). The RefSeq gene starts at the first exon (start 43,044,295) of the primary transcript and ends at the last exon (43,125,364 end) of the primary transcript which more or less makes sense. From a meta analysis perspective Ensembl predicts the gene is longer because it has been annotated based on the beginning exon of ANY transcript to the ending transcript of ANY exon to obtain the longest possible window for the gene coordinates. B.) shows the two transcripts side by side, BRCA1-232 and BRCA-230, from Ensembl increasing the width of the gene. The second gene transcript, BRCA1-203, is what is represented by the RefSeq gene in NCBI genome browser above (4).

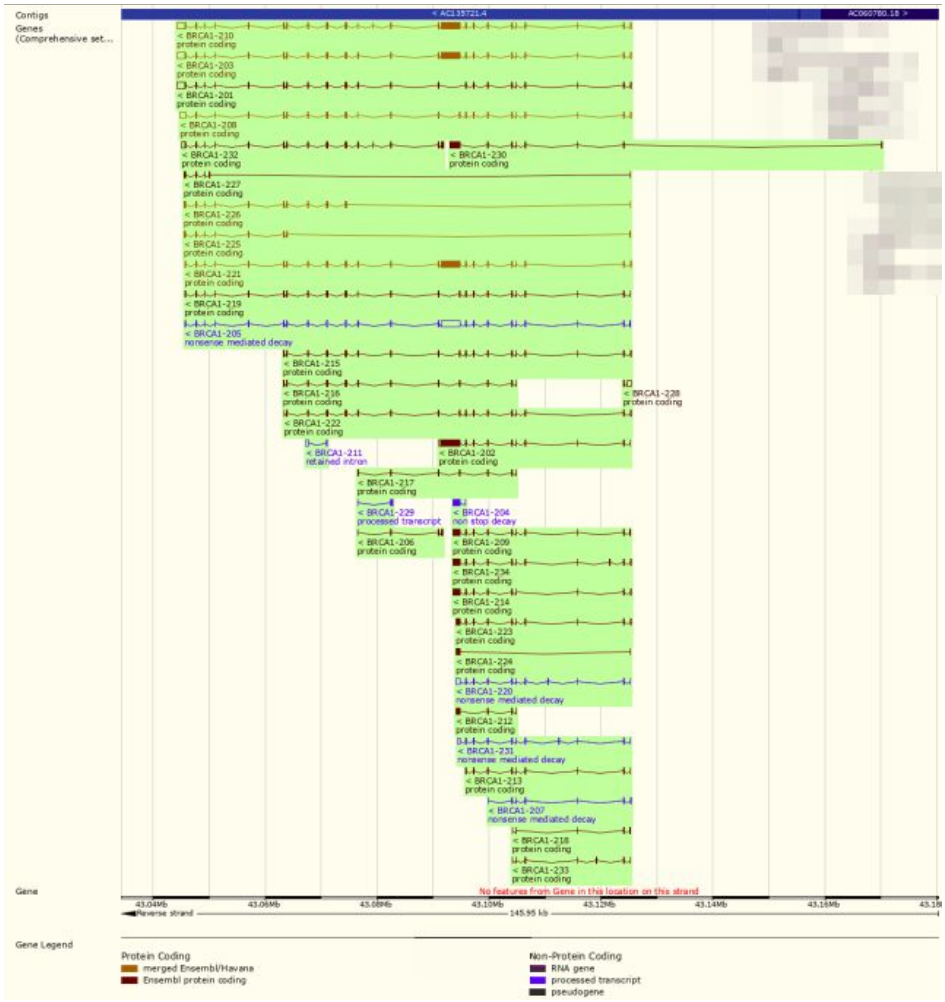
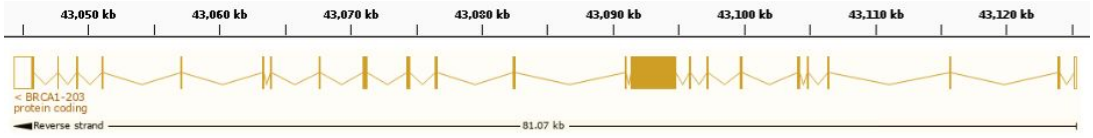


Figure 4: According to Ensembl a comprehensive list of the 34 transcripts for BRCA1 (Chromosome 17: 43,044,295-43,170,245) (1). Notice that BRCA1 is below the contig, thus indicating that it is on the reverse strand. The shaded boxes represent coding exons, while the unshaded represent non-coding exons (5'/3' UTRs) (1). In the Gene legend (below) red represents Ensembl protein-coding (which comes from the Ensembl automatic annotation pipeline **or** manual curation by the VEGA/Havana project), gold represents merged Ensembl/Havana, while blue, brown, and gray represent non-coding transcripts (1).

A



B

```
mRNA            join(1..94,1250..1348,9586..9639,18832..18909,
                20409..20497,21104..21243,25485..25590,28076..28121,
                29443..29519,30505..33930,34333..34421,42790..42961,
                48751..48877,50844..51034,54127..54437,57670..57757,
                61414..61491,61992..62032,68230..68313,74248..74302,
                76171..76244,77662..77722,79563..81070)
                /gene="BRCA1"
```

Figure 5: A.) shows an expanded view of the BRCA1 transcript referred to in ClinVar (NM_007294.4, labeled as BRCA1-203 in the figure as coordinates from NCBI and Ensembl match exactly for transcript variant 1) has 23 exons, 22 coding exons, a transcript length of 7,088 bps, and a translation length of 1,863 residues. The coding regions are represented by shaded orange boxes, noncoding untranslated regions (UTRs) are represented as blank white boxes, while the lines in between the boxes represent introns (1). The first exon is completely in the 5’ UTR (reverse strand) while the rest of the 22 exons are included in the CDS while the last exon, 23 to the left represents the 3’ UTR (reverse strand). The 5' Guanosine-triphosphate Cap or 3’ Poly A tail are not seen here. B.) shows all exon regions and introns of the BRCA1 gene (NM_007294.4 transcript, includes both coding and noncoding exons) (2).

