SI Appendix

Intragenic DNA methylation and BORIS mediated cancer-specific splicing contributes to Warburg effect

Smriti Singh, Sathiya Pandi Narayanan, Kajal Biswas, Amit Gupta, Neha Ahuja, Sandhya Yadav, Rajendra Kumar Panday, Atul Samaiya, Shyam K Sharan, Sanjeev Shukla

SI Appendix Materials and Methods

Cell culture

The human breast cancer cell lines MCF7 and MDA-MB-231 were obtained from American Type Culture Collection (ATCC). MCF7 cell line was cultured at 37ºC, 5% CO_2 in ATCC recommended growth medium DMEM (Gibco, 11995-065) whereas MDA-MB-231 cell line was cultured at 37ºC without CO_2 in ATCC recommended growth medium L-15 (Gibco, 11415-064). Both the cell lines were supplemented with 10% Fetal Bovine Serum (Invitrogen, 16000044), 100 units/ml of Penicillin and Streptomycin (Invitrogen, 15140122), 2 mmol/l L-glutamine (Sigma, G7513). The human mammary epithelial cells (HMECs) were obtained from Lonza (CC-2551) and cultured in the mammary epithelial cell growth medium (Lonza, CC-3151& CC-4136) according to the manufacturer's instructions.

Breast cancer sample collection

Tumor and adjacent normal tissue pairs were collected from patients undergoing surgery for breast cancer at Navodaya Cancer Hospital and Bansal Hospital, Bhopal, India. The study was approved by the Institute Ethics Committee of Indian Institute of Science Education and Research Bhopal. The informed consent was obtained from all the patients. The tissue samples were snap frozen immediately after surgery and stored at -80°C until use. The tissues for RNA isolation were collected in RNA later (Thermo Fisher Scientific, AM7024)

after surgery, snap frozen and stored at -80°C until use. Clinical characteristics of patients used in the study are presented in *SI Appendix* Table S6.

Plasmids

The PKM2 overexpression plasmid was constructed by amplifying a 1.5-kb PKM fragment from MCF7 cDNA using Platinum Taq High-Fidelity DNA Polymerase (Thermo Fisher Scientific, 11304029) and primers PKM2 Fw (5′TTTTGCGGCCGCATGTCACCGGAAGCCCAA-3′) and PKM2 Rev (5′-TTTTGTCGACTCACGGCACAGGAACAAC-3′). The product was cloned between the NotI and SalI sites of pFLAG-CMV-6c expression Vector (Sigma).

RNA interference

The MCF7 cells ($3X10^5$) were seeded in six-well culture plates. After 24 hours, gene specific shRNA (Sigma, Mission human genome shRNA library) and shRNA against eGFP (Sigma, SHC005) as a control (2µg/well) were transfected using Lipofectamine-3000 reagent (Thermo Fisher Scientific, L3000-008) as per the manufacturer's instructions. After 36 hours of transfection, cells were selected with 0.8 µg/ml puromycin (Sigma, P9620) for 72 hours. For ChIP, the transfection with gene specific as well as control shRNA was done in the 100mm culture dish (8µg/plate).

Oligo sequence of shRNAs

eGFP shControl	5'-CCGGTACAACAGCCACAACGTCTATCTCGAGATAGACGTTGTGGCTGTTGTATTTTT-3'
shBORIS	5'-CCGGGTCCGACACAGGCGCTATAAACTCGAGTTTATAGCGCCTGTGTCGGACTTTTTG-3'
shDNMT3B	5'-CCGGCCATGCAACGATCTCTCAAATCTCGAGATTTGAGAGATCGTTGCATGGTTTTTG-3'
shDNMT3A	5'-CCGGCCACCAGAAGAAGAGAAGAATCTCGAGATTCTTCTCTTCTTCTGGTGGTTTTTG-3'
shDNMT1	5'-CCGGCGACTACATCAAAGGCAGCAACTCGAGTTGCTGCCTTTGATGTAGTCGTTTTT-3'
shCTCF	5'-GTACCGGTATGATTTCCCATCGACATTTCTCGAGAAATGTCGATGGGAAATCATATTTTTTG-3'

Quantitative RT-PCR

Total RNA was extracted from cultured MCF7 cells, MDA-MB-231 cells and human breast tissue sample using the RNeasy kit (Qiagen, 74106) and Trizol (Ambion, 15596018) respectively according to manufacturer's instruction. RNA was quantified using Nanodrop (Thermo Fisher Scientific, ND8000). 1µg of total RNA was reverse transcribed by iScript cDNA synthesis kit (BioRad, 17088) as per the manufacturer's instructions. Amplification reactions were prepared in duplicate using SYBR green (Affymetrix, 75665) and amplification was performed on light cycler 480 II (Roche) according to manufacturer's instruction. Primers were designed using IDT PrimerQuest tool (https://www.idtdna.com/). The primers used in this study are mentioned in the *SI Appendix* Table S1. The average cycle thresholds from three independent biological replicate samples were calculated and normalized to housekeeping control gene *RPS16*. Normalization was performed using *RPS16* as a normalization control using the formula: $[2^{\wedge(\text{Ct control} - \text{Ct target})}]$. In addition to control gene normalization, constitutive exon normalization was performed for exon-level expression analysis. Student's t-test was used to compare gene/exon expression between two different groups. $P < 0.05$ was considered as statistically significant.

Chromatin immunoprecipitation (ChIP)

ChIP assays were performed as described previously (1). Briefly, sonicated chromatin (25 µg) was immunoprecipitated by adding the antibody of interest followed by overnight incubation at 4 °C. The following antibodies were used for ChIP: Anti-BORIS (Millipore ABE631, batch no. VP1403050), Anti- RNA Pol II (Millipore 17672, batch no. 2080890), Anti- CTCF (Millipore 1710044, batch no. 2377389), Normal Rabbit IgG (Millipore 12370, batch no. 2295402), Normal mouse IgG (Calbiochem NI03, batch no. D00151120). Immunoprecipitated fractions and 5% input were analyzed by quantitative real-time PCR in duplicate using the SYBR Green Master Mix (Affymetrix, 75665) and specific primers

(mentioned in the *SI Appendix* Table S1) across the exonic regions. H19 Imprint Control Region (ICR) (2) was used as a positive control for CTCF ChIP. Each experiment was performed at least three times. Normalization was performed to input using the formula: [$2^{\wedge(Ct\ input - CT\ immunoprecipitation)}$]. Resultant values were further normalized relative to the rabbit or mouse Ig control IP values for the primer set. Student's t-test was used to identify the significance between two different groups. $P< 0.05$ was considered statistically significant.

Methylated DNA immunoprecipitation (MeDIP)

Genomic DNA was isolated from MCF7 cells and breast tumor tissue sample using DNA isolation kit (Sigma, G1N70) and DNAzol (Thermo Fisher Scientific, 10503-027) respectively, according to manufacturer's instruction. MeDIP assays were performed as described previously (1). Briefly, 2 µg of sonicated DNA was processed and incubated with 5-Methyl cytosine antibody (Active Motif 39649, batch no. 20414013) and Normal mouse IgG antibody (Calbiochem NI03, batch no. D00151120) for overnight at 4^0C. Immunoprecipitated fractions and 5% input were analyzed by quantitative real-time PCR in duplicate using the SYBR Green Master Mix (Affymetrix, 75665) and specific primers (mentioned in the *SI Appendix* Table S1) across the exonic regions. Normalization was performed to input using the formula: [$2^{\wedge(Ct\ input - CT\ immunoprecipitation)}$]. Resultant values were further normalized relative to the mouse Ig control IP values for the primer set. Student's t-test was used to identify the significance between two different groups. $P< 0.05$ was considered statistically significant.

Western blot

The MCF7 cells/MDA-MB-231 cells/ HMEC cells/homogenized tissues were lysed in NP-40 buffer (150mM NaCl, 1% NP-40, 50mM Tris-HCl pH8.0, 1X Protease inhibitor). The following antibodies were used for western blotting: Anti- PKM1 (Cell Signaling Technology 7067S, batch no.1), Anti-PKM2 (Cell Signaling Technology 4053S, batch no. 3), Anti-

BORIS (Millipore ABE631, batch no. VP1403050), Anti-DNMT1 (Abcam ab87656, batch no. GR215476-4), Anti- DNMT3A (Abcam ab2850, batch no. GR218985-4), Anti-DNMT3B (Abcam ab13604, batch no. GR236376-4), and Anti Flag (Sigma F7425). Anti-GAPDH (Cell Signaling Technology 5174S, batch no. 6), Anti- alpha Tubulin (Abcam ab126165, batch no. GR106527-20) and Anti-Ran (BD Biosciences 610340, batch no. 3004982) were used as loading controls for protein assays.

Generation of BORIS binding mutant clone using CRISPR/Cas9:

MCF7 cells were subjected to sgDNA directed Cas9 mediated DNA double strand break in BORIS binding site. The sgDNA sequence (5'-TGGTAATGGGCGCCAGGCGG-3') was cloned at *Bbs*I site into pSpCas9 (BB) (Addgene plasmid ID: 42230). During double strand break repair, two alleles can be repaired by two independent NHEJ (non-homologous end joining) events in the same cell and thus generates a compound heterozygote (3, 4) where two alleles have two different mutations. The correct clone was confirmed by Sanger sequencing and cotransfected into MCF7 cells along with GFP expressing plasmid. After 36 hours of transfection, GFP positive cells were sorted using cell sorter into 96-well plate. DNA was isolated from individual clones and 600 bp region near the sgDNA targeting sequence was amplified using the primers (5' TCACAGGCACTTGGTGAAGGAC-3' and 5'-GATCCAAATGTTCTGACATTCC-3'). Sanger sequencing was performed from individual clones and analyzed for deletion. The clones that showed deletion in BORIS binding site were then checked for integrity at three top off-target sites of the sgDNA identified by sgDNA designing website (http://crispr.mit.edu). For further experiments using CRISPR deletion clone, clones that showed no modification after Cas9-sgDNA transfection was used as a control.

Cell viability assay

The MCF7 cells ($3X10^5$) were transfected with BORIS shRNA and shRNA against eGFP (Sigma, SHC005) as a control in six-well culture plates. After selection of transfected cells with puromycin, shBORIS transfected cells were transfected with PKM2 overexpression plasmid. After 24 hours, cells ($4x10^3$) were seeded in 96 well culture plates and were cultured for 24 hours, 48 hours, 60 hours, and 72 hours (in triplicate for each condition). Cell growth was determined by measuring the conversion of MTT Tetrazolium salt (Sigma, M2128) to MTT formazan. In brief, 20µl of MTT stock solution (2mg/ml) was added to each well and incubated for two to three hours. After the incubation time, formazan crystals formed in the cells were solubilized using lysis buffer (Isopropanol, 0.1%SDS, 4mMHCl). Numbers of viable cells were measured by a plate reader (BioTek Eon,11-120-611) at an optical density of 570 nm.

Wound healing assay

The MCF7 cells ($3X10^5$) were seeded in six-well culture plates. After 24hours, shBORIS and shControl (2µg/well) were transfected using Lipofectamine reagent (Thermo Fisher Scientific, L3000-008) as per the manufacturer's instructions. After selection of transfected cells with puromycin, shBORIS transfected cells were transfected with PKM2 overexpression plasmid. After 24hours, the cell monolayer was scratched using a sterile 10µl pipette tip and washed with 1XPBS two times to remove cellular debris. Three wounded areas in each well were marked on the bottom of plates and images were captured at 24 hours, 48 hours, 72 hours, and 96 hours with an inverted microscope. The gap lengths were calculated from the images using ImageJ software.

Colony forming assay

The MCF7 cells ($3X10^5$) were seeded in six-well culture plates. After 24 hours, shBORIS and shControl (2µg/well) were transfected using Lipofectamine reagent (Thermo Fisher Scientific, L3000-008) as per the manufacturer's instructions. The colony forming assay was

done as described earlier (5). Briefly, after 48 hours of transfection, cells were trypsinized, and $1x10^3$ cells were seeded in the new 12-well plate. After 24 hours of seeding, cells were selected with 0.8µg/ml puromycin for seven days. At the time of staining, cells were fixed using methanol and acetic acid (3:1) for five minutes. After fixation, cells were washed with 1x PBS three times. After washing cells were stained with 0.05% crystal violet stain.

Lactate assay

The MCF7 cells ($3X10^5$) were transfected with respective shRNA and shRNA against eGFP (Sigma, SHC005) as a control in six-well culture plates. After selection of transfected cells with puromycin, shBORIS transfected cells were transfected with PKM2 overexpression plasmid. The cells were trypsinized 48 hours after transfection of PKM2 overexpression plasmid, and cell counting was performed using trypan blue (Himedia, TC193). An equal number of cells were homogenized in the presence of lactate assay buffer and centrifuged at 13,000g for 10 minutes. Similarly, for the lactate quantification after Aza treatment, the MCF7 cells ($3X10^5$) were treated with 10µM 5-Aza-2'-deoxycytidine (Sigma, 83656) for 6 days in six-well cell culture plates. After treatment, cell counting was performed using trypan blue and equal numbers of cells were homogenized in the presence of lactate assay buffer. For lactate quantification in tumor samples, 50 mg of tissue sections were crushed using liquid nitrogen and homogenized with lactate assay buffer. Lactate quantification was performed using commercially available lactate assay kit (Sigma, MAK064) in a 96-well plate as per the manufacturer's instruction. Lactate level was measured with a plate reader (BioTek Eon, 11-120-611) at an optical density of 570 nm.

Glucose uptake assay

The MCF7 cells ($3X10^5$) were transfected with respective shRNA and shRNA against eGFP (Sigma, SHC005) as a control in six-well cell culture plates. After selection of transfected cells with puromycin, shBORIS transfected cells were transfected with PKM2 overexpression

plasmid. The cells were trypsinized 48 hours after transfection of PKM2 overexpression plasmid, and cell counting was performed using trypan blue (Himedia, TC193). An equal number of cells were homogenized in the presence of glucose assay buffer and centrifuged at 13,000g for 10 minutes. Similarly, the MCF7 cells ($3X10^5$) were also treated with 10µM 5-Aza-2'-deoxycytidine (Sigma, 83656) in six-well cell culture plates for 6 days. After treatment, cell counting was performed using trypan blue and equal numbers of cells were homogenized in the presence of glucose assay buffer. Glucose level quantification was performed using commercially available glucose assay kit (Abcam ab65333) in a 96-well plate as per the manufacturer's instruction. Glucose level was measured with a plate reader (BioTek Eon, 11-120-611) at an optical density of 570 nm.

