**1. Brief description of its function in your own words.**

This gene is tumor suppressor, which encodes a phosphoprotein that can combine with other tumor suppressors and repair the damage of DNA.

**2. What type of tissues express this gene?**

40% of breast cancer and 80% ovarian cancer are related to mutation of this gene.

**3. On what chromosome is this gene located?**

Chr 17

**4. The exact location of this gene on the genome according to GRCh38.p7 assembly.**

I didn’t find GRCh38.p7.

Instead I found GRCh38.p12:

Start: 43044295

End: 43125483

**5. Names of the 3 genes neighboring this gene on the chromosome.**

Right: RND2 VAT1 IFI35

Left: NBR2 BRCA1P1 NBR1

**6. How many papers have been published on this gene?**

5 papers

**7. What genetic diseases is associated with mutations of thig gene?**

Breast-ovarian cancer, familial 1

Familial cancer of breast

FANCONI ANEMIA, COMPLEMENTATION GROUP S

Hereditary breast and ovarian cancer syndrome

Pancreatic cancer 4

**8. What proteins interact with the BRCA1 protein? Name 2 with references.**

HIV-1 Tat: <https://www.ncbi.nlm.nih.gov/protein/9629358>

and

HIV-1 Vpr: <https://www.ncbi.nlm.nih.gov/protein/28872817>

**9. What molecular pathways is the BRCA1 protein part of? Name 2.**

[ATF-2 transcription factor network, organism-specific biosystem](https://www.ncbi.nlm.nih.gov/biosystems/138006?Sel=geneid:672#show=genes)

[Breast cancer, organism-specific biosystem](https://www.ncbi.nlm.nih.gov/biosystems/1435207?Sel=geneid:672#show=genes)

**10. How many alternatively spliced transcripts (variants) does this gene have?**

Variants count in different DB are different.

In ClinVar, we can find 7178 variants.

In GRCh38, we can find 14489 variants.

The subset of these two DB is 4568 variants.

In NCBI dbVar, we can find 362 variants related studies.

**11. How many protein isoforms does this gene encode?**

5 different protein isoforms

**12. For the transcript called NM\_007294.3, how many exons and introns?**

23 exons

22 introns

**13. Retrieve the following sequences for the transcript NM\_007294.3. First report their length and then paste the FASTA formatted sequence.**

**a. Spanning genomic DNA sequence**

Length: 81189

Too long. Using separate file.

**b. mRNA sequence**

Length: 5592

Too long. Using separate file.