	A	B	C	D	E	F	G
1	Gene name	Gene description	Gene start (bp)	Gene end (bp)	Source (gene)	Transcript type	Gene type
2	LINC00671	long intergenic non-protein coding RNA 671 [Source:HGNC Symbol;Acc:HGNC:44339]	42874670	42898704	havana	lncRNA	lncRNA
3	NBR2	neighbor of BRCA1 lncRNA 2 [Source:HGNC Symbol;Acc:HGNC:20691]	43125551	43153671	havana	lncRNA	lncRNA
4	AC060780.1	novel transcript	43148368	43171037	havana	lncRNA	lncRNA
5	LINC00910	long intergenic non-protein coding RNA 910 [Source:HGNC Symbol;Acc:HGNC:44361]	43338741	43389199	havana	lncRNA	lncRNA
6	MIR2117HG	MIR2117 host gene [Source:HGNC Symbol;Acc:HGNC:51999]	43444707	43451200	havana	lncRNA	lncRNA
7	LINC02594	long intergenic non-protein coding RNA 2594 [Source:HGNC Symbol;Acc:HGNC:53935]	43679341	43706709	havana_tagene	lncRNA	lncRNA
8	AC003098.1	novel transcript, antisense to chromosome 17 open reading frame 105	43782804	43784682	havana	lncRNA	lncRNA
9	AC007993.3	novel transcript, antisense to FAM215A	43914433	43923001	havana	lncRNA	lncRNA
10	FAM215A	family with sequence similarity 215 member A [Source:HGNC Symbol;Acc:HGNC:17505]	43917194	43917982	ensembl_havana	lncRNA	lncRNA
11	AC007993.2	novel transcript	43927563	43932622	havana	lncRNA	lncRNA
12	LINC01976	long intergenic non-protein coding RNA 1976 [Source:HGNC Symbol;Acc:HGNC:52803]	43936363	43938959	havana	lncRNA	lncRNA
13	AC023855.1	novel transcript, antisense to HDAC5	44115912	44120595	havana	lncRNA	lncRNA
14	ASB16-AS1	ASB16 antisense RNA 1 [Source:HGNC Symbol;Acc:HGNC:25442]	44175968	44186722	ensembl_havana	lncRNA	lncRNA
15	AC004596.1	novel transcript, antisense to ATXN7L3	44198882	44216565	havana	lncRNA	lncRNA
16	AC003102.1	novel transcript, antisense to UBTF	44221401	44223710	havana	lncRNA	lncRNA
17	AC003043.2	novel transcript	44276368	44281182	havana	lncRNA	lncRNA
18	RUNDC3A-AS1	RUNDC3A antisense RNA 1 [Source:HGNC Symbol;Acc:HGNC:51344]	44299574	44315315	havana	lncRNA	lncRNA
19	AC003043.1	novel transcript	44328613	44331462	havana	lncRNA	lncRNA
20	AC103703.1	novel transcript, sense intronic to GPATCH8	44486153	44488815	havana	lncRNA	lncRNA
21	LINC01180	long intergenic non-protein coding RNA 1180 [Source:HGNC Symbol;Acc:HGNC:49558]	44646364	44650349	havana	lncRNA	lncRNA
22	AC091152.2	novel transcript, antisense to C17orf104	44673689	44676257	havana	lncRNA	lncRNA
23	AC005180.2	novel transcript	44793199	44794474	havana	lncRNA	lncRNA
24	AC005180.1	novel transcript	44794747	44797783	havana	lncRNA	lncRNA
25	AC015936.1	novel transcript	44947912	44948939	havana	lncRNA	lncRNA
26	AC015936.2	novel transcript	44982514	44982772	havana	lncRNA	lncRNA
27	AC142472.1	novel transcript	45146730	45148470	havana	lncRNA	lncRNA
28	AC138150.1	novel transcript, antisense to HEXIM1 and HEXIM2	45150400	45161510	havana	lncRNA	lncRNA
29	AC138150.2	novel transcript, antisense to HEXIM2	45168800	45171584	havana	lncRNA	lncRNA
30	AC008105.3	novel transcript, antisense to FMNL1	45190931	45222222	havana	lncRNA	lncRNA
31	AC008105.1	novel transcript, antisense to FMNL1	45238028	45241734	havana	lncRNA	lncRNA
32	MAP3K14-AS1	MAP3K14 antisense RNA 1 [Source:HGNC Symbol;Acc:HGNC:44359]	45247916	45269824	havana	lncRNA	lncRNA
33	AC003070.2	novel transcript	45371402	45372057	havana	lncRNA	lncRNA
34	AC003070.1	novel transcript, antisense to ARHGAP27	45396932	45397477	havana	lncRNA	lncRNA
35	AC091132.2	novel transcript, antisense to PLEKHM1	45452844	45464063	havana	lncRNA	lncRNA
36	AC091132.1	novel transcript	45533963	45534710	havana	lncRNA	lncRNA
37	AC091132.4	novel transcript	45545804	45563230	havana	lncRNA	lncRNA
38	AC091132.5	novel transcript	45549781	45558730	havana	lncRNA	lncRNA
39	AC126544.2	novel transcript	45586452	45588379	havana	lncRNA	lncRNA
40	AC217774.1	novel transcript	45731703	45732977	havana	lncRNA	lncRNA
41	AC217774.2	novel transcript	45733353	45734668	havana	lncRNA	lncRNA
42	MAPT-AS1	MAPT antisense RNA 1 [Source:HGNC Symbol;Acc:HGNC:43738]	45799390	45895680	havana	lncRNA	lncRNA
43	MAPT-IT1	MAPT intronic transcript 1 [Source:HGNC Symbol;Acc:HGNC:43741]	45895783	45898798	havana	lncRNA	lncRNA
44	CR936218.2	novel transcript, antisense to MAPT	45907670	45910779	havana	lncRNA	lncRNA
45	CR936218.1	novel transcript, antisense to a novel protein	46035313	46035770	havana	lncRNA	lncRNA
46	KANSL1-AS1	KANSL1 antisense RNA 1 [Source:HGNC Symbol;Acc:HGNC:43740]	46193576	46196723	havana	lncRNA	lncRNA
47	FAM215B	family with sequence similarity 215 member B [Source:HGNC Symbol;Acc:HGNC:43639]	46559830	46562795	havana	lncRNA	lncRNA
48	MIR8781	microRNA 6781 [Source:HGNC Symbol;Acc:HGNC:50185]	42823880	42823943	mirbase	miRNA	miRNA
49	MIR2117	microRNA 2117 [Source:HGNC Symbol;Acc:HGNC:37311]	43444808	43444885	mirbase	miRNA	miRNA
50	MIR6782	microRNA 6782 [Source:HGNC Symbol;Acc:HGNC:50270]	44207771	44207839	mirbase	miRNA	miRNA
51	MIR8783	microRNA 6783 [Source:HGNC Symbol;Acc:HGNC:50159]	44934618	44934681	mirbase	miRNA	miRNA
52	MIR6784	microRNA 6784 [Source:HGNC Symbol;Acc:HGNC:49988]	45114367	45114433	mirbase	miRNA	miRNA
53	MIR4315-1	microRNA 4315-1 [Source:HGNC Symbol;Acc:HGNC:38342]	45475363	45475435	mirbase	miRNA	miRNA

