Supplemental Tables and Figures

Table S1. Summary of HOTAIR-associated proteins identified by mass spectrometry. HOTAIR RNA was incubated with heavy nuclear extracts and anti-Luc was incubated with light nuclear extracts. Proteins enriched in the HOTAIR sample (defined as a heavy:light ratio greater than 1.5) are listed below for three SILAC experiments performed with HeLa and MDA-MB-231 nuclear extracts. Data are filtered to include only those proteins detected with greater than 2 peptides, and with variability in the heavy-light ratio < 100%.

HeLa Biological Replicate 1					
Protein Names	**Gene Name**	**Ratio H/L**	**Ratio H/L Normalized**	**Ratio H/L Variability [%]**	**Ratio H/L Count**
70 kDa subunit of Ku antigen	G22P1;XRCC6	7.3021	**6.8781**	21.774	3
Helix-destabilizing protein;Heterogeneous nuclear ribonucleoprotein A1	HNRNPA1;HNRPA1;HNRNPA1L;HNRNPA1L2	9.5458	**6.7093**	16.121	8
Heterogeneous nuclear ribonucleoprotein A3	HNRNPA3;HNRPA3	6.17	**5.6672**	23.293	4
Heterogeneous nuclear ribonucleoproteins A2/B1	HNRNPA2B1;HNRPA2B1	8	**5.606**	21.23	6
Heterogeneous nuclear ribonucleoprotein L protein	HNRNPL;HNRPL	5.0307	**3.6248**	8.1323	9
Heterogeneous nuclear ribonucleoprotein A0	HNRNPA0;HNRPA0	4.1232	**3.3561**	9.5231	2
Heterogeneous nuclear ribonucleoprotein M	HNRNPM;HNRPM;NAGR1;ORF	2.3986	**1.9798**	80.249	7
AU-rich element RNA-binding factor;Heterogeneous nuclear ribonucleoprotein D-like	HNRPDL;JKTBP	2.2996	**1.6112**	41.978	3
Heterogeneous nuclear ribonucleoprotein U	HNRNPU;HNRPU;SAFA;	1.9583	**1.5162**	25.394	16
HeLa Biological Replicate 2					
Helix-destabilizing protein;Heterogeneous nuclear ribonucleoprotein A1protein	HNRNPA1;HNRPA1;HNRNPA1L;HNRNPA1L2	6.8975	**10.913**	16.812	7
70 kDa subunit of Ku antigen	G22P1;XRCC6	7.3942	**10.372**	78.343	2
Heterogeneous nuclear ribonucleoproteins A2/B1	HNRNPA2B1;HNRPA2B1	4.7424	**6.287**	15.337	8
Heterogeneous nuclear ribonucleoprotein L	HNRNPL;HNRPL;P/OKcl.14	4.7268	**6.2504**	4.7842	8
Heterogeneous nuclear ribonucleoprotein A3	HNRNPA3;HNRPA3	4.1527	**5.5332**	13.97	6
Heterogeneous nuclear ribonucleoprotein U	HNRNPU;HNRPU;SAFA	2.5149	**3.3232**	19.865	9
APOBEC1-binding protein 1;Heterogeneous nuclear ribonucleoprotein A/B	ABBP1;HNRNPAB;HNRPAB	2.1145	**2.7897**	2.9725	2
Heterogeneous nuclear ribonucleoprotein R	HNRNPR;HNRPR	1.366	**1.9118**	15.199	6
Myb-binding protein 1A	MYBBP1A;P160	1.1714	**1.7825**	10.332	7

CBF5 homolog;Dyskerin;H/ACA ribonucleoprotein complex subunit 4	DKC1;NOLA4	1.1863	**1.5983**	23.639	3
H/ACA ribonucleoprotein complex subunit 1;Nucleolar protein family A member 1	GAR1;NOLA1	1.1433	**1.5628**	11.69	2
MDA-MB-231					
Helix-destabilizing protein;Heterogeneous nuclear ribonucleoprotein A1	HNRNPA1;HNRPA1;HNRNPA1L;HNRNPA1L2	6.2277	**4.3151**	24.609	7
Heterogeneous nuclear ribonucleoproteins A2/B1	HNRNPA2B1;HNRPA2B1	5.5376	**3.9475**	17.297	6
Heterogeneous nuclear ribonucleoprotein U;	HNRNPU;HNRPU;SAFA;	3.2451	**2.3364**	11.947	6
Heterogeneous nuclear ribonucleoprotein H	HNRNPH1;HNRPH	2.4091	**1.8741**	23.087	2
AU-rich element RNA-binding factor;Heterogeneous nuclear ribonucleoprotein D-like	HNRPDL;JKTBP	2.4225	**1.6819**	4.9353	3

Table S3. Primers used in experiments.

