



A comprehensive promoter landscape identifies a novel promoter for CD133 in restricted tissues, cancers, and stem cells

Ramakrishna Sompallae^{1†}, Oliver Hofmann^{1,2†}, Christopher A. Maher³, Craig Gedye⁴, Andreas Behren⁴, Morana Vitezic^{5,6}, Carsten O. Daub^{5,7}, Sylvie Devalle⁸, Otavia L. Caballero⁹, Piero Carninci^{5,7}, Yoshihide Hayashizaki^{5,10}, Elizabeth R. Lawlor¹¹, Jonathan Cebon⁴ and Winston Hide^{1,2*}

¹ Department of Biostatistics, Harvard School of Public Health, Boston, MA, USA

² Harvard Stem Cell Institute, Faculty of Arts and Sciences, Holyoke Center, Cambridge, MA, USA

³ The Genome Institute, Washington University School of Medicine, St. Louis, MO, USA

⁴ Ludwig Institute for Cancer Research, Heidelberg, VIC, Australia

⁵ RIKEN Omics Science Center, RIKEN Yokohama Institute, Kanagawa, Japan

⁶ Department of Cell and Molecular Biology, Karolinska Institute, Stockholm, Sweden

⁷ RIKEN Center for Life Science Technologies, Division of Genomic Technologies, Yokohama, Kanagawa, Japan

⁸ Ludwig Institute for Cancer Research Ltd., New York Branch at Memorial Sloan-Kettering Cancer Center, New York, NY, USA

⁹ Ludwig Collaborative Laboratory, Department of Neurosurgery, Johns Hopkins University, Baltimore, MD, USA

¹⁰ RIKEN Preventive Medicine and Diagnosis Innovation Program, Wako, Saitama, Japan

¹¹ Departments of Pediatrics and Pathology, University of Michigan, Ann Arbor, MI, USA

Edited by:

David M. Thomas, Peter MacCallum Cancer Centre, Australia

Reviewed by:

Xiu-bao Chang, Mayo Clinic Arizona, USA

Paola Parrella, IRCCS Casa Sollievo Della Sofferenza, Italy

*Correspondence:

Winston Hide, Department of Biostatistics, Harvard School of Public Health, 655 Huntington Ave., Boston, MA 02115, USA
e-mail: whide@hsph.harvard.edu

† These authors have contributed equally to this work.

PROM1 is the gene encoding prominin-1 or CD133, an important cell surface marker for the isolation of both normal and cancer stem cells. PROM1 transcripts initiate at a range of transcription start sites (TSS) associated with distinct tissue and cancer expression profiles. Using high resolution Cap Analysis of Gene Expression (CAGE) sequencing we characterize TSS utilization across a broad range of normal and developmental tissues. We identify a novel proximal promoter (P6) within CD133⁺ melanoma cell lines and stem cells. Additional exon array sampling finds P6 to be active in populations enriched for mesenchyme, neural stem cells and within CD133⁺ enriched Ewing sarcomas. The P6 promoter is enriched with respect to previously characterized PROM1 promoters for a HMGI/Y (HMGA1) family transcription factor binding site motif and exhibits different epigenetic modifications relative to the canonical promoter region of PROM1.

Keywords: PROM1 protein, human, AC133 antigen, transcription start site, promoter regions, genetic, melanoma, cancer stem cells

INTRODUCTION

Surface markers play an important role in the purification of stem and progenitor cells. CD133, (PROM1) is a transmembrane protein (Corbeil et al., 2001; Shmelkov et al., 2005) that is widely used as a cell-surface marker for stem cell and cancer stem cell populations (Bussolati et al., 2005; Lee et al., 2005; Tirino et al., 2008; Zhang et al., 2008). Originally identified in hematopoietic progenitor cells by the AC133 antibody (Miraglia et al., 1997; Yin et al., 1997; Fargeas et al., 2003) CD133 displays restricted expression in both adult human (Shmelkov et al., 2004; Florek et al., 2005) and adult mouse tissues (Mizrak et al., 2007). CD133⁺ cells have also been broadly identified in non-hematopoietic tissues during differentiation *in vitro* and *in vivo* (Bussolati et al., 2005; Lee et al., 2005; Snippert et al., 2009) and have been isolated from brain and other cancers that possess stem-cell properties. For some tumor types (e.g., brain, liver, and Ewing sarcoma) CD133⁺ cell populations have been reported

to be more tumorigenic than CD133⁻ cell populations in xenograft assays (Singh et al., 2004; Jiang et al., 2010; Tang et al., 2011; von Levetzow et al., 2011). However, in colon cancers CD133⁺ and CD133⁻ populations have been found to be equally capable of tumor initiation in xenografts (Shmelkov et al., 2008), and both cell fractions have substantial tumor initiating activity in melanoma, lung, and ovarian cancer (Meng et al., 2009; Shackleton, 2010; Stewart et al., 2011), making CD133 a controversial marker for cancer stem cells (LaBarge and Bissell, 2008). Recently, a functional role for CD133 in suppression of neuroblastoma differentiation has been described (Takenobu et al., 2010), further complicating the understanding of its role and value as a suitable surface marker.

To date, using sampling across limited types of tissues and cancers, five TATA-less promoters (P1–P5) have been identified in the 5' upstream region of PROM1. These promoters differentially regulate expression of PROM1 in adult tissues and cancer cell types (Shmelkov et al., 2004). The most distally located promoter, P5, is present at 46 kb from the start codon followed by P1, P2, P3, and P4 promoters which are present at a distance

RIKEN Omics Science Center ceased to exist as of April 1st, 2013, due to RIKEN reorganization.

of -10 , -8 , -7.8 , and -6 kb, respectively (**Figure 1**). Promoters P1 and P2 drive PROM1 transcription in kidney, liver, pancreas, placenta, lung, spleen, and colon, but can also exhibit tissue restriction with P1 activity in small intestine and prostate whereas P2 is active in brain and ovary (**Table 1**); P3 is rarely active and has only been reported in skeletal muscle, P4 and P5 activity appears to be restricted to testis (Shmelkov et al., 2004). Further characterization has shown that P1, P2, and P3 promoters contain stretches of CpG islands under epigenetic regulation (Pleshkan et al., 2008; Tabu et al., 2008; Pellacani et al., 2011) under transcriptional control of Sp1 and Myc (Gopisetty et al., 2012). Collectively these findings suggest that PROM1 expression

is tightly regulated in adult tissues through the choice of specific promoters across different cell types. However, additional relationships between choice of promoter, regulatory elements, and expression restriction in normal and malignant tissues have yet to be determined.

As cancer cells can acquire the properties of stem cells, and contain a stem-like population marked with CD133, a comprehensive understanding of the differential utilization of PROM1 promoters that regulate the expression of CD133 may illustrate the differences in its expression within populations of cells with stem-like phenotypes. In order to more broadly determine aspects of PROM1 regulation and to identify key regulatory

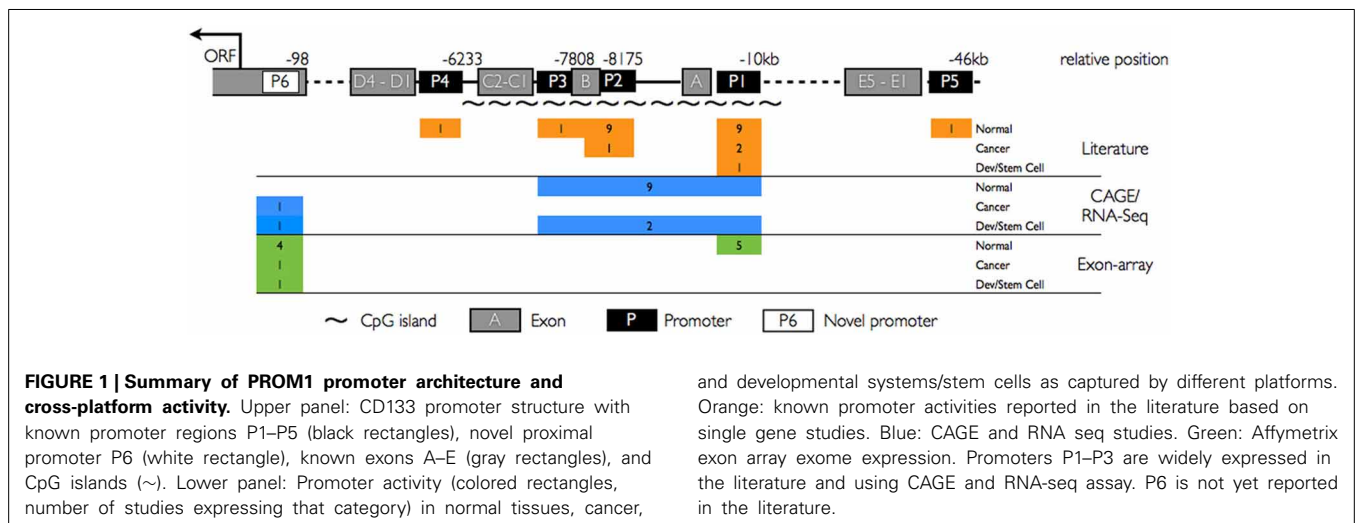


Table 1 | PROM1 alternative promoters and regulation.