Figure 6: Biomart data, examples of some noncoding genes around chromosome 17q21.31 (1)

	hgnc_symbol	percentage_gene_gc_content
1	NAGS	63.44
2	SPPL2C	61.18
3	SLC25A39	61.11
4	PPY	60.67
5	FAM171A2	60.58
6	UBTF	59.52
7	ADAM11	59.33
8	ATXN7L3	59.15
9	RUNDC3A	59.03
10	FMNL1	58.77

Figure 7. Top 10 genes in band 17q21.31 with respect to GC content.

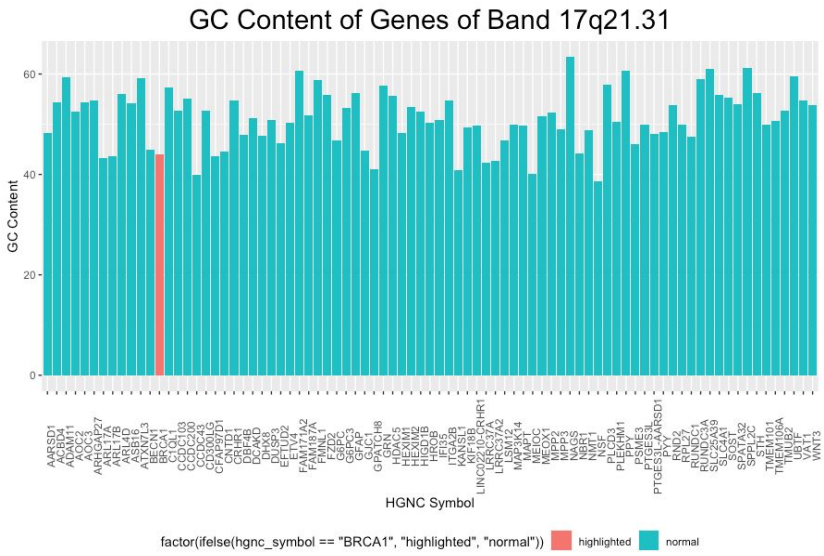
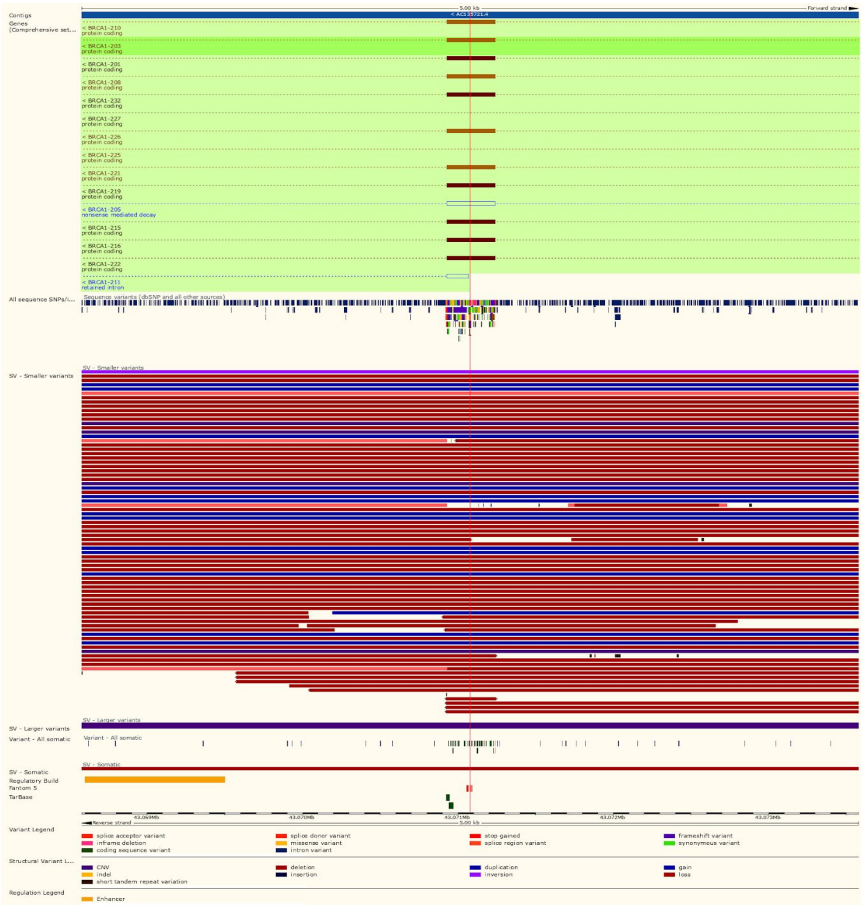