Experiment	Primer Sequence
ChIP Primers	
JAM2 F ChIP	ACCTGACTTCCAGCACGAGT
JAM2 R ChIP	CCAACTCCTTTCTTCCCCTC
PCDH10 F ChIP	ACCAGGCTCTGTTCTGTTCG
PCDH10 R ChIP	TCTTGGGTCATAGGGGTCTG
HOXD10 F ChIP	GCTGAGGCGCTTTAATGAAC
HOXD10 R ChIP	GGTCCCAGAAACTCTGACCA
GAPDH F ChIP	TACTAGCGGTTTTACGGGCG
GAPDH R ChIP	TCGAACAGGAGGAGCAGAGAGCGA
For shRNA construction	
sh hnRNPA1B1 F	CCGGCAGAAATACCATACCATCAATCTCGAGATTGATGGTATGGTATTT CTGTTTTTG
sh hnRNPA1B1 R	AATTCAAAAACAGAAATACCATACCATCAAT CTCGAGATTGATGGTATGGTATTTCTG
sh hnRNPA1B1 F #2	CCGGGCTTCTTCCTATTTGCCATGGCTCGAGCCATGGCAAATAGGAAGA AGCTTTTTG
sh hnRNPA1B1 R #2	AATTCAAAAA GCTTCTTCCTATTTGCCATGG CTCGAG CCATGGCAAATAGGAAGAAGC
For RNA and hnRNPB1 pulldowns	
U1 F	ATACTTACCTGGCAGGGGAG
U1 R	CAGGGGGAA AGC GCG AAC GCA
7SL F	ATCGGGTGTCCGCACTAAGTT
7SL R	CAGCACGGGAGTTTTGACCT
PCDHB5 F	AGGTGTGTTTGACCGGAGAC
PCDHB5 R	TCCCTATTTCTTCACCAGCG
PCDH10 F	CCCGTCTACACTGTGTCCCT
PCDH10 R	GGAGTACACGACCTCACCGT
JAM2 F	CACCAACAGCTCATACACAAT
JAM2 R	GACACCTGCGATATCCAACA
HOTAIR F	AGACCCTCAGGTCCCTAATATC
HOTAIR R	CCCTACTGCAGGCTTCTAAATC
anti-luciferase F	CGGTACTTCGTCCACAAACA
anti-luciferase R	CGGAAAGACGATGACGGAAA
SDHA F	TGGGAACAAGAGGGCATCTG
SDHA R	CCACCACTGCATCAAATTCATG
HPRT F	TGACACTGGCAAAACAATGCA
HPRT R	GGTCCTTTTCACCAGCAAGCT
5.8S F	GGTGGATCACTCGGCTCGT
5.8S R	GCAAGTGCGTTCGAAGTGTC

FOXB1 F	CAAGCAGAGATAACTGGCTACA
FOXB1 R	CAGCACAACATGGCAAGAAC
HSF5 F	CTTATCGATCAGCCGCTCTT
HSF5 R	TGGTGAAGCTGGTGGTTT
ACTB F	ACCATGGATGATGATATCGCC
ACTB R	GCCTTGCACATGCCGG
GAPDH F	CCGGGAAACTGTGGCGTGATGG
GAPDH R	AGGTGGAGGAGTGGGTGTCGCTGTT
JAM2 F	GACGTCGGCACGAGGTTTAG (primer for RNA:RNA interaction exp.)
JAM2 R	CGGGAAAGGAGCAGAGAGTC (primer for RNA:RNA interaction exp.)
GFP F	GAACCGCATCGAGCTGAA
GFP R	TGCTTGTCGGCCATGATATAG
For generation of T7 in vitro transcripts	
HOTAIR	F: GCGTAATACGACTCACTATAGGGGACTCGCCTGTGCTCTGGAGC
	R: CGTCATCGCTGAATACAGTTACATTGGATCC
Anti-HOTAIR	F: GCGTAATACGACTCACTATAGGGGAAAAT GCATCCAGATATTAATATATCTACATTTTAA
	R: GTAAAGTCGACACCGGGGCCCAGATC
Anti-Luc	F: GCGTAATACGACTCACTATAGGG TTACACGGCGATCTTTCCGC
	R: CGTCATCGCTGAATACAGTTACATTGGATCC
HOTAIR 361-2148	F: GCGTAATACGACTCACTATAGGGCGTCCTGGC AGAGAAAAGGCTGAA
	R: CGTCATCGCTGAATACAGTTACATTGGATCC
HOTAIR 1-360	F: GCGTAATACGACTCACTATAGGGGACTCGCCTGTGCTCTGGAGC
	R: CGTCATCGCTGAATACAGTTACATTGGATCC
JAM2	F: GCGTAATACGACTCACTATAGGG AAGCTTTAAGGAGTTTATATGGAAACC
	R: GGGGCGGGAAAGGAGCAGAGAGTC
GFP	F: GCGTAATACGACTCACTATAGGGATGGTGAGCAAGGGCGAGGAGC
	R: CTTGTACAGCTCGTCCATGCCGAGA