Promoter	Genomic location	Distance from start codon ^c	Tissue	Transcription factors	References
P1 ^{a, b}	16085637 – 16087537	–10008	Fetal liver, liver, kidney, pancreas, placenta, lung, spleen, colon, small intestine, and prostate Glioblastoma, lung cancer	OCT4, SOX2 (during hypoxia), Sp1/Myc	Shmelkov et al., 2004; Iida et al., 2011; Gopisetty et al., 2012
P2 ^{a, b}	16085338 – 16085704	–8175	Brain, ovary, kidney, liver, pancreas, placenta, lung, spleen and colon Glioblastoma	Sp1/Myc	Shmelkov et al., 2004; Tabu et al., 2008; Gopisetty et al., 2012
P3 ^a	16084913 – 16085337	–7808	Skeletal muscle	Sp1/Myc	Shmelkov et al., 2004; Gopisetty et al., 2012
P4	16082163 – 16083762	–6233	Testis	Not studied	Shmelkov et al., 2004; Tabu et al., 2008
P5	16122394 – 16123893	–46364	Testis	ETS	Shmelkov et al., 2004; Tabu et al., 2008, 2010
P6	16077254 – 16077627	–98	Cancer, stem cells, and retina	HMG	–

^aContains CpG island.

^bHypomethylated in glioblastomas.

^cStart codon at location 16077529 in Exon 2 (16077741 – 16077309).

elements associated with PROM1 expression in cancer and stem cells, we have used Cap Analysis of Gene Expression [CAGE, (Kodzius et al., 2006)] to perform an exhaustive assessment of the landscape of the PROM1 upstream promoter region. CAGE precisely defines the location of transcription start sites (TSS) by sequencing from the 5' end of capped, full-length mRNA. In addition to TSS identification, CAGE can measure transcript abundance, allowing comparison of promoter activity between samples.

To gain additional insight into PROM1 promoter activity in the context of cancer cells with stem-like properties we have performed high-coverage CAGE sequencing of five melanoma cell lines directly derived from patient biopsies (Table S2), sorted by CD133⁺ into a small minority of cells from the total CD133⁻ population (Gedye et al., 2009). The promoter activity in these populations was compared with a panel of CAGE libraries derived from 72 tissues and cell types, including 13 normal tissues, 25 cancer tissues, and 34 developmental states (Table S1).

We identify a previously unknown promoter that shows differential expression and regulation of PROM1 mRNA in restricted tissues, stem-like cells within cancer cell lines and stem cells.

MATERIALS AND METHODS

CELL CULTURE

To isolate melanoma cells from fresh human melanoma explants, freshly excised human melanoma specimens were inspected by pathologists and fragments removed for cell line establishment without disturbing surgical margins. The melanoma cell lines were derived from metastatic melanoma tissue and used before passage 10. Description of the cell lines and associated gene-expression data have been reported previously (Behren et al., 2013). Patient consent was collected and ethical approval for the use of the cell lines has been granted by the Austin Health Human Research Ethics Committee (HREC). Cell lines were cultured in our standard media ("RF10") comprising RPMI 1640 supplemented with 2 mM Glutamax®, 25 mM HEPES, 50 μM 2-mercaptoethanol (Hamburger and Salmon, 1977), 100 U/mL penicillin, 100 μg/mL streptomycin (all from Invitrogen, Mulgrave, Australia) plus 10% fetal calf serum (FCS; from CSL, Melbourne, Australia). Tissue fragments were mechanically dissociated and passed through a cell strainer, remaining fragments were subjected to enzymatic digestion in a collagenase/DNAse/serum-free digestion media mixture overnight at 37°C and single cell suspension plated out the next morning. Once established the cell lines were HLA-typed by the Red Cross in Melbourne to ensure the match with donor tissue and were tested for *mycoplasma* contamination. Harvested cells were washed, counted and plated into 96 well round bottom plates at 10⁴–10⁵ per well. After pelleting by centrifuge the cells were washed once with PBS cells and blocked in 50 μL PBS/10% normal human serum for 10 min. The plate was gently vortexed to resuspend cells and 1 μL of AC133-PE antibody (Miltenyi Biotec, Bergisch Gladbach, Germany) was added to each well prior to incubation at 4°C for 15 min. Cells were washed and resuspended and immediately analyzed on a FACSCalibur flow cytometer (Becton Dickson, San Jose, CA). An anti-CD4-PE antibody was used at same concentration as Isotype control. The five

different melanoma cell cultures derived from biopsy specimens of patients with malignant melanoma were evaluated for CD133 expression by immuno-histochemistry (IHC) and flow cytometry (Table S2).

CELL SEPARATION AND RNA EXTRACTION

Cell separation was performed by magnetic bead antibody labeled in the cell separation MACS buffer prepared according to the manufacturer's instructions (PBS pH 7.4 with 0.5% BSA and 0.5 mM EDTA). MACS columns were refrigerated for at least 1 h prior to use. Positive selection of cells was performed using LS columns followed by depletion with LD columns. 10⁷ cells were resuspended in 80 μL MACS buffer; 20 μL FcR blocking reagent +20 μl of directly conjugated CD133 beads added, mixed well, and incubated for 30 min at 4°C. Labeled cells were washed, resuspended, and applied to the column. To increase purity columns were run in serial. Cells were passed through LS columns and were then applied to LD columns as "pre-depleted" cells. The cell population selected by the first LS column was then applied to a second LS column to increase enrichment. After separation all fractions were stained as described and purity of subpopulation measured by flow cytometry.

RNA was extracted from 10⁷ purified CD133⁺ or CD133⁻ cells using TriReagent following manufacturer's protocol (Molecular Research Center, Inc., Cincinnati, OH, USA). Briefly, cells were homogenized in Trireagent, RNA collected in the aqueous phase after addition of chloroform and precipitated by isopropanol addition. RNA was quality checked by gel electrophoresis and quantified using a nanodrop.

GENE EXPRESSION TRANSCRIPT ANALYSIS

CAGE was performed as described previously (Kodzius et al., 2006; Kawaji et al., 2009). Total RNA extracted from CD133⁺ and CD133⁻ melanoma cells was used to synthesize the cDNA. RNA and cDNA pools were treated with RNase I to cleave all ssRNA, leaving only full length cDNA/RNA hybrids for capture with biotin-streptavidin interactions in an cDNA/RNA hybrid enrichment process called as cap-trapping. In this process full length cDNAs are then ligated with specific linker oligos containing MmeI restriction sites and the second strand cDNA is synthesized. Double-stranded cDNAs are digested with MmeI creating ~20nt of cDNA sequence attached to a 5' linker. After ligation of the second linker XmaJI to MmeI-cleaved 3' ends of cDNA, fragments are subjected PCR amplification and restriction site digestion to obtain CAGE sequencing tags (see Kodzius et al., 2006, for details). The resulting CAGE tags were then concatenated and cloned into pZER0-2 plasmids (Invitrogen) for sequencing. Sequence reads were extracted, filtered and aligned to the hg18 genome build using Nexalign (Lassmann, <http://genome.gsc.riken.jp/osc/english/dataresource/>), following the methods described in (Kawaji et al., 2009). TSS in the upstream PROM1 gene region based on CAGE were identified from clustered sequence reads using HPeak (Qin et al., 2010) and mapped to known PROM1 promoters extracted from GenBank gene records (AY275524, AY438641, AY438640), resulting in the confirmation of known promoters P1–4 and identification of a novel promoter, P6. Start sites identified by HPeak were compared

to *in silico* TSS predictions in the oPOSSUM (Ho Sui et al., 2007) and SwitchGear [UCSC Genome Browser track, (Karolchik et al., 2008)] database.