Figure 8: GC Content of Genes on band 17q21.31 (BRCA1 in orange)

A



B

Show All entries — Show/hide columns

Gene	Transcript (strand)	Allele (Tr. allele)	Consequence Type	Position in transcript	Position in CDS	Position in protein
ENSG00000012048 HGNC: BRCA1	ENST00000352993.7 (-) biotype: protein_coding	A (T)	missense variant	1530 (out of 3668)	1411 (out of 2166)	471 (out of 721)
ENSG00000012048 HGNC: BRCA1	ENST00000352993.7 (-) biotype: protein_coding	C (G)	missense variant	1530 (out of 3668)	1411 (out of 2166)	471 (out of 721)
ENSG00000012048 HGNC: BRCA1	ENST00000357654.9 (-) biotype: protein_coding	A (T)	missense variant	4950 (out of 7088)	4837 (out of 5592)	1613 (out of 1863)
ENSG00000012048 HGNC: BRCA1	ENST00000357654.9 (-) biotype: protein_coding	C (G)	missense variant	4950 (out of 7088)	4837 (out of 5592)	1613 (out of 1863)
ENSG00000012048 HGNC: BRCA1	ENST00000468300.5 (-) biotype: protein_coding	A (T)	missense variant	1719 (out of 3273)	1525 (out of 2100)	509 (out of 699)
ENSG00000012048 HGNC: BRCA1	ENST00000468300.5 (-) biotype: protein_coding	C (G)	missense variant	1719 (out of 3273)	1525 (out of 2100)	509 (out of 699)
ENSG00000012048 HGNC: BRCA1	ENST00000471181.7 (-) biotype: protein_coding	A (T)	missense variant	5132 (out of 7270)	4900 (out of 5655)	1634 (out of 1884)
ENSG00000012048 HGNC: BRCA1	ENST00000471181.7 (-) biotype: protein_coding	C (G)	missense variant	5132 (out of 7270)	4900 (out of 5655)	1634 (out of 1884)
ENSG00000012048 HGNC: BRCA1	ENST00000478531.5 (-) biotype: protein_coding	A (T)	missense variant	1627 (out of 1972)	1525 (out of 1870)	509 (out of 623)
ENSG00000012048 HGNC: BRCA1	ENST00000478531.5 (-) biotype: protein_coding	C (G)	missense variant	1627 (out of 1972)	1525 (out of 1870)	509 (out of 623)
ENSG00000012048 HGNC: BRCA1	ENST00000484087.5 (-) biotype: protein_coding	A (T)	missense variant	1150 (out of 1495)	1150 (out of 1495)	384 (out of 498)
ENSG00000012048 HGNC: BRCA1	ENST00000484087.5 (-) biotype: protein_coding	C (G)	missense variant	1150 (out of 1495)	1150 (out of 1495)	384 (out of 498)
ENSG00000012048 HGNC: BRCA1	ENST00000491747.6 (-) biotype: protein_coding	A (T)	missense variant	1624 (out of 2379)	1525 (out of 2280)	509 (out of 759)
ENSG00000012048 HGNC: BRCA1	ENST00000491747.6 (-) biotype: protein_coding	C (G)	missense variant	1624 (out of 2379)	1525 (out of 2280)	509 (out of 759)
ENSG00000012048 HGNC: BRCA1	ENST00000493795.5 (-) biotype: protein_coding	A (T)	missense variant	4928 (out of 5732)	4696 (out of 5451)	1566 (out of 1816)
ENSG00000012048 HGNC: BRCA1	ENST00000493795.5 (-) biotype: protein_coding	C (G)	missense variant	4928 (out of 5732)	4696 (out of 5451)	1566 (out of 1816)
ENSG00000012048 HGNC: BRCA1	ENST00000493919.5 (-) biotype: protein_coding	A (T)	missense variant	1619 (out of 1948)	1387 (out of 1716)	463 (out of 572)
ENSG00000012048 HGNC: BRCA1	ENST00000493919.5 (-) biotype: protein_coding	C (G)	missense variant	1619 (out of 1948)	1387 (out of 1716)	463 (out of 572)
ENSG00000012048 HGNC: BRCA1	ENST00000591534.5 (-) biotype: protein_coding	A (T)	missense variant	412 (out of 1282)	310 (out of 1065)	104 (out of 354)
ENSG00000012048 HGNC: BRCA1	ENST00000591534.5 (-) biotype: protein_coding	C (G)	missense variant	412 (out of 1282)	310 (out of 1065)	104 (out of 354)
ENSG00000012048 HGNC: BRCA1	ENST00000644379.1 (-) biotype: protein_coding	A (T)	missense variant	1224 (out of 2571)	1225 (out of 1979)	409 (out of 659)
ENSG00000012048 HGNC: BRCA1	ENST00000644379.1 (-) biotype: protein_coding	C (G)	missense variant	1224 (out of 2571)	1225 (out of 1979)	409 (out of 659)
LRG_292 HGNC: BRCA1	LRG_292t1.1 (+) biotype: LRG_gene	G (G)	missense variant	5069 (out of 7207)	4837 (out of 5592)	1613 (out of 1863)
LRG_292 HGNC: BRCA1	LRG_292t1.1 (+) biotype: LRG_gene	T (T)	missense variant	5069 (out of 7207)	4837 (out of 5592)	1613 (out of 1863)
ENSG00000012048 HGNC: BRCA1	ENST00000461221.5 (-) biotype: nonsense_mediated_decay	A (T)	3 prime UTR variant, NMD transcript variant	4938 (out of 5693)	-	-
ENSG00000012048 HGNC: BRCA1	ENST00000461221.5 (-) biotype: nonsense_mediated_decay	C (G)	3 prime UTR variant, NMD transcript variant	4938 (out of 5693)	-	-
ENSG00000012048 HGNC: BRCA1	ENST00000586385.5 (-) biotype: protein_coding	A (T)	intron variant	-	-	-
ENSG00000012048 HGNC: BRCA1	ENST00000586385.5 (-) biotype: protein_coding	C (G)	intron variant	-	-	-
ENSG00000012048 HGNC: BRCA1	ENST00000591849.5 (-) biotype: protein_coding	A (T)	intron variant	-	-	-
ENSG00000012048 HGNC: BRCA1	ENST00000591849.5 (-) biotype: protein_coding	C (G)	intron variant	-	-	-