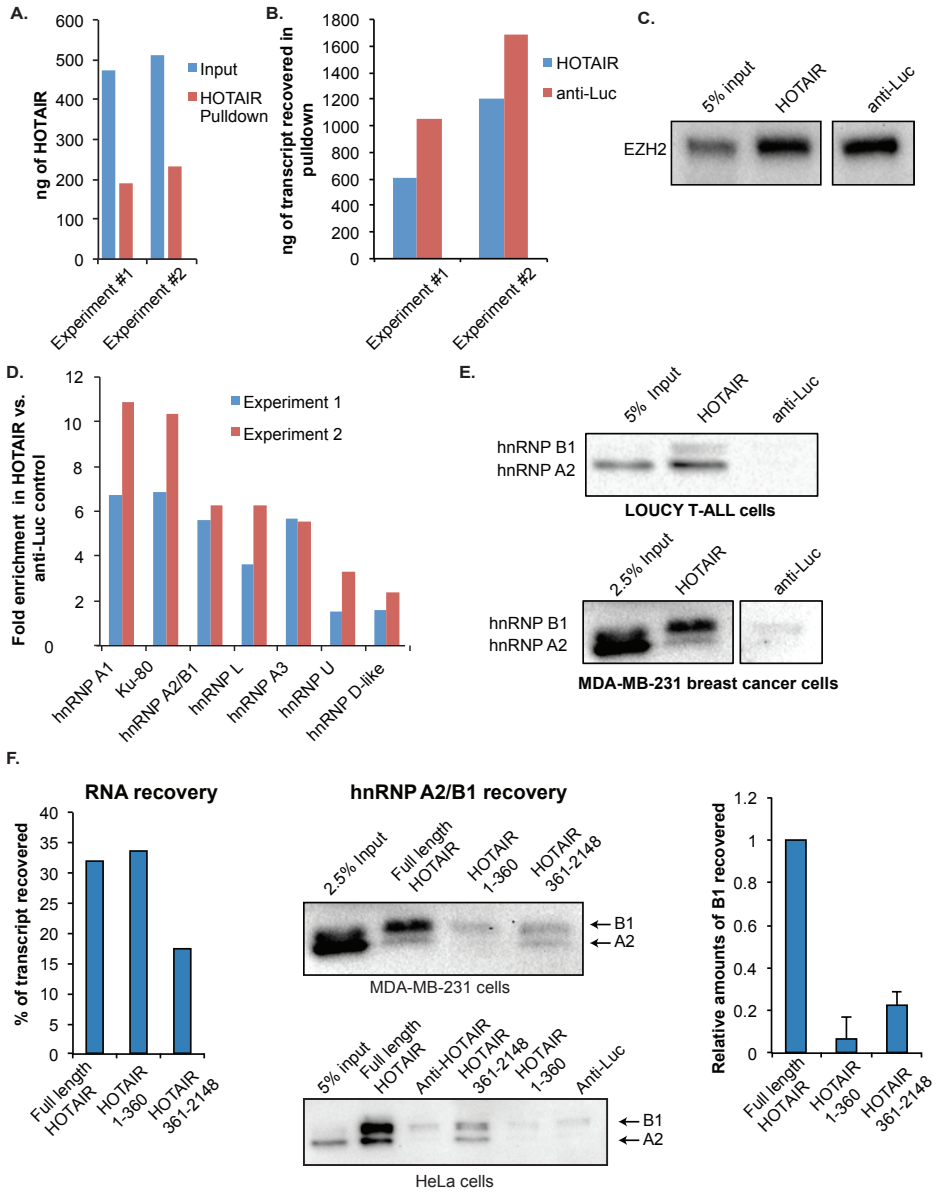

Fig. S1. In vitro HOTAIR pulldowns. (A) RT-qPCR confirms the recovery of tagged HOTAIR RNA IVT in HeLa nuclear extracts (n=2). (B) RT-qPCR was used to assess the recovery of the HOTAIR versus anti-Luc control transcripts from in vitro pulldown experiments in HeLa extracts. (C) Western analysis reveals no enrichment of EZH2 in HOTAIR pulldowns in HeLa nuclear extracts compared to pulldowns performed with an anti-Luc control (Lanes from single blot). (D) Proteins more than 1.5 fold enriched in HOTAIR versus anti-Luc pulldowns across two mass spectrometry experiments with HeLa nuclear extracts. (E) hnRNP A2 and B1 associate with HOTAIR in nuclear extracts from multiple cell lines. (F) Truncated versions of HOTAIR IVTs were tested for their ability to recover the A2 and B1 proteins from MDA-MB-231 nuclear extract. RT-qPCR was used to confirm RNA recovery. Western blot analysis was used to assess the association of the A2 and B1 proteins. The bar graph depicts Western blot quantification of the average enrichment of the B1 protein with full length and 5' or 3' truncated HOTAIR across two experiments performed in HeLa and MDA-MB-231 cells (n=2, average +/- SD).

