PNAS
www.pnas.org

Supporting Information for

Conversion of Sox2-dependent Merkel cell carcinoma to a differentiated neuron-like phenotype by T antigen inhibition

Alexis Harold, Yutaka Amako, Junichi Hachisuka, Yulong Bai, Meng Yen Li, Linda Kubat, Jan Gravemeyer, Jonathan Franks, Julia R. Gibbs, Hyun Jung Park, Elena Ezhkova, Jürgen C. Becker and Masahiro Shuda

This PDF file includes:

Other supplementary materials for this manuscript include the following:

Email: mas253@pitt.edu

www.pnas.org/cgi/doi/10.1073/pnas.1907154116

SI Experimental procedures

Cell Lines

MCV-positive cell lines including MS-1, MKL-2, CVG-1, WaGa and MKL-1 were cultured in RPMI medium supplemented with non-essential amino acid (NEAA) and 10% fetal bovine serum (FBS). MCV LT is truncated in all MCC tumors by premature stop codons or deletions (1). Among the three MCV+ cell lines we used for the keratinocyte coculture studies, MS-1 encodes a truncated LT having the highest molecular size (428 aa), whereas MKL-2 expresses the LT with the lowest molecular size (274 aa). The LT expressed in CVG-1 has an intermediate molecular size (330 aa), which is identical to MKL-1 cells (2). Human keratinocytes (KRT) were cultured in a keratinocyte growth medium (Gibco). 293FT cells (Invitrogen) were cultured in DMEM medium supplemented with NEAA and 10% FBS.

Plasmids

For lentiviral shRNA construction, we used pENTR1A e7SK-Pro vector to drive shRNA expression under the enhanced 7SK polymerase III (e7SK) promoter described previously (2, 3) Various shRNA-coding sequences were generated by annealing two oligonucleotide (**Table S1**) and inserted into pENTR1A e7SK-Pro construct by *AgeI* and *EcoRI* restriction sites. To target MCV T antigen expression, we used an shpanT shRNA that targets common exon 1 region of T antigen transcripts (4, 5). The shRNA expression cassette cloned in entry vectors was recombined into pMuLE Lenti Dest-eGFP (Addgene plasmid #62175), pMuLE Lenti Dest-Neo (Addgene plasmid #62178) and pLenti Dest-Puro (2) by using LR-clonase II (Invitrogen). For lentiviral viral T antigen expression, pLVX EF.puro vectors that express codon-optimized sT, 339LT, and 339LT.$_{LFCDK}$ were used (5). Sanger sequencing was performed by MCLAB (San Francisco, CA) to confirm the cloned DNA sequences in the vector. To generate inducible HA-E6, and HA-E7 constructs, we first cloned attR1-ccdB-attR2 cassette into Afe I site of the pTRE.MCS EF.Puro-rTet (2) to generate pTRE.DEST EF.Puro-rTet. Then, pENTR D-TOPO HPV16 HA-E7 and pENTR D-TOPO HA-E7.$_{LYCYK}$ vectors (kind gift from Dr. Saumendra Sarkar) were cloned into pTRE.DEST EF.Puro-rTet by Gateway LR Clonase Enzyme mix (Invitrogen) to generate pTRE HA-E7 EF.Puro-rTet and HA-E7.$_{LYCYK}$ EF.Puro-rTet vectors.

Lentivirus production and infection

For lentivirus production, a lentiviral construct expressing shRNAs or MCV T antigens was cotransfected with psPAX2 (Addgene plasmid# 12260) and pMD2.G (Addgene plasmid# 12259) into 293FT cells by Lipofectamine 2000 (Invitrogen). At 16h of transfection, cell culture medium containing plasmid DNAs and transfection reagent was removed, and fresh 10% FBS DMEM was added. At 72 h post transfection, the culture supernatant was harvested, and the lentivirus was aliquoted into small tubes.

Five hundred µL of the lentivirus was infected into MCV-positive MCC cells in the presence of 1 µg/mL polybrene for 2 days, and then infected cells were selected with puromycin (1 µg/mL) or G418 (300µg/mL) for 4 days. For qRT-PCR and immunoblots, cells were recovered from puromycin selection for 2 days prior to sampling. Keratinocytes infected with 500 µL of pLVX EF lentiviruses encoding codon-optimized sT, 339LT, or 339LT.$_{LFCDK}$ or empty vector control in the presence of 6 µg/mL polybrene for 2 days, were selected with puromycin (2 µg/mL) for 5 days. Established stable cell lines were used for immunoblot and qRT-PCR experiments. MCC cell lines infected with 500 µL of pLVX EF empty, 339LT, 339LT.$_{LFCDK}$, pTRE empty, HA-E6, or HA-E7 in the presence of 1 µg/mL polybrene for 2 days selected with puromycin (1 µg/mL) for 5 days to establish stable MCC cells expressing shpanT-resistant, codon-optimized 339LT proteins used in Fig. 5 and SI Appendix, Fig. S5.

RNA extraction and quantitative reverse transcription-polymerase chain reaction (qRT-PCR)

RNA was extracted from MCC cells cultured in a various condition by Purezol (BioRad) according to the manufacturer's instruction. Five hundred ng of total RNA were used for reverse transcription by iScript cDNA Synthesis Kit (BioRad). After cDNA was diluted five-fold with nuclease free water (Ambion), PCR reaction was set up with 5 µL of cDNA and PowerUp SYBR Green Master Mix (Applied Biosystems), and the real time PCR was performed on QuantStudio 3 (Applied Biosystems) as described (2). 18S ribosomal RNA was used for normalization except for the experiments presented in Figure S3B that were normalized to ACTB gene expression. All primer sequences used for qRT-PCR are listed in **SI Appendix, Table S1**. The deduced cycle of threshold (Ct) values were used to calculate the abundance of mRNA in relation to 18S rRNA or ACTB by the comparative Ct method (ΔΔ Ct method), and a relative mRNA abundance to the control samples (control shRNA and empty vector) was determined by $2^{-\Delta\Delta Ct}$.