Caspase 3/7 assay

The MCF7 cells ($3X10^5$) were seeded in six-well culture plates. After 24 hours, shBORIS and shControl (2µg/well) were transfected using Lipofectamine reagent (Thermo Fisher Scientific, L3000-008) as per the manufacturer's instructions. After selection of transfected cells with puromycin, shBORIS transfected cells were transfected with PKM2 overexpression plasmid. After 72 hours of transfection, cells were trypsinized and seeded (10^3) in 96 well plate. After 24 hours, Caspase 3/7 activation in shControl and shBORIS transfected cells was measured using the Caspase-Glo 3/7 assay (Promega, G8090) as recommended by the manufacturer. Luminescence readings were taken using Glomax multi-detection system (Promega). Hydrogen peroxide (20µM) treated MCF7 cells were used as a positive control. Similarly, Mut BBS MCF7 cells were transfected with PKM2 overexpression plasmid and Caspase- Glo 3/7 assay was performed in Wt BBS and Mut BBS MCF7 cells.

Human Transcriptome Array 2.0 array profiling

Total RNA samples were isolated from shControl and shBORIS transfected cells using RNeasy kit (Qiagen, 74106). RNA integrity was analyzed using Bioanalyzer (Agilent).

Biotinylated cDNAs were prepared from a minimal 100ng of total RNA using Life Technologies (Ambion) WT-plus RNA Amplification system. Following the amplification, cDNA fragmentation was done, and 5.5µg of fragmented cDNAs were hybridized for 16 hr at 45^0C on Affymetrix GeneChip Human Transcriptome Array 2.0 (HTA2.0) chips. The chips were washed and stained in the Affymetrix Fluidics Station 450. After hybridization, the fluorescence intensity of the arrays was scanned using the Affymetrix Scanner 7G. The raw data was analyzed using Affymetrix expression console and Transcriptome console software.

Human Transcriptome Array 2.0 array data analysis

Raw data were normalized by SST-RMA method using Expression Console software. SST-RMA normalized HTA2.0 array data files were used for the global alternative splicing analysis. Transcriptome Analysis Console was used to identify the alternative splicing pattern between shBORIS and shControl transfected MCF7 cells. Splicing Index (linear) ≤ -2 or Splicing Index (linear) ≥ +2; $P< 0.05$. For breast cancer HTA2.0 array profile (GSE76250) (6, 7), we used expression profile of 165 tumor samples and 33 normal breast tissue samples for splicing analysis. Splicing pattern was identified by comparing normal breast versus breast tumor tissues with splicing cut-off index ≤ -1.2 and ≥ +1.2; $P< 0.05$. Splicing Index (SI) is the measure of exon inclusion and exclusion levels in alternative splicing analysis and was defined as the ratio of normalized exon intensity under two conditions. Normalized exon intensity (NI) is the ratio of exon intensity to the gene intensity. In brief, the SI was calculated using the following formula (8), (https://tools.thermofisher.com/content/sfs/brochures/id_altsplicingevents_technote.pdf);

$$\text{Normalized exon intensity (NI)} = \frac{\text{Exon intensity}}{\text{Gene level}}$$

$$\text{Splicing Index (SI)} = \log_2 \frac{\text{Sample 1 NI}}{\text{Sample 2 NI}}$$

Where exon intensity is the expression level of the probe, and gene intensity represents the overall transcript cluster of a gene. The positive SI means inclusion whereas negative SI

means exclusion. In addition to these measures, following criteria's were followed during splicing analysis: (1), Probe Selection Region (PSR)/Junction must be expressed in at least one condition, and (2), a gene must contain at least one PSR.

TCGA data analysis

The Level3 RNA-SeqV2 data for gene and exon level expression from normal and breast cancer tissues were downloaded from The Cancer Genome Atlas (TCGA) data portal (https://tcga-data.nci.nih.gov/tcga/). Downloaded level3 seq files were integrated to give gene as well as exon expression profile separately. Gene expression values were extracted from gene level data for *BORIS*. Fragments Per Kilobase of transcript per Million mapped reads (FPKM) and Transcripts per million (TPM) values were used to plot the *BORIS* expression in breast cancer samples. First we calculated reads per kilobase from TCGA data and count up all the reads per kilobase values in a sample and divided by $1x10^6$ to calculate the transcripts per million reads. *PKM* isoform-specific expression values were obtained from exon level data. Based on the isoform exon information, we assigned uc002att.1, uc002atv.1, uc002atw.1, uc002atx.1, uc010ukj.1, uc010ukk.1, and uc010bit.1 to *PKM1* and uc002atr.1, uc002ats.1, uc002atu.1, uc002aty.1, uc010biu.1, uc010uki.1, and uc002atz.1 to *PKM2* (9). These values were summed up to give *PKM1* and *PKM2* expression levels respectively.

ChIP-Sequencing data analysis

The ChIP-sequencing data were downloaded from GEO data portal. Downloaded ChIP-seq raw files (.SRA) (GSM1817668 and GSM1817669) were converted into FASTQ files using SRA tool kit. Sequencing reads were aligned to hg19 assembly using BOWTIE2 tool. MACS tool was used to call the peaks. IGV was used to visualize the peaks from the .bigwig file. Cistrome db tool was used to identify the enriched motifs in the BORIS ChIP seq data.

Nucleosome occupancy assay

Nucleosome occupancy assay was performed according to the previously described protocol (10) in shControl and shDNMT3B transfected MCF7 cells. Cells were cross-linked with 1% formaldehyde for 10 minutes at room temperature. The reaction was quenched by adding glycine to a final concentration of 0.125M at room temperature for 5 min with rotation. The cells were washed twice in chilled PBS and harvested by spinning down at 1200 rpm for 2minutes at 4^{o}C. The pellet was resuspended in cold buffer A (10mM HEPES pH 7.9, 10mM KCl, 1.5mM MgCl2, 0.34M Sucrose, 10% Glycerol, 1X Protease inhibitor, 1mM DTT) to which Triton X-100 was added to a final concentration of 0.1%. The cells were then incubated for 10 min in ice and were spun down at 1300 x g for 5 min at 4^{o}C. The pellet was again resuspended in buffer B (0.3mM EDTA, 0.2mM EGTA, 1mM DTT, 1X Protease inhibitor) and after incubating in ice for 30 minutes; the cells were spun down at 1700xg for 5 min at 4^{o}C. Then, the pellet is resuspended in MNase buffer (50mM Tris-HCl pH 7.4, 25mM KCl, 4mM MgCl2, 12.5% glycerol, 1mM CaCl2). The samples were divided into equal volumes (~500 µl), and 5% of the samples were taken from each tube as input. 0U/ml and 400U/ml (2 µl) of MNase were added accordingly in different tubes and was incubated for 10 min at 37^{o}C in water bath. The digestion reaction was stopped by adding equivalent volume of stop buffer (100mM EDTA, 100mM EGTA). The DNA was spun down at 16000 x g for 2 minutes. The pellet, as well as the input DNA, was resuspended in 150 µl of elution buffer (50mM Tris-HCl pH 8.0, 10mMEDTA, 1% SDS, 50mM NaHCO3). The DNA was then treated with 1 ml RNaseA (1mg/ml, Ambion) at 37^{o}C for 30 minutes. Reverse-crosslinking was done by adding 1 ml proteinase K (20mg/ml, Ambion) and incubating at 65^{o}C for 4 h or overnight. Eluted DNA was purified with QIAquick PCR purification (Qiagen). The digested DNA sample (400U/ml of MNase treated sample) was loaded on 1.2% agarose gel with 0.5% Tris-EDTA buffer, and the mononucleosome-sized fragments were cut out and were further

gel purified using Gel Extraction Kit (Qiagen). DNA was subsequently checked by quantitative PCR using primers for the regions to be assayed for nucleosome density.

Immunofluorescence

Immunofluorescence was performed as described previously (11). Wt BBS and Mut BBS MCF7 cells were plated in 12 well plate. After 24 hours, cells were washed with 1X PBS followed by fixation with 4% formaldehyde for 10 minutes at room temperature. Cells were then washed with PBS three times for 5 minutes at room temperature, and permeabilized by PBS containing 0.1% Triton-X100 (PBST). The cells were then blocked by 1% goat serum, followed by overnight incubation with PKM1 and PKM2 antibody in the blocking solution (PBST). The cells were washed with PBS and incubated with the fluorescein-conjugated anti rabbit secondary antibody (Invitrogen, A11036, batch no. 997761) for 1 hour, followed by washing with PBS. Then cells were stained with DAPI (Invitrogen, D1306, batch no. 1673432) for 5 minutes and washed with PBS. The coverslips were mounted inverted on a glass slide with fluoroshield mounting media (Sigma, F6182). Cells were then analyzed by ApoTome Axio microscope.

References:

1. Shukla S*, et al.* (2011) CTCF-promoted RNA polymerase II pausing links DNA methylation to splicing. *Nature* 479(7371):74-79.
2. Bell AC & Felsenfeld G (2000) Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene. *Nature* 405(6785):482-485.
3. Wang H*, et al.* (2013) One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell* 153(4):910-918.
4. Ran FA*, et al.* (2013) Genome engineering using the CRISPR-Cas9 system. *Nat Protoc* 8(11):2281-2308.
5. Narayanan SP*, et al.* (2015) Integrated genomic analyses identify KDM1A's role in cell proliferation via modulating E2F signaling activity and associate with poor clinical outcome in oral cancer. *Cancer Lett* 367(2):162-172.
6. Liu YR*, et al.* (2016) Comprehensive Transcriptome Profiling Reveals Multigene Signatures in Triple-Negative Breast Cancer. *Clin Cancer Res* 22(7):1653-1662.
7. Jiang YZ*, et al.* (2016) Transcriptome Analysis of Triple-Negative Breast Cancer Reveals an Integrated mRNA-lncRNA Signature with Predictive and Prognostic Value. *Cancer Res* 76(8):2105-2114.
8. Chen L (2011) Statistical and Computational Studies on Alternative Splicing. *Handbook of Statistical Bioinformatics*, eds Lu HH-S, Schölkopf B, & Zhao H (Springer Berlin Heidelberg, Berlin, Heidelberg), pp 31-53.

9. Desai S*, et al.* (2014) Tissue-specific isoform switch and DNA hypomethylation of the pyruvate kinase PKM gene in human cancers. *Oncotarget* 5(18):8202-8210.
10. Bryant GO (2012) Measuring nucleosome occupancy in vivo by micrococcal nuclease. *Methods Mol Biol* 833:47-61.
11. Das AM, Eggermont AM, & ten Hagen TL (2015) A ring barrier-based migration assay to assess cell migration in vitro. *Nat Protoc* 10(6):904-915.

SI Appendix Figures

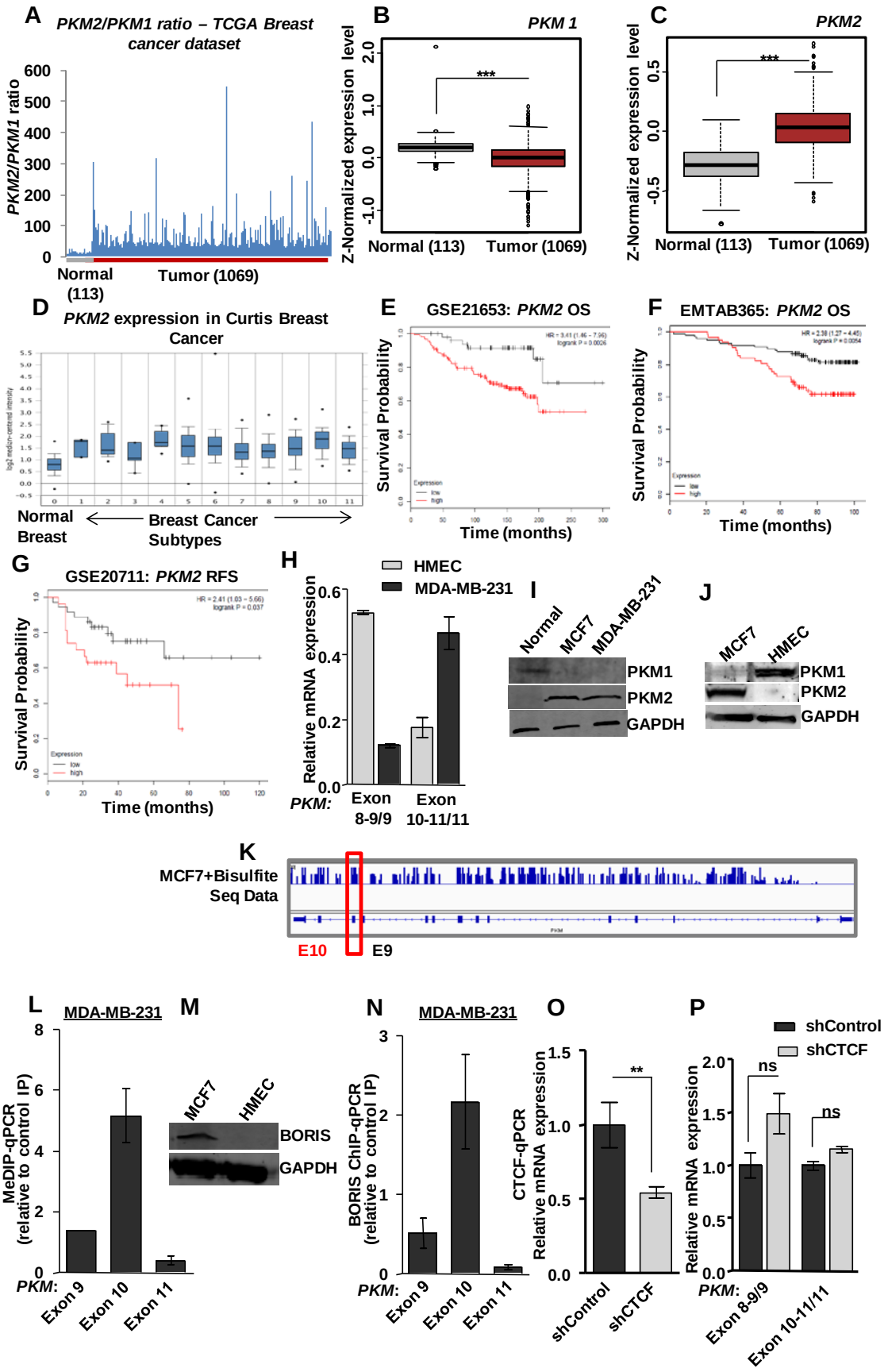

Fig. S1. Increased *PKM2* isoform expression correlates with higher DNA methylation at exon 10 and leads to poor survival outcome in breast cancer.