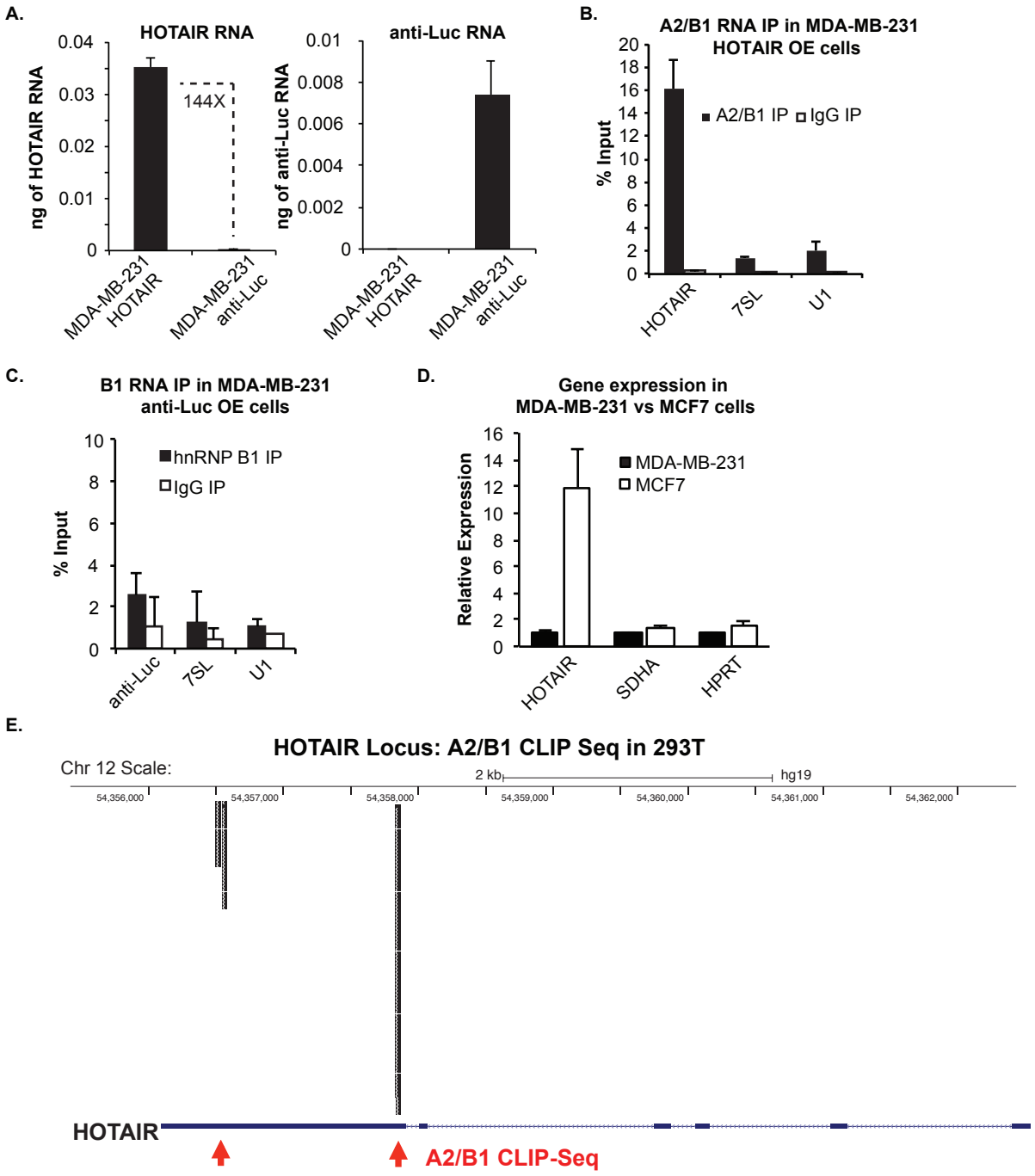

Fig. S2. RNA immunoprecipitations of the A2/B1 proteins. (A) RT-qPCR confirms overexpression of HOTAIR and anti-Luc in MDA-MB-231. The graphs depict the average +/- SD of each transcript detected following independent viral transductions (n=2). Standard curves using HOTAIR and anti-Luc RNA transcripts were used to calculate the expression of each transcript. The average expression of two housekeeping genes, *HPRT* and *SDHA*, was used to normalize for RNA input. (B) RNA immunoprecipitation from HOTAIR-overexpressing MDA-MB-231 cells was performed with a joint A2/B1 antibody. The graphs depict the average recovery of

transcripts across two independent pulldowns +/- SD as determined by RT-qPCR. (C) Recovery of anti-Luc and the control RNAs, 7SL and U1, was assessed by RT-qPCR following RIP of B1 versus an IgG control in anti-Luc-overexpressing MDA-MB-231 cells. The graph depicts the average recovery from two independent experiments +/- SD, and reveals no significant difference between B1 and the IgG control. (D) RT-qPCR was used to assess the level of HOTAIR and the two housekeeping genes, *SDHA* and *HPRT*, in parental MDA-MB-231 versus MCF7 cells (n=3, graphs depict average +/-SD). Equal amounts of total RNA was used as input into the reverse transcription reactions. (E) A2/B1 CLIP-Seq enrichment at 5' end of HOTAIR in 293T. Publicly-available A2/B1 CLIP-Seq data from (Huelga et al. 2012) were visualized in the UCSC Genome Browser. Red arrows indicate regions of direct RNA-protein interaction. A2/B1 CLIP-Seq data confirms recovery of the HOTAIR endogenous transcript.