Immunoblots

Cells were lysed in RIPA buffer (50mM Tris-HCl, 150 mM NaCl, 0.1% SDS, 0.5% Triton X, 0.5% sodium deoxycholate, ph 7.0) containing proteinase inhibitor cocktail (Roche). The lysates were then sonicated to shear the genomic DNA and the concentration was determined using the colorimetric DC-Protein Assay Kit (Bio-Rad). After the protein lysates were mixed with Laemmli buffer (4X) and heat-denatured (100˚C for 6 minutes), the 30 ~ 50 µg of total protein was loaded and resolved in SDS-polyacrylamide gels and transferred onto the nitrocellulose membrane (Amersham). After blocked in 5% skim milk in TBS, membranes were probed with primary antibodies for overnight. Membranes were washed three times with TBS containing 0.5% tween 20 and reacted with anti-mouse IgG or anti-rabbit IgG secondary antibodies linked to IRDye 800CW or IRDye 680RD (LI-COR, Lincoln, NE) at 1:10,000 dilution in 5% skim milk TBS for 1h. Primary antibodies used are CM2B4(6), Sox2 (Cell Signaling), Atoh1 (Proteintech), Keratin 20 (Cell Signaling), CM8E6 (7), HA (Biolegend), GAPDH (SantaCruz), Hsc70 (SantaCruz).

Immunofluorescence

For immunofluorescence, cells were seeded on a coverslip and grown at least for overnight. Cells were fixed with 10% buffered formalin for 10 min and permeabilized in 0.3% tritonX100/PBS for 6 min. After blocked in 5% FBS in PBS for 1 h, cells were reacted with primary antibodies diluted in 1% FBS/PBS for overnight. After washing three times with PBS containing 0.5% tween20, cells were reacted with anti-rabbit IgG linked to Alexa Fluor 488 (Invitrogen). Primary antibodies used for dual immunofluorescence staining are rabbit anti-neurofilament L (Cell Signaling) and mouse anti-Sox2 (Proteintech) antibodies. Sox2 expression in the nucleus was quantitated by Image J. DAPI was used to locate the nuclei of neurofilament L (NFL)-positive MCC cells. MCC cells were differentiated from keratinocytes based on the size of nucleus and NFL positivity. A background value was determined through measuring a space empty of both MCC cells and keratinocytes.

Transmission electron microscopy

The specimens were fixed in cold 2.5% glutaraldehyde in 0.01 M PBS. The specimens were rinsed in PBS, post-fixed in 1% osmium tetroxide with 1% potassium ferricyanide, rinsed in PBS, and dehydrated through a graded series of ethanol and propylene oxide and embedded in Poly/Bed® 812 (Luft formulations) (Polysceinces, Warrington, PA). Semi-thin (300 nm) sections were cut on a Leica Reichart Ultracut, stained with 0.5% toluidine blue in 1% sodium borate and examined under a light microscope. Ultrathin sections (65 nm) were stained with uranyl acetate and Reynold's lead citrate, and were examined on a JEOL 1011 transmission electron microscope (JEOL USA, Peabody, MA) (grant # 1S10RR019003-01 NIH for Simon Watkins) with a side mount AMT 2k digital camera (Advanced Microscopy Techniques, Danvers, MA).

WST-8 cell proliferation assays

To determine the effect of Sox2 and Atoh1 knockdown for cell proliferation, MCC cells (MS-1, MKL-2 and CVG-1) were transduced with 600µL of shSox2.1-puro, shSox2.2-puro, shAtoh1-puro, as well as control shRNA (shCtrl-puro) lentiviruses for 2 days in the presence of 1 µg/mL polybrene. The selection of transduced cells was done with puromycin for 4 days. At day 6 post infection, cells were counted and 2.5 x 10^4 cells were seeded in a 96 well plate (day 0) and cell proliferation was measured by WST-8 (Wako) at days 1, 3, 5, 7, 9, and 11. Eight µL of the WST-8 was added to each 96 well containing cells in 95 µL of the cell culture medium, and cells were incubated for 2 h at 37˚C CO_2 incubator. The WST-8 formazan was measured from sextuplicated samples at 440 nm with a reference filter at 600 nm using a Synergy 2TM Multi-Mode Microplate Reader (Biotek). The data is presented as an average ± standard deviation (SD).

Whole cell patch clamp recording

The MCC cells seeded on a coverslip were transferred to a recording chamber and continuously perfused with oxygenated (95% O_2 and 5% CO_2) extracellular bath solution containing: 117 mM NaCl, 3.6 mM KCl, 2.5 mM $CaCl_2$, 1.2 mM $MgCl_2$, 1.2 mM NaH_2PO_4, 25 mM $NaHCO_3$, 11 mM glucose. Cells were visualized using a microscope with infrared differential interference contrast (IR-DIC) optics (BX-51WI, Olympus, Tokyo, Japan). The shpanT-puro-transduced MCC cells that displayed grade 3 differentiation were selected for recording, while small, round cells without cell processes were selected as undifferentiated cells from MS-1 cells transduced with shCtrl-puro. Patch pipettes made from borosilicate thin-walled glass capillaries (Warner Instruments, G150F-6) using a PC-10 micropipette puller (Narishige International) had a tip resistance of 8–12 MΩ. The composition of the pipette solution was 135 mM potassium gluconate, 5 mM KCl, 0.5 mM $CaCl_2$, 5 mM EGTA, 5 mM Hepes, 5 mM ATP-Mg, pH 7.2. All experiments were conducted at room temperature (19°C). A current clamp recording was made by using Axopatch 200B (Molecular Devices). Data were digitized using a Digidata 1322A (Molecular Devices) and stored and analyzed using pClamp 10 software (Molecular Devices). The current injection was made to test whether the recorded neurons were excitable. To test the involvement of the voltage-gated calcium channel, we used artificial cerebrospinal fluid (ACSF, 117 mM NaCl, 3.6 mM KCl, 2.5mM $CaCl_2$, 1.2 mM $MgCl_2$, 1.2 mM NaH_2PO_4, 25 mM $NaHCO_3$, 11 mM glucose) and the low calcium ACSF (117 mM NaCl, 3.6 mM KCl, 0.02 mM $CaCl_2$, 3.7mM $MgCl_2$, 1.2 mM NaH_2PO_4, 25 mM $NaHCO_3$, 11 mM glucose).

