Supplemental Materials and Methods

Panhandle PCR

To detect the 3′unknown gene fused to the 5′ *NUP98* gene, we adapted the cDNA panhandle PCR strategy.[1] Briefly, total RNA extracted from the patient's leukemic cells was reverse-transcribed to first-strand cDNAs using oligonucleotides with the *NUP98* exon 7 coding sequence at the 5′ ends and random hexamers at the 3′ ends (5′-GCTTGTTTGGGTCTTCTCCAGCCACTTCCANNNNNN-3′). NUP98-1S primer (5′-CTAACAGGAAGGGCCCACAG-3′) was used for generating sense second-strand templates for panhandle PCR, and stem-loop templates contained the fusion point of the chimeric transcript in the loop were made by intrastrand annealing. Then, the *NUP98*-2S (5′-CCTCTTGGTACAGGAGCCTTT-3′) primer was added to amplify the chimeric transcript. Finally, the *NUP98*-3S (5′-GCAGGATTTGGAACAGCTCTTGGT-3′) and *NUP98*-4S (5′-ACTACGACAGCCACTTTGGG-3′) primer pair was used for nested PCR. The PCR products were purified and cloned into the pCR.TOPO Blunt cloning vector (Thermo Fisher Scientific) and sequenced.

Plasmid construction

The entire coding region of the *NUP98-HBO1* fusion cDNA was generated by RT-PCR using the patient's BM, and wild-type *HBO1* cDNA using normal BM. The *NUP98-HBO1-ΔMYST* mutant was generated by PCR to remove the carboxy-terminal 864-bp fragment. The integrity of the amplified sequence was confirmed by DNA sequencing. The cDNA with an amino-terminal FLAG tag was subcloned into pMXs-IRES-EGFP (pMXs.IG) or pMYs-IRES-EGFP (pMYs.IG).

Immunoblotting

Immunoblot analysis was performed as reported previously.[2] Anti-FLAG M2 antibody (F3165, Sigma-Aldrich) was used as a primary antibody.

Immunofluorescence staining

HeLa cells were transiently transfected with pMXs-HBO1-IRES-EGFP or pMXs-NUP98-HBO1-IRES-EGFP vectors. Cells were fixed with 4% paraformaldehyde in phosphate buffered saline (PBS) 24 hours after transfection, after which 15 mM glycine in PBS was added. Following PBS washing, cells were stained with monoclonal ANTI-FLAG M2-Cy3 (M2) antibody (Sigma) and monoclonal anti-PML Protein antibody (EPR16792, Alexa Fluor 488) (Abcam). Cells were then washed 3 times with PBS, mounted, and examined using fluorescence microscopy (BZ-X700, Keyence).

Dot Blot Analysis

Dot Blot analysis was performed as reported previously.[3] Anti-5-methylcytosine (5-mc) antibody (clone 33D3, Millipore) was used as a primary antibody.

qRT-PCR

Total RNA was isolated using RNeasy Micro or Mini Kit (Qiagen) and converted to cDNA using RevertraAce (Toyobo). The cDNA was amplified using an Applied Biosystems Step One Plus thermal cycler (Applied Biosystems). For the experiment in Figure 5D, TaqMan master mix and the following gene expression assays (Applied Biosystems) were used; HOXA9,

Hs00266821_m1; GAPDH, Hs03929097_g1. For other experiments, SYBR Green master mix and specific primers for each gene (Supplemental Table 2) were used.

ChIP assay followed by qPCR for human cells

Human cord blood CD34$^+$ cells were infected with retrovirus produced using PLAT-F-packaging cells. After 48 hours, GFP$^+$ cells that expressed NUP98-HBO1 were sorted and cultured on Tst4/min feeder cells in the presence of 100 ng/mL SCF, TPO, and FLT3-L. Then, 2×10^6 cells were fixed with paraformaldehyde (0.5% final concentration) for 10 minutes. The sample was neutralized with glycine, then the cell pellet was preserved at −80 °C until usage. Sheared chromatin fraction was prepared according to published protocol (https://ethanomics.files.wordpress.com/2012/01/chip_covaris4.pdf). The chromatin fraction was mixed with mouse IgG bound by an anti-Histone antibody (2.5 μL antibody was attached with 20 μL mouse IgG). The following antibodies were used: anti-acetyl-Histone H3 (Lys14) antibody (MABE351, Merck), anti-acetyl-Histone H4 (Lys8) antibody (17-10099, Millipore), and anti-acetyl-Histone H4 (Lys12) antibody (17-10121, Millipore). The reaction was performed at 4 °C, overnight. The sample was sequentially washed with "Low salt", "High salt", and "LiCl" buffers. DNA was extracted from the sample according to the protocol. Finally, the DNA was eluted with 17 μL buffer. The DNA was further diluted 20-fold with H_2O, and 1 μL of diluted DNA was subjected to PCR amplification.