RAPID AMPLIFICATION OF cDNA ENDS (5'-RACE)

To confirm the novel PROM1 TSS, 5'-RACE PCR was performed according to the manufacturer's protocol (Invitrogen). RNA was prepared from a CD133⁺ LM-Mel-34 melanoma cell line by ligating the RNA with a 5' RACE adapter 5'-GCU-GAU-GGC-GAU-GAA-UGA-ACA-CUG-CGU-UUG-CUG-GCU-UUG-AUG-AAA-3'); and a single-stranded cDNA was generated. Two CD133 specific anti-sense primers were chosen from exon2 for a nested PCR to enhance specificity and to obtain a sufficient amplification product. Primers were designed using Primer3 software (Rozen and Skaletsky, 2000), and checked for uniqueness by querying against the human genome using BLAST (Altschul et al., 1990) (Table 2). Amplification of 5'-RACE cDNA was performed with nested reverse primers of PROM1 and adapter specific primers with 1 μ l of the first-strand cDNA reaction. Amplified products were separated on an agarose gel and visualized by ethidium bromide staining.

5' RACE products were cloned into the pcDNA3.1 TA cloning vector and transformed into bacteria. The clones of each transformation were subjected to colony PCR and the sequencing of inserts was carried out with RACE adapter primers and specific reverse primers. Sanger sequence products of RACE PCR amplified fragments were separated and aligned to the human genome using BLAT (Kent, 2002) together with CAGE mapped TSS to confirm unique mapping to the target sequence.

EXON ARRAYS

Human embryonic stem cell-derived neural crest stem cells (hESC-NCSC), human adult bone marrow-derived mesenchymal stem cells as well as hESC-NCSC differentiated for 6 weeks *in vitro* together with three independent CD133-FACS-sorted cell populations from STA-ET-8.2 Ewing sarcoma cells (Table S3) were profiled by Affymetrix Human Exon 1.0 (HuEx) arrays as previously described (Jiang et al., 2010). HuEx arrays generated from four primary Ewing sarcomas (tumor RNA graciously provided by Tissue Biorepositories at Children's Hospital Los Angeles and the Children's Oncology Group) were also included for analysis. HuEx data for additional adult tissues was obtained from an Affymetrix tissue panel (http://www.affymetrix.com/support/technical/sample_data/exon_array_data.affx). HuEx data was RMA normalized using BioConductor (package *affy*) and probe intensities for

probes covering PROM1 promoter regions P1–P6 were compared to identify differences in TSS utilization.

PROMOTER ANALYSIS

To characterize regulatory motifs of the novel PROM1 promoter P6; an additional set of proximal promoters (–300/+100 bp) was selected from a total of 149 TSS found to be co-upregulated with P6 in at least four out of five CD133⁺ of the melanoma cell lines. The 149 TSS were selected based upon a significant difference in CAGE peaks in LM-MEL14/34/42/47/62 from five different patients, CD133⁺ over CD133[–] as determined by HPeak, ($p \leq 0.05$, Table S4). The TSS set was tested for nucleotide motif enrichment using MEME (Bailey et al., 2010) (motif width 4–21 nucleotides, both strands, any number of repetitions, p -value ≤ 0.05) and compared to a random background distribution of 10,000 CAGE-based proximal promoters taken from the FANTOM4 collection (Kawaji et al., 2009). Significant motifs were tested for overlap with the JASPAR 2009 Core Transcription Factor collection (Sandelin et al., 2004) using TomTom (Bailey et al., 2009). Additional experimentally determined transcription factor binding sites from the ENCODE TF ChIP-seq collection (ENCODE Project Consortium, 2004) were retrieved from the UCSC Genome Browser (hg18, update 2010-06-24, track *wgEncodeRegTfbsClustered*).

EPIGENETIC CHANGES AND RNA-seq

CpG island information and differential methylation in the Encyclopedia of DNA Elements (ENCODE) project were retrieved from the UCSC Genome Browser's summary track (Ernst and Kellis, 2010).

RESULTS

EXISTING PROMOTER LANDSCAPE OF PROM1 AND A NOVEL PROMOTER IN MELANOMA CELL LINES

To determine PROM1 promoter utilization across diverse tissues at high resolution we analyzed CAGE tags obtained from CD133⁺ melanoma cell lines and an additional 72 samples grouped as cancer, normal adult tissues and developmental stages from the public FANTOM4 data set (Table S1). Our CAGE analysis confirmed known promoters and identified a novel, sixth promoter (P6) close to the translational start codon (AUG), strongly upregulated in CD133⁺ melanoma cells and with weak expression in normal colon and small intestine libraries (Figure 2A). We determined that TSS marked by CAGE tags from cancer, normal adult tissues, and developmental stages were consistent with previous reports showing P1–P3 to be widely expressed canonical promoters. CAGE tags revealed utilization of all three promoters in normal tissues, whereas promoter utilization cancer samples are biased toward P1/P2, and developing tissues are biased toward P2.

Table 2 | Primers used for validating CD133 novel TSS.

Location	Forward primer (5'-3') (specific to RACE adapter)	Reverse primer (5'-3') (specific to CD133)	Product size
Outer	GCTGATGGCGATGAATGAACACTG	CAAGCCTTAGGAGCATCTGTGGAT	177 bp
Inner	AACACTGCGTTTGCTGGCTTTGATG	GCTAGCAAGATCCTCCAACATGA	62 bp