Single cell RNA sequencing

We constructed the two single cell transcriptome libraries from shpanT-puro- or shCtrl-puro-transduced CVG-1 cells cocultured with primary human keratinocytes. Single nuclei are stained using trypan blue for visualization, counting and viability assessment using the Countess II Automated Cell Counter and diluted as needed to a final concentration of 700-1200 nuclei/µL in 1x PBS + 0.04% BSA. A minimum of two reproducible counts with SD <25% are required before droplet preparation and <5% of input should be viable cells. Nuclei are added to a master mix containing reverse transcription reagents and primer, and are then transferred to the ChromiumChip with gel beads and partitioning oil for the preparation of nanoliter-scale Gel Bead-In-Emulsions (GEMs). The emulsion undergoes reverse transcription to produce cDNA with a cellular 10X barcode and UMI, and is recovered with Dynabeads MyOne SILANE magnetic beads according to the manufacturer's directions. Subsequently, the cDNA undergoes 8-14 cycles of amplification before clean-up using SPRIselect Reagent (Beckman). Total cDNA yield is calculated following a quality assessment (High Sensitivity D5000 ScreenTape, Agilent) and fluorometric quantitation (Qubit). Libraries are constructed by enzymatic fragmentation, end repair, and A-tailing according to the manufacturer's instructions. A double-sided SPRIselect Reagent (Beckman) clean up prepares samples for adapter ligation and another round of magnetic bead clean-up before sample

indexing PCR for 5-16 cycles, depending on the cDNA input for library construction. A further double-sided size selection produces libraries of 400-600bp, which are quantified for paired-end sequencing (26bp x 98bp) using a NextSeq 500 System.

Single-cell RNA sequencing data analysis

Briefly, we processed the raw sequencing data using the 10X Genomics software package CellRanger (version 3.0.0). The reads were aligned to a concatenation of the human and MCV transcriptomes (in gtf and FASTA). In particular, we used "cellranger mkred" to construct the concatenated transcriptome and "cellranger count" to quantify genes defined in the transcriptome, as described in [Russell et al.]. We generated the human transcriptome based on genome assembly GRCh38.87. The MCV transcriptome was generated from the reverse-complement of the viral RNA sequences as encoded in the reverse-genetics plasmids. The aligned deep sequencing data are available on the GEO repository (Accession# will be obtained upon acceptance of the paper). Further, CellRanger creates a cell-gene matrix consisted of 19962 rows of genes and 36026 columns of cells. For cell-wise filtering, since there is no common set of external RNA controls in the study design, we used library size, the proportion of expressed genes, and mitochondrial proportion as the internal controls by assuming that most cells were in normal condition and their RNAs were well sequenced and quantified. A Scater package was used to implement cell-wise cleaning (8). Cells (85 cells) with log-transformed library size 2 median absolute deviations (MADs) (9) below the median and cells (112 cells) with log-transformed number of expressed genes 2 MADs below median were identified as low quality and filtered out. Cells (1,746 cells) with a mitochondrial proportion 2 MADs higher than the median were identified as broken or dead and removed. For gene-wise filtering, we kept genes that are expressed (UMI > 1) in at least 1% of all cells. After the filtering, the expression matrix was reduced to hold 11,851 genes and 25,836 cells. We further implemented differential expression analysis using MAST (10) approach in the Seurat (9) package. MAST generalized linear models that adapted for zero-inflated single cell gene expression data, which is able to handle the feature of high sparsity of our data appropriately. We considered genes with Benjamini-Hochberg-adjusted *p*-values of less than 0.01 as differentially expressed genes. Further, we selected genes that have more than or equal to 0.15 log2 fold change between the average expression of two groups to be further considered as genes with fold changes. Please note that, although these genes come out with significant (BH *p*-value < 0.01) fold change values ($\log_2$FC 0.15), the signals were too weak to show a biological implication.