ChIP assay followed by qPCR for murine cells

Cells were fixed with 1% formaldehyde in PBS and glycine was subsequently added to the cell suspension. Fixed cells were washed twice with PBS containing 1 mM Phenylmethanesulfonyl fluoride (PMSF) and protease inhibitor, then lysed with sodium dodecyl sulfate (SDS) lysis buffer (50 mM Tris-HCl pH8.1, 10 mM EDTA, and 1% SDS). The lysates were sonicated using an Ultra-schalloprozessor UP50H (Hielscher) and centrifuged at highest speed, then the supernatants containing fragmented chromatin were harvested for immunoprecipitation. The supernatants were diluted to 10-times volume with ChIP dilution buffer (16.7 mM Tris-HCL pH8.0, 167 mM NaCl, 1.2 mM EDTA pH8.0, 0.01% SDS, and 1.1% TritonX-100), and 2 μg of anti-acetyl-Histone H3 (06-559, Millipore), anti-acetyl-Histone H4 (39925, Active Motif), or anti-Histone H3 (ab1791, Abcam) antibodies mixed with 20 μL Dynabeads protein G (Invitrogen) was added, followed by incubation overnight at 4 °C. The antibody-chromatin complexes were washed 4 times by 150 mM NaCl wash buffer (20 mM Tris-HCL pH8.0, 2 mM EDTA pH8.0, 1% TritonX-100, and 0.1% SDS), 500 mM NaCl wash buffer, and TE buffer with 0.1% TritonX-100, and then eluted by elution buffer (1% SDS, 0.1 M $NaHCO_3$). The eluted samples were decrosslinked in high salt condition (200 mM NaCl) and incubated overnight at 65 °C. Next, 10 μL of 0.5 M EDTA, 20 μL of Tris-HCl pH 6.5, and 2 μL of 20 mg/mL proteinase K (Wako) were added, followed by incubation for 1 hour at 45 °C. DNA was extracted by two sequential phenol/chloroform extraction steps and collected by isopropanol precipitation with Etachinmate (Nippongene). Immunoprecitipated DNA was analyzed using a StepOnePlus Real-Time PCR System (ThermoFisher).

References

1. Taketani T, Taki T, Shibuya N, et al. The HOXD11 gene is fused to the NUP98 gene in acute myeloid leukemia with t(2;11)(q31;p15). *Cancer Res*. 2002;62(1):33-37.
2. Harada H, Harada Y, Tanaka H, Kimura A, Inaba T. Implications of somatic mutations in the AML1 gene in radiation-associated and therapy-related myelodysplastic syndrome/acute myeloid leukemia. *Blood*. 2003;101(2):673-680.
3. Hayashi Y, Zhang Y, Yokota A, et al. Pathobiological Pseudohypoxia as a Putative Mechanism Underlying Myelodysplastic Syndromes. *Cancer Discov*. 2018;8(11):1438-1457.