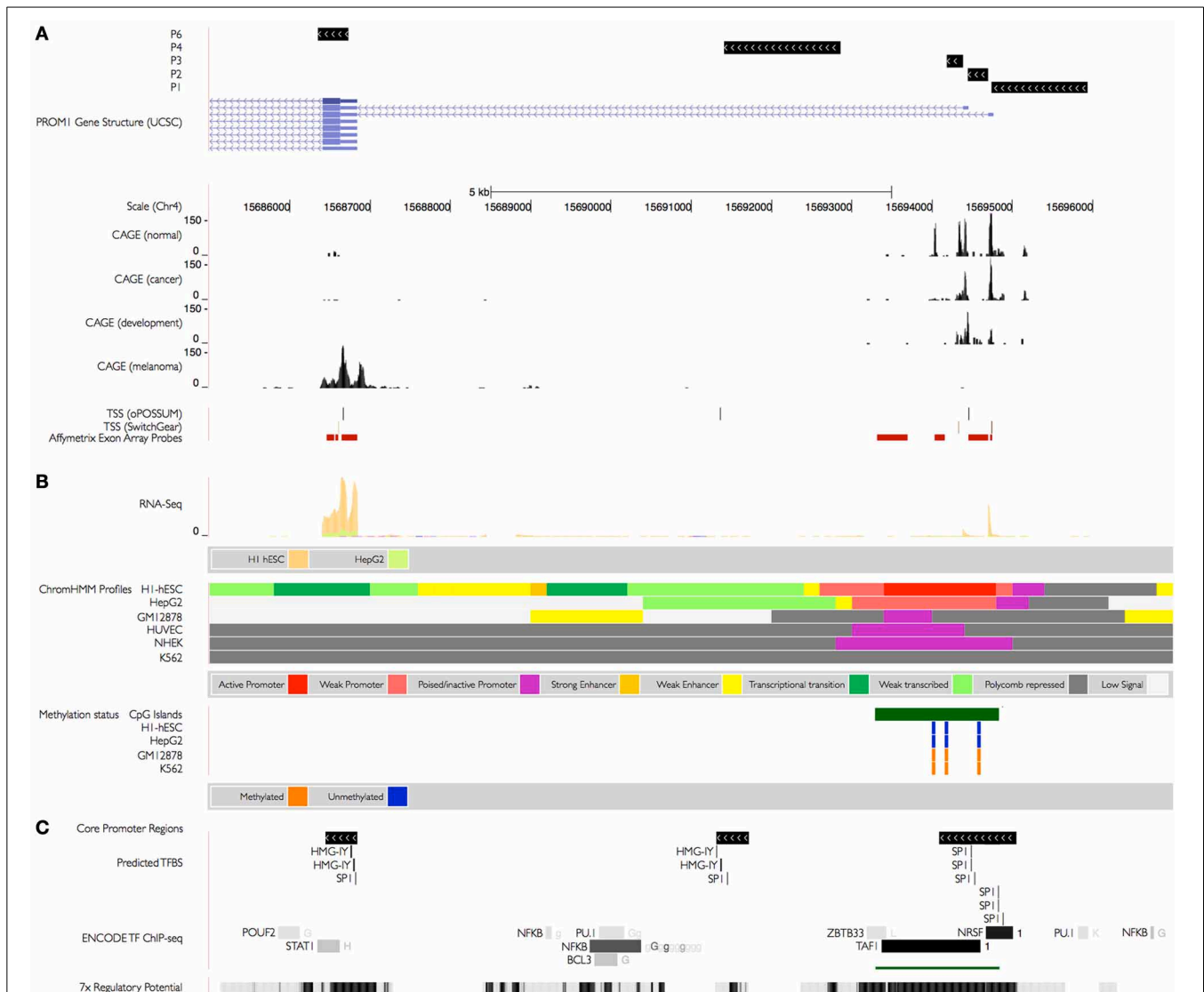


FIGURE 2 | PROM1 regulatory landscape. (A) Previously described PROM1 TSS (Promoter 1–3, 4) supported by Capped Analysis of Gene Expression (CAGE) assays from the FANTOM3 collection 57 containing 72 experimental samples of which 13 from normal tissues, 25 from different cancers, and 34 are from developmental states (Table S3). Non-pathogenic tissue CAGE tags are distributed evenly whereas those obtained from tumors are biased toward promoter 1 and 2; transcription in embryonic tissues is biased toward promoter 2. Promoter 6 is supported by CAGE tags obtained from melanoma cell lines and normal colon, small intestine, and rectum. TSS for promoters 1–4 and promoter 6 are supported by predictive models from the SwitchGear and the oPOSSUM database. Exon array probes used to characterize differential exon usage (see Figure 4) are highlighted in red. (B) Epigenetic changes in PROM1 promoters. Data from ENCODE tier 1 human cell lines (H1-hESC: embryonic stem cell line, HepG2: liver carcinoma, GM12878: lymphoblastoid, HUVEC: umbilical vein endothelial cells, NHKE: kidney epithelium, K562: myelogenous leukemia). Promoters 1–3 are supported by active (weak) promoters in H1-hESC (HepG2) cells respectively, with stronger expression in H1-hESC as

indicated by matching RNA-seq. PROM1 expression from all promoters is polycomb-repressed in GM12878, HUVEC, NHKE, and K562 resulting in no discernible expression. Expression in H1-hESC cells is enhanced at promoter 6, indicating an independent TSS unrelated to transcription driven by canonical promoters P1–3. The known CpG-island (REF) associated with the canonical promoters is unmethylated in both H1-hESC and HepG2 and methylated in GM12878 and K562, further repressing expression in those cells. (C) Regulatory potential. A strong regulatory potential is detected around the proximal promoter and promoter 1–3 (ESPERR, REF), with confirmed binding sites for transcription factors POUF2, STAT1, NFKB, and others in the ENCODE tier 1 samples (1: H1-hESC; G: GM12878; g: GM12891/GM12892/GM15510/GM18505/GM18526/GM18951/GM19099/GM19193; H: HeLa-S3; K: K562; L: HepG2). Binding sites for transcription factor found to be enriched in the proximal promoters (+300/–100 nucleotides) of 149 TSS found to be co-activated in CD133⁺-melanoma cells classify the PROM1 promoters into SP1-rich (promoters 1–3, in agreement with the CpG-island) or HMG-IY-rich (promoter 6, promoter 4).

These promoter profiles are consistent with previous reports (Table 1). The CAGE tissue panel includes one low coverage testis library with insufficient CAGE tags to support a previously reported testis-specific P4 TSS (Shmelkov et al., 2004).

VALIDATION OF NOVEL PROMOTER

To confirm the initiation of transcripts at P6 in CD133⁺ sorted cells from the melanoma cell lines we used 5'-RACE PCR. mRNA isolated from CD133⁺ cells derived from the melanoma cell line

LM-MEL-34 was used to amplify the 5' end of the PROM1 transcript with the help of a pair of forward primers targeted to a RACE adapter and the other pair of reverse primers specifically targeted to 5' exon of PROM1 (Table 2, Figure 3A). Gel electrophoresis of amplified products from these cells showed a stronger band with the expected size of 62 bp (Figure 3B). There was no sign of non-specific product in the negative control. Further, the amplified product was cloned into pcDNA3.1 plasmid for sequencing. The sequenced insert region was then mapped to the PROM1 promoter facilitating the TSS identification and the approximate location of promoter elements (Figure 3C). 5' RACE PCR efficiently detected the initiation RNA transcripts at P6.

CROSS PLATFORM EVIDENCE FOR A NOVEL P6 PROMOTER

The novel P6 promoter is supported by *in silico* TSS predictions from the oPOSSUM and SwitchGear databases. In order to further explore its utilization we studied human exon array expression data from mesenchymal stem cells (bone marrow MSC), neural crest stem cells (undifferentiated NCSC), Ewing sarcoma (primary tumors and sorted cells, STA-ET-8.2), and an Affymetrix panel of 12 adult tissues. Affymetrix Human Exon 1.0 (HuEx) arrays have an average of four probes per known exon, and seven probes cover the 5' region of PROM1, four of which showed significant expression, one matching the P1–P2 location and three cover the novel P6 promoter region. Using expression information from these probes we classified promoter utilization across different cell types, testing for stronger expression (higher probe intensity) at P6 than at upstream promoters P1–P2. We found increased expression of probes at the P6 promoter compared to P1–P2 in CD133⁺ Ewing sarcoma cells, mesenchymal stem cells, and undifferentiated NCSC, all of which express high levels of CD133. In contrast, CD133⁻ sorted Ewing sarcoma cells, CD133⁻ mesenchymal stem cells, and differentiated neural crest stem cells show no significant difference in intensities between

probes covering P1–P2 and P6 (Figure 4). We found higher expression at P6 in four tissues (colon, pancreas, kidney, and testis), with inconsistent replicate patterns or no discernible difference between P1–P2 and P6 in the other tissues (Figure S1).

NOVEL PROMOTER REGULATION

TRANSCRIPTION FACTOR BINDING SITE (TFBS) MOTIF ENRICHMENT

We performed an enrichment analysis of proximal promoters from 149 TSS in genes in the CD133 enriched CD133⁺ melanoma cell lines which were found to be co-upregulated with P6 in at least four of five CD133⁺ cell lines (see methods). We identified three significant motifs characterized as binding sites for AZF/HMG-I/Y, Sp1, and Klf4, two of which (AZF/HMG-I/Y and Sp1) are present in the P6 core promoter region (Figure 2C) and are evolutionarily conserved between human, chimp, and mouse.