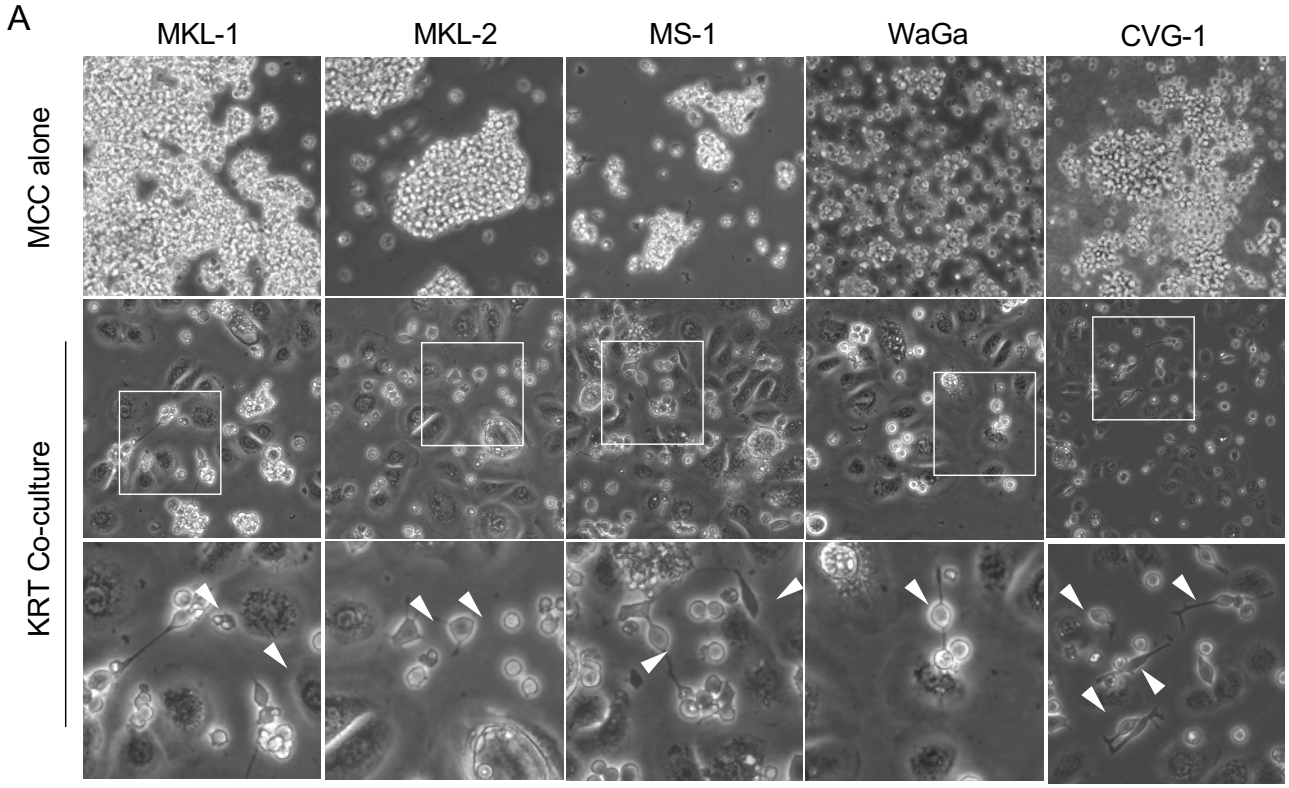
DNA methylation microarray analysis

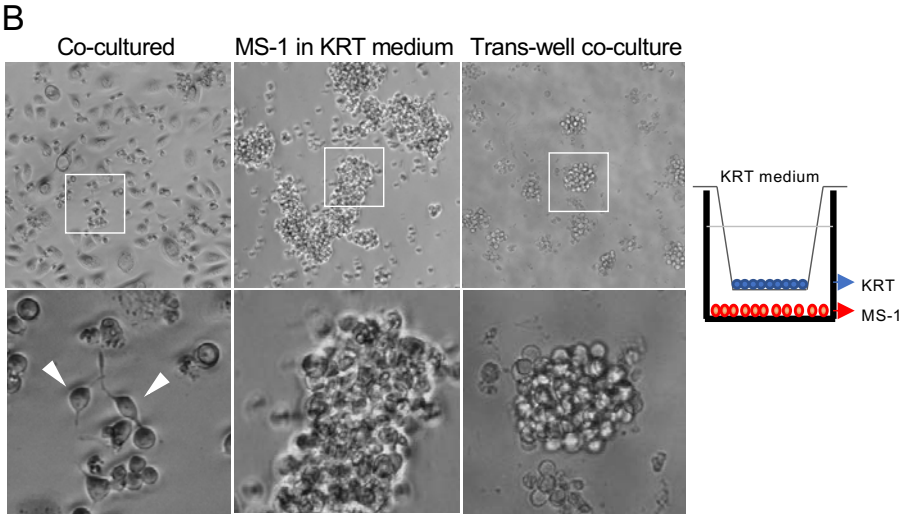
Infinium Methylation EPIC BeadChips (Illumina) were used to measure DNA methylation. It enables the analysis of >850,000 CpG sites per sample at single nucleotide resolution. The data analysis was performed in R version 3.4.4 (https://www.R-project.org/). Raw IDAT files were imported and further processed following the cross-package workflow (11), incorporating the Bioconductor packages minfi (12) and limma (13). Data were normalized using functional normalization (14) implemented in minfi. Probes with a detection p-value above 0.01 in at least one sample were removed. Additionally, probes located on sex chromosomes, with SNPs at CpG sites or cross-reactive probes were filtered out from the analysis. For visualizing the percentage of methylation, methylation levels were converted to beta-values ($\beta = M/(M+U)$), the ratio between methylated and the sum of methylated and unmethylated signal. For multidimensional scaling, methylation levels were converted to M-values ($M = \log_2(M/U)$), the logit transformation of the beta-values. No batch effects were detected with multidimensional scaling. Differentially methylated probes and differentially methylated regions were called as described and considered to be significant if *p*-value <= 0.05. Annotation information provided by Illumina correspond to human genome (hg19) reference coordinates.

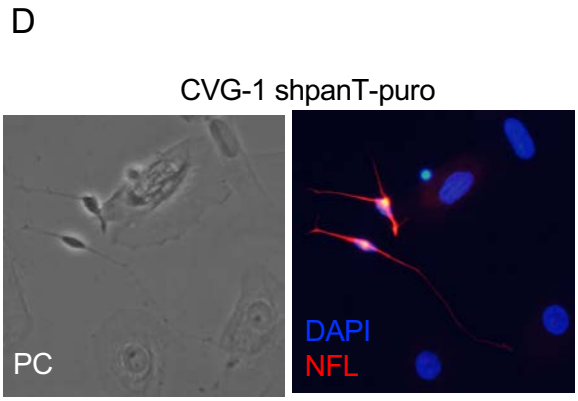
Chromatin immunoprecipitation (ChIP) assay