Supplemental Figure 1

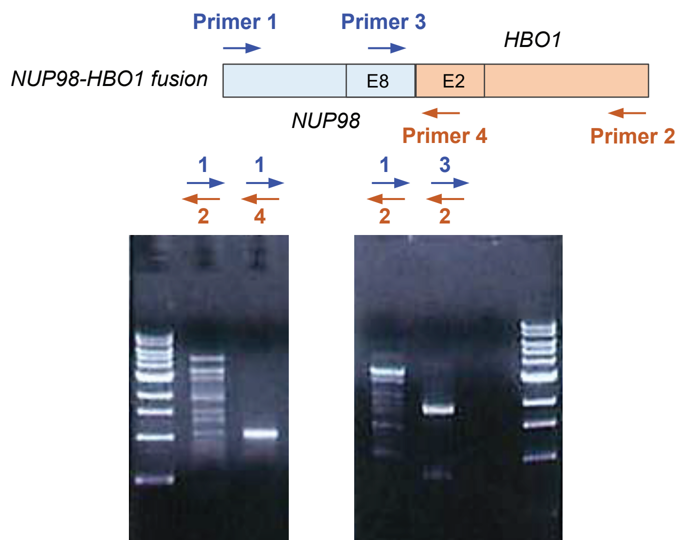

Figure 1. Identification of a novel HBO1 -fusion in a CMML patient. The fusion gene expression in the patient's cells was confirmed by RT-PCR. The KOD plus system (Toyobo) and the indicated primers (Supplemental Table 2) were used. The following parameters were used: 94°C for 2 min, followed by 35 cycles at 94°C for 15 sec, 60°C for 30 sec, and 68°C for 3 min (for the left); and 94°C for 2 min, followed by 35 cycles at 94°C for 15 sec, 62°C for 30 sec, and 68°C for 3 min (for the right).

Supplemental Figure 2

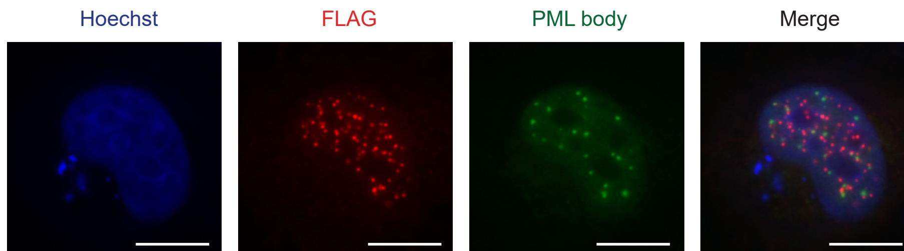

Figure 2. Localization of the NUP98-HBO1 fusion protein. HeLa cells were transfected with the indicated vectors 24 hours before immunofluorescence staining. Scale bar, 20 µm; original magnification, ×1000.

Supplemental Figure 3

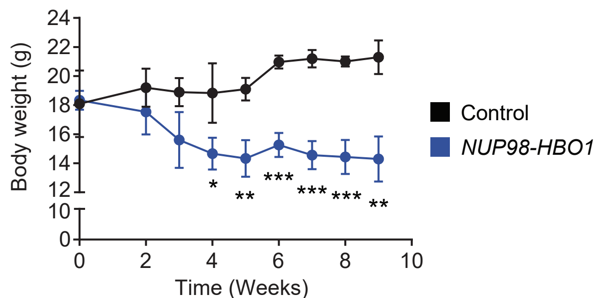

Figure 3. Progressive weight loss in NUP98-HBO1 BMT mice. Body weight of the control (n = 3) and *NUP98-HBO1* BMT mice (n = 5) after transplantation. Data are presented as mean ± s.d.

Supplemental Figure 4

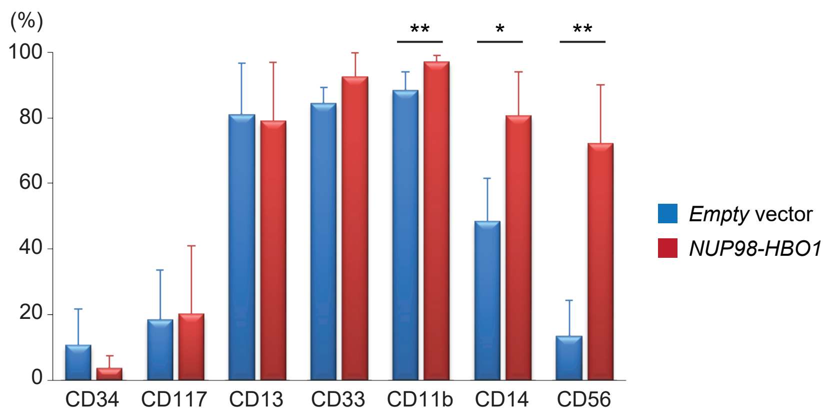

Figure 4. Expression pattern of cell surface marker on the NUP98-HBO1-transduced CB CD34^{+} cells. The indicated cell surface markers were analyzed by flow cytometry. Data are presented as mean ± s.d. *$P < 0.05$, **$P < 0.01$.