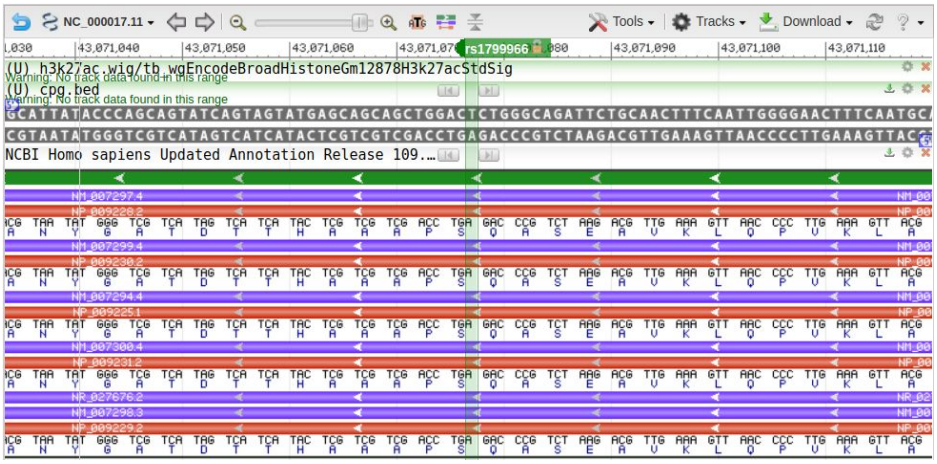
C

Gene and Transcript consequences

Show All entries — Show/hide columns (8 hidden) — Filter

Transcript (strand)	Allele (Tr. allele)	Consequence Type	Position in transcript	Position in CDS	Position in protein	AA	Codons
ENST00000352993.7 (-) biotype: protein_coding	A (T)	missense variant	1530 (out of 3668)	1411 (out of 2166)	471 (out of 721)	S/C	AGT/TGT
ENST00000352993.7 (-) biotype: protein_coding	C (G)	missense variant	1530 (out of 3668)	1411 (out of 2166)	471 (out of 721)	S/G	AGT/GGT
ENST00000357654.9 (-) biotype: protein_coding	A (T)	missense variant	4950 (out of 7088)	4837 (out of 5592)	1613 (out of 1863)	S/C	AGT/TGT
ENST00000357654.9 (-) biotype: protein_coding	C (G)	missense variant	4950 (out of 7088)	4837 (out of 5592)	1613 (out of 1863)	S/G	AGT/GGT
ENST00000468300.5 (-) biotype: protein_coding	A (T)	missense variant	1719 (out of 3273)	1525 (out of 2100)	509 (out of 699)	S/C	AGT/TGT
ENST00000468300.5 (-) biotype: protein_coding	C (G)	missense variant	1719 (out of 3273)	1525 (out of 2100)	509 (out of 699)	S/G	AGT/GGT
ENST00000471181.7 (-) biotype: protein_coding	A (T)	missense variant	5132 (out of 7270)	4900 (out of 5655)	1634 (out of 1884)	S/C	AGT/TGT
ENST00000471181.7 (-) biotype: protein_coding	C (G)	missense variant	5132 (out of 7270)	4900 (out of 5655)	1634 (out of 1884)	S/G	AGT/GGT

D



E

Gene: BRCA1, BRCA1 DNA repair associated (minus strand)