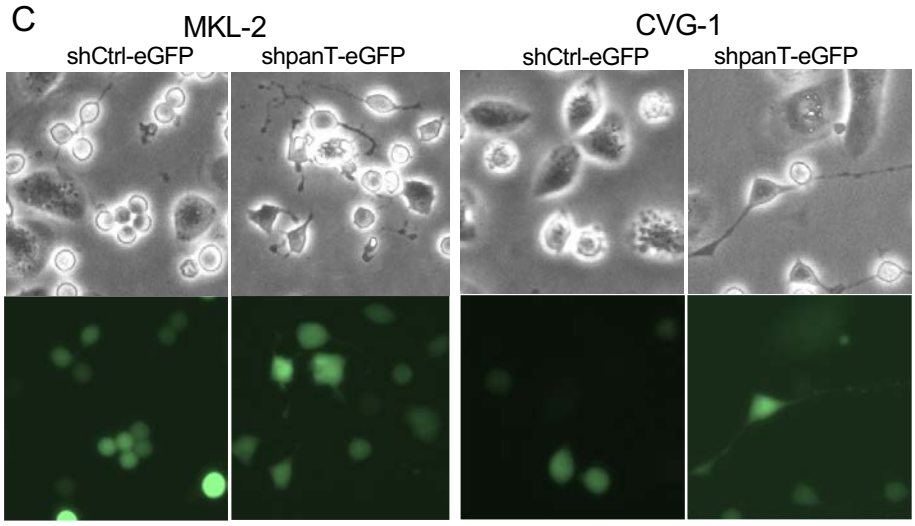
A total of 6 x 10^6 MKL-2 cells was cross-linked by 1% formaldehyde (Thermo Fisher Scientific; Rockford, IL) for 10 minutes at room temperature. The crosslinking reaction was quenched by adding 125 mM glycine for 5 minutes and washed with PBS twice. Cells were sequentially incubated for 10 mins in lysis buffer 1 (50mM HEPES pH=7.5, 140mM NaCl, 1mM EDTA, 10% glycerol, 0.5% NP-40, 0.25% Triton X-100), and then lysis buffer 2 (10mM Tris-HCl pH=7.5, 200mM NaCl, 1mM EDTA, 0.5mM EGTA). Cell pellets were resuspended in lysis buffer 3 (10mM Tris-HCl pH=8.0, 100mM NaCl, 1mM EDTA, 0.5mM EGTA, 0.1% Na-deoxycholate, 0.5% N-laurylsarcosine, 1% Triton X-100, 1X protease inhibitor cocktail (Roche)). Chromatin was sonicated into 300 to 500 bp fragments using Bioruptor (Diagenode). Cell lysate was incubated with a Sox2 antibody (AF2018, R&D Systems) overnight at 4˚C. The immunoprecipitates were sequentially washed with low salt (20mM Tris-HCl pH=8.0, 150mM NaCl, 2mM EDTA pH=8.0, 0.1% SDS, 1% Triton X-100), high salt (20mM Tris-HCl pH=8.0, 500mM NaCl, 2mM EDTA pH=8.0, 0.1% SDS, 1% Triton X-100), LiCl (10mM Tris-HCl pH=8.0, 250mM LiCl, 1mM EDTA pH=8.0, 0.1% Na-deoxycholate, 1% NP-40), and Tris-EDTA buffers for 10 mins each at 4˚C. The immunoprecipitated DNA was eluted by elution buffer (50mM Tris-HCl pH=8.0, 10mM EDTA, 1% SDS) and reverse-crosslinked by overnight incubation at 65˚C, followed by RNase A (Sigma-Aldrich) and proteinase K (Roche Diagnostics) treatment. DNA was purified using ChIP DNA Clean and Concentrator kit for qPCR (Zymo Research; Irvine, CA). qPCR of ChIP samples was performed by LightCycler® 480 SYBR Green I Master Mix (Roche Diagnostics). Primer sequences are available in **SI Table 1**.

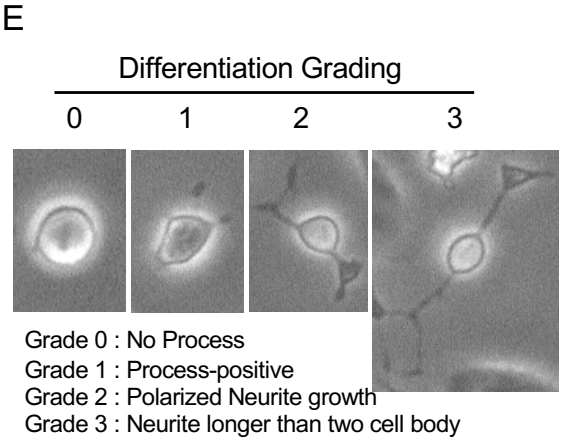
SI Figures


A
MKL-1
MKL-2
MS-1
WaGa
CVG-1
MCC alone
KRT Co-culture


B
Co-cultured
MS-1 in KRT medium
Trans-well co-culture
KRT medium
KRT
MS-1


D
CVG-1 shpanT-puro
PC
DAPI
NFL


C
MKL-2
CVG-1
shCtrl-eGFP
shpanT-eGFP
shCtrl-eGFP
shpanT-eGFP


E
Differentiation Grading
0
1
2
3
Grade 0 : No Process
Grade 1 : Process-positive
Grade 2 : Polarized Neurite growth
Grade 3 : Neurite longer than two cell body

(continued)