EPIGENETIC MODIFICATIONS OF PROM1 PROMOTER OBSERVED IN ENCODE DATA

Epigenetic modifications are key factors for regulation of gene transcription. Since HMG-I/Y has a role in regulating chromatin structure we explored epigenetic modifications of the PROM1 promoter landscape in the publicly available ENCODE consortium dataset of genome signals. Hypermethylation of CpG islands close to P1–3 has been linked to CD133 expression in glioma stem cells and cell lines (Gopisetty et al., 2012). Differential methylation of the same region observed in ENCODE cell lines indicate an unmethylated CpG island in ENCODE cell lines H1-hESC (embryonic) and HepG2 (liver) and methylation of the same CpG island in K562 (blood, leukemic) and GM12878 (blood, lymphoblastoid), in agreement with their RNA-seq expression status in the same ENCODE cell lines (Figure 2B). The novel P6 promoter region does not overlap known CpG islands; in particular, RNA-seq data indicates independent transcription from P6 in H1-hESC. We explored histone modifications as an

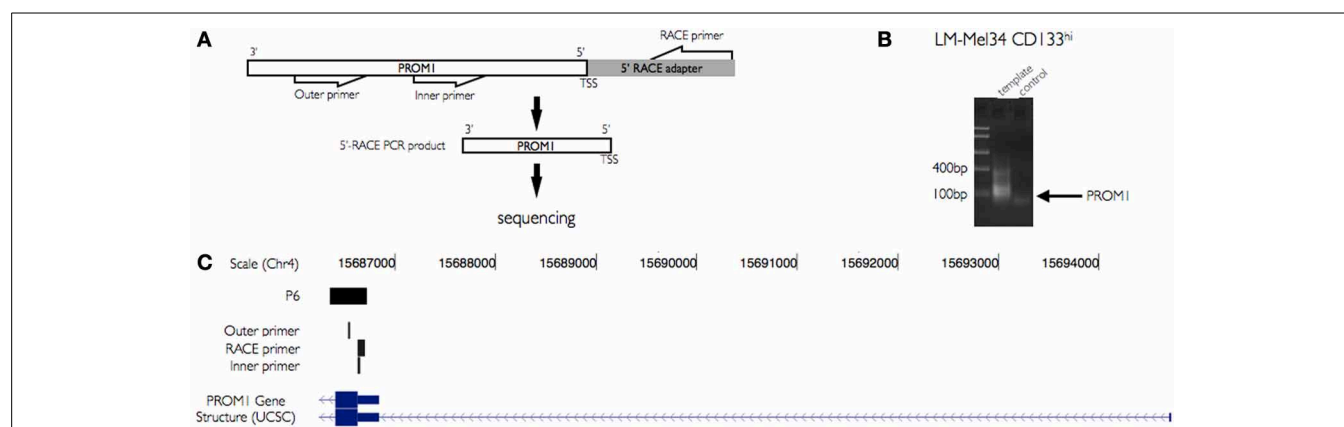
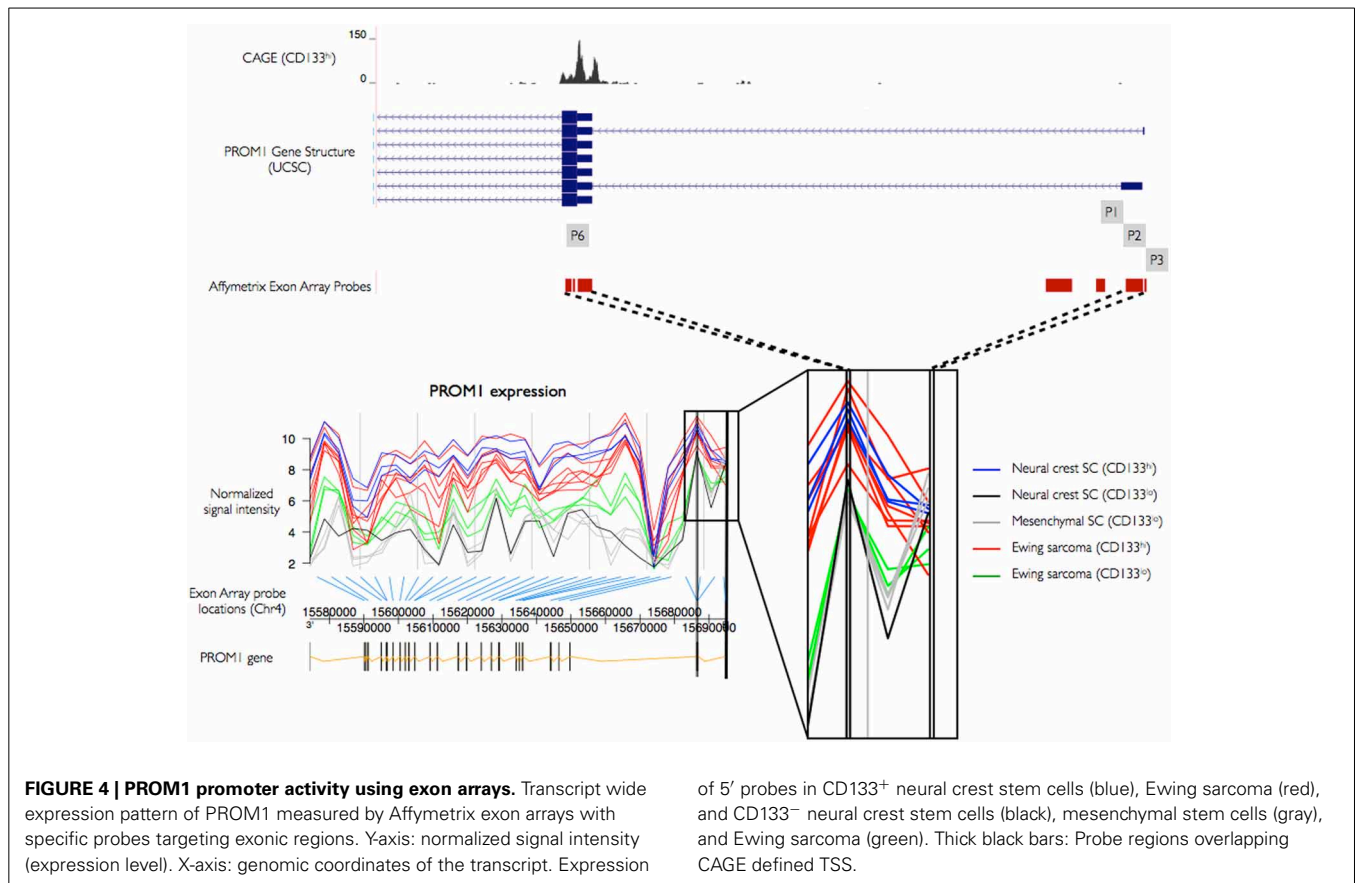


FIGURE 3 | Validation of novel PROM1 TSS using 5' RACE PCR.

(A) Schema illustrating primer design for PROM1 TSS validation. PROM1 TSS is amplified using a pair of outer and inner primer sets targeted specifically to 5' Rapid amplification of cDNA ends (RACE) adapter and PROM1 sequences. (B) The 5' end regions of PROM1 mRNA were amplified from CD133⁺ and CD133⁻ populations of LM-Mel-34 cell lines. Gel electrophoresis of amplified products shows difference in expression

level in CD133⁺ and CD133⁻ populations. Amplified 5' ends of PROM1 RNA were cloned in to plasmid pcDNA3.1 and then sequenced. (C) UCSC genome browser view of sequenced 5' end region of PROM1 mapped to the promoter. The tracks shown here illustrates CAGE captured TSS regions from CD133⁺ melanoma cell lines and the RACE PCR identified 5' end of the PROM1 transcript. The other tracks show outer and inner RACE PCR primers specific to PROM1.



alternative regulatory mechanism using summary information generated by ChromHMM (Figure 2B, Figure S1), an algorithm that characterizes chromatin states by integrating multiple ChIP-seq histone modification data sets (Ernst et al., 2011). Observed histone changes support an enhancer upstream of P6 as well as transcriptional activity around P1–3 and P6, whereas P1–3 appears poised or repressed in all other ENCODE cell lines represented, mostly due to H3K27me3 and H3K36me3 silencing. In summary, an ensemble of histone methylation marks, RNA pol II binding sites and sequence conservation observed in the PROM1 promoter region support the likelihood of transcript initiation at the P6 promoter.

DISCUSSION

Although CD133 is widely used as a stem cell marker, its significance and relationship to cancer cells with stem-like properties is controversial (Wu and Wu, 2009; Campos et al., 2011). Previous studies have established five alternative promoters (P1–P5) which drive CD133 expression in normal tissues and cancer cell lines. This transcriptional complexity raises questions in relation to differential regulation of the alternative promoters, a genetic feature which has recently been widely reported (Davuluri, 2008; Pal et al., 2011). To provide a coherent overview of PROM1 promoter choice and regulation of expression of CD133 in disease and development we have performed a comprehensive assessment of TSS activity using genome wide assay of transcription initiation.