Molecule type	Change	Amino acid[Codon]	SO Term
BRCA1 transcript variant 1	NM_007294.4:c.4837A>T	S [AGT] > C [TGT]	Coding Sequence Variant
BRCA1 transcript variant 1	NM_007294.4:c.4837A>G	S [AGT] > G [GGT]	Coding Sequence Variant
BRCA1 transcript variant 2	NM_007300.4:c.4900A>T	S [AGT] > C [TGT]	Coding Sequence Variant
BRCA1 transcript variant 2	NM_007300.4:c.4900A>G	S [AGT] > G [GGT]	Coding Sequence Variant
BRCA1 transcript variant 3	NM_007297.4:c.4696A>T	S [AGT] > C [TGT]	Coding Sequence Variant
BRCA1 transcript variant 3	NM_007297.4:c.4696A>G	S [AGT] > G [GGT]	Coding Sequence Variant
BRCA1 transcript variant 4	NM_007298.3:c.1525A>T	S [AGT] > C [TGT]	Coding Sequence Variant
BRCA1 transcript variant 4	NM_007298.3:c.1525A>G	S [AGT] > G [GGT]	Coding Sequence Variant
BRCA1 transcript variant 5	NM_007299.4:c.1525A>T	S [AGT] > C [TGT]	Coding Sequence Variant
BRCA1 transcript variant 5	NM_007299.4:c.1525A>G	S [AGT] > G [GGT]	Coding Sequence Variant
BRCA1 transcript variant 6	NR_027676.2:n.5014A>T	N/A	Non Coding Transcript Variant

Figure 9: A.) shows the transcript highlighted in light green (BRCA1-203, also known as NM_007294.4) was the original transcript the reference paper discussed as being of uncertain significance. In ClinVar, rs1799966 on BRCA1-203 is described as a missense variant c.4837A>T (p.Ser1613Cys) (1). Overall, ensembl lists this variant as having 15 transcripts, 2504 sample genotypes, 11 phenotypes, and is mentioned in 62 citations (2). B.) shows a table listing how each transcript is affected by rs1799966 (under the consequence column). Here, the transcript ENST00000357654.9 is BRCA1-203 (NM_007294.4). Notice that the table lists where rs1799966 takes place in each transcript, CDS, and protein. C.) is a close up highlighting rs1799966 NC_000017.11:g.43071077T>A (NP_009225.2:p.Ser1613Cys). Alternative to large image in Figure 1 A.), D.) is a figure from NCBI showing the direction, the main transcripts, as well as the codon AGT (reverse direction for BRCA1, the bottom strand of the top two grey strands for negative coding strand) for Serine. E.) shows transcript variants in the dbSNP database with corresponding amino acid change (our SNP: rs1799966 (c.4837A>T (p.Ser1613Cys)) (NM_007294.4), arrow pointing to it in red) (3).