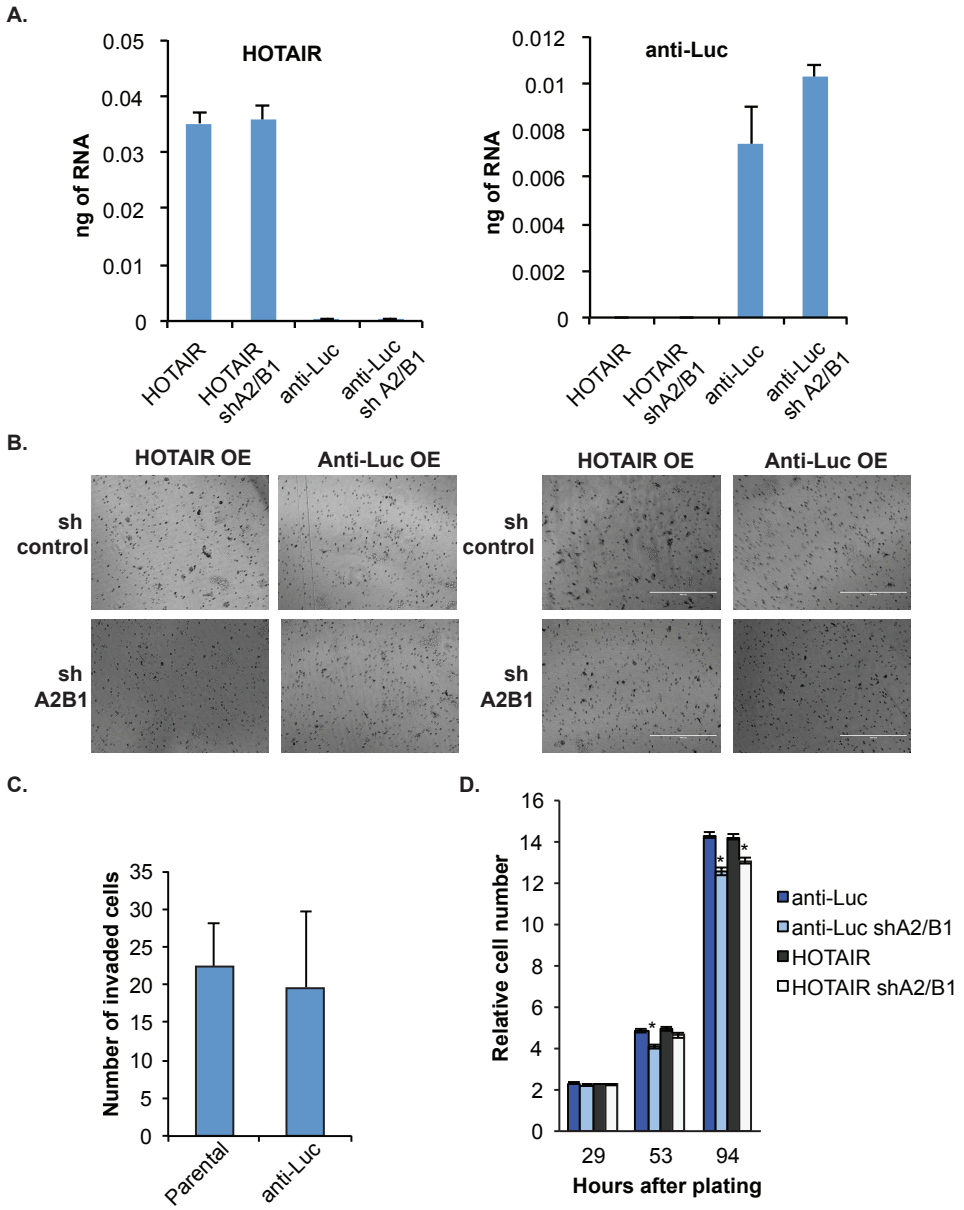

Fig. S3. Invasion and migration of MDA-MB-231 cells. (A) The expression of HOTAIR and anti-Luc was determined following knockdown of A2/B1 in MDA-MB-231 overexpression cells. Bars depict the average expression +/- SD of two biological replicates. (B) Representative 10X images of the invasion of MDA-MB-231 cells through Matrigel. Images on the left depict the average invasion for each cell line quantified in Fig. 3B. Cells stain darker than the visible membrane pores. Images on the right depict the average invasion from an independent invasion experiment with three replicates per cell line. (C) Overexpression of anti-Luc does not affect invasion. The graphs depict the invasion of two biological replicates of Matrigel invasion assays performed with unmodified parental MDA-MB-231 or the anti-Luc overexpression line. (D) hnRNP A2/B1 knockdown has minimal effects on cellular proliferation. Proliferation was assayed by Cell-Titer Glo and normalized to the cell number at the first time point for each cell line. (* = $p < 0.01$, two tailed t-test versus anti-Luc expressing cells, n=4).