Supplemental Figure 5

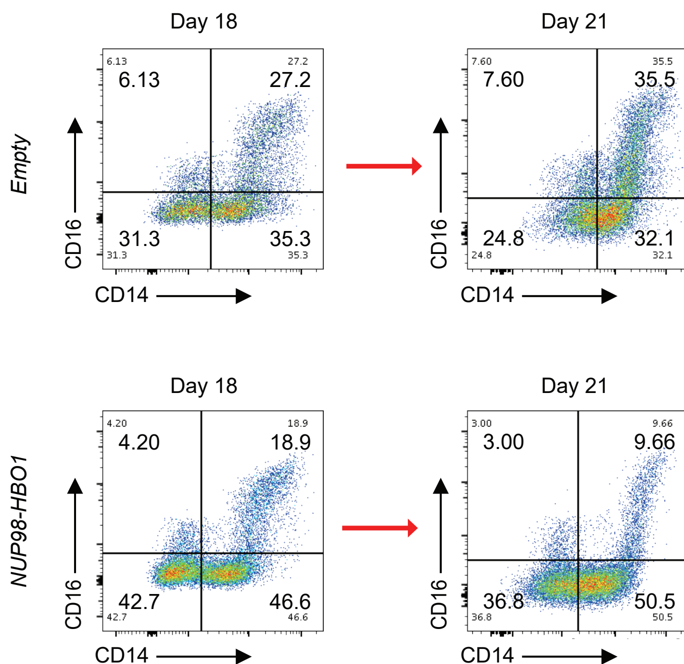

Figure 5. Expansion of CD14+ CD16- fraction of *NUP98-HBO1*-transduced human CB CD34^{+} cells. Flow cytometric analysis of CD14/CD16 expression in *NUP98-HBO1*-transduced human CB CD34^{+} cells at day 18 and day 21 after transduction. CD14 and CD16 expression pattern in the whole cultured cell population is shown.

Supplemental Figure 6

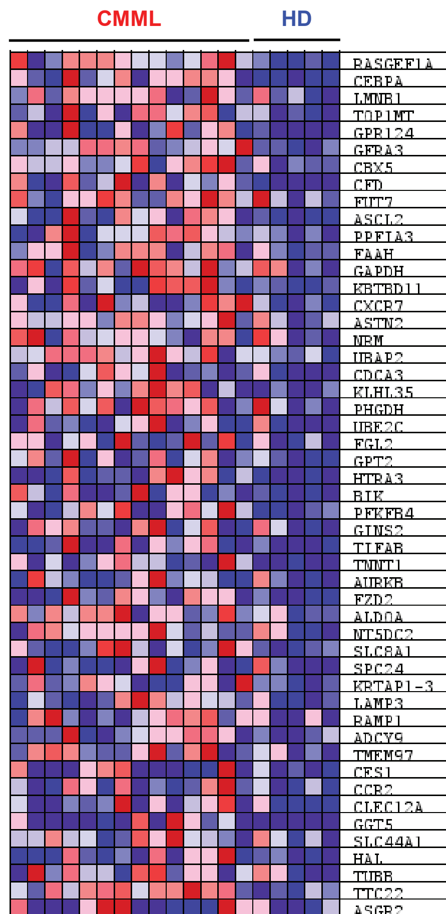

Figure 6. NUP98-HBO1-induced gene signature in human CB CD34⁺ cells. Top 50 leading edge genes from the GSEA of NUP98-HBO1-induced genes in Figure 4, A and B.

Supplemental Figure 7

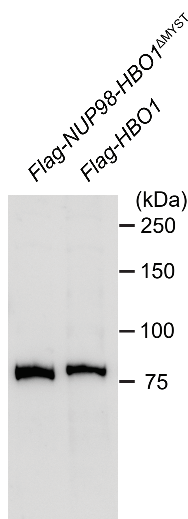

Figure 7. Expression of protein product in 293T cells transduced with the indicated vectors.