**c. Corresponding protein sequence**

Length: 1861

MDLSALRVEEVQNVINAMQKILECPICLELIKEPVSTKCDHIFCKFCMLKLLNQKKGPSQ

CPLCKNDITKRSLQESTRFSQLVEELLKIICAFQLDTGLEYANSYNFAKKENNSPEHLKD

EVSIIQSMGYRNRAKRLLQSEPENPSLQETSLSVQLSNLGTVRTLRTKQRIQPQKTSVYI

ELGSDSSEDTVNKATYCSVGDQELLQITPQGTRDEISLDSAKKAACEFSETDVTNTEHHQ

PSNNDLNTTEKRAAERHPEKYQGSSVSNLHVEPCGTNTHASSLQHENSSLLLTKDRMNVE

KAEFCNKSKQPGLARSQHNRWAGSKETCNDRRTPSTEKKVDLNADPLCERKEWNKQKLPC

SENPRDTEDVPWITLNSSIQKVNEWFSRSDELLGSDDSHDGESESNAKVADVLDVLNEVD

EYSGSSEKIDLLASDPHEALICKSERVHSKSVESNIEDKIFGKTYRKKASLPNLSHVTEN

LIIGAFVTEPQIIQERPLTNKLKRKRRPTSGLHPEDFIKKADLAVQKTPEMINQGTNQTE

QNGQVMNITNSGHENKTKGDSIQNEKNPNPIESLEKESAFKTKAEPISSSISNMELELNI

HNSKAPKKNRLRRKSSTRHIHALELVVSRNLSPPNCTELQIDSCSSSEEIKKKKYNQMPV

RHSRNLQLMEGKEPATGAKKSNKPNEQTSKRHDSDTFPELKLTNAPGSFTKCSNTSELKE

FVNPSLPREEKEEKLETVKVSNNAEDPKDLMLSGERVLQTERSVESSSISLVPGTDYGTQ

ESISLLEVSTLGKAKTEPNKCVSQCAAFENPKGLIHGCSKDNRNDTEGFKYPLGHEVNHS

RETSIEMEESELDAQYLQNTFKVSKRQSFAPFSNPGNAEEECATFSAHSGSLKKQSPKVT

FECEQKEENQGKNESNIKPVQTVNITAGFPVVGQKDKPVDNAKCSIKGGSRFCLSSQFRG

NETGLITPNKHGLLQNPYRIPPLFPIKSFVKTKCKKNLLEENFEEHSMSPEREMGNENIP

STVSTISRNNIRENVFKEASSSNINEVGSSTNEVGSSINEIGSSDENIQAELGRNRGPKL

NAMLRLGVLQPEVYKQSLPGSNCKHPEIKKQEYEEVVQTVNTDFSPYLISDNLEQPMGSS

HASQVCSETPDDLLDDGEIKEDTSFAENDIKESSAVFSKSVQKGELSRSPSPFTHTHLAQ

GYRRGAKKLESSEENLSSEDEELPCFQHLLFGKVNNIPSQSTRHSTVATECLSKNTEENL

LSLKNSLNDCSNQVILAKASQEHHLSEETKCSASLFSSQCSELEDLTANTNTQDPFLIGS

SKQMRHQSESQGVGLSDKELVSDDEERGTGLEENNQEEQSMDSNLGEAASGCESETSVSE

DCSGLSSQSDILTTQQRDTMQHNLIKLQQEMAELEAVLEQHGSQPSNSYPSIISDSSALE

DLRNPEQSTSEKAVLTSQKSSEYPISQNPEGLSADKFEVSADSSTSKNKEPGVERSSPSK

CPSLDDRWYMHSCSGSLQNRNYPSQEELIKVVDVEEQQLEESGPHDLTETSYLPRQDLEG

TPYLESGISLFSDDPESDPSEDRAPESARVGNIPSSTSALKVPQLKVAESAQSPAAAHTT

DTAGYNAMEESVSREKPELTASTERVNKRMSMVVSGLTPEEFMLVYKFARKHHITLTNLI

TEETTHVVMKTDAEFVCERTLKYFLGIAGGKWVVSYFWVTQSIKERKMLNEHDFEVRGDV

VNGRNHQGPKRARESQDRKIFRGLEICCYGPFTNMPTDQLEWMVQLCGASVVKELSSFTL

GTGVHPIVVVQPDAWTEDNGFHAIGQMCEAPVVTREWVLDSVALYQCQELDTYLIPQIPH

SHY

**d. First exon (Hint: easier to find in Ensembl.org, see below)**

Length: 213

[G](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs1372236983;vf=377763296)TACCTTG[A](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs539845430;vf=364592921)TTT[C](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs1182524199;vf=372409435)[G](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs1261058390;vf=374626702)TAT[T](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs1042701541;vf=371266683)CTGAGAGGCTGCTGCT[T](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs569113962;vf=365440366)AGCGG[T](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs113323025;vf=363116423)A[G](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs1175395946;vf=372208430)CCC[CT](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs1403825048;vf=378654881)TGGTTTC[C](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs916153090;vf=367841577)GTGG

CAACGG[A](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs993065651;vf=369922608)AAAGCGC[G](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs1360659853;vf=377436016)GGAATTACAG[A](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs1380213941;vf=377987722)[T](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs1297242495;v=rs8176075;v=rs1318107972;vf=375648167;vf=362671054;vf=376235253)[A](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs1318107972;vf=376235253)[A](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs759523510;vf=366149975)[A](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs1223309082;vf=373559242)[TTAAA](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs1261582095;vf=374642056)[A](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs1326294116;vf=376468084)C[T](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs1211421601;v=rs35436937;vf=373225211;vf=362807997)[G](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs35436937;vf=362807997)[C](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs1260285579;vf=374604040)[G](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs1484915984;vf=380941724)ACTG[C](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs938827457;vf=368455073)G[C](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs148196794;vf=363459804)G[G](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs190861568;vf=363784764)C[G](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs1438870433;vf=379641810)[T](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs1253346809;vf=374407927)GAG[C](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs548750620;vf=364850548)T[C](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs1468059405;vf=380465807)G