(A) Ratio of *PKM1* and *PKM2* isoform in normal and breast tumor profiles (TCGA RNASeqV2.level 3; n=1182), indicating higher *PKM2*/*PKM1* ratio in tumor profiles. **(B-C)** mRNA expression level of (B) *PKM1* isoform and (C) *PKM2* in normal and breast tumor profiles (Fold change: -0.76; TCGA RNASeqV2.level 3; n=1182). **(D)** *PKM2* expression in Curtis breast cancer profile analyzed in Oncomine. See *SI Appendix* Table S2. (**E-G**) Kaplan-Meier survival analysis of *PKM2* expression in breast cancer profiles, (E) GSE21653 (HR=3.41; p=0.0026) (F) EMTAB365 (HR=2.38; p=0.0054) (G), GSE20711 (HR=2.41; p=0.037). Red color represents high expression, black color represents low expression of *PKM*2 (OS=Overall Survival, RFS=Recurrence Free Survival, HR=Hazard ratio). **(H)** *RPS16* normalized qRT-PCR data in MDA-MB-231breast cancer cells and HMEC cells (Fig. 1*C*) using the indicated exon junction specific primers. **(I)** Western blot for PKM1 and PKM2 proteins expression in MCF7 and MDA-MB-231breast cancer cells. **(J)** Western blot for PKM1 and PKM2 protein expression in MCF7 and HMEC cells. GAPDH acts as a loading control. **(K)** WGBS data analysis (GSM1328112) shows higher methylation at exon 10 region compared to the neighboring exons of *PKM* gene. **(L)** Methylated DNA immunoprecipitation (MeDIP) in MDA-MB-231 genomic DNA using 5methyl-cytosine antibody followed by qPCR relative to input and mouse IgG. **(M)** Western blot for BORIS protein expression in MCF7 and HMEC cells. GAPDH acts as a loading control. **(N)** BORIS chromatin immunoprecipitation (ChIP) in MDA-MB-231 followed by qPCR relative to input and mouse IgG. **(O)** *RPS16* normalized qRT-PCR in MCF7 cells transfected with shRNA against CTCF and control shRNA to detect *CTCF* mRNA levels. **(P)** *RPS16* normalized qRT-PCR data from MCF7 cells transfected with shCTCF versus shControl transfected cells using indicated exon junction spanning primers. Three independent experiments were conducted. Graphs show mean values ± SD. *P* as calculated using two-tailed Student's t-test, * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, ns=non-significant.

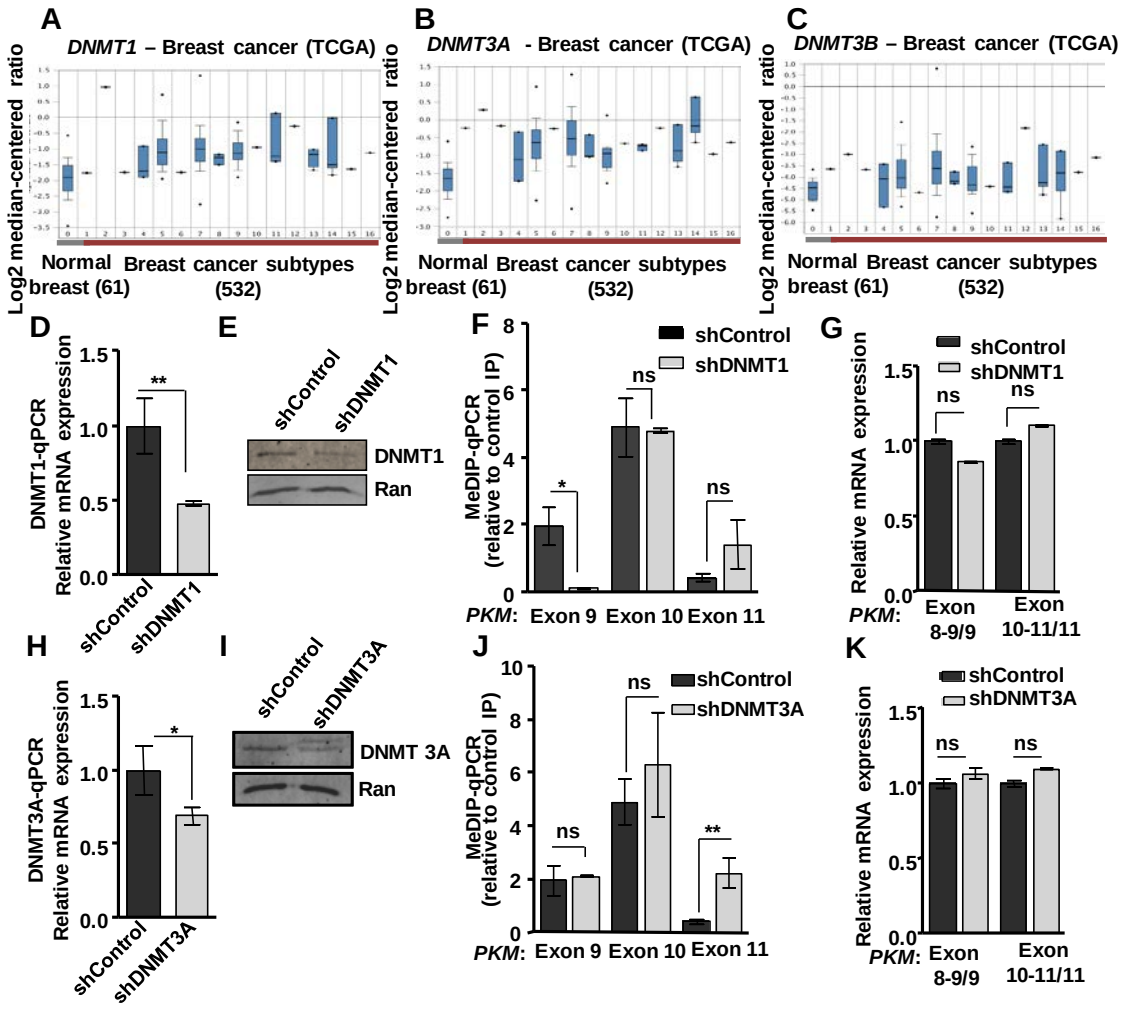

Fig. S2. *DNMT* expression in breast cancer and their role in *PKM* alternative splicing.

(A, B, C) Gene expression data analysis of DNA Methyl Transferases (*DNMT1*, *DNMT3A*, *DNMT3B*) in different breast cancer subtypes (A), (B), (C) analyzed from TCGA breast cancer profile using Oncomine. (Breast cancer subtypes, fold change, statistical significance are presented in *SI Appendix* Table S3). **(D)** *RPS16* normalized qRT-PCR in MCF7 cells transfected with shRNA against *DNMT1* and control shRNA to detect *DNMT1* mRNA levels. **(E)** Western blot for DNMT1 expression in cells from shControl and shDNMT1 transfected MCF7 cells. Ran acts as a loading control. **(F)** MeDIP and qPCR relative to input and mouse IgG in MCF7 cells transfected with shRNA against *DNMT1* versus shControl cells. **(G)** *RPS16* normalized qRT-PCR data from MCF7 cells transfected with shDNMT1 versus shControl transfected cells using indicated exon junction specific primers. **(H)** *RPS16* normalized qRT-PCR in MCF7 cells transfected with shRNA against *DNMT3A* and control shRNA to detect *DNMT3A* mRNA levels. **(I)** Western blot for DNMT3A expression in cells from shDNMT3A transfected MCF7 cells. Ran acts as a loading control. **(J)** MeDIP and qPCR relative to input and mouse IgG in MCF7 cells transfected with shRNA against DNMT3A versus shControl cells. **(K)** *RPS16* normalized qRT-PCR data from MCF7 cells transfected with shDNMT3A versus shControl transfected cells using indicated exon junction spanning primers. Three independent experiments were conducted, with the mean values ±

SD. *P* as calculated using two-tailed Student's t-test, * *P* <0.05, ** *P* <0.01, *** *P* <0.001, ns=non-significant.

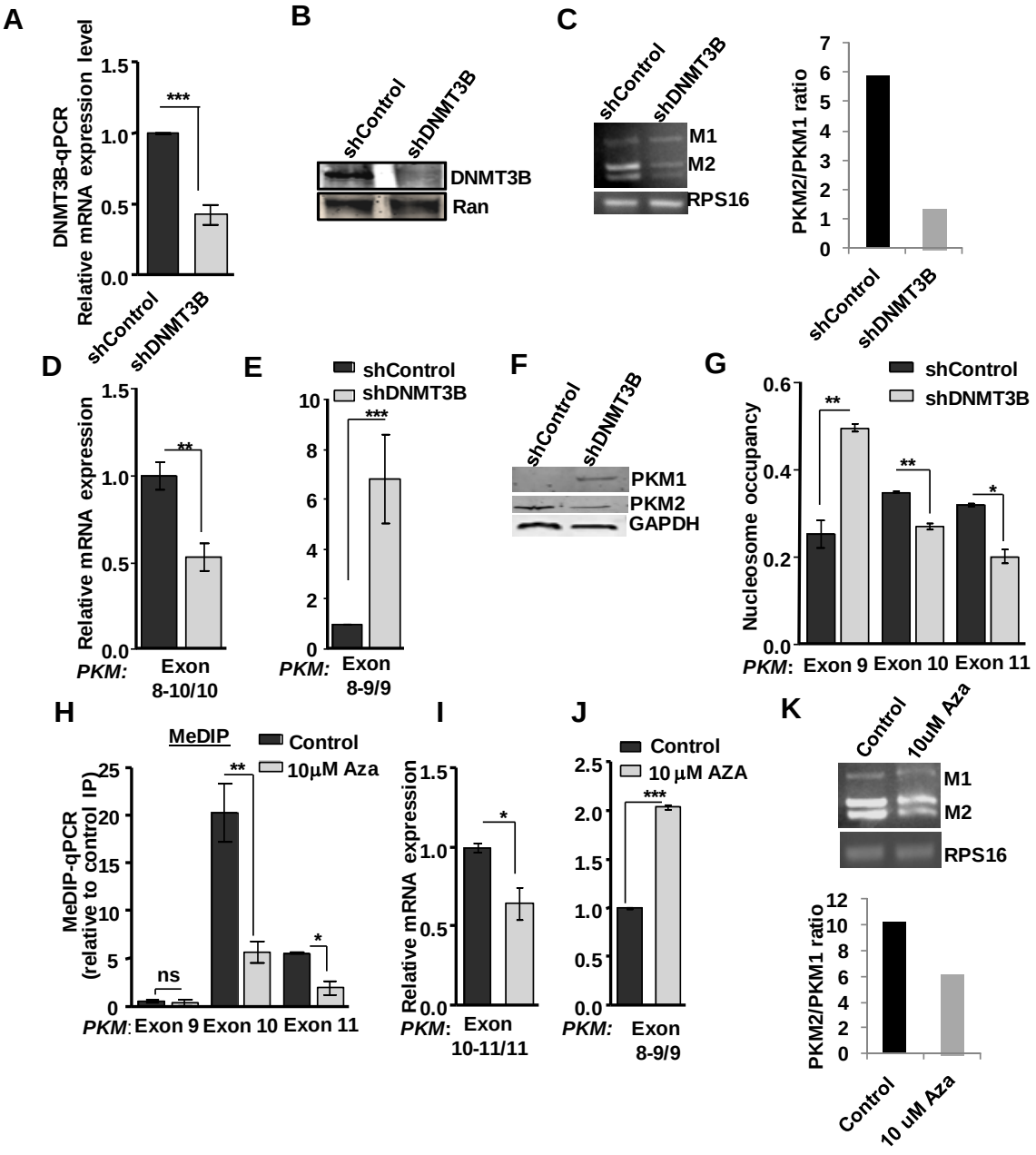

Fig. S3. Role of *DNMT* 3B in *PKM* alternative splicing.

(A) *RPS16* normalized qRT-PCR in MCF7 cells transfected with shRNA against *DNMT3B* and control shRNA to detect *DNMT3B* mRNA levels. **(B)** Western blot for DNMT3B expression in shControl and shDNMT3B transfected MCF7 cells. Ran acts as a loading control. **(C)** Semi-quantitative PCR of *PKM* gene in shControl and shDNMT3B cells were followed by PstI digestion in order to distinguish the amplicons. **(D)** *RPS16* normalized qRT-PCR from shDNMT3B versus shControl transfected MCF7cells using exon8-10 forward junction and exon 10 reverse primers in order to check the splicing of *PKM* gene. **(E)** *RPS16* normalized qRT-PCR from shDNMT3B versus shControl transfected MCF7cells using exon8-9 forward junction and exon 9 reverse primers in order to check the splicing of *PKM*

gene. **(F)** Western blot for PKM1 and PKM2 protein level expression in shDNMT3B and shControl MCF7 cells. **(G)** MNase assay was performed in DNMT3B depleted cells. Half of the sample was treated with micrococcal nuclease while the other half remains untreated. Mononucleosomal DNA was extracted from an agarose gel, and qRT-PCR was performed with exon specific primers, normalized to control IP and untreated sample (n=3). **(H)** MeDIP in MCF7 cells treated with 10µm Aza for 7 days and qRT-PCR relative to input and control IgG. **(I-J)** *RPS16* normalized qRT-PCR data from MCF7 cells after 10µm Aza treatment for 7 days using (I) exon10-11 forward junction and exon 11 reverse (J) exon8-9 forward junction and exon 9 reverse primers in order to check the splicing of *PKM* gene. **(K)** Semi-quantitative PCR of *PKM* gene in control and Aza treated MCF7 cells followed by PstI digestion in order to distinguish the amplicons. Three independent experiments were conducted, with the mean values ± SD. *P* as calculated using two-tailed Student's t-test, * *P* <0.05, ** *P* <0.01, *** *P* <0.001, ns=non-significant.