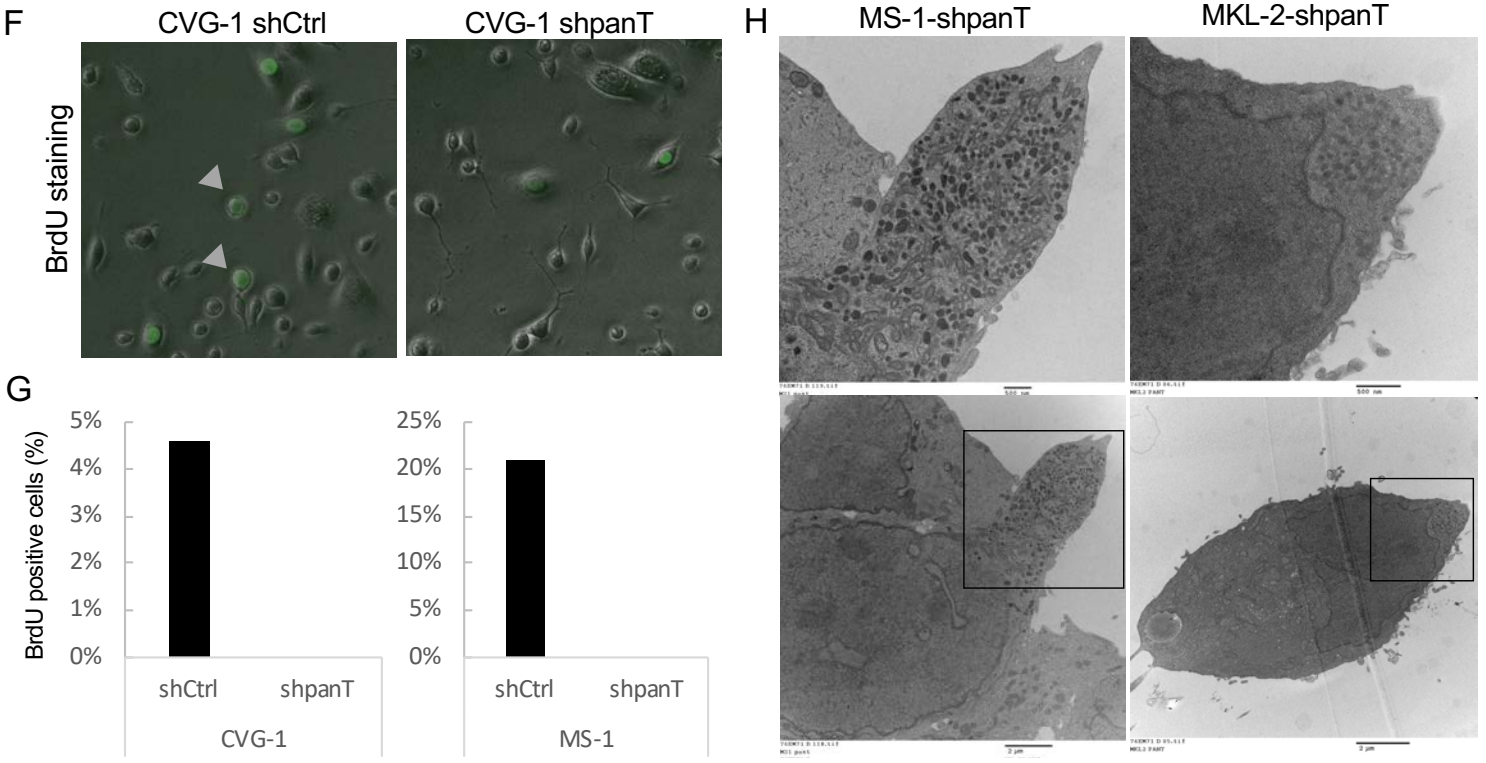
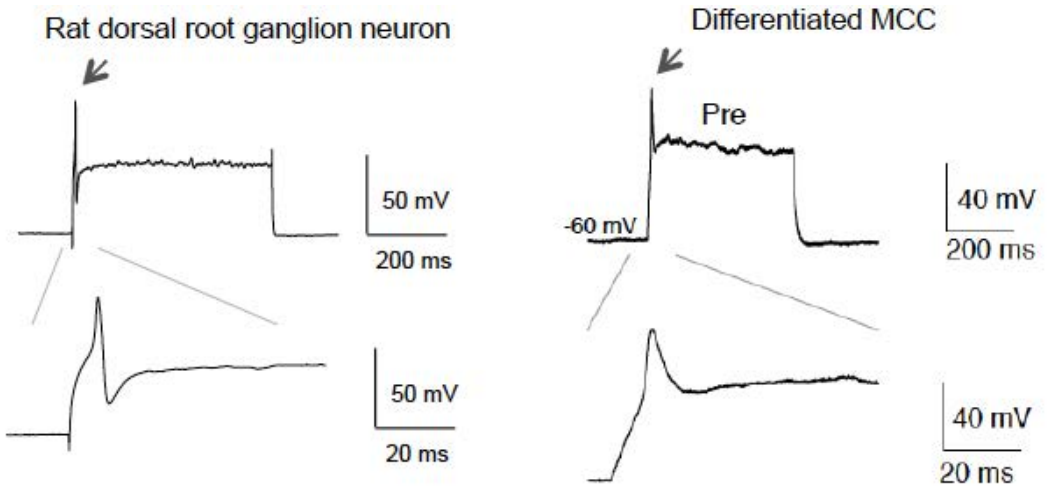


Fig. S1. T antigen knockdown in MCV-positive MCC cells induces neurite-positive cell formation under keratinocyte coculture. (*A*) Five MCV-positive MCC cell lines MKL-1, MKL-2, MS-1, WaGa and CVG-1 were cultured in KRT medium (top) or cocultured with human primary keratinocytes (middle and bottom panels). MCC cells exhibiting projections were observed only when cocultured with KRT (bottom arrowheads). (Top and middle row magnification: 10x; bottom row: 40x) (*B*) Requirement of the direct cell-to-cell attachment for MCC cell differentiation. The differentiated MS-1 cells were observed only when they were in a direct attachment (arrowheads). Phase contrast images from MS-1 cells cultured with KRT (top), MS-1 grown in KRT growth medium and MS-1 cells that were trans-well cocultured with KRT were shown. Bottom images are magnified images above (boxes). (Magnification: top row 10x and bottom row: 40x) (*C*) Transduction efficiency of three MCV-positive MCC cell lines MS-1, MKL-2, CVG-1 infected either with shCtrl-eGFP or with shpanT-eGFP lentiviruses. GFP-fluorescence images show transduction efficiency of shRNA. (Magnification: 40x) (*D*) NFL immunostaining for differentiated shpanT-puro transduced CVG-1 cells. The elongated neurite structure in differentiated cells (grade 3) are neurofilament L (NFL)-positive axons. (Magnification: 40x) (*E*) The neurite formation was graded from 0 to 3. Grade 0 represents no process, grade 1 cellular process-positive, grade 2 polarized projections, and grade 3 with projection longer than two cell body size. A right panel demonstrates images from CVG-1 cells cocultured with KRT, which represent each differentiation grading. (Magnification: 40x) (*F*) Representative BrdU staining in CVG-1 cells. To determine cell cycle commitment of MCC cells cocultured with KRT, CVG-1 cells transduced with shpanT or shCtrl were labeled with BrdU for 4h and stained with BrdU antibody (green) by immunofluorescence staining. Gray arrowheads indicate BrdU-incorporated CVG-1 cells while other BrdU-positive cells are cocultured KRTs with larger cytoplasm. (Magnification: 40x) (*G*) T antigen knockdown ablates S phase cell cycle commitment in MS-1 and CVG-1 cells. Graphs indicate quantitation of BrdU-positive MCC cells. The percentage of BrdU positive cells was determined in MS-1 and CVG-1 cells. (*H*) Accumulation of the electron-dense vesicles in the polarized cellular projections in shpanT-transduced MS-1 and MKL-2 cocultured with KRT.