Using transcriptional initiation events from a representative panel of 72 developmental, cancer, and normal CAGE (De Hoon and Hayashizaki, 2008; Hoskins et al., 2010; Kurosawa et al., 2010) libraries we have characterized PROM1 promoter utilization, confirming the activity of four out of five known promoters (P1–P4) and one novel alternate promoter (P6). P1 and P3 are consistently utilized in normal tissues, cancer, and development while P2 is active in developmental samples. A novel, proximal promoter P6 was identified in our high-resolution CAGE assays of a CD133⁺ subpopulation derived from melanoma cell lines and independently confirmed using 5'-RACE PCR. Initiation from P6 results in a significantly shorter 5' untranslated region (UTR).

EXPRESSION OF P6

The discovery of P6 by genome wide CAGE assay is supported by existing *in silico* predictions and by comparison with exon array probes overlapping regions of PROM1 TSS. PROM1 transcripts initiate at P6 in cancer tissues, CD133⁺ melanoma cells, adult tissues, and stem cell enriched populations, including CD133⁺ selected cells from four primary Ewing sarcoma samples for which stem cell behavior has been established (Jiang et al., 2010).

TRANSCRIPTIONAL REGULATION OF P6

Comparison of 149 core promoters of TSS found to be consistently co-expressed with the novel PROM1 P6 promoter in at least four out of five CD133⁺ melanoma cell lines identified

motif enrichment for Sp1 binding sites present in all PROM1 promoters, in agreement with their importance in the CpG islands located around P1–3 (Gopisetty et al., 2012). A second enriched motif, HMGI/Y, was found in the P6 promoter and also the testis-specific promoter P4. HMG family proteins are ubiquitously expressed nuclear proteins which regulate transcription and chromatin structure (Reeves and Beckerbauer, 2001) and have role in differentiation, tumor progression, and malignancy (Wisniewski and Schwanbeck, 2000) by controlling genes involved in tumor initiation, invasion, cell proliferation, and angiogenesis (Reeves et al., 2001). HMGI/Y (HMG1A) is usually expressed at low levels in adult tissues, but found at high expression levels in embryonic and neoplastic tissues (Chiappetta et al., 1996), its aberrant expression has been associated with tumorigenesis (Tkachenko et al., 1997) and high expression is a requirement for the production of CXC ligand 1, a major effector of tumor growth (Amiri et al., 2006). Both isoforms (HMGI and HMGY) are expressed in neuroblastic tumors, with higher levels in less differentiated tumor (Giannini et al., 2000). In gliomas, HMGI/Y expression correlates with malignancy, proliferation, and invasion (Pang et al., 2011). High levels of HMGI/Y are found in more aggressive tumors and correlate with poor prognosis and are associated with a stem-like state (Shah and Resar, 2012). In addition, ENCODE ChIP-seq data indicates binding of two transcription factors (POUF2 and NF-kb) known to interact with HMGI/Y (Reeves et al., 2001) immediately downstream of P6 (POUF2) and in an upstream enhancer (NF-kb). Both Sp1 and HMGI/Y are expressed in CD133⁺ melanoma cell lines. We did not find them to be differentially expressed when compared to CD133-depleted cells, although HMGI/Y undergoes extensive post-translational modifications which influence its binding properties (Bianchi and Agresti, 2005).

EPIGENETIC LANDSCAPE

Given the role of HMGI/Y in modifying chromatin structure we explored epigenetic changes of the upstream PROM1 region in ENCODE cell line data. Based on ChIP-qPCR analysis, methylation is thought to affect CD133 expression only in cell lines but not in primary tissues (Pellacani et al., 2011), although methylation of P2 is thought to be tissue specific (Pleshkan et al., 2008). As expected, the CpG island close to P1–P3 was found to be differentially methylated between different ENCODE cell lines, with h-ESC and hepG2 being free of methylation and leukemic (K562) and lymphoblastoid (GM12878) cell lines showing methylation.

In addition, P1–3 were found to be polycomb-repressed in all surveyed cell line types with the exception of h1-ESC and, to a lesser extent, hepG2 which exhibit active promoters. Interestingly, h1-ESC showed signs of transcriptional transition between P1–3 and P6 which in combination with the active enhancer region upstream of P6 might explain increased transcriptional activity found in h1-ESC RNA-seq data around P6.

SUMMARY

By combining comprehensive bioinformatics analysis of genome-wide exon array and exhaustively and consistently sequenced

CAGE samples across a broad range of cell and tissue types and a series of melanoma cell lines, it has been possible to reveal a strong association between a specific new promoter and clonogenic CD133⁺ cells. Together, these findings provide evidence of multiple regulatory events contributing to the diversity of PROM1 expression and indicate a potential role for HMGI/Y in combination with epigenetic changes to initiate transcription from P6 in less differentiated cells or stem cells, resulting in an upregulation of CD133. This study provides one of the few links between expression of a stem cell marker and its likely regulation.

AUTHOR CONTRIBUTIONS

Ramakrishna Sompallae: Collection and/or assembly of data, Data analysis and interpretation, Manuscript writing; Oliver Hofmann: Collection and/or assembly of data, Data analysis and interpretation, Manuscript writing; Craig Gedye: Conception and design, Provision of study material or patients, Manuscript writing; Andreas Behren: Study design and conception, provision of study material or patients, Manuscript writing; Morana Vitezic: Data analysis and interpretation; Sylvie Devallé: Collection and/or assembly of data; Otavia L. Caballero: Collection and/or assembly of data, Elizabeth R. Lawlor: Provision of study material or patients, Collection and/or assembly of data, Manuscript writing, Final approval of manuscript; Jonathan Cebon: Provision of study material, Conception and design, Financial support, Review of manuscript; Winston Hide: Conception and design, Financial support, Data analysis and interpretation, Manuscript writing, Final approval of manuscript.

ACKNOWLEDGMENTS

This research was supported by a Melanoma Research Alliance Team Science Award, in part by Operational Infrastructure Support Program Funding of the Victorian State Government and a research grant for RIKEN OSC from MEXT to Yoshihide Hayashizaki. MA received post graduate scholarship support from Cancer Council Victoria, University of Melbourne, and the Natural Sciences and Engineering Research Council of Canada. Jonathan Cabon was supported by NHMRC Practitioner Fellowships (487905). Andreas Behren is the recipient of a Postdoctoral Fellowship from the Cure Cancer Australia Foundation (1034798).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://www.frontiersin.org/journal/10.3389/fgene.2013.00209/abstract>

Figure S1 | PROM1 promoter activity using exon arrays in additional tissue panels. Transcript wide expression pattern of PROM1 measured by Affymetrix exon arrays with specific probes targeting exonic regions. **Left panel:** colon, pancreas, testis, and kidney expressing P6. **Right panel:** spleen, prostate, muscle, and thyroid expressing P1-P2.

Table S1 | Tissues surveyed for PROM1 promoter activity.

Table S2 | Melanoma cell cultures used in the study and the patient disease stage at collection.

Table S3 | Exon arrays used in the study.

Table S4 | CD133hi_set_4_5.