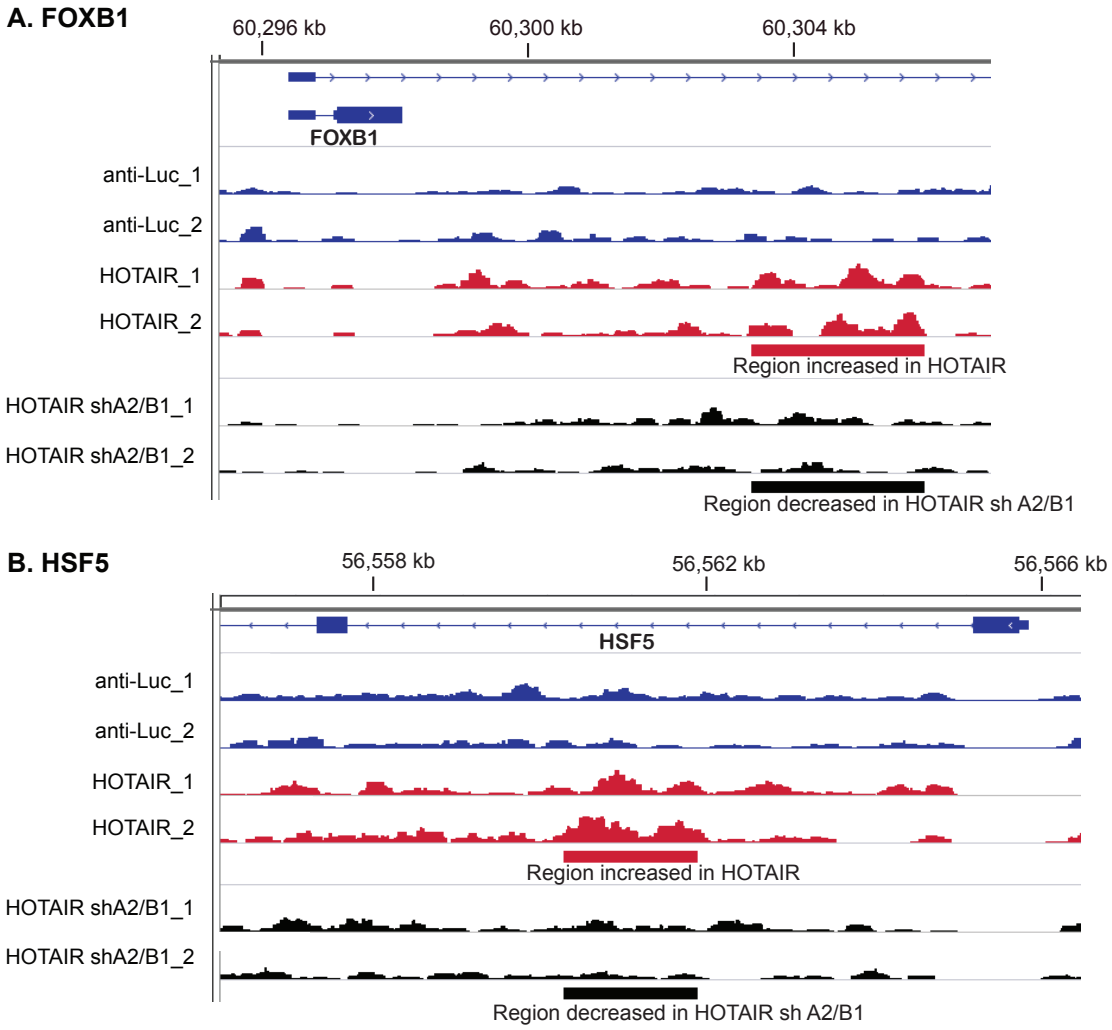

Fig. S4. HOTAIR-mediated H3K27me3 is regulated by A2/B1 at the *FOXB1* and *HSF5* genes in MDA-MB-231 cells. Pileup tracks display the H3K27me3 signal (normalized per million reads) at each genomic position from two independent ChIP-Seq experiments for each cell line. Red horizontal bars indicate regions significantly increased in H3K27me3 in HOTAIR vs anti-Luc overexpression cells. Black horizontal bars indicate regions significantly decreased in H3K27me3 in the HOTAIR sh A2/B1 versus HOTAIR samples. All differential analyses were performed using the SICER program, with a FDR of < 0.01. All tracks are plotted with the same data range on the y-axis.