A



B

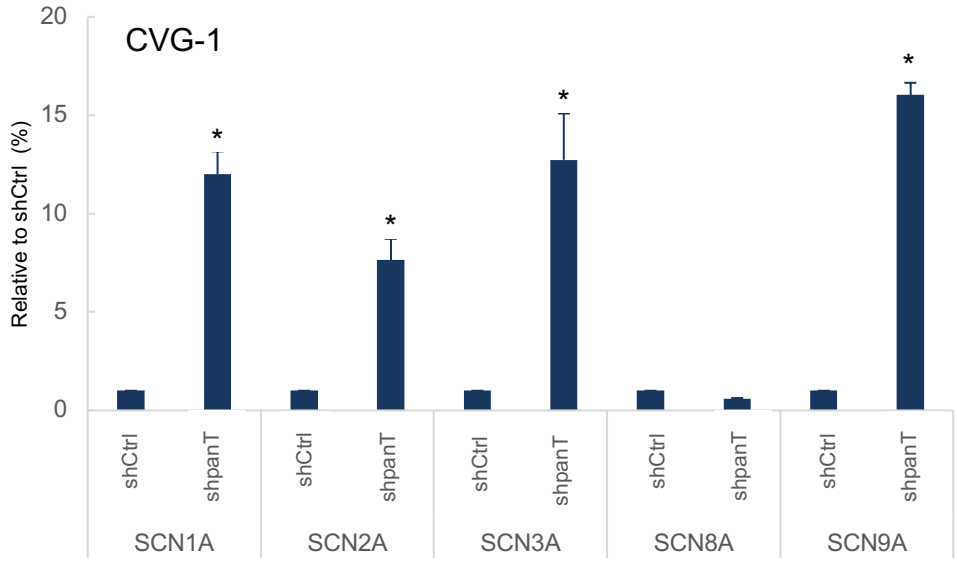


Fig. S2. (*A*) Comparison of a typical action potential between Rat dorsal root ganglion neuron and differentiated MCC. Current injection into Rat DRG neuron (350 pA, 500 ms) or into differentiated MCC cell line MS-1 (120 pA, 400 ms) induces an action potential (arrows). DRG recording was performed previously and used as a control for comparison. (*B*) Expression of voltage-gated sodium channel genes *SCN1A*, *SCN2A*, *SCN3A*, and *SCN9A* is elevated by T antigen knockdown in CVG-1. qRT-PCR was performed with β-actin as a normalization control. Relative abundance to shCtrl was determined by ΔΔCt method ($2^{-\Delta\Delta Ct}$). Data are shown as average ± SD *$p <$ 0.05 (Student *t* test).