REFERENCES

- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410. doi: 10.1016/S0022-2836(05)80360-2
- Amiri, K. I., Ha, H. C., Smulson, M. E., and Richmond, A. (2006). Differential regulation of CXC ligand 1 transcription in melanoma cell lines by poly(ADP-Ribose) polymerase-1. *Oncogene* 25, 7714–7722. doi: 10.1038/sj.onc.1209751
- Bailey, T. L., Bodén, M., Buske, F. A., Frith, M., Grant, C. E., Clementi, L., et al. (2009). MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res.* 37, W202–W208. doi: 10.1093/nar/gkp335
- Bailey, T. L., Bodén, M., Whittington, T., and Machanick, P. (2010). The value of position-specific priors in motif discovery using MEME. *BMC Bioinformatics* 11:179. doi: 10.1186/1471-2105-11-179
- Behren, A., Anaka, M., Lo, P.-H., Vella, L. J., Davis, I. D., Catimel, J., et al. (2013). The Ludwig institute for cancer research Melbourne melanoma cell line panel. *Pigment Cell Melanoma Res.* 26, 597–600. doi: 10.1111/pcmr.12097
- Bianchi, M. E., and Agresti, A. (2005). HMG proteins: dynamic players in gene regulation and differentiation. *Curr. Opin. Genet. Dev.* 15, 496–506. doi: 10.1016/j.gde.2005.08.007
- Bussolati, B., Bruno, S., Grange, C., Buttiglieri, S., Deregiibus, M. C., Cantino, D., et al. (2005). Isolation of renal progenitor cells from adult human kidney. *Am. J. Pathol.* 166, 545–555. doi: 10.1016/S0002-9440(10)62276-6
- Campos, B., Zeng, L., Daotrong, P. H., Eckstein, V., Unterberg, A., Mairbaurl, H., et al. (2011). Expression and regulation of AC133 and CD133 in glioblastoma. *Glia* 59, 1974–1986. doi: 10.1002/glia.12139
- Chiappetta, G., Avantaggiato, V., Visconti, R., Fedele, M., Battista, S., Trapasso, F., et al. (1996). High level expression of the HMGI (Y) gene during embryonic development. *Oncogene* 13, 2439–2446.
- Corbeil, D., Fargeas, C. A., and Huttner, W. B. (2001). Rat prominin, like its mouse and human orthologues, is a pentaspan membrane glycoprotein. *Biochem. Biophys. Res. Commun.* 285, 939–944. doi: 10.1006/bbrc.2001.5271
- Davuluri, R. (2008). The functional consequences of alternative promoter use in mammalian genomes. *Trends Genet.* 24, 167–177. doi: 10.1016/j.tig.2008.01.008
- De Hoon, M., and Hayashizaki, Y. (2008). Deep cap analysis gene expression (CAGE): genome-wide identification of promoters, quantification of their expression, and network inference. *Biotechniques* 44, 627–628, 630, 632. doi: 10.2144/000112802
- ENCODE Project Consortium. (2004). The ENCODE (ENCyclopedia of DNA Elements) project. *Science (New York, NY)* 306, 636–640. doi: 10.1126/science.1105136
- Ernst, J., and Kellis, M. (2010). Discovery and characterization of chromatin states for systematic annotation of the human genome. *Nat. Biotechnol.* 28, 817–825. doi: 10.1038/nbt.1662
- Ernst, J., Kheradpour, P., Mikkelsen, T. S., Shores, N., Ward, L. D., Epstein, C. B., et al. (2011). Mapping and analysis of chromatin state dynamics in nine human cell types. *Nature* 473, 43–49. doi: 10.1038/nature09906
- Fargeas, C. A., Corbeil, D., and Huttner, W. B. (2003). AC133 antigen, CD133, prominin-1, prominin-2, etc.: prominin family gene products in need of a rational nomenclature. *Stem Cells (Dayton, Ohio)* 21, 506–508. doi: 10.1634/stemcells.21-4-506
- Florek, M., Haase, M., Marzocco, A. M., Freund, D., Ehninger, G., Huttner, W. B., et al. (2005). Prominin-1/CD133, a neural and hematopoietic stem cell marker, is expressed in adult human differentiated cells and certain types of kidney cancer. *Cell Tissue Res.* 319, 15–26. doi: 10.1007/s00441-004-1018-z
- Gedye, C., Quirk, J., Browning, J., Svobodová, S., John, T., Sluka, P., et al. (2009). Cancer/testis antigens can be immunological targets in clonogenic CD133+ melanoma cells. *Cancer Immunol. Immunother.* 58, 1635–1646. doi: 10.1007/s00262-009-0672-0
- Giannini, G., Kim, C. J., Di Marcotullio, L., Manfioletti, G., Cardinali, B., Cerignoli, F., et al. (2000). Expression of the HMGI(Y) gene products in human neuroblastoma correlates with differentiation status. *Br. J. Cancer* 83, 1503–1509. doi: 10.1054/bjoc.2000.1494
- Gopisetty, G., Xu, J., Sampath, D., Colman, H., and Puduvali, V. K. (2012). Epigenetic regulation of CD133/PROM1 expression in glioma stem cells by Sp1/Myc and promoter methylation. *Oncogene* 32, 3119–3129. doi: 10.1038/nc.2012.331
- Hamburger, A. W., and Salmon, S. E. (1977). Primary bioassay of human tumor stem cells. *Science (New York, NY)* 197, 461–463. doi: 10.1126/science.560061
- Ho Sui, S. J., Fulton, D. L., Arenillas, D. J., Kwon, A. T., and Wasserman, W. W. (2007). oPOSSUM: integrated tools for analysis of regulatory motif over-representation. *Nucleic Acids Res.* 35, W245–W252. doi: 10.1093/nar/gkm427
- Hoskins, R. A., Landolin, J. M., Brown, J. B., Sandler, J. E., Takahashi, H., Lassmann, T., et al. (2010). Genome-wide analysis of promoter architecture in *Drosophila melanogaster*. *Genome Res.* 21, 182–192. doi: 10.1101/gr.112466.110
- Iida, H., Suzuki, M., Goitsuka, R., and Ueno, H. (2011). Hypoxia induces CD133 expression in human lung cancer cells by up-regulation of OCT3/4 and SOX2. *Int. J. Oncol.* 40, 71–79. doi: 10.3892/ijo.2011.1207
- Jiang, X., Gwyne, Y., Russell, D., Cao, C., Douglas, D., Hung, L., et al. (2010). CD133 expression in chemo-resistant ewing sarcoma cells. *BMC Cancer* 10:116. doi: 10.1186/1471-2407-10-116
- Karolchik, D., Kuhn, R. M., Baertsch, R., Barber, G. P., Clawson, H., Diekhans, M., et al. (2008). The UCSC genome browser database: 2008 update. *Nucleic Acids Res.* 36, D773–D779. doi: 10.1093/nar/gkm966
- Kawaji, H., Severin, J., Lizio, M., Waterhouse, A., Katayama, S., Irvine, K. M., et al. (2009). The FANTOM web resource: from mammalian transcriptional landscape to its dynamic regulation. *Genome Biol.* 10, R40. doi: 10.1186/gb-2009-2010-4-r40
- Kent, W. J. (2002). BLAT—the BLAST-like alignment tool. *Genome Res.* 12, 656–664. doi: 10.1101/gr.229202
- Kodzius, R., Kojima, M., Nishiyori, H., Nakamura, M., Fukuda, S., Tagami, M., et al. (2006). CAGE: cap analysis of gene expression. *Nat. Methods* 3, 211–222. doi: 10.1038/nmeth0306-211
- Kurosawa, J., Nishiyori, H., and Hayashizaki, Y. (2010). Deep cap analysis of gene expression. *Methods Mol. Biol. (Clifton, NJ)* 687, 147–163. doi: 10.1007/978-1-60761-944-4_10
- LaBarge, M. A., and Bissell, M. J. (2008). Is CD133 a marker of metastatic colon cancer stem cells? *J. Clin. Invest.* 118, 2021–2024. doi: 10.1172/JCI36046
- Lee, A., Kessler, J. D., Read, T. A., Kaiser, C., Corbeil, D., Huttner, W. B., et al. (2005). Isolation of neural stem cells from the postnatal cerebellum. *Nat. Neurosci.* 8, 723–729. doi: 10.1038/nn1473
- Meng, X., Li, M., Wang, X., Wang, Y., and Ma, D. (2009). Both CD133+ and CD133– subpopulations of A549 and H466 cells contain cancer-initiating cells. *Cancer Sci.* 100, 1040–1046. doi: 10.1111/j.1349-7006.2009.01144.x
- Miraglia, S., Godfrey, W., Yin, A. H., Atkins, K., Warnke, R., Holden, J. T., et al. (1997). A novel five-transmembrane hematopoietic stem cell antigen: isolation, characterization, and molecular cloning. *Blood* 90, 5013–5021.
- Mizrak, D., Brittan, M., and Alison, M. (2007). CD133: molecule of the moment. *J. Pathol.* 214, 3–9. doi: 10.1002/path.2283
- Pal, S., Gupta, R., Kim, H., Wickramasinghe, P., Baubet, V., Showe, L. C., et al. (2011). Alternative transcription exceeds alternative splicing in generating the transcriptome diversity of cerebellar development. *Genome Res.* 21, 1260–1272. doi: 10.1101/gr.120535.111
- Pang, B., Fan, H., Zhang, I. Y., Liu, B., Feng, B., Meng, L., et al. (2011). HMGA1 expression in human gliomas and its correlation with tumor proliferation, invasion and angiogenesis. *J. Neurooncol.* 106, 543–549. doi: 10.1007/s11060-011-0710-6
- Pellacani, D., Packer, R. J., Frame, F. M., Oldridge, E. E., Berry, P. A., Labarthe, M. C., et al. (2011). Regulation of the stem cell marker CD133 is independent of promoter hypermethylation in human epithelial differentiation and cancer. *Mol. Cancer* 10, 94. doi: 10.1186/1476-4598-10-94
- Pleshkan, V. V., Vinogradova, T. V., and Sverdlov, E. D. (2008). Methylation of the prominin 1 TATA-less main promoters and tissue specificity of their transcript content. *Biochim. Biophys. Acta* 1779, 599–605. doi: 10.1016/j.bbagr.2008.06.002
- Qin, Z. S., Yu, J., Shen, J., Maher, C. A., Hu, M., Kalyana-Sundaram, S., et al. (2010). HPeak: an HMM-based algorithm for defining read-enriched regions in ChIP-seq data. *BMC Bioinformatics* 11:369. doi: 10.1186/1471-2105-11-369
- Reeves, R., and Beckerbauer, L. (2001). HMGI/Y proteins: flexible regulators of transcription and chromatin structure. *Biochim. Biophys. Acta* 1519, 13–29. doi: 10.1016/S0167-4781(01)00215-9
- Reeves, R., Edberg, D. D., and Li, Y. (2001). Architectural transcription