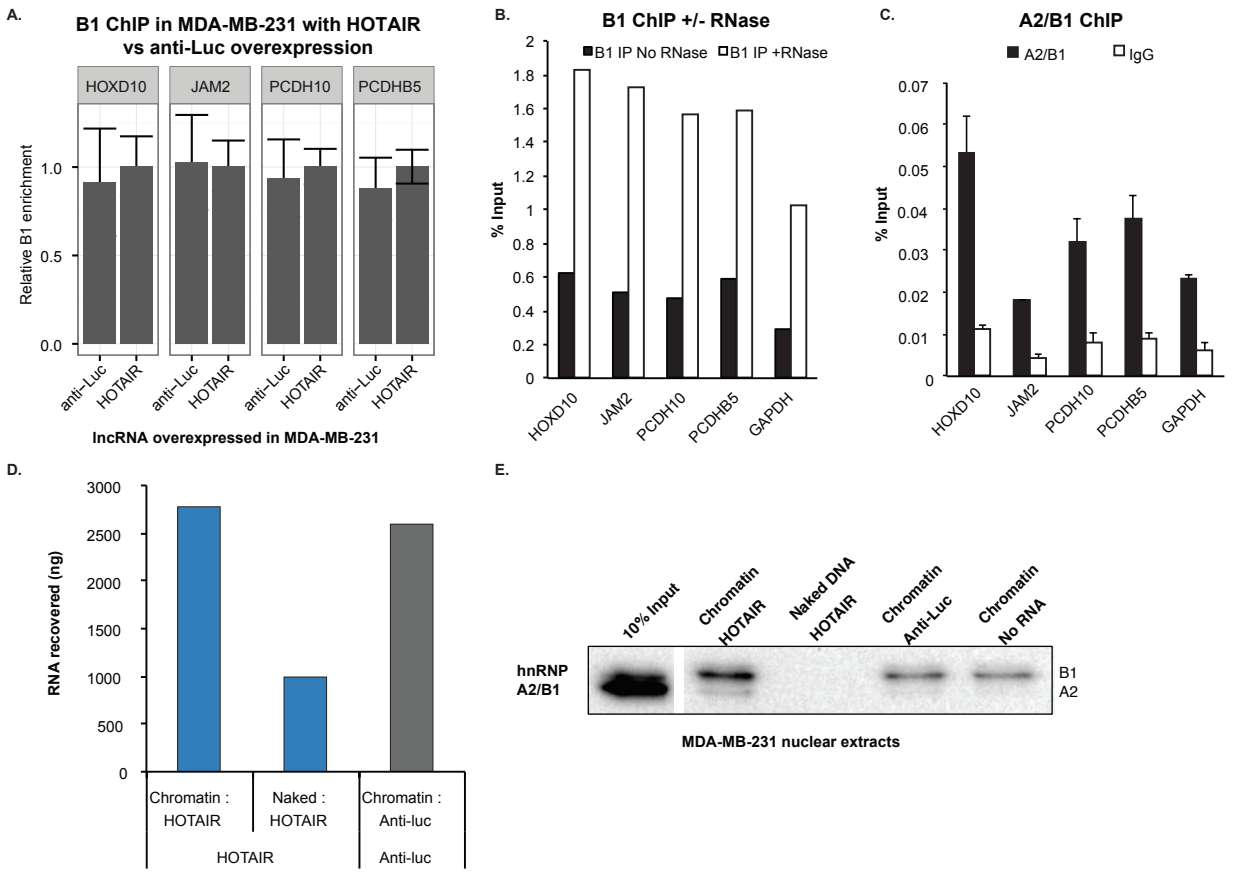

Fig. S5. B1 localizes to chromatin and interacts with chromatin-bound HOTAIR. (A) Following B1 ChIP, the B1 enrichment at four HOTAIR genomic targets was determined in cells overexpressing either HOTAIR or the control anti-Luc lncRNA. Bars depict the qPCR analysis of the B1 enrichment +/- SD from two independent experiments. HOTAIR overexpression does not affect B1 localization to these loci. (B) B1 ChIP was performed on MDA-MB-231 cell lysates +/- RNase treatment. The B1 enrichment at four HOTAIR targets and the GAPDH control loci was assessed by qPCR. (C) ChIP of A2/B1 at HOTAIR-target genes in MDA-MB-231 cells. An antibody recognizing both A2 and B1 isoforms was used for chromatin immunoprecipitation. IgG was used as a negative control antibody. Graphs show the average percent input recovered across biological replicates +/- SD. (D) Recovery of RAT-tagged in vitro transcripts following streptavidin purification of a biotinylated DNA substrate. RT-qPCR was used to assess the tethering of the HOTAIR and anti-Luc control RNA to either chromatin or naked DNA via the LexA-PP7 fusion protein. (E) Naked DNA or chromatin with tethered HOTAIR or anti-Luc in vitro transcripts were incubated with nuclear extract from MDA-MB-231 cells. Western analysis using the pan-A2/B1 antibody was used to determine the level of B1 and A2 recruitment to the chromatin and naked DNA substrates.