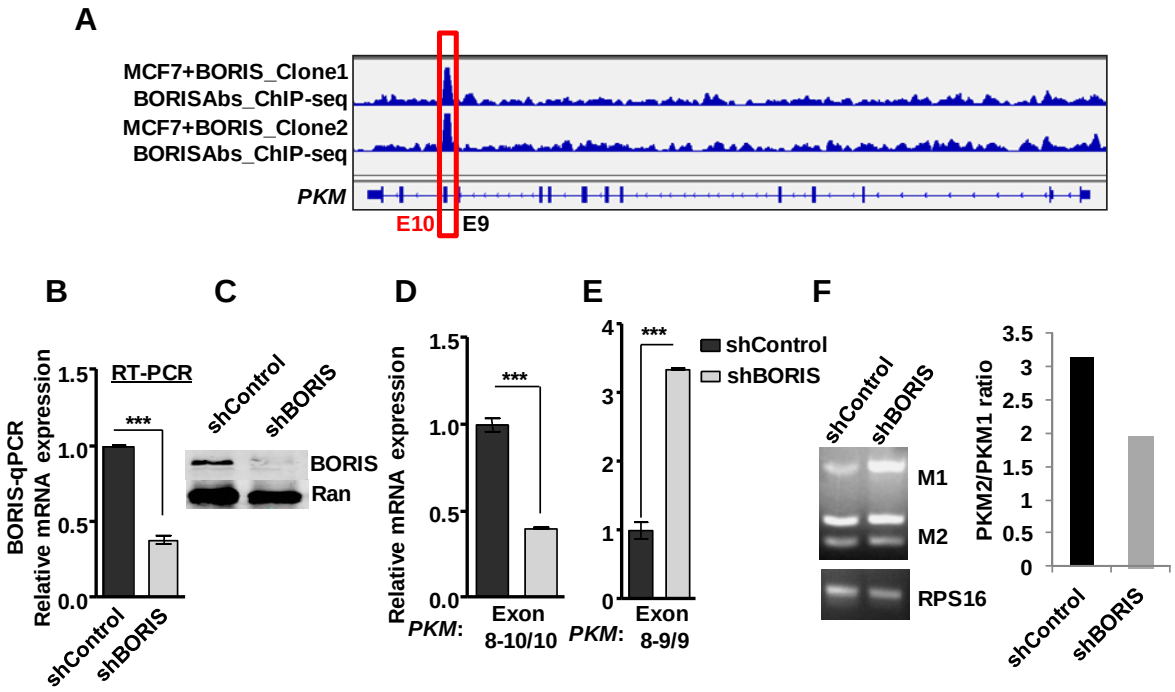

Fig. S4. *BORIS* regulates *PKM* splicing.

(A) BORIS ChIP-seq analysis (GSM1817668 and GSM1817669) in MCF7 cells to show the enrichment of BORIS at the exon 10 of *PKM* gene. Red color box indicates the enrichment of BORIS at the exon 10 over the exon 9 of *PKM* gene. **(B)** *RPS16* normalized qRT-PCR in MCF7 cells transfected with shRNA against BORIS and control shRNA to detect *BORIS* mRNA levels. **(C)** Protein expression of BORIS in shBORIS and shControl transfected MCF7 cells. Ran acts as a loading control. **(D)** *RPS16* normalized qRT-PCR from shBORIS versus shControl transfected MCF7cells using exon8-10 forward junction and exon 10 reverse primers in order to check the splicing of *PKM* gene. **(E)** *RPS16* normalized qRT-PCR from shBORIS versus shControl transfected MCF7cells using exon8-9 forward junction and exon 9 reverse primers in order to check the splicing of *PKM* gene. **(F)** Semi-quantitative PCRs of *PKM* gene in shControl and shBORIS cells were followed by PstI digestion in order

to distinguish the amplicons. Three independent experiments were conducted, and graph shows mean values ± SD. p as calculated using two-tailed Student's t-test, *P <0.05, ** P <0.01, *** P <0.001, ns=non-significant.

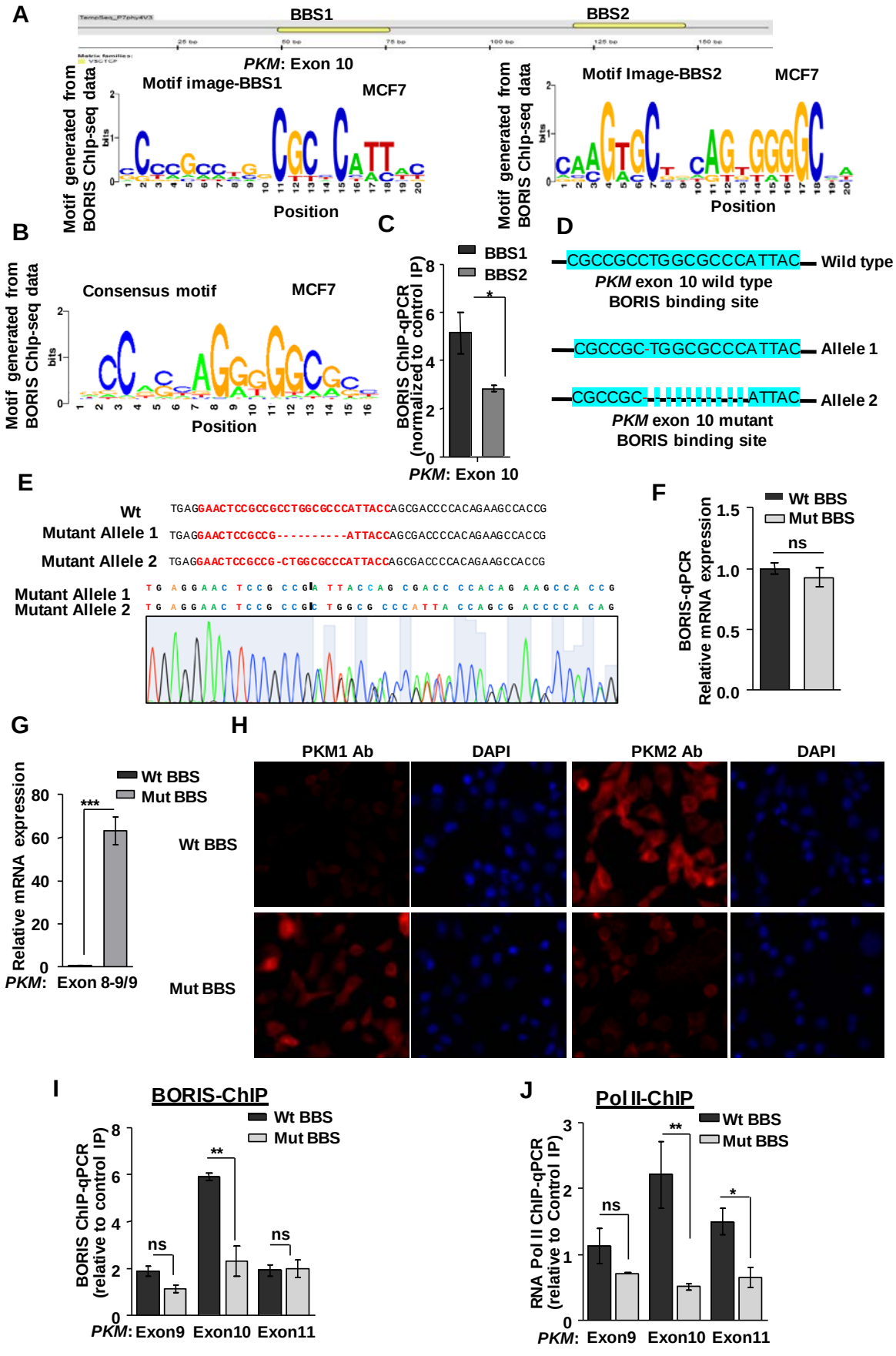
A
BBS1
BBS2
PKM: Exon 10
Motif image-BBS1
MCF7
Motif generated from BORIS ChIp-seq data
Position
Motif Image-BBS2
MCF7
Motif generated from BORIS ChIp-seq data
Position
B
Consensus motif
MCF7
Motif generated from BORIS ChIp-seq data
Position
C
BBS1
BBS2
BORIS ChIP-qPCR (normalized to control IP)
PKM: Exon 10
D
CGCCGCCTGGCGCCCATTAC
Wild type
PKM exon 10 wild type BORIS binding site
CGCCGC-TGGCGCCCATTAC
Allele 1
CGCCGC----------ATTAC
Allele 2
PKM exon 10 mutant BORIS binding site
E
Wt
TGAGGAACTCCGCCGCCTGGCGCCCATTACCAGCGACCCCACAGAAGCCACCG
Mutant Allele 1
TGAGGAACTCCGCCG----------ATTACCAGCGACCCCACAGAAGCCACCG
Mutant Allele 2
TGAGGAACTCCGCCG-CTGGCGCCCATTACCAGCGACCCCACAGAAGCCACCG
Mutant Allele 1
Mutant Allele 2
F
Wt BBS
Mut BBS
ns
BORIS-qPCR Relative mRNA expression
G
Wt BBS
Mut BBS
***
Relative mRNA expression
PKM: Exon 8-9/9
H
PKM1 Ab
DAPI
PKM2 Ab
DAPI
Wt BBS
Mut BBS
I
BORIS-ChIP
Wt BBS
Mut BBS
BORIS ChIP-qPCR (relative to control IP)
ns
**
ns
PKM: Exon9 Exon10 Exon11
J
Pol II-ChIP
Wt BBS
Mut BBS
RNA Pol II ChIP-qPCR (relative to Control IP)
ns
**
*
PKM: Exon9 Exon10 Exon11

Fig. S5. CRISPR-Cas9 based mutation of *BORIS* binding site (BBS) and its effect on *PKM* splicing.

(A) Image showing two BORIS binding sites at the exon 10 region of *PKM* gene using Genomatix software suite. The motif enrichment analysis of BORIS ChIP-seq data shows specific motifs for BBS1 (CGCCGCCTGGCGCCCATTAC) and BBS2 (TCAAGTGCTGCAGTGGGGCCATA) at *PKM* exon 10 region. **(B)** The motif enrichment analysis of BORIS ChIP-seq data showed consensus sequence motif for BORIS with 4506 hits throughout the genome. **(C)** BORIS chromatin immunoprecipitation (ChIP) in MCF7 cells followed by qPCR relative to input and mouse IgG. **(D)** Sequence of BORIS mutated clones and wild type MCF7 cells. **(E)** Top panel shows the BORIS binding site sequence of Wt allele and two mutant alleles of compound heterozygote MCF7 clone. BORIS binding site is marked by bold red letters. Dashed region marks deletion of bases. Bottom panel shows the sequencing histogram of compound heterozygote MCF7 clone consisting of different mutations in two mutant alleles. Region of deletion is marked by dashed line. **(F)** BORIS mRNA expression level in wild type and BORIS binding site mutated MCF7 cells. **(G)** *RPS16* normalized qRT-PCR data from wild-type BBS and mutated BBS in *PKM* gene using exon 9, specific junction primer in MCF7 cells. **(H)** Analysis of PKM1 and PKM2 protein level in wild type BBS and mutated BBS MCF7 cells using immunofluorescence. Representative images were acquired by ApoTome Axio microscope (scale bar 20μm). **(I-J)** ChIP in wild-type BBS and mutated BBS MCF7 cells, as in Fig. 4*C* and *D* using (I) BORIS and (J) RNA Pol II antibody, followed by qRT-PCR relative to input and control IgG. Three independent experiments were conducted. Graphs show mean values ± SD. *P* as calculated using two-tailed Student's t-test, * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, ns=non-significant.

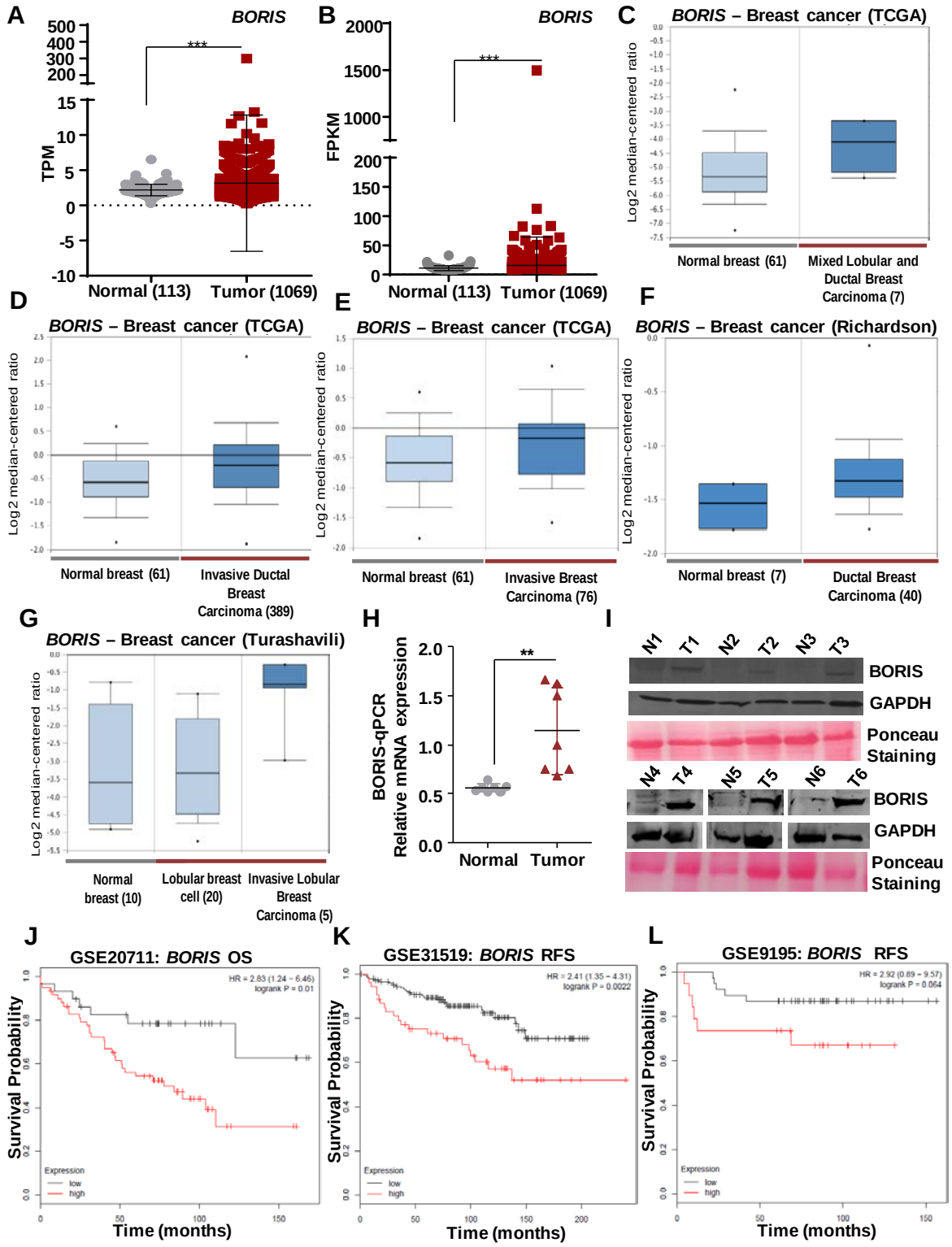

Fig. S6. *BORIS* overexpression results in poor survival outcome in breast cancer.