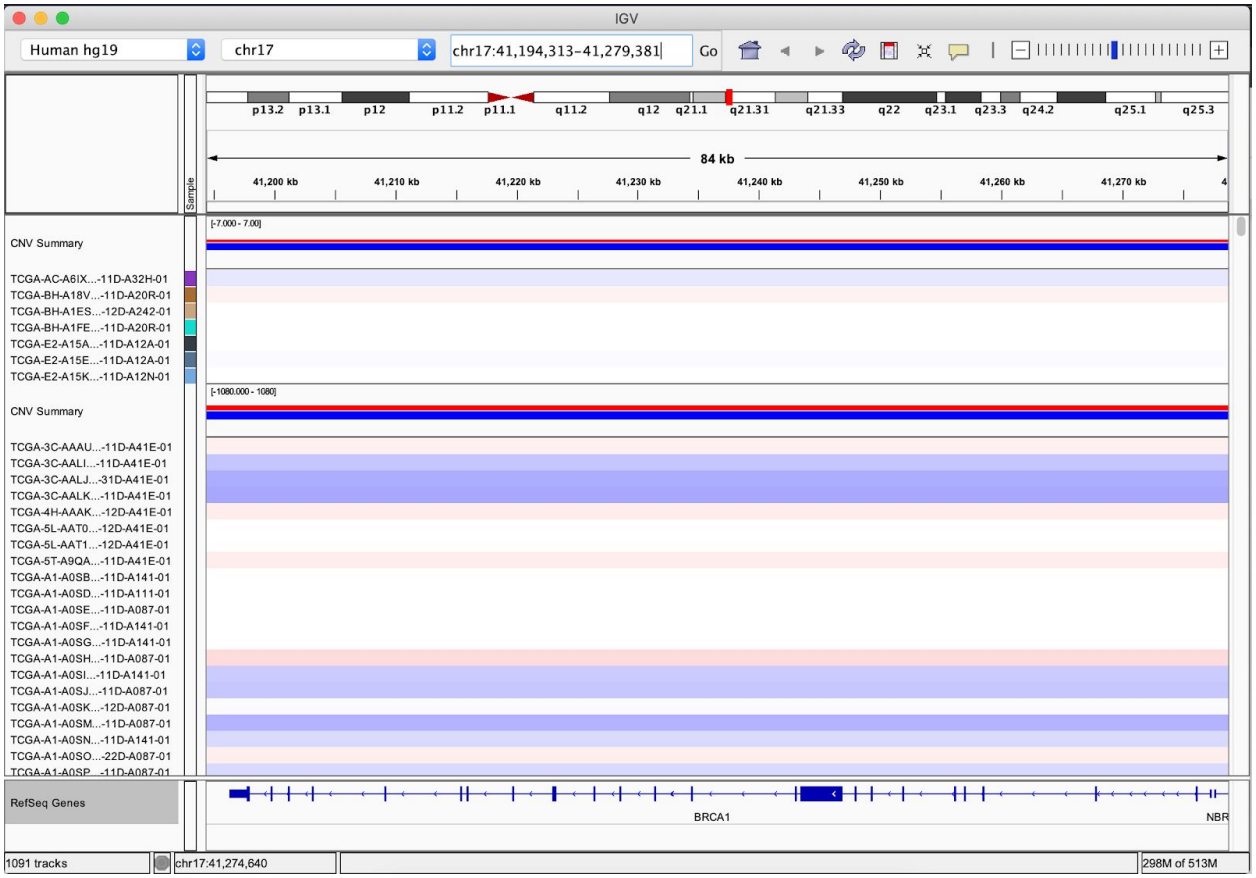


Figure 10: Firehose (The Cancer Genome Analysis (TCGA) group) is an available track for hg19 on IGV. We can utilize this track to show CNVs (gain or loss of genomic material) using the BRCA-TM/BRCA-TP CopyNumber minus germline track (TP is primary tumor, while TM is metastatic tumor). Notice that the top red/blue line (metastatic breast cancer) is mostly blue, representing a heavy loss of CNV. The bottom is slightly more blue than red (primary breast cancer), representing a slightly lower loss of CNV. Each line on the track represents an individual patient, and red represents gain of genomic material, while blue represents loss of genomic material. Notice that the top line is mostly blue, representing a loss of BRCA1 for cancer patients. This makes sense because patients with BRCA1 breast cancer, for example, usually have a nonfunctioning or truncated BRCA1 protein.

Variants linked to rs1799966 in 1000GENOMES:phase_3:TSI [back to top]

Show 10 entries | Show/hide columns | Filter

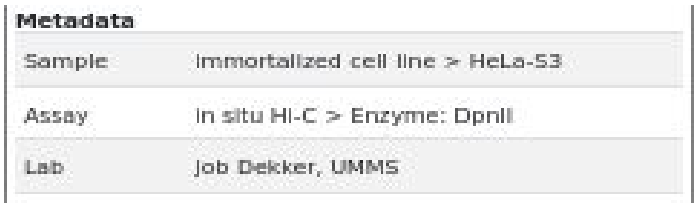
Variant	Location	Distance (bp)	r^2	D'	Associated phenotype(s)	Consequence Type	Located in gene(s)	Gene phenotype(s)
rs8176218	17:43071522-43071524	445	1.000	1.000	-	intron variant	BRCA1	119 phenotypes
rs8176220	17:43070445	632	1.000	1.000	-	intron variant	BRCA1	119 phenotypes
rs8176231	17:43068206	2871	1.000	1.000	-	intron variant	BRCA1	119 phenotypes
rs8176233	17:43067787	3290	1.000	1.000	-	intron variant	BRCA1	119 phenotypes
rs8176234	17:43067763	3314	1.000	1.000	-	intron variant	BRCA1	119 phenotypes
rs3785546	17:43064916	6161	1.000	1.000	-	intron variant	BRCA1	119 phenotypes
rs12940378	17:43077369	6292	1.000	1.000	-	intron variant	BRCA1	119 phenotypes
rs8176205	17:43077744-43077751	6667	1.000	1.000	-	intron variant	BRCA1	119 phenotypes
rs8176204	17:43077746	6669	1.000	1.000	-	intron variant	BRCA1	119 phenotypes
rs3092994	17:43063808	7269	1.000	1.000	-	intron variant	BRCA1	119 phenotypes

Showing 1 to 10 of 146 entries << < 1 2 3 4 5 > >>

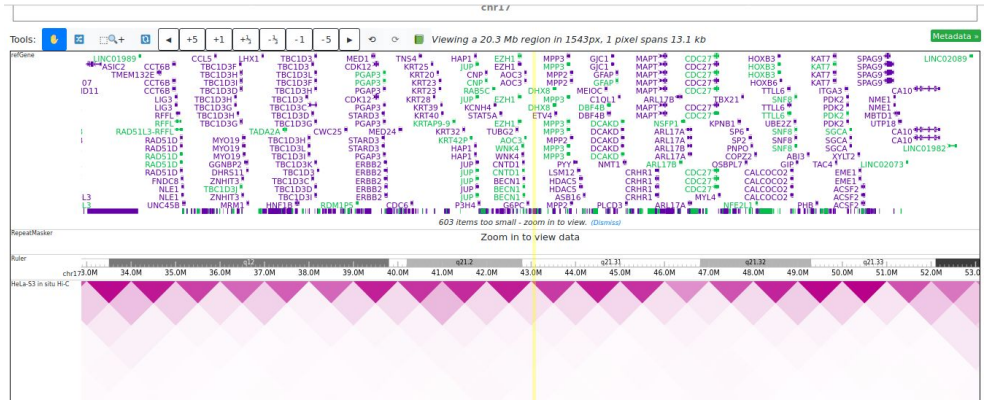
Figure 11: SNPs in LD with rs1799966 (2)

Figure 12: Transcription factors in the immediate vicinity of rs1799966.