Supplemental Figure 8

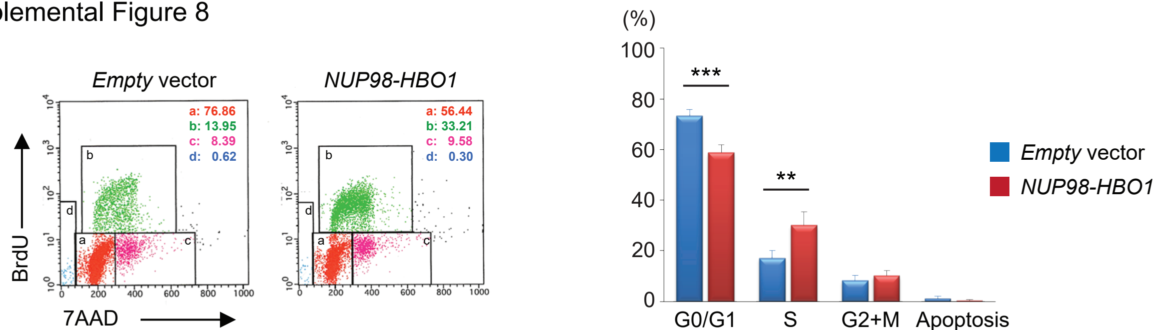

Figure 8. Cell cycle activation in *NUP98-HBO1*-transduced human CB CD34⁺ cells. Representative quantitative cell cycle analysis allowed the discrimination of cell subsets that were undergoing G0/G1 (a), S (b) or G2 + M (c) phases of the cell cycle, or apoptosis (d). Representative FACS plot and frequency of each stage are shown ($n = 4$ / group). Data are presented as mean ± s.d. $^{**}P < 0.01$, $^{***}P < 0.001$.

Supplemental Figure 9

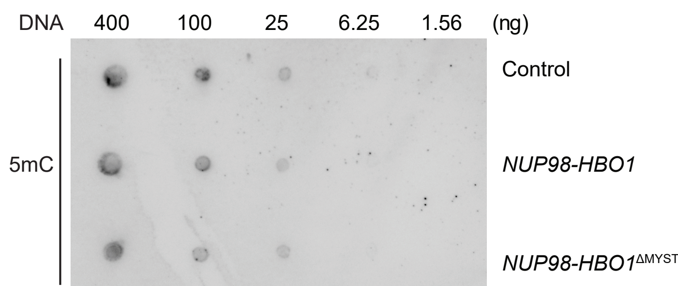

Figure 9. Impact of NUP98-HBO1 fusion on DNA methylation status. Dot blot assay for 5-methylcytosine (5mC) using genomic DNA of murine BM HSC/Ps transduced with the indicated vectors.

Supplemental Figure 10

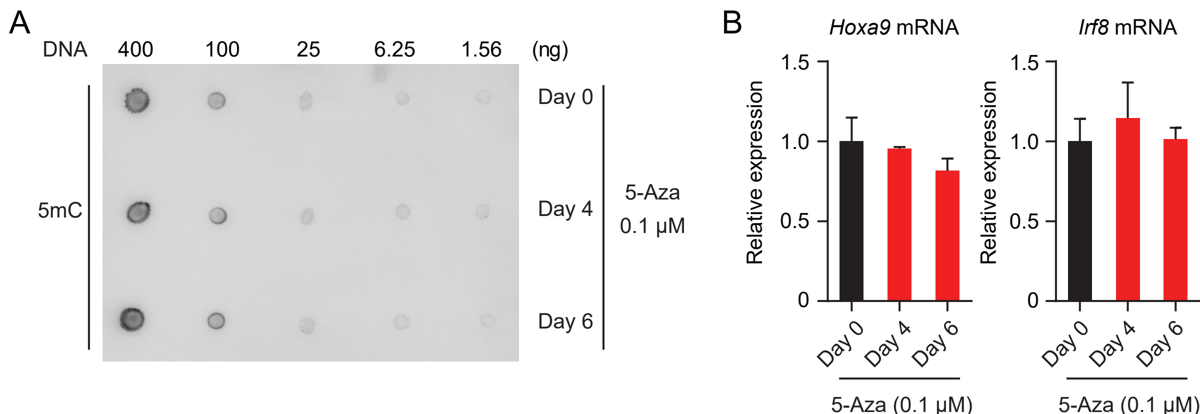

Figure 10. DNA methylation status of NUP98-HBO1 transduced cells after 5-Aza treatment. (A) Dot blot assay for 5mC using genomic DNA of *NUP98-HBO1*-transduced murine BM cells. Samples were obtained at the indicated time points after *in vitro* 5-Aza treatment. 5-Aza was added every other day to maintain the indicated concentration. (B) *Hoxa9* mRNA and *Irf8* mRNA expression in the same samples used in (A). Data are presented as mean ± s.d. from triplicates.