(A-B) *BORIS* expression in breast cancer expression profiles from TCGA (Fold change: 1.69; RNAseqV2 level3 n=1182). (A) *BORIS* expression is plotted in units of Transcript per Million reads (TPM). TPM is one of the relative quantification values to identify the relative expression of transcript levels per cell. (B) *BORIS* expression is plotted in units of Fragments Per Kilobase of transcript per Million mapped reads (FPKM). FPKM is a relative

quantification method of expression of transcript per cell. **(C-G)** *BORIS* expression in breast cancer expression profiles (C-E), TCGA data from different breast cancer subtypes analyzed in Oncomine, (F) Richardson breast cancer profile (G) Turashvili breast cancer profile, analyzed in Oncomine. (Fold change, and statistical significance are presented in *SI Appendix* Table S5). **(H)** *BORIS* relative mRNA expression, normalized to the gene *RPS16*. **(I)** Protein expression of BORIS in matched normal and breast tumor tissues. GAPDH and Ponceau staining as a loading control. (Breast cancer patient details are presented in *SI Appendix* Table S6). **(J-L)** Kaplan-Meier survival analysis of *BORIS* expression in (J) GSE20711 (HR=2.83; p=0.01), (K) GSE31519 (HR=2.41; p=0.0022) (L), GSE9195 (HR=2.92; p=0.064). Red color shows high expression of *BORIS* and black color shows low expression of *BORIS* in all the breast cancer cohorts. (OS=Overall Survival, RFS=Recurrence Free Survival, HR=Hazard ratio).

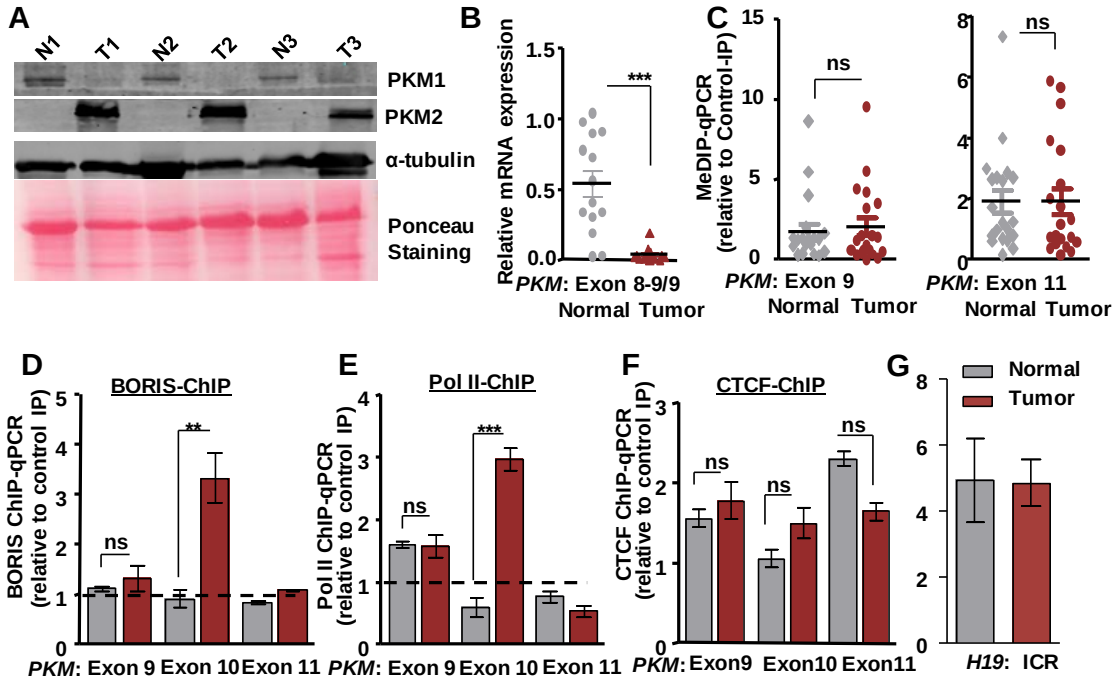

Fig. S7. Assay of PKM expression, DNA methylation and CTCF binding in breast cancer patients.

(**A)** Protein expression of PKM1 and PKM2 in normal and breast tumor tissues by western blot analysis. Tubulin and Ponceau act as a loading control. **(B)** *RPS16* normalized qRT-PCR data from normal and breast tumor tissue samples using the indicated exon junction-specific primers (n=14) (*SI Appendix* Table S6). **(C)** MeDIP in normal and breast tumor tissues genomic DNA and qPCR of *PKM* exon 9 and exon 11 region, relative to input and mouse IgG (n=20). **(D-E)** ChIP in normal and breast tumor tissues, as in Fig. 5*C* and *D*, using (D) BORIS and (E) RNA Pol II antibody and qPCR with the indicated exonic primers, relative to input and control IgG (n=3). **(F)** ChIP in normal and breast tumor tissues using CTCF antibody and qPCR with the indicated exonic primers, relative to input and control IgG (n=3). **(G)** ChIP in normal and breast tumor tissues using CTCF antibody and qPCR with H19 Imprinting control region (ICR) primer relative to input and control IgG. H19 ICR is the

positive control for the CTCF binding (2). Three independent experiments were conducted. Graphs show mean values ± SD. *P* as calculated using two-tailed Student's t-test, * *P* <0.05, ** *P* <0.01, *** *P* <0.001, ns=non-significant.

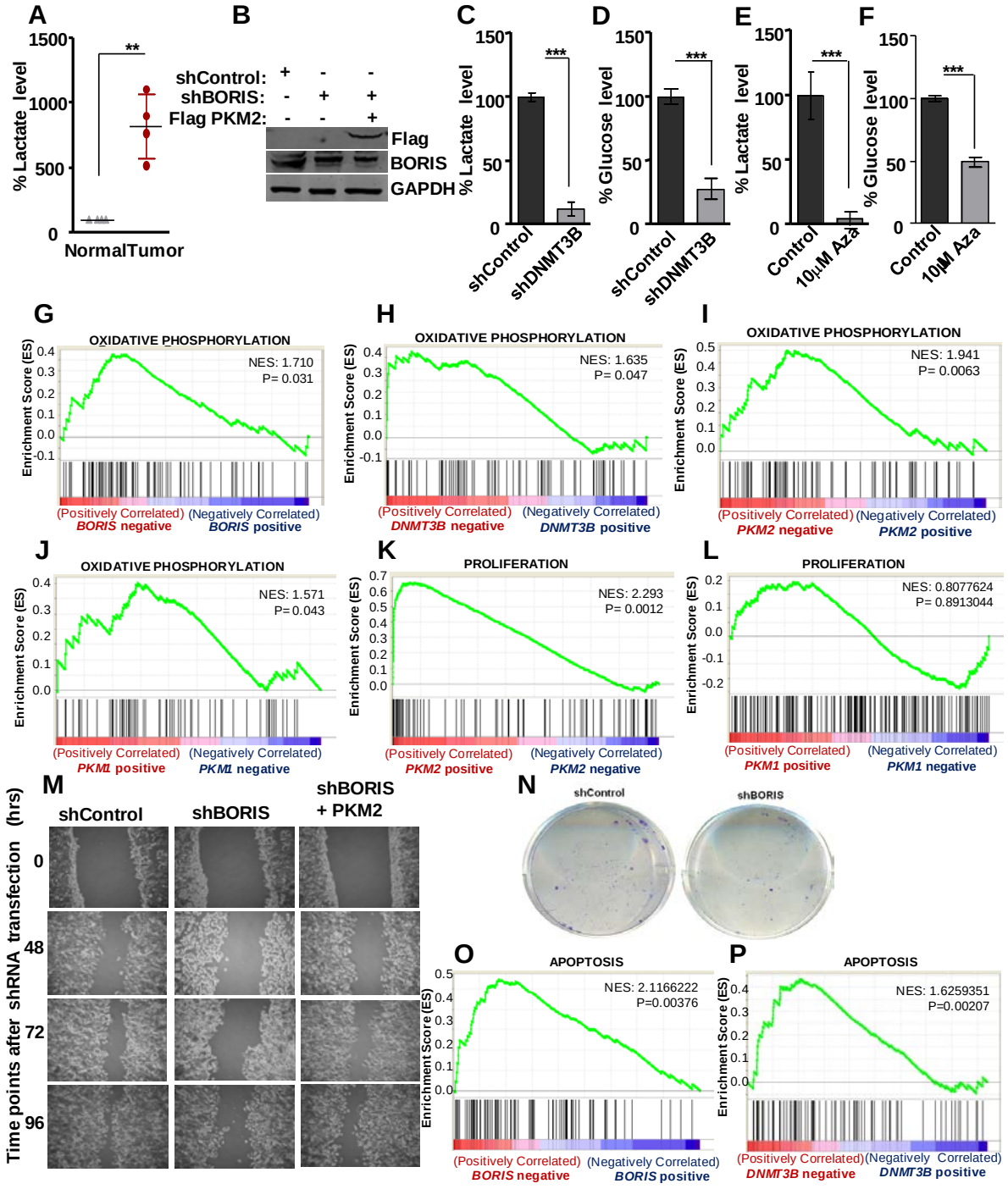

Fig. S8. Role of BORIS on proliferation and Warburg effect in breast cancer.

(A) Percentage of lactate production in normal versus tumor tissue (n=4) using lactate assay kit. **(B)** Protein expression of BORIS and Flag tagged PKM2 in shControl and shBORIS transfected MCF7 cells. GAPDH acts as a loading control. **(C-D)** Percentage of lactate production and glucose uptake after DNMT3B depletion in MCF7 cells (n=3). **(E-F)** Percentage of lactate production and glucose uptake after 10μm Aza treatment for 6 days

(n=3). Three independent experiments were conducted. Graphs show mean values ± SD. *P* as calculated using two-tailed Student's t-test, * *P* <0.05, ** *P* <0.01, *** *P* <0.001, ns=non-significant. **(G-J)** GSEA enrichment plot for oxidative phosphorylation (OXPHOS) in the comparison of (G) *BORIS*-positive versus negative (n=1000), (H) *DNMT3B*-positive versus negative (n=1000), (I) *PKM2*-positive versus negative (n=1010), (D) *PKM1*-positive versus negative (n=1010). The plot shows that OXPHOS pathway (n=135 genes) is up-regulated in *BORIS, DNMT3B* and *PKM2* negative as compared to *BORIS, DNMT3B* and *PKM2* positive breast cancer profiles, whereas is up-regulated in *PKM1* positive as compared to *PKM1* negative breast cancer profiles. Each of the black peaks represents a gene in the OXPHOS pathway. Red and blue colors represent positive and negative Pearson correlation with respective gene expression. **(K-L)** GSEA enrichment plot for Proliferation in the comparison of (K) *PKM2* positive versus negative (n=1069) (L) *PKM1* positive versus negative (n=1069) breast cancer profiles. The plot shows that proliferation pathway (n=178 genes) is significantly up-regulated only in *PKM2* positive as compared to *PKM2* negative breast cancer profiles. Each of the black peaks represents a gene in the proliferation pathway. Red and blue colors represent positive and negative Pearson correlation with respective gene expression (See *SI Appendix* Table S7). **(M)** Representative image of wound healing assay in shBORIS versus shControl transfected MCF7 cells at different time points (n=3). Overexpression of PKM2 partially rescued the wound area in shBORIS transfected cells. **(N)** Colony forming assay in shBORIS versus shControl transfected MCF7 cells. **(O-P)** GSEA enrichment plot for Apoptosis in the comparison of (O) *BORIS*-positive versus negative (n=1010) (P) *DNMT3B* positive versus negative (n=1010) in breast cancer profiles. The plot shows that Apoptosis pathway (n=88 genes) is overall more active in *BORIS, DNMT3B* negative as compared to *BORIS, DNMT3B* positive breast cancer profiles. Each of the black peaks represents a gene in the proliferation pathway. Red and blue colors represent positive and negative Pearson correlation with respective gene expression.

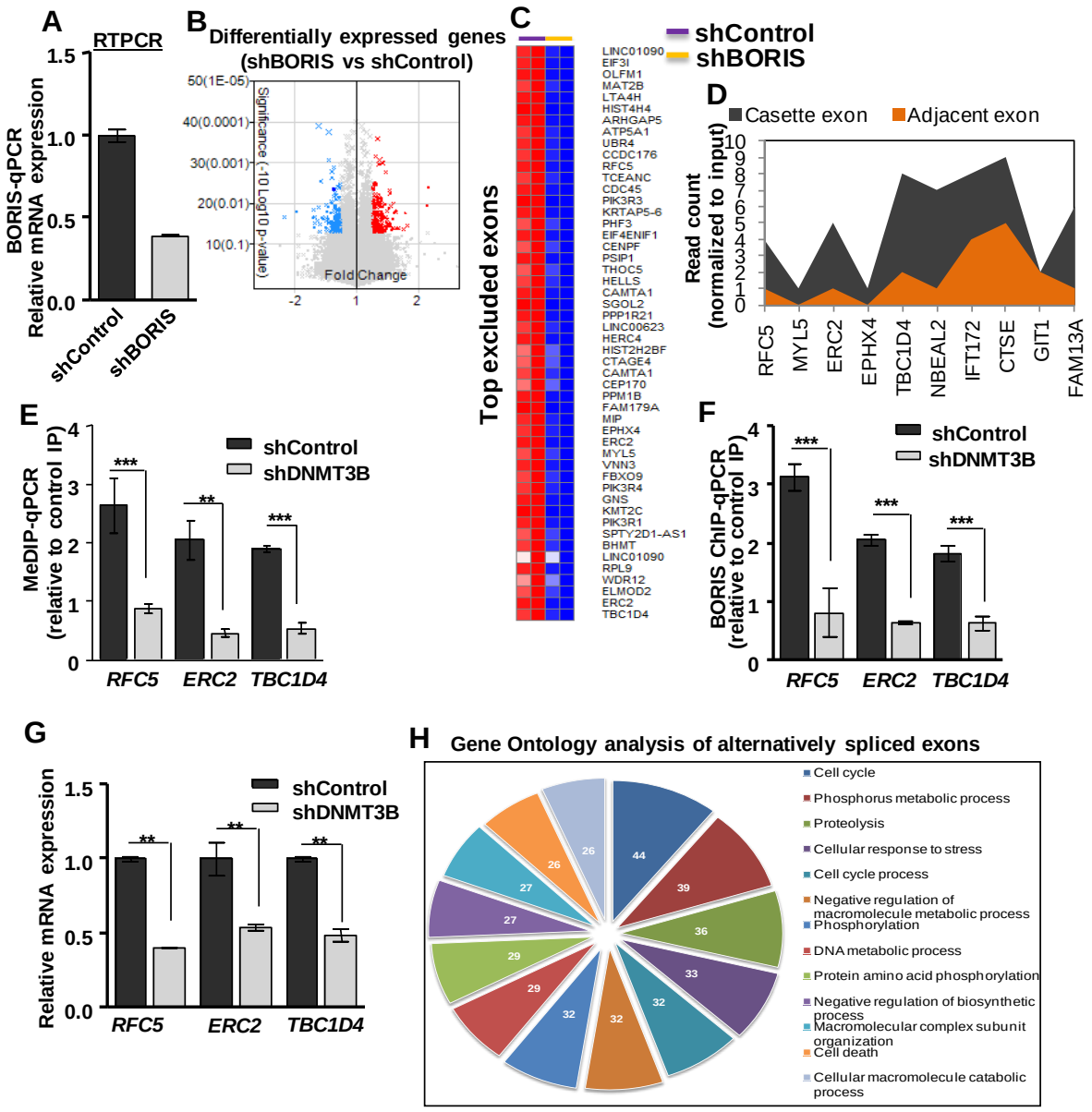

Fig. S9. Global identification of BORIS mediated alternative splicing in breast cancer cells.