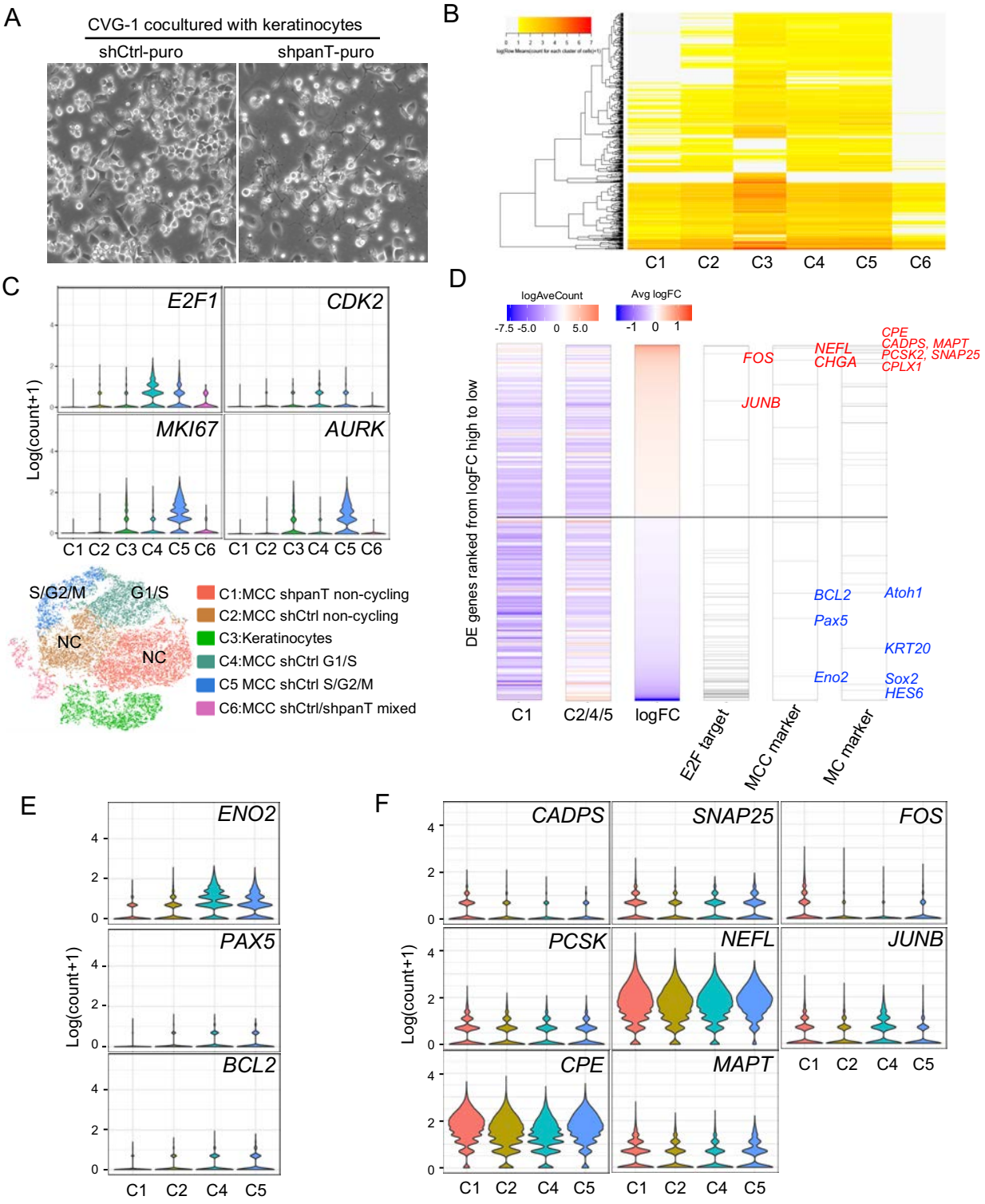


Fig. S3. Single cell transcriptome analysis of shpanT- and shCtrl-transduced CVG-1 cells cocultured with human keratinocytes (*A*) CVG-1 cells transduced with shCtrl-puro or with shpanT-puro that were cocultured with KRT. The mixture of CVG-1 cells and keratinocytes was directly subjected to 10x Genomics single cell sequencing. (Magnification: 10x) (*B*) Unique gene expression patterns for each cluster (C1~C6) obtained by the k-clustering algorithm. (*C*) Violin plots for additional cell cycle markers and E2F target genes in clusters defined the tSNE below it. (*D*) Differentially expressed (DE) E2F target, MCC markers, and Merkel cell (MC) marker genes between C1 (shpanT) and C2/4/5 (shCtrl) are ranked by log-transformed fold change (logFC) according to the SI Dataset 1. Leftmost two bars show the expression level of DE genes. The middle bar shows the log2-fold change. Positive values indicate highly expressed genes in C1. Genes induced by shpanT (in C1) are shown in red and those reduced in blue. (*E*) Expression of MCC/Merkel cell marker genes in the clusters defined in (*C*). (*F*) Expression of synaptic release machinery genes and neurite/axonal markers.

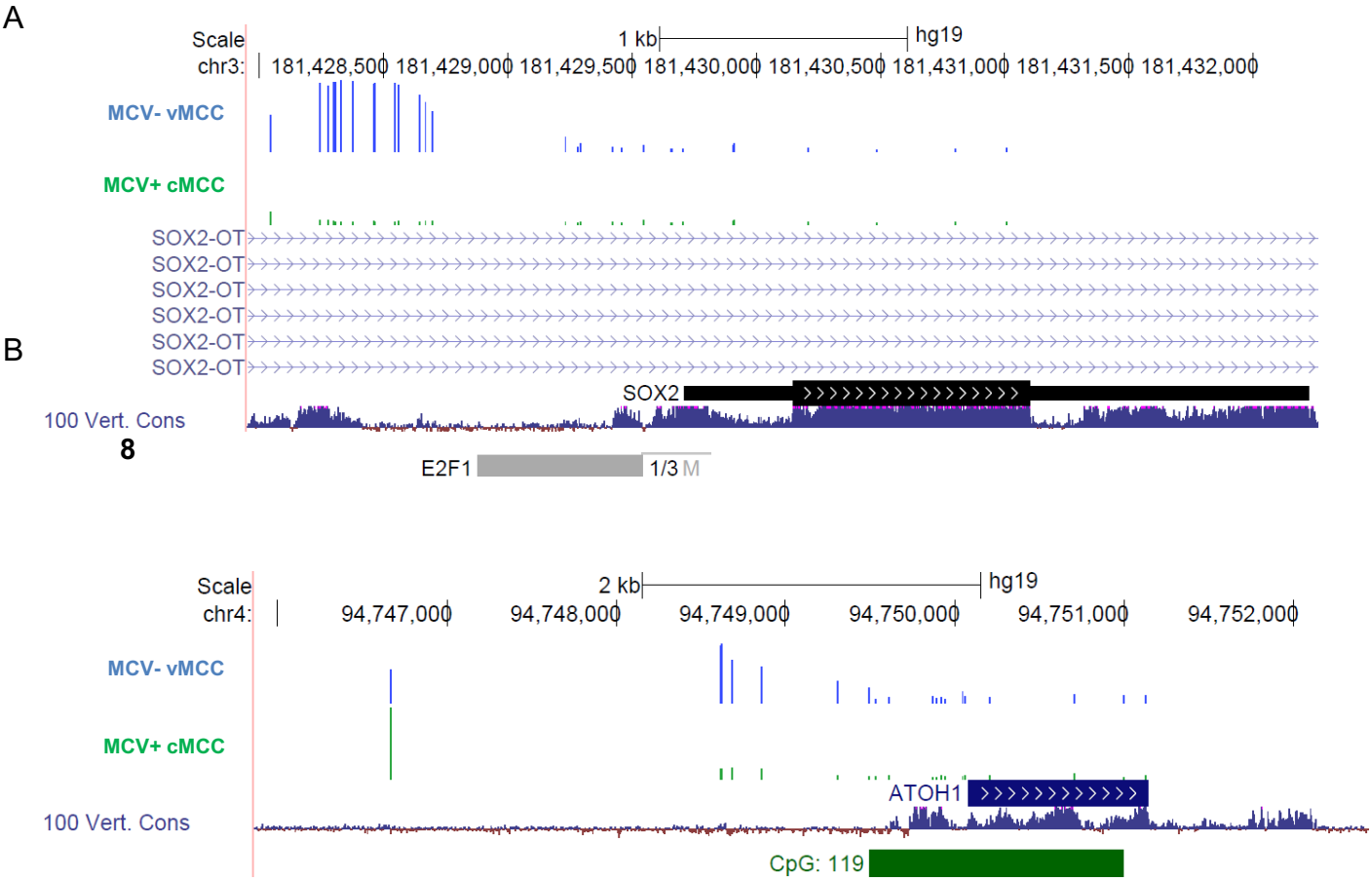


Fig. S4. Methylation patterns of (*A*) *Sox2* and (*B*) *ATOH1* promoters determined by methylation array. The MCV+ MCC cell lines with classic morphology (MCV+ cMCC, n=3 green) and variant MCV- MCC cell lines with variant morphology (MCV- vMCC, n=2 blue) indicate the degree of methylation in percent over the classic and variant MCC cell lines (15, 16). Each bar represents single CpG site. A green box represents the location of CpG islands. The 100 vertebrates Conservation (100 Vert.Cons) track displays the conservation of DNA sequence within vertebrates. A high degree of conservation outside of the exon and the presence of CpG island suggest the possible presence of promoter region. The potential E2F binding site is shown below the promoter alignment. Closed and colored boxes indicate exons. Black indicates that the feature has a corresponding entry in the Protein Data Bank. Dark blue indicates that the transcript has been reviewed or validated by either RefSeq, SwissProt or CCDS staff.