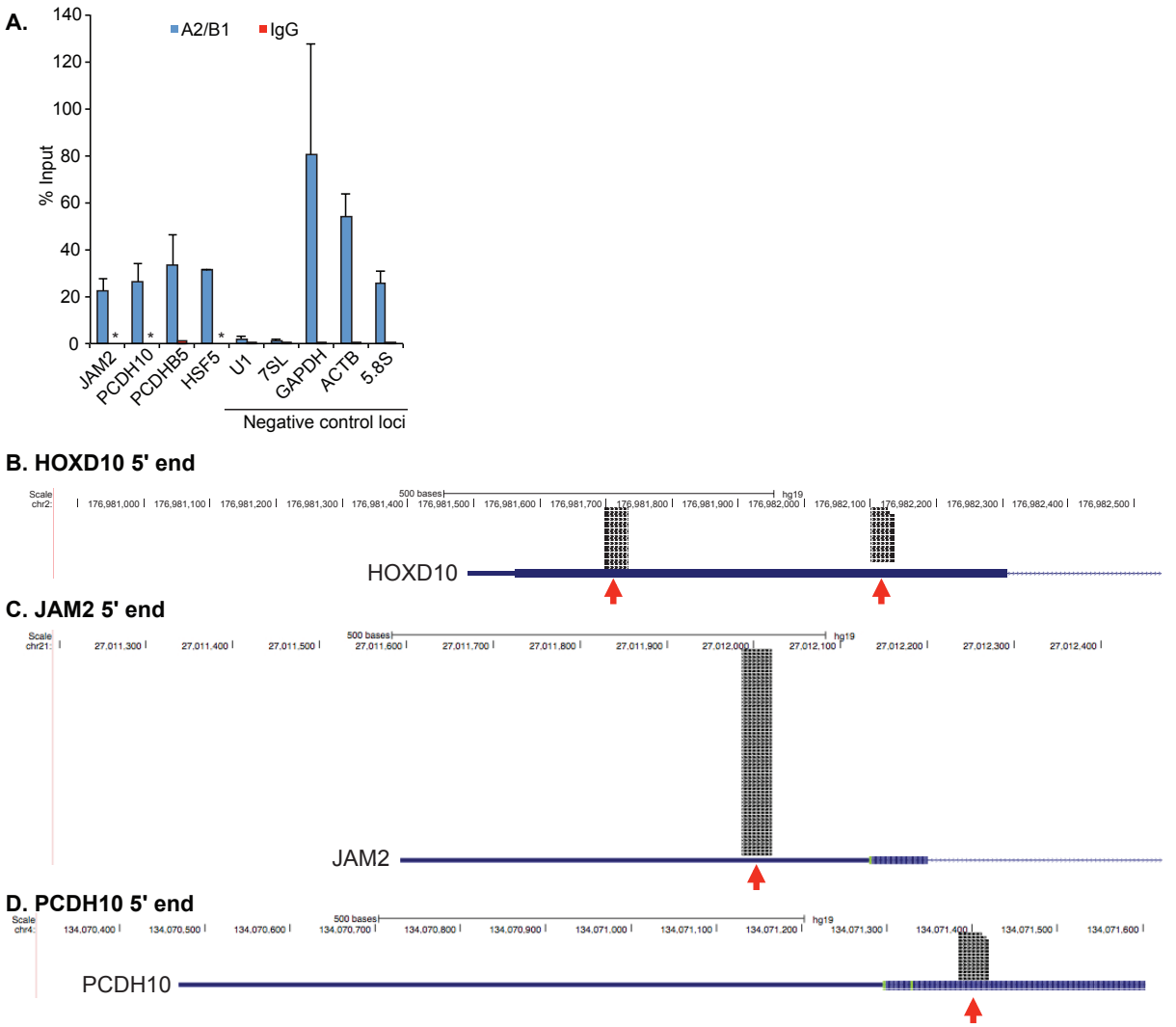

Fig. S6. hnRNP A2/B1 binds to the RNA transcripts of HOTAIR target genes and control transcripts. (A) Native RNA immunoprecipitation of A2/B1 in MDA-MB-231 cells overexpressing HOTAIR. RT-qPCR was used to assess the enrichment of HOTAIR targets and control RNA transcripts. Control IPs were performed with IgG. Graphs show the average percent input recovered across two biological replicates +/- SD. * indicates IgG samples in which no signal could be detected by qRT-PCR. HOXD10 RNA was not detectable in the input samples. (B-D) A2/B1 CLIP-Seq enrichment at the 5' end of targets of HOTAIR. Publicly-available A2/B1 293T CLIP-Seq data from (Huelga et al. 2012) were visualized in the UCSC Genome Browser. Red arrows indicate regions of direct RNA-protein interaction. A2/B1 CLIP tags were observed at the 5’ end of the HOTAIR targets: HOXD10 (B), JAM2 (C), and PCDH10 (D).