(A) *RPS16* normalized qRT-PCR in MCF7 cells transfected with shRNA against BORIS and control shRNA to detect *BORIS* mRNA levels. **(B)** Scatter plot shows the RMA normalized expression levels of differentially expressed genes in shBORIS versus shControl transfected MCF7 cells. Red color dots represent upregulation, blue color dots represent down-regulation and gray color represent unaltered genes (Fold change (≥+1.2 and ≤-1.2); $P<0.05$; n=2) (See Dataset S1). **(C)** Clustering analysis of top 50 spliced events (excluded) and top 30 altered gene events (top 15 up regulated and 15 down regulated) in shBORIS versus shControl transfected MCF7 cells (n=2). **(D)** BORIS ChIP-Seq (GSE70764) analysis to show the enrichment of BORIS at the alternative exons versus adjacent exons of candidate genes. These candidate genes showed exclusion of alternative exon in shBORIS cells as compared to the shControl cells (See Dataset S1). **(E)** MeDIP qPCR (relative to input and IgG) in MCF7 cells transfected with shControl and shDNMT3B cells to detect methylation level at

alternative exons of genes from shBORIS versus shControl HTA2.0 transcriptome profile. **(F)** BORIS ChIP-qPCR in MCF7 cells transfected with shControl and shDNMT3B cells to detect BORIS binding at alternative exons of genes from shBORIS versus shControl HTA2.0 transcriptome profile. **(G)** *RPS16* normalized qRT-PCR from shDNMT3B versus shControl transfected MCF7cells using indicated exon junction specific primers to detect a change in alternative splicing of genes from shBORIS versus shControl HTA2.0 transcriptome profile (See Dataset S1). **(H)** Gene Ontology analysis of alternatively spliced genes upon BORIS depletion in MCF7 cells (See *SI Appendix* Table S8). The graph shows the top GO functions regulated in molecular and biological process category. Three independent experiments were conducted. Graphs show mean values ± SD. *P* as calculated using two-tailed Student's t-test, * *P* <0.05, ** *P* <0.01, *** *P* <0.001, ns=non-significant, RMA-robust multi average.

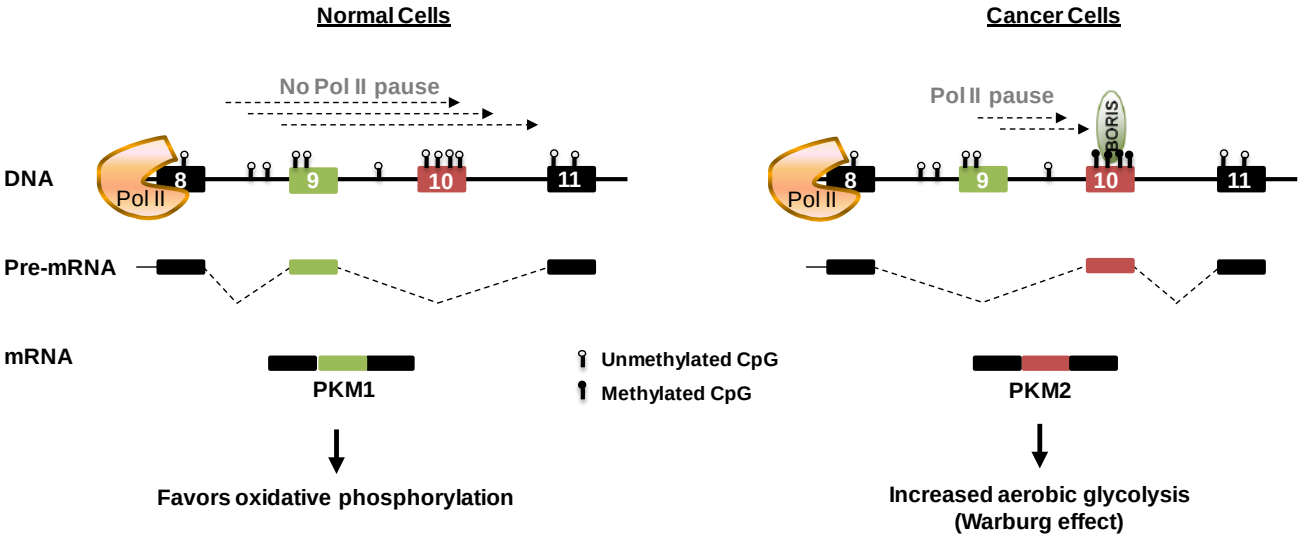

Fig. S10. Model for BORIS-mediated regulation of *PKM* alternative splicing.

In cancer cells, BORIS preferentially binds to methylated *PKM* exon 10 DNA as compared to unmethylated exon 10 in normal cells. In the absence of BORIS-mediated Pol II enrichment, exon 10 is excluded from spliced *PKM* transcripts in normal cells. BORIS binds to and promotes local Pol II enrichment at *PKM* exon 10 when the target DNA is methylated on cytosine. The Pol II enrichment thereby provides sufficient time for spliceosome assembly prior to reactivation of elongation. Consequently, exon 10 is included in the spliced *PKM* transcripts in cancer cells. The increased *PKM2* expression in cancer cells promotes Warburg effect and thereby tumorigenesis whereas increased *PKM1* expression in normal cells leads to Oxidative phosphorylation and consequently normal cell growth.

SI Appendix Tables

Table S1: List of primer sequences utilized for qPCR.

S.No.	Primers	Sequence
1	PKM Ex8 Fw	GCTGGAGAGCATGATCAAGAA
2	PKM Ex8 Rev	CTGTTTCTCCAGACAGCATGA
3	PKM E11 Fw	CCATCATTGCTGTGACCCGGAAT
4	PKM E11 Rev	CATTCATGGCAAAGTTCACCCGGA
5	PKM Ex10 Fw	TAGATTGCCCGTGAGGCAGAGGCT
6	PKM E10 Rev	TGCCAGACTTGGTGAGGACGATTA
7	PKM E8-E9 Fw	ATGCAGCACCTGATAGCTCGTGA
8	PKM E9 Fw	GTTCCACCGCAAGCTGTTTGAAGA
9	PKM E9 Rev	TGCCAGACTCCGTCAGAACTATCA
10	PKM E8-10 Fw	ATGCAGCACCTGATTGCCCGTGA
11	PKM E10-11 Fw	TCACCAAGTCTGGCAGGTCTG
12	RPS16 SET5 Fw	AAACGCGGCAATGGTCTCATCAAG
13	RPS16 SET5 Rev	TGGAGATGGACTGACGGATAGCAT
14	beta-ACTIN E3 Fw	TTCTACAATGAGCTGCGTGTGGCT
15	beta-ACTIN E3 Rev	TCATCTTCTCGCGGTTGGCCTT
16	DNMT3A EX7 Fw	GCCAAGGTCATTGCAGGAA
17	DNMT3A EX7 Rev	CGTACTCTGGCTCGTCATC
18	DNMT3B EX5 Fw	AACAGCATCGGCAGGAA
19	DNMT3B EX5 Rev	GATACTCTGAACTGTCTCCATCTC
20	DNMT1 EX4 Fw	TGCTTACAACCGGGAAGTGAATGG
21	DNMT1 EX4 Rev	TTGGCATCTGCCATTCCCACTCTA
22	BORIS Ex1 (211) Fw	GAAGAGCAGGAGAAGAA
23	BORIS Ex1 (211) Rev	CCTTTGTCTTTCTTTGATTT
24	CTCF (121) Fw	ATGTTATATTTGTCATGCTCGGTTTA
25	CTCF (121) Rev	TTTCGGGCTATGACTGTGTC
26	RFC5 E5 Fw	GTGATCTTGGATGAAGCAGAC
27	RFC5 E5 Rev	CTCTTCTCAAGGCATTCTGG
28	RFC5 E6 Fw	TAATTGAGAAATTCACAGAAAATACCA
29	RFC5 E6 Rev	TTCTCTTCTTCCACGACATGTTC
30	RFC5 Cons Fw	GGTTGAAAAATACCGGCCA
31	RFC5 Cons Rev	TGGTACTCAGAATGTCCTGATGA
32	ERC2 E15 Fw	CAGGAAGCACTACTTGCAGC
33	ERC2 E15 Rev	CTGCTGCTTTAATTGATGTACTAGTCG
34	ERC2 E16 Fw	AGAATGAAGTTGATGGCAGACA
35	ERC2 E16 Rev	CCTGTGATTGGAATGTTGCG
36	ERC2 Cons Fw	GAATTGCACCGAAGAAGCC
37	ERC2 Cons Rev	CTTCATTTCGATGACAGTCTGG
38	TBC1D4 E7 Fw	CAGTGACCAGGAAGAAAATGAAC
39	TBC1D4 E7 Rev	CACGTGTGTCTTCTGCTTGG

40	TBC1D4 E8 Fw	AATAGTACAATCCCAGAAAATGCAA
41	TBC1D4 E8 Rev	CCTTGAGAAGATATTTTCCAGGG
42	TBC1D4 Cons Fw	AGAGCCAAGCTGGTGATACAG
43	TBC1D4 Cons Rev	CTGAACTCTTTCAAAGATGTCAGC
44	RFC5 E5-6 Jn Rev	TTCTCAATTACTCTTCTCAAGGCAATT
45	ERC2 E15-16 Jn Rev	ATTCTGTTCTGGGTCTGCTGCTTTAA
46	TBC1D4 E7-8 Jn Rev	TATTTGAAATAGTAGAAGGGCCTTCC
47	H19 ICR Fw	CCCATCTTGCTGACCTCAC
48	H19 ICR Rev	AGACCTGGGACGTTTCTGTG
49	PKM2 OE Fw	TTTTGCGGCCGCATGTCACCGGAAGCCCAA
50	PKM2 OE Rev	TTTTGTCGACTCACGGCACAGGAACAAC

Table S2: ***PKM2*** **fold change in breast cancer profiles**

Annotation number	***PKM2*** **– Breast cancer (Curtis)**	**Fold change (compared to normal)**	***P*** **value**
0	Normal (61)		
1	Benign Breast Neoplasm (3)	1.695	0.035
2	Breast Carcinoma (14)	1.705	3.37E-5
3	Breast Phyllodes Tumor (5)	1.143	0.117
4	Ductal Breast Carcinoma insitu (10)	1.755	9.66E-7
5	Invasive Breast Carcinoma (21)	1.999	9.72E-6
6	Invasive Ductal Breast Carcinoma (1,556)	2.116	6.19E-73
7	Invasive Ductal and Invasive Lobular Breast Carcinoma (90)	1.646	4.84E-20
8	Invasive Lobular Breast Carcinoma (148)	1.789	4.03E-39
9	Medullary Breast Carcinoma (32)	2.392	3.84E-14
10	Mucinous Breast Carcinoma (46)	2.068	1.83E-22
11	Tubular Breast Carcinoma (67)	1.813	1.77E-20

Table S3: DNMT's fold change in breast cancer profiles

Annotation number	***DNMT*1 – Breast cancer (TCGA)**	**Fold change (compared to normal)**	***P* value**
0	Normal (61)		
1	Apocrine Breast Carcinoma (1)	ns	Ns
2	Breast Large Cell Neuroendocrine Carcinoma (1)	ns	Ns
3	Ductal Breast Carcinoma (1)	ns	Ns
4	Intraductal Cribriform Breast Adenocarcinoma (3)	1.398	0.01
5	Invasive Breast Carcinoma (76)	1.886	2.26E-16
6	Invasive Cribriform Breast Carcinoma (1)	ns	Ns
7	Invasive Ductal Breast Carcinoma (392)	1.93	7.62E-21
8	Invasive Ductal and Lobular Carcinoma (3)	1.45	3.43E-10
9	Invasive Lobular Breast Carcinoma (36)	1.63	1.84E-14
10	Invasive Papillary Breast Carcinoma (1)	ns	Ns
11	Male Breast Carcinoma (3)	2.285	6.00E-02
12	Metaplastic Breast Carcinoma (1)	ns	Ns
13	Mixed Lobular and Ductal Breast Carcinoma (7)	1.555	2.18E-05
14	Mucinous Breast Carcinoma (4)	1.967	0.047
15	Papillary Breast Carcinoma (1)	ns	Ns
16	Pleomorphic Breast Carcinoma (1)	ns	Ns
Annotation number	***DNMT*3A - Breast cancer (TCGA)**	**Fold change (compared to normal)**	***P* value**
0	Normal (61)		
1	Apocrine Breast Carcinoma (1)	ns	Ns
2	Breast Large Cell Neuroendocrine Carcinoma (1)	ns	Ns
3	Ductal Breast Carcinoma (1)	ns	Ns
4	Intraductal Cribriform Breast Adenocarcinoma (3)	1.794	0.054

5	Invasive Breast Carcinoma (76)	1.923	1.38E-23
6	Invasive Cribriform Breast Carcinoma (1)	ns	Ns
7	Invasive Ductal Breast Carcinoma (392)	2.201	1.27E-42
8	Invasive Ductal and Lobular Carcinoma (3)	1.815	0.021
9	Invasive Lobular Breast Carcinoma (36)	1.587	7.51E-13
10	Invasive Papillary Breast Carcinoma (1)	ns	Ns
11	Male Breast Carcinoma (3)	1.892	6.16E-07
12	Metaplastic Breast Carcinoma (1)	ns	Ns
13	Mixed Lobular and Ductal Breast Carcinoma (7)	1.933	4.66E-04
14	Mucinous Breast Carcinoma (4)	3.282	0.002
15	Papillary Breast Carcinoma (1)	ns	Ns
16	Pleomorphic Breast Carcinoma (1)	ns	Ns
Annotation number	***DNMT*3B – Breast cancer (TCGA)**	**Fold change (compared to normal)**	***P* value**
0	Normal (61)		
1	Apocrine Breast Carcinoma (1)	ns	Ns
2	Breast Large Cell Neuroendocrine Carcinoma (1)	ns	Ns
3	Ductal Breast Carcinoma (1)	ns	Ns
4	Intraductal Cribriform Breast Adenocarcinoma (3)	1.248	0.304
5	Invasive Breast Carcinoma (76)	1.75	1.15E-09
6	Invasive Cribriform Breast Carcinoma (1)	ns	Ns
7	Invasive Ductal Breast Carcinoma (392)	2.259	3.43E-30
8	Invasive Ductal and Lobular Carcinoma (3)	1.373	0.046
9	Invasive Lobular Breast Carcinoma (36)	1.413	4.45E-04
10	Invasive Papillary Breast Carcinoma (1)	ns	Ns
11	Male Breast Carcinoma (3)	1.314	2.17E-01

12	Metaplastic Breast Carcinoma (1)	ns	Ns
13	Mixed Lobular and Ductal Breast Carcinoma (7)	1.647	3.40E-02
14	Mucinous Breast Carcinoma (4)	1.447	0.264
15	Papillary Breast Carcinoma (1)	ns	Ns
16	Pleomorphic Breast Carcinoma (1)	ns	Ns

Table S4: List of splicing factors binding at exon 10 region of *PKM* gene.