[C](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs1198439560;vf=372859390)[T](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs759750310;vf=366157017)[G](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs1426929209;vf=379305049)[A](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs886052977;vf=367026724)GA[C](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs527449526;vf=364233720)[T](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs1326597773;vf=376476709)[T](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs963494793;v=rs34191881;vf=369121999;vf=362779338)[C](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs1016395214;v=rs34191881;vf=370553919;vf=362779338)[C](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs1327413886;vf=376499695)[T](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs776813709;vf=366686534)G[G](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs986120323;vf=369734882)[A](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs1336253769;vf=376749097)[C](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs766347954;vf=366362132)[G](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs187992882;vf=363714897)G[G](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs1210920972;vf=373211321)[G](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs544857747;vf=364738007)[G](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs1359109190;vf=377392025)[A](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs1271328462;vf=374915419)[C](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs1218048314;vf=373411715)[A](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs1326068348;vf=376461509)GG[C](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs143160357;vf=363337092)TGT[G](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs1382379016;vf=378050839)GG[G](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs1158334695;vf=371728701)T[T](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs1220534986;vf=373480766)T[C](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs529629382;vf=364296642)[T](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs1365946346;vf=377585952)[C](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs1341356647;vf=376893473)A[G](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs1402792661;vf=378625623)[A](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs1192234384;vf=372683089)TA[A](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs1280398288;vf=375171291)[C](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs1454233540;vf=380077112)TG[G](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs1253690897;vf=374417358)[G](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs776540382;vf=366678090)[C](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs549332987;vf=364867475)[C](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs1360586320;vf=377433944)[C](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs900972535;vf=367429971)[C](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs1031507808;vf=370964033)[T](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs1283883558;vf=375269777)GC[G](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs1205965118;vf=373071160)C

TCA[G](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs1057523081;vf=371667924)GAG[G](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs1057521869;vf=371667860)[C](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs1262271602;vf=374661454)C[T](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs544342552;vf=364722977)T[C](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs1437669505;vf=379607186)AC[C](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs1057524754;vf=371668022)CTCTGC[T](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs1370244781;vf=377707475)[C](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs774839252;vf=366625715)TG[G](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs1057522752;vf=371667910)[G](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs1057520890;vf=371667806)TA[A](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs886052976;vf=367026723)A[G](http://uswest.ensembl.org/Homo_sapiens/ZMenu/TextSequence?db=core;factorytype=Location;g=ENSG00000012048;r=17:43044295-43125483;t=ENST00000357654;v=rs886039588;vf=367025650)