Supplemental Figure 11

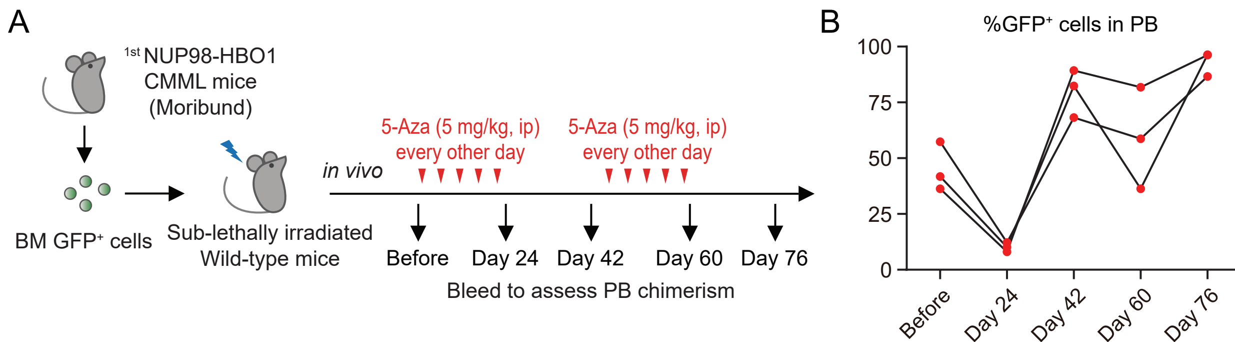

Figure 11. Efficacy of two cycles of 5-Aza treatment in the NUP98-HBO1-derived CMML model. (A) Schematic of *in vivo* treatment. (B) PB chimerism of the *NUP98-HBO1* BMT mice with 5-Aza treatment ($n = 3$).

Supplemental Table 1

Complete blood count at the initial diagnosis

Test	Laboratory data
WBC (10^9/L)	36.19
neutrophils (%)	13
monocytes (%)	37
blasts (%)	14
RBC (10^9/L)	2.41
Hb (g/dL)	7.8
HT (%)	25.9
MCV (fL)	107.5
MCH (pg)	32.4
MCHC (%)	30.1
Plts (10^9/L)	70

Supplemental Table 2

Genes	Forward primer sequence (5'-3')	Reverse primer sequence (5'-3')	Note
HOXA9 (a)	CATGTAACAACTTGGTGGCACC	CGCGTTATTGTTCTGCTGGACG	Figure 5C
HOXA9 (b)	ATCGACTCCAGTCCCATTCCTG	GGGAGTTTCCTTCCAATTAACG	Figure 5C
HOXB1	TGGACCTCCATTTGCCATTCTG	GATGGAAGTATGGGGAAGGCTG	Figure 5C
Hoxa9	GCCTTCTCCGAAAACAATGCCG	TTCCGAGTGGAGCGAGCATGTA	Figure 6E
Irf8	TGACACCAACCAGTTCATCCGAGA	CACCAGAATGAGTTTGGAGCGCAA	Figure 6E
Actb	ACACCCGCCACCAGTTC	TACAGCCCGGGGAGCAT	Figure 6E
Hoxa9 (a)	TGCCTTGGTTATCCTTGAAACAG	TTCCTCCCGGGTTAATTTGTAG	Figure 6F
Hoxa9 (b)	GGCGCCGGCAACTTATTAG	CCGACCCGCCGAAATTATG	Figure 6F
Irf8 (c)	AAAAGCCAGACGGGTGAGAG	TTAGGTGAGGGTCGAGAGGG	Figure 6G
Irf8 (d)	ACCGACTGGCTTTGATTCTG	GCCACTGGACAATCAGAGAAG	Figure 6G
Irf8 (e)	GATGGGTCTGTGATGGATTTCC	ATCTGCACCAGCCATGTTTG	Figure 6G
Irf8 (f)	TCAGCCTGCGATACAAAAC	TGGTTTTAGCCACACAGCAG	Figure 6G
NUP98-HBO1 Fw/Rv	CGCGCTTTCTGAAACAAAGA	ATTCCTACTGCTGGGGTTCG	Supplemental Figure 1
NUP98 exon 8 Fw	ACCATAGGACAGCCAAGCAC		Supplemental Figure 1
HBO1 exon 2 Rv		GGAATCTTCGGTTCCATCTG	Supplemental Figure 1