- factor HMGI(Y) promotes tumor progression and mesenchymal transition of human epithelial cells. *Mol. Cell Biol.* 21, 575–594. doi: 10.1128/MCB.21.2.575-594.2001
- Rozen, S., and Skaletsky, H. (2000). Primer3 on the WWW for general users and for biologist programmers. *Methods Mol. Biol.* 132, 365–386. doi: 10.1385/1-59259-192-2:365
- Sandelin, A., Alkema, W., Engström, P., Wasserman, W. W., and Lenhard, B. (2004). JASPAR: an open-access database for eukaryotic transcription factor binding profiles. *Nucleic Acids Res.* 32, D91–D94. doi: 10.1093/nar/gkh012
- Shackleton, M. (2010). Normal stem cells and cancer stem cells: similar and different. *Sem. Cancer Biol.* 20, 85–92. doi: 10.1016/j.semcancer.2010.04.002
- Shah, S. N., and Resar, L. M. (2012). High mobility group A1 and cancer: potential biomarker and therapeutic target. *Histol. Histopathol.* 27, 567–579.
- Shmelkov, S. V., St Clair, R., Lyden, D., and Rafii, S. (2005). AC133/CD133/Prominin-1. *Int. J. Biochem. Cell Biol.* 37, 715–719. doi: 10.1016/j.biocel.2004.08.010
- Shmelkov, S. V., Butler, J. M., Hooper, A. T., Hormigo, A., Kushner, J., Milde, T., et al. (2008). CD133 expression is not restricted to stem cells, and both CD133+ and CD133– metastatic colon cancer cells initiate tumors. *J. Clin. Invest.* 118, 2111–2120. doi: 10.1172/JCI34401
- Shmelkov, S. V., Jun, L., St Clair, R., McGarrigle, D., Derderian, C. A., Usenko, J. K., et al. (2004). Alternative promoters regulate transcription of the gene that encodes stem cell surface protein AC133. *Blood* 103, 2055–2061. doi: 10.1182/blood-2003-06-1881
- Singh, S. K., Hawkins, C., Clarke, I. D., Squire, J. A., Bayani, J., Hide, T., et al. (2004). Identification of human brain tumour initiating cells. *Nature* 432, 396–401. doi: 10.1038/nature03128
- Snippert, H. J., van Es, J. H., van den Born, M., Begthel, H., Stange, D. E., Barker, N., et al. (2009). Prominin-1/CD133 marks stem cells and early progenitors in mouse small intestine. *Gastroenterology* 136, 2187–2194 e1. doi: 10.1053/j.gastro.2009.03.002
- Stewart, J. M., Shaw, P. A., Gedye, C., Bernardini, M. Q., Neel, B. G., and Ailles, L. E. (2011). Phenotypic heterogeneity and instability of human ovarian tumor-initiating cells. *Proc. Natl. Acad. Sci. U.S.A.* 108, 6468–6473. doi: 10.1073/pnas.1005529108
- Tabu, K., Sasai, K., Kimura, T., Wang, L., Aoyanagi, E., Kohsaka, S., et al. (2008). Promoter hypomethylation regulates CD133 expression in human gliomas. *Cell Res.* 18, 1037–1046. doi: 10.1038/cr.2008.270
- Tabu, K., Kimura, T., Sasai, K., Wang, L., Bizen, N., Nishihara, H., et al. (2010). Analysis of an alternative human CD133 promoter reveals the implication of Ras/ERK pathway in tumor stem-like hallmarks. *Mol. Cancer* 9, 39. doi: 10.1186/1476-4598-9-39
- Takenobu, H., Shimozato, O., Nakamura, T., Ochiai, H., Yamaguchi, Y., Ohira, M., et al. (2010). CD133 suppresses neuroblastoma cell differentiation via signal pathway modification. *Oncogene* 30, 97–105. doi: 10.1038/onc.2010.383
- Tang, K. H., Ma, S., Lee, T. K., Chan, Y. P., Kwan, P. S., Tong, C. M., et al. (2011). CD133(+) Liver tumor-initiating cells promote tumor angiogenesis, growth, and self-renewal through Neurotensin/Interleukin-8/CXCL1 signaling. *Hepatology* 55, 807–820. doi: 10.1002/hep.24739
- Tirino, V., Desiderio, V., d'Aquino, R., De Francesco, F., Pirozzi, G., Graziano, A., et al. (2008). Detection and characterization of CD133+ cancer stem cells in human solid tumours. *PLoS ONE* 3:e3469. doi: 10.1371/journal.pone.0003469
- Tkachenko, A., Ashar, H. R., Meloni, A. M., Sandberg, A. A., and Chada, K. K. (1997). Misexpression of disrupted HMGI architectural factors activates alternative pathways of tumorigenesis. *Cancer Res.* 57, 2276–2280.
- von Levetzow, C., Jiang, X., Gwye, Y., von Levetzow, G., Hung, L., Cooper, A., et al. (2011). Modeling initiation of ewing sarcoma in human neural crest cells. *PLoS ONE* 6:e19305. doi: 10.1371/journal.pone.0019305
- Wisniewski, J. R., and Schwanbeck, R. (2000). High mobility group I/Y: multifunctional chromosomal proteins causally involved in tumor progression and malignant transformation (Review). *Int. J. Mol. Med.* 6, 409–419.
- Wu, Y., and Wu, P. Y. (2009). CD133 as a marker for cancer stem cells: progresses and concerns. *Stem Cells Dev.* 18, 1127–1134. doi: 10.1089/scd.2008.0338
- Yin, A. H., Miraglia, S., Zanjani, E. D., Almeida-Porada, G., Ogawa, M., Leary, A. G., et al. (1997). AC133, a novel marker for human hematopoietic stem and progenitor cells. *Blood* 90, 5002–5012.
- Zhang, M., Song, T., Yang, L., Chen, R., Wu, L., Yang, Z., et al. (2008). Nestin and CD133: valuable stem cell-specific markers for determining clinical outcome of glioma patients. *J. Exp. Clin. Cancer Res.* 27, 85. doi: 10.1186/1756-9966-27-85

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 08 July 2013; accepted: 30 September 2013; published online: 29 October 2013.

Citation: Sompallae R, Hofmann O, Maher CA, Gedye C, Behren A, Vitezic M, Daub CO, Devalle S, Caballero OL, Carninci P, Hayashizaki Y, Lawlor ER, Cebon J and Hide W (2013) A comprehensive promoter landscape identifies a novel promoter for CD133 in restricted tissues, cancers, and stem cells. *Front. Genet.* 4:209. doi: 10.3389/fgene.2013.00209

This article was submitted to *Cancer Genetics*, a section of the journal *Frontiers in Genetics*.

Copyright © 2013 Sompallae, Hofmann, Maher, Gedye, Behren, Vitezic, Daub, Devalle, Caballero, Carninci, Hayashizaki, Lawlor, Cebon and Hide. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.