Sequence Position	Linked SR protein	cDNA Position	Enhancer motif score
3	SC35	TGCCCGTG	81.81
3	SF2/ASF (IgM-BRCA1)	TGCCCGT	71.23
9	SRp55	TGAGGC	77.19
14	SF2/ASF (IgM-BRCA1)	CAGAGGC	82.31
14	SF2/ASF	CAGAGGC	82.18
18	SC35	GGCTGCCA	86.42
20	SF2/ASF (IgM-BRCA1)	CTGCCAT	72.92
30	SRp40	CCACTTG	84.25
43	SRp40	TTTGAGG	81.56
45	SRp55	TGAGGA	74.82
49	SC35	GAACTCCG	82.97
54	SF2/ASF (IgM-BRCA1)	CCGCCGC	75.54
55	SF2/ASF (IgM-BRCA1)	CGCCGCC	71.61
57	SF2/ASF (IgM-BRCA1)	CCGCCTG	71.23
58	SF2/ASF (IgM-BRCA1)	CGCCTGG	79.61
63	SC35	GGCGCCCA	76.77
65	SF2/ASF (IgM-BRCA1)	CGCCCAT	81.38
65	SF2/ASF	CGCCCAT	74.61
71	SRp40	TTACCAG	84.31
79	SC35	GACCCCAC	77.63
82	SF2/ASF (IgM-BRCA1)	CCCACAG	75.54
83	SRp40	CCACAGA	83.59
84	SRp55	CACAGA	75.4
84	SF2/ASF (IgM-BRCA1)	CACAGAA	82.46
84	SF2/ASF	CACAGAA	81.48
85	SRp40	ACAGAAG	80
90	SC35	AGCCACCG	82.42
92	SRp40	CCACCGC	82.34
93	SC35	CACCGCCG	75.91
95	SF2/ASF (IgM-BRCA1)	CCGCCGT	84.69
95	SF2/ASF	CCGCCGT	80.72
96	SF2/ASF (IgM-BRCA1)	CGCCGTG	80
98	SF2/ASF (IgM-BRCA1)	CCGTGGG	71.92

104	SC35	GTGCCGTG	76.21
104	SF2/ASF	GTGCCGT	75.01
107	SF2/ASF (IgM-BRCA1)	CCGTGGA	75.23
114	SC35	GGCCTCCT	77.5
120	SRp40	CTTCAAG	86.29
127	SRp55	TGCTGC	76.16
129	SF2/ASF (IgM-BRCA1)	CTGCAGT	71.92
132	SF2/ASF (IgM-BRCA1)	CAGTGGG	77.15
132	SF2/ASF	CAGTGGG	76.47
137	SC35	GGGCCATA	75.23
138	SC35	GGCCATAA	77.5
142	SRp40	ATAATCG	79.1
150	SRp40	CCTCACC	84.61
151	SF2/ASF (IgM-BRCA1)	CTCACCA	81.77
151	SF2/ASF	CTCACCA	78.51
158	SRp40	AGTCTGG	81.02
159	SC35	GTCTGGCA	76.83

Table S5: ***BORIS*** **fold change in breast cancer profiles**

BORIS **- Breast cancer (TCGA)**	**Fold change (compared to normal)**	***P*** **value**
Normal (61)		
Invasive Ductal Breast Carcinoma (389)	1.271	2.85E-05
BORIS **- Breast cancer (TCGA)**	**Fold change (compared to normal)**	***P*** **value**
Normal (61)		
Invasive Breast Carcinoma (76)	1.223	0.003
BORIS **- Breast cancer (TCGA)**	**Fold change (compared to normal)**	***P*** **value**
Normal (61)		
Mixed lobular and Ductal Breast Carcinoma (7)	1.699	0.022
BORIS **- Breast cancer (Richardson)**	**Fold change (compared to normal)**	***P*** **value**
Normal (7)		
Ductal Breast Carcinoma (40)	1.233	0.002

Table S6: Clinical characteristics of patients.

S.No.	Patient No.	Histopathology
1	Patient 1	Duct Carcinoma
2	Patient 2	Infiltrating Duct Carcinoma
3	Patient 3	Duct Carcinoma
4	Patient 4	Grade III infiltrating duct carcinoma
5	Patient 5	Grade III invasive Duct Carcinoma
6	Patient 6	Invasive Duct Carcinoma
7	Patient 7	Invasive Duct Carcinoma
8	Patient 8	Grade II Duct Carcinoma
9	Patient 9	Invasive Lobular Carcinoma
10	Patient 10	Invasive Duct Carcinoma
11	Patient 11	Infiltrating Duct Carcinoma
12	Patient 12	Duct Carcinoma Breast Nuclear Grade III
13	Patient 13	Infiltrating Duct Carcinoma
14	Patient 14	Invasive Duct Carcinoma
15	Patient 15	Infiltrating Duct Carcinoma
16	Patient 16	Infiltrating Duct Carcinoma
17	Patient 17	Invasive Duct Carcinoma
18	Patient 18	Infiltrating Duct Carcinoma Grade III T4N0
19	Patient 19	Invasive Duct Carcinoma Grade III
20	Patient 20	Infiltrating Duct Carcinoma T2N0Mx
21	Patient 21	Infiltrating Duct Carcinoma Grade II
22	Patient 22	Infiltrating Duct Carcinoma T2N0M0
23	Patient 23	Infiltrating Duct Carcinoma Grade III T3N0Mx
24	Patient 24	Infiltrating Duct Carcinoma Grade II
25	Patient 25	Infiltrating Duct Carcinoma Grade II T3N0Mx

Table S7: Genesets used for GSEA analysis

Genesets used for GSEA analysis		
KEGG_OXIDATIVE_PHOSPHORYLATION	KEGG_APOPTOSIS	CHIANG_LIVER_CANCER_SUBCLASS_PROLIFERATION_UP
> Oxidative phosphorylation	> Apoptosis	> Top 200 marker genes up-regulated in the 'proliferation' subclass of hepatocellular carcinoma (HCC); characterized by increased proliferation, high levels of serum

		AFP [GeneID=174], and chromosomal instability.
ATP12A	AIFM1	ABCC1
ATP4A	AKT1	AFP
ATP4B	AKT2	ANLN
ATP5A1	AKT3	ARHGAP18
ATP5B	APAF1	ARID3A
ATP5C1	ATM	ASPM
ATP5D	BAD	ASRGL1
ATP5E	BAX	ATP1A1
ATP5F1	BCL2	AURKA
ATP5G1	BCL2L1	AURKB
ATP5G1P5	BID	B3GNT5
ATP5G2	BIRC2	B4GALT5
ATP5G3	BIRC3	BACE2
ATP5H	CAPN1	BARD1
ATP5I	CAPN2	BCAT1
ATP5J	CASP10	BIRC5
ATP5J2	CASP3	BUB1B
ATP5L	CASP6	C11orf93
ATP5O	CASP7	C4orf7
ATP6	CASP8	CCDC109B
ATP6AP1	CASP9	CCNA2
ATP6V0A1	CFLAR	CCNB1
ATP6V0A2	CHP	CCNB2
ATP6V0A4	CHP2	CCNE1
ATP6V0B	CHUK	CD24
ATP6V0C	CSF2RB	CDC20
ATP6V0D1	CYCS	CDC6
ATP6V0D2	DFFA	CDC7
ATP6V0E1	DFFB	CDCA5
ATP6V0E2	ENDOD1	CDCA7
ATP6V1A	ENDOG	CDCA7L
ATP6V1B1	EXOG	CDK1
ATP6V1B2	FADD	CDKN3
ATP6V1C1	FAS	CENPE
ATP6V1C2	FASLG	CENPF
ATP6V1D	IKBKB	CEP55
ATP6V1E1	IKBKG	CHST11
ATP6V1E2	IL1A	CKAP2L
ATP6V1F	IL1B	CKAP4
ATP6V1G1	IL1R1	CMTM3
ATP6V1G2	IL1RAP	CSDA

ATP6V1G3	IL3	CTBP2
ATP6V1H	IL3RA	CYBA
ATP8	IRAK1	DBN1
COX1	IRAK2	DDR1
COX10	IRAK3	DEPDC1
COX11	IRAK4	DEPDC1B
COX15	LOC651610	DLGAP5
COX17	MAP3K14	DSCC1
COX2	MYD88	DTL
COX3	NFKB1	DUSP9
COX4I1	NFKBIA	E2F8
COX4I2	NGF	ECT2
COX5A	NTRK1	ELF4
COX5B	PIK3CA	ELOVL7
COX6A1	PIK3CB	ETV4
COX6A2	PIK3CD	EZH2
COX6B1	PIK3CG	FAM118A
COX6B2	PIK3R1	FAM164A
COX6C	PIK3R2	FANCI
COX6CP3	PIK3R3	FBXO5
COX7A1	PIK3R5	FEN1
COX7A2	PPP3CA	FHOD3
COX7A2L	PPP3CB	FLVCR1
COX7B	PPP3CC	FMNL2
COX7B2	PPP3R1	FOXM1
COX7C	PPP3R2	FUNDC1
COX8A	PRKACA	G6PD
COX8C	PRKACB	GALNT7
CYC1	PRKACG	GLIS2
CYTB	PRKAR1A	GPD1L
LHPP	PRKAR1B	H19
LOC100133737	PRKAR2A	HDAC2
LOC642502	PRKAR2B	HELLS
LOC644310	PRKX	HIST1H4C
LOC727947	RELA	HJURP
ND1	RIPK1	HK2
ND2	TNF	HMGB2
ND3	TNFRSF10A	HN1
ND4	TNFRSF10B	IGF2BP3
ND4L	TNFRSF10C	KIF11
ND5	TNFRSF10D	KIF14
ND6	TNFRSF1A	KIF18B

NDUFA1	TNFSF10	KIF20A
NDUFA10	TP53	KIF23
NDUFA11	TRADD	KIF2C
NDUFA2	TRAF2	KIF4A
NDUFA3	XIAP	LAMB1
NDUFA4		LDLRAD3
NDUFA4L2		LEPREL4
NDUFA5		LHFPL2
NDUFA6		LMNB1
NDUFA7		LRRC1
NDUFA8		MAD2L1
NDUFA9		MAPK13
NDUFAB1		03-Mar
NDUFB1		MARCKS
NDUFB10		MARCKSL1
NDUFB2		MCM2
NDUFB3		MECOM
NDUFB4		MEP1A
NDUFB5		MKI67
NDUFB6		MMD
NDUFB7		MMP12
NDUFB8		MMP9
NDUFB9		MTMR2
NDUFC1		NCAPG
NDUFC2		NCEH1
NDUFS1		NCK2
NDUFS2		NDC80
NDUFS3		NEK2
NDUFS4		NT5DC2
NDUFS5		NUF2
NDUFS6		NUSAP1
NDUFS7		OIP5
NDUFS8		ORC6
NDUFV1		PAFAH1B3
NDUFV2		PAG1
NDUFV3		PAPLN
PPA1		PBK
PPA2		PDE9A
SDHA		PELI1
SDHB		PIGAP1
SDHC		PKDCC
SDHD		PKM2

TCIRG1		PLBD1
UQCR10		PLP2
UQCR11		PM20D2
UQCRB		PNMA1
UQCRC1		PRC1
UQCRC2		PRKCD
UQCRFS1		PRR11
UQCRH		PTP4A3
UQCRHL		PTTG1
UQCRQ		RACGAP1
		RAD51AP1
		RFC4
		RMI2
		S100P
		SALL2
		SALL4
		SASS6
		SEL1L3
		SELM
		SGOL2
		SHCBP1
		SKA1
		SLAMF8
		SLC12A3
		SLC16A3
		SLC1A5
		SLC38A1
		SLC39A10
		SLC7A7
		SMC4
		SOX4
		SOX9
		SPHK1
		SYNJ2
		TMED3
		TMEM51
		TMEM65
		TNFRSF21
		TOP2A
		TPX2
		TRIP13
		TRNP1
		TTF2
		TTK

		TUBA4A
		UBE2C
		UGCG
		VEGFB
		WASF1
		WSB1
		ZNF532
		ZNFX1-AS1
		ZWINT

Table S8: Gene ontology analysis of alternatively spliced genes upon *BORIS* depletion