A



B



C

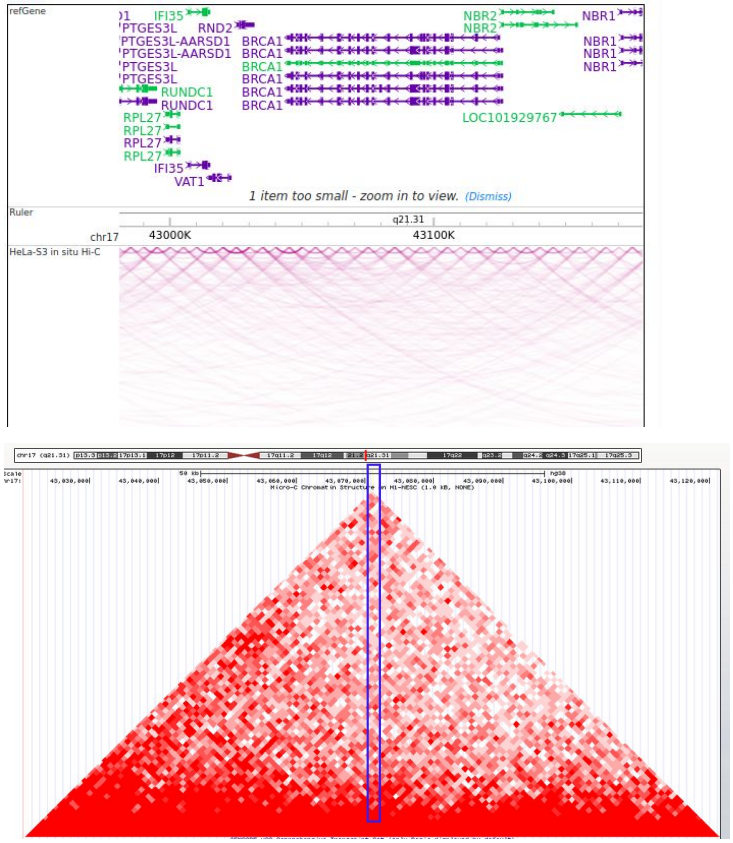


Figure 13: B) I tried to get a better picture of our variant but the yellow line for a single base does not stay. This yellow line represents the entire gene. C) Another image (upper) shows looping between flanking regions of our gene and in our gene while data UCSC track (lower) also shows this near our variant highlighted in blue.

A

Search Result for BRCA1

KeyName	Species	NONCODE ID
lnc-BRCA1-1-5_dup1	human	NONHSAT053833.2
lnc-BRCA1-1-7_dup1	human	NONHSAT053842.2
lnc-BRCA1-1-6_dup1	human	NONHSAT053843.2
lnc-BRCA1-1-3_dup1	human	NONHSAT053835.2
lnc-BRCA1-1-8_dup1	human	NONHSAT053834.2
lnc-BRCA1-2-1_dup1	human	NONHSAT176229.1
lnc-BRCA1-1-4	human	NONHSAT176235.1
lnc-BRCA1-3:7	human	NONHSAT237912.1
lnc-BRCA1-3:1	human	NONHSAT237915.1
lnc-BRCA1-3:10	human	NONHSAT237913.1
lnc-BRCA1-3:12	human	NONHSAT237914.1
lnc-BRCA1-3:11	human	NONHSAT237916.1

B

NONCODE TRANSCRIPT ID	NONHSAT237912.1
NONCODE Gene ID	NONHSAG021875.3
Chromosome	chr17
Start Site	43148367
End Site	43170316
Strand	-
Exon Number	4
CNCI Score	-0.0094208
Length	1782
Assembly	hg38
Other transcript Versions	None

Figure 14: NONCODE output. A) All human lncRNAs associated with BRCA1. B.) General information for NONHSAT237912.1.

Release Version: 20200227123210

Search:

Population	Group	Sample Size	Ref Allele	Alt Allele
African American	Sub	2206	T=0.7471	C=0.2529
African	Sub	2258	T=0.7462	C=0.2538
African Others	Sub	52	T=0.71	C=0.29
European	Sub	121692	T=0.668705	C=0.331295
Total	Global	137946	T=0.663252	C=0.336748
Other	Sub	8794	T=0.6540	C=0.3460
Other Asian	Sub	68	T=0.63	C=0.37
Latin American 1	Sub	24	T=0.62	C=0.38
Asian	Sub	236	T=0.619	C=0.381
East Asian	Sub	168	T=0.613	C=0.387
Latin American 2	Sub	92	T=0.60	C=0.40
South Asian	Sub	4850	T=0.5082	C=0.4918

Figure 15: dbSNP Population Frequencies (T allele) (1)

chr17:41223094
Name: rs1799966
Observed: A/G/T
Mol type: genomic
Class: single
Function: ncRNA,missense
Alleles: A, G, T,
Allele freqs: 0.650117, 0.349820, 0.000063,
Submitters: 1000GENOMES, ABI, AFFY, BCM-HGSC-SUB, BGI, BIC_BRODY, BUSHMAN, CANCER-GENOME, CGM_KYOTO, CLINSEQ_SNP, COMPLETE_GENOMICS, CSHL-HAPMAP, DDI, EGP_SNPS, ENSEMBL, EVA-GONL, EVA_DECODE, EVA_EXAC, EVA_FINRISK, EVA_GENOME_DK, EVA_MGP, EVA_SVP, EVA_UK10K_ALSPAC, EVA_UK10K_TWINSUK, EXOME_CHIP, F_HASHEMI, GMI, HAMMER_LAB, HGBASE, ILLUMINA, JMKIDD_LAB, KRIBB_YJKIM, NHLBI-ESP, OSTRANDER, PERLEGEN, PJP, SEATTLESEQ, SNP500CANCER, SSAHASNP, SSMP, TISHKOFF, WEILL_CORNELL_DGM, WIAF-CSNP,
https://www.ncbi.nlm.nih.gov/projects/SNP/snp_ref.cgi?rs=rs1799966

Figure 16: IGV dbSNP 1.4.7 allele frequency for rs1799966