**e. 5’ Untranslated region (UTR)**

‘N’ denote the 5’ UTR

Length: 232

NTACCTTGNTTTNNTATNCTGAGAGGCTGCTGCTNAGCGGNANCCCNNTGGTTTCNGTGG

CAACGGNAAAGCGCNGGAATTACAGNNNNNNNNNNNCNNNNACTGNGNGNCNNGAGNTNG

NNNNGANNNNNNGNNNNGNNNNNNGGNTGTNGGNTNTNNNANNTANNTGNNNNNNNGCNC

TCANGAGNNCNTNACNCTCTGCNNTGNNTANANTTCNTNNNNNCAGAAANNN

**f. 3’ untranslated region (UTR)**

‘N’ denote the 3’ UTR

Length: 1383

NTNNAGCCAGCCNNAGGTANNNNGNNACAGNACCCNAANAANNAGCTTACAAAGT

GGNCNTNNCNGNCCCTGGGAGNTCCTCTNNCTNTTCNGTCNTNNTANNNNNCTGNCTACT

AANTATTTTNNGTACATNAGCCTNNAANNNACTTCTNNCNNTGCNAGNGTCCCTTAAAGA

TTNNCTGCTNNAAGTCTCCCTTGGAAATNNGCCNTGAGNACAAAATTNTGNTNNTTTTTC

ACCTGAGAAGATTNTAANNCCNTTTAAANNCCANCANTTGAGCAAGATGCTGATTNATTA

TTTATCNGCNNTANTCTTTCTATNNANGCTNTTNTTGGCTTAGGNNTGGNAGCACAGANT

GGCTTNGCCTCANGANAANNNNNNNNNNNNNNNNGTNTACTTNTNTAAAACCNTNTGTTC

ACAANNNNNNAGNNTCAGNCCCTTCANTGGAAGGAGNGTNCTTGNNANNNNNTATNTGAC

NTANNNTCANAANANTCCTTGGGNAGTTNTCAAANNTTGGNNTGNAANANNGNGGAGNAA

NNTCTGAGGCAGGNNTNANAAATGANNNGNNAANTTGAAANCTNNGNANGGTGGNNNANN

CNTGTAATCCCAGCANTTNNGGAGGNCNNGNNGNGNAGANCACTGNAGGTCAGGNGNTNN

ANACNAGCNNGGCCANCANGNTGAAACCCNNTCTCTANTNAAAATANANAAATTNNCNNG

NCNTGNTGNNGGNCANNTNTNATNCCANCNANNCNGGTGGCTNAGNCAGGANAATCACTT

CNNCCNNGGAGGTNGANNTTNNNNNGAGNCANGATCNTACNANNGCANTCCAGCCTGNGT

GANANNNNNACTGTNNNNNNAAAAAAANNNNNNNNNNNNNNAAANNAANNTNGNANNNNN

TTCTNAAAGTNTNAGNTATANTNNCNAGATTNNTAAANNNTGTGTTCTANAACAGCAGNN

NNNTTTCANGANCNNNNTTCCAAAGACAGTCTNNTAATTNCTNATTNGTAANAAGTNAAN

TGTTTATTNTTGTAGCTNTGNTATANAATCNANTCCNNTTAAANTNTAAGANNTNTGGCN

TGANTATTTCNTATCTNTAANATNACAGATCCCACCANGANNNAAGCTGTTNCTTTCTNT

GAGNTGATNTTTTNNCTTTGNTCNNNGNTGCTGAAACCATACANCTTCNTNAATANTNTT

NNTTGNTNANGGAAGAAAANNTNTTTTTCATAAANCCNNNANCCAGNACTGNNTANNGCN

GTTGGAAGNACTANNNCTTCCCNANNNCNNNNNGTGNGCANNNNCAGTNANGANTTGATN

GTACNAANTANNTTNTNTNANTGTNNNGCTNTNAACNCTGCANATNANNTTGGNNGCANN

CACTTCNN

**g. 1000bp promoter sequence upstream of transcription start site (TSS)**

Length: 1000

GTACCTTGATTTCGTATTCTGAGAGGCTGCTGCTTAGCGGTAGCCCCTTGGTTTCCGTGGCAACGGAAAAGCGCGGGAATTACAGATAAATTAAAACTGCGACTGCGCGGCGTGAGCTCGCTGAGACTTCCTGGACGGGGGACAGGCTGTGGGGTTTCTCAGATAACTGGGCCCCTGCGCTCAGGAGGCCTTCACCCTCTGCTCTGGGTAAAGGTAGTAGAGTCCCGGGAAAGGGACAGGGGGCCCAAGTGATGCTCTGGGGTACTGGCGTGGGAGAGTGGATTTCCGAAGCTGACAGATGGGTATTCTTTGACGGGGGGTAGGGGCGGAACCTGAGAGGCGTAAGGCGTTGTGAACCCTGGGGAGGGGGGCAGTTTGTAGGTCGCGAGGGAAGCGCTGAGGATCAGGAAGGGGGCACTGAGTGTCCGTGGGGGAATCCTCGTGATAGGAACTGGAATATGCCTTGAGGGGGACACTATGTCTTTAAAAACGTCGGCTGGTCATGAGGTCAGGAGTTCCAGACCAGCCTGACCAACGTGGTGAAACTCCGTCTCTACTAAAAATACAAAAATTAGCCGGGCGTGGTGCCGCTCCAGCTACTCAGGAGGCTGAGGCAGGAGAATCGCTAGAACCCGGGAGGCGGAGGTTGCAGTGAGCCGAGATCGCGCCATTGCACTCCAGCCTGGGCGACAGAGCGAGACTGTCTCAAAACAAAACAAAACAAAACAAAACAAAAAACACCGGCTGGTATGTATGAGAGGATGGGACCTTGTGGAAGAAGAGGTGCCAGGAATATGTCTGGGAAGGGGAGGAGACAGGATTTTGTGGGAGGGAGAACTTAAGAACTGGATCCATTTGCGCCATTGAGAAAGCGCAAGAGGGAAGTAGAGGAGCGTCAGTAGTAACAGATGCTGCCGGCAGGGATGTGCTTGAGGAGGATCCAGAGATGAGAGCAGGTCACTGGGAAAGGTTAGGGGCGGGGAGGCCTTGATTGGTGTTG