Gene Ontology analysis of alternative spliced genes				
Category	**Term**	**Count**	**%**	***P* value**
GOTERM_BP_FAT	cell cycle	44	9.4	9.80E-07
GOTERM_BP_FAT	phosphorus metabolic process	39	8.3	4.90E-03
GOTERM_BP_FAT	proteolysis	36	7.7	5.70E-02
GOTERM_BP_FAT	cellular response to stress	33	7	1.80E-05
GOTERM_BP_FAT	cell cycle process	32	6.8	4.30E-05
GOTERM_BP_FAT	negative regulation of macromolecule metabolic process	32	6.8	3.50E-03
GOTERM_BP_FAT	phosphorylation	32	6.8	1.20E-02
GOTERM_BP_FAT	DNA metabolic process	29	6.2	8.80E-05
GOTERM_BP_FAT	protein amino acid phosphorylation	29	6.2	5.80E-03
GOTERM_BP_FAT	negative regulation of biosynthetic process	27	5.7	2.90E-03
GOTERM_BP_FAT	macromolecular complex subunit organization	27	5.7	3.70E-02
GOTERM_BP_FAT	cell death	26	5.5	6.60E-02
GOTERM_BP_FAT	death	26	5.5	6.90E-02
GOTERM_BP_FAT	cellular macromolecule catabolic process	26	5.5	7.20E-02
GOTERM_BP_FAT	response to DNA damage stimulus	23	4.9	2.10E-04
GOTERM_BP_FAT	cell cycle phase	23	4.9	8.60E-04
GOTERM_BP_FAT	chromosome organization	23	4.9	6.00E-03
GOTERM_BP_FAT	negative regulation of nitrogen compound metabolic process	22	4.7	2.30E-02
GOTERM_BP_FAT	proteolysis involved in cellular protein catabolic process	22	4.7	8.30E-02
GOTERM_BP_FAT	apoptosis	22	4.7	8.50E-02
GOTERM_BP_FAT	cellular protein catabolic process	22	4.7	8.60E-02
GOTERM_BP_FAT	programmed cell death	22	4.7	9.50E-02

GOTERM_BP_FAT	negative regulation of gene expression	21	4.5	3.20E-02
GOTERM_BP_FAT	negative regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	21	4.5	3.60E-02
GOTERM_BP_FAT	modification-dependent macromolecule catabolic process	21	4.5	9.20E-02
GOTERM_BP_FAT	modification-dependent protein catabolic process	21	4.5	9.20E-02
GOTERM_BP_FAT	M phase	20	4.3	7.30E-04
GOTERM_BP_FAT	cell projection organization	20	4.3	2.60E-03
GOTERM_BP_FAT	regulation of phosphate metabolic process	20	4.3	3.90E-02
GOTERM_BP_FAT	regulation of phosphorus metabolic process	20	4.3	3.90E-02
GOTERM_BP_FAT	mitotic cell cycle	19	4	6.20E-03
GOTERM_BP_FAT	cellular component morphogenesis	19	4	1.20E-02
GOTERM_BP_FAT	cytoskeleton organization	19	4	2.90E-02
GOTERM_BP_FAT	cell proliferation	19	4	2.90E-02
GOTERM_BP_FAT	protein complex biogenesis	19	4	9.00E-02
GOTERM_BP_FAT	protein complex assembly	19	4	9.00E-02
GOTERM_BP_FAT	DNA repair	18	3.8	9.10E-04
GOTERM_BP_FAT	cell morphogenesis	18	3.8	9.20E-03
GOTERM_BP_FAT	negative regulation of transcription	18	3.8	7.50E-02
GOTERM_BP_FAT	regulation of phosphorylation	18	3.8	8.30E-02
GOTERM_BP_FAT	cell division	17	3.6	3.50E-03
GOTERM_BP_FAT	response to abiotic stimulus	17	3.6	2.50E-02
GOTERM_BP_FAT	regulation of cell cycle	16	3.4	2.20E-02
GOTERM_BP_FAT	DNA replication	15	3.2	3.40E-04
GOTERM_BP_FAT	skeletal system development	15	3.2	3.20E-02
GOTERM_BP_FAT	regulation of protein kinase activity	15	3.2	5.60E-02
GOTERM_BP_FAT	regulation of kinase activity	15	3.2	6.90E-02
GOTERM_BP_FAT	protein kinase cascade	15	3.2	8.80E-02
GOTERM_BP_FAT	regulation of transferase activity	15	3.2	9.20E-02
GOTERM_BP_FAT	positive regulation of developmental process	14	3	2.50E-02
GOTERM_BP_FAT	neuron development	14	3	9.10E-02
GOTERM_BP_FAT	regulation of locomotion	13	2.8	3.60E-03
GOTERM_BP_FAT	mitosis	13	2.8	1.00E-02
GOTERM_BP_FAT	nuclear division	13	2.8	1.00E-02
GOTERM_BP_FAT	M phase of mitotic cell cycle	13	2.8	1.20E-02
GOTERM_BP_FAT	organelle fission	13	2.8	1.40E-02

GOTERM_BP_FAT	cell projection morphogenesis	13	2.8	2.20E-02
GOTERM_BP_FAT	cell part morphogenesis	13	2.8	3.00E-02
GOTERM_BP_FAT	MAPKKK cascade	12	2.6	7.20E-03
GOTERM_BP_FAT	regulation of cell motion	12	2.6	1.00E-02
GOTERM_BP_FAT	neuron projection development	12	2.6	6.10E-02
GOTERM_BP_FAT	regulation of cell adhesion	11	2.3	2.50E-03
GOTERM_BP_FAT	regulation of cell migration	11	2.3	1.10E-02
GOTERM_BP_FAT	positive regulation of protein kinase activity	11	2.3	5.70E-02
GOTERM_BP_FAT	positive regulation of cell differentiation	11	2.3	6.60E-02
GOTERM_BP_FAT	positive regulation of kinase activity	11	2.3	6.90E-02
GOTERM_BP_FAT	positive regulation of transferase activity	11	2.3	8.40E-02
GOTERM_BP_FAT	chromosome segregation	10	2.1	2.00E-04
GOTERM_BP_FAT	cell projection assembly	10	2.1	2.40E-04
GOTERM_BP_FAT	positive regulation of cellular component organization	10	2.1	4.10E-02
GOTERM_BP_FAT	posttranscriptional regulation of gene expression	10	2.1	8.70E-02
GOTERM_BP_FAT	negative regulation of cell differentiation	10	2.1	9.80E-02
GOTERM_BP_FAT	regulation of organelle organization	10	2.1	1.00E-01
GOTERM_BP_FAT	regulation of cell morphogenesis	9	1.9	1.80E-02
GOTERM_BP_FAT	negative regulation of cellular protein metabolic process	9	1.9	8.70E-02
GOTERM_BP_FAT	cell cycle checkpoint	8	1.7	8.20E-03
GOTERM_BP_FAT	regulation of cellular component biogenesis	8	1.7	6.90E-02
GOTERM_BP_FAT	dephosphorylation	8	1.7	9.60E-02
GOTERM_BP_FAT	negative regulation of locomotion	7	1.5	4.30E-03
GOTERM_BP_FAT	integrin-mediated signaling pathway	7	1.5	8.40E-03
GOTERM_BP_FAT	regulation of ossification	7	1.5	1.40E-02
GOTERM_BP_FAT	DNA damage response, signal transduction	7	1.5	1.60E-02
GOTERM_BP_FAT	positive regulation of cell migration	7	1.5	2.50E-02
GOTERM_BP_FAT	positive regulation of cell motion	7	1.5	3.80E-02
GOTERM_BP_FAT	positive regulation of locomotion	7	1.5	3.80E-02
GOTERM_BP_FAT	cell cycle arrest	7	1.5	4.60E-02
GOTERM_BP_FAT	one-carbon metabolic process	7	1.5	6.50E-02
GOTERM_BP_FAT	regulation of DNA metabolic process	7	1.5	6.90E-02

GOTERM_BP_FAT	ossification	7	1.5	7.10E-02
GOTERM_BP_FAT	protein ubiquitination	7	1.5	8.10E-02
GOTERM_BP_FAT	bone development	7	1.5	9.20E-02
GOTERM_BP_FAT	DNA integrity checkpoint	6	1.3	9.90E-03
GOTERM_BP_FAT	negative regulation of cell migration	6	1.3	1.40E-02
GOTERM_BP_FAT	negative regulation of cell motion	6	1.3	2.10E-02
GOTERM_BP_FAT	regulation of action potential	6	1.3	2.90E-02
GOTERM_BP_FAT	regulation of gene expression, epigenetic	6	1.3	4.50E-02
GOTERM_BP_FAT	regulation of cell-matrix adhesion	5	1.1	4.40E-03
GOTERM_BP_FAT	mitotic sister chromatid segregation	5	1.1	1.20E-02
GOTERM_BP_FAT	sister chromatid segregation	5	1.1	1.40E-02
GOTERM_BP_FAT	determination of left/right symmetry	5	1.1	2.10E-02
GOTERM_BP_FAT	osteoblast differentiation	5	1.1	2.10E-02
GOTERM_BP_FAT	determination of symmetry	5	1.1	2.30E-02
GOTERM_BP_FAT	determination of bilateral symmetry	5	1.1	2.30E-02
GOTERM_BP_FAT	regulation of cell-substrate adhesion	5	1.1	2.80E-02
GOTERM_BP_FAT	phagocytosis	5	1.1	3.30E-02
GOTERM_BP_FAT	DNA damage checkpoint	5	1.1	3.30E-02
GOTERM_BP_FAT	sulfur compound biosynthetic process	5	1.1	3.70E-02
GOTERM_BP_FAT	regulation of action potential in neuron	5	1.1	4.70E-02
GOTERM_BP_FAT	regulation of cyclin-dependent protein kinase activity	5	1.1	4.70E-02
GOTERM_BP_FAT	nucleotide-excision repair	5	1.1	5.00E-02
GOTERM_BP_FAT	gene silencing	5	1.1	5.30E-02
GOTERM_BP_FAT	DNA-dependent DNA replication	5	1.1	5.90E-02
GOTERM_BP_FAT	response to UV	5	1.1	6.20E-02
GOTERM_BP_FAT	JNK cascade	5	1.1	6.20E-02
GOTERM_BP_FAT	positive regulation of cell adhesion	5	1.1	6.50E-02
GOTERM_BP_FAT	double-strand break repair	5	1.1	7.10E-02
GOTERM_BP_FAT	regulation of DNA replication	5	1.1	7.10E-02
GOTERM_BP_FAT	stress-activated protein kinase signaling pathway	5	1.1	7.50E-02
GOTERM_BP_FAT	cellular response to extracellular stimulus	5	1.1	7.80E-02
GOTERM_BP_FAT	lipid modification	5	1.1	9.70E-02

GOTERM_BP_FAT	sister chromatid cohesion	4	0.9	2.30E-03
GOTERM_BP_FAT	gene silencing by miRNA	4	0.9	8.40E-03
GOTERM_BP_FAT	posttranscriptional gene silencing	4	0.9	1.50E-02
GOTERM_BP_FAT	posttranscriptional gene silencing by RNA	4	0.9	1.50E-02
GOTERM_BP_FAT	protein polyubiquitination	4	0.9	2.70E-02
GOTERM_BP_FAT	cilium assembly	4	0.9	3.30E-02
GOTERM_BP_FAT	regulation of bone mineralization	4	0.9	3.90E-02
GOTERM_BP_FAT	regulation of biomineral formation	4	0.9	4.60E-02
GOTERM_BP_FAT	cilium morphogenesis	4	0.9	5.00E-02
GOTERM_BP_FAT	gene silencing by RNA	4	0.9	5.00E-02
GOTERM_BP_FAT	insulin receptor signaling pathway	4	0.9	6.60E-02
GOTERM_BP_FAT	regulation of anti-apoptosis	4	0.9	7.10E-02
GOTERM_BP_FAT	axon ensheathment	4	0.9	8.90E-02
GOTERM_BP_FAT	ensheathment of neurons	4	0.9	8.90E-02
GOTERM_BP_FAT	regulation of osteoblast differentiation	4	0.9	9.40E-02
GOTERM_BP_FAT	mitotic cell cycle checkpoint	4	0.9	9.40E-02
GOTERM_BP_FAT	negative regulation of cell adhesion	4	0.9	9.40E-02
GOTERM_BP_FAT	apoptotic cell clearance	3	0.6	8.90E-03
GOTERM_BP_FAT	substrate adhesion-dependent cell spreading	3	0.6	1.20E-02
GOTERM_BP_FAT	regulation of DNA replication initiation	3	0.6	2.00E-02
GOTERM_BP_FAT	gene silencing by miRNA, production of miRNAs	3	0.6	3.60E-02
GOTERM_BP_FAT	dsRNA fragmentation	3	0.6	4.10E-02
GOTERM_BP_FAT	DNA unwinding during replication	3	0.6	4.70E-02
GOTERM_BP_FAT	positive regulation of actin filament bundle formation	3	0.6	4.70E-02
GOTERM_BP_FAT	DNA replication initiation	3	0.6	6.10E-02
GOTERM_BP_FAT	peptide biosynthetic process	3	0.6	6.80E-02
GOTERM_BP_FAT	DNA geometric change	3	0.6	7.50E-02
GOTERM_BP_FAT	negative regulation of endothelial cell proliferation	3	0.6	7.50E-02
GOTERM_BP_FAT	DNA duplex unwinding	3	0.6	7.50E-02
GOTERM_BP_FAT	regulation of stress fiber formation	3	0.6	8.20E-02
GOTERM_BP_FAT	virus-host interaction	3	0.6	9.00E-02
GOTERM_BP_FAT	regulation of interleukin-1 beta production	3	0.6	9.00E-02
GOTERM_BP_FAT	regulation of actin filament bundle formation	3	0.6	9.80E-02

GOTERM_BP_FAT	photoreceptor cell maintenance	3	0.6	9.80E-02
GOTERM_BP_FAT	maintenance of sister chromatid cohesion	2	0.4	7.40E-02
GOTERM_BP_FAT	maintenance of mitotic sister chromatid cohesion	2	0.4	7.40E-02
GOTERM_BP_FAT	regulation of transcriptional preinitiation complex assembly	2	0.4	9.70E-02
GOTERM_BP_FAT	negative regulation of smooth muscle cell migration	2	0.4	9.70E-02
GOTERM_BP_FAT	osteoblast fate commitment	2	0.4	9.70E-02

Dataset S1: List of alternatively spliced events upon BORIS knockdown in MCF7 cells

The excel file includes alternative splicing events and differentially expressed genes in shControl vs. shBORIS transfected MCF7 cells identified from Affymetrix HTA 2.0 array data using Transcriptome analysis console (Affymetrix). Algorithm options and default filtering criteria used for the identification of events are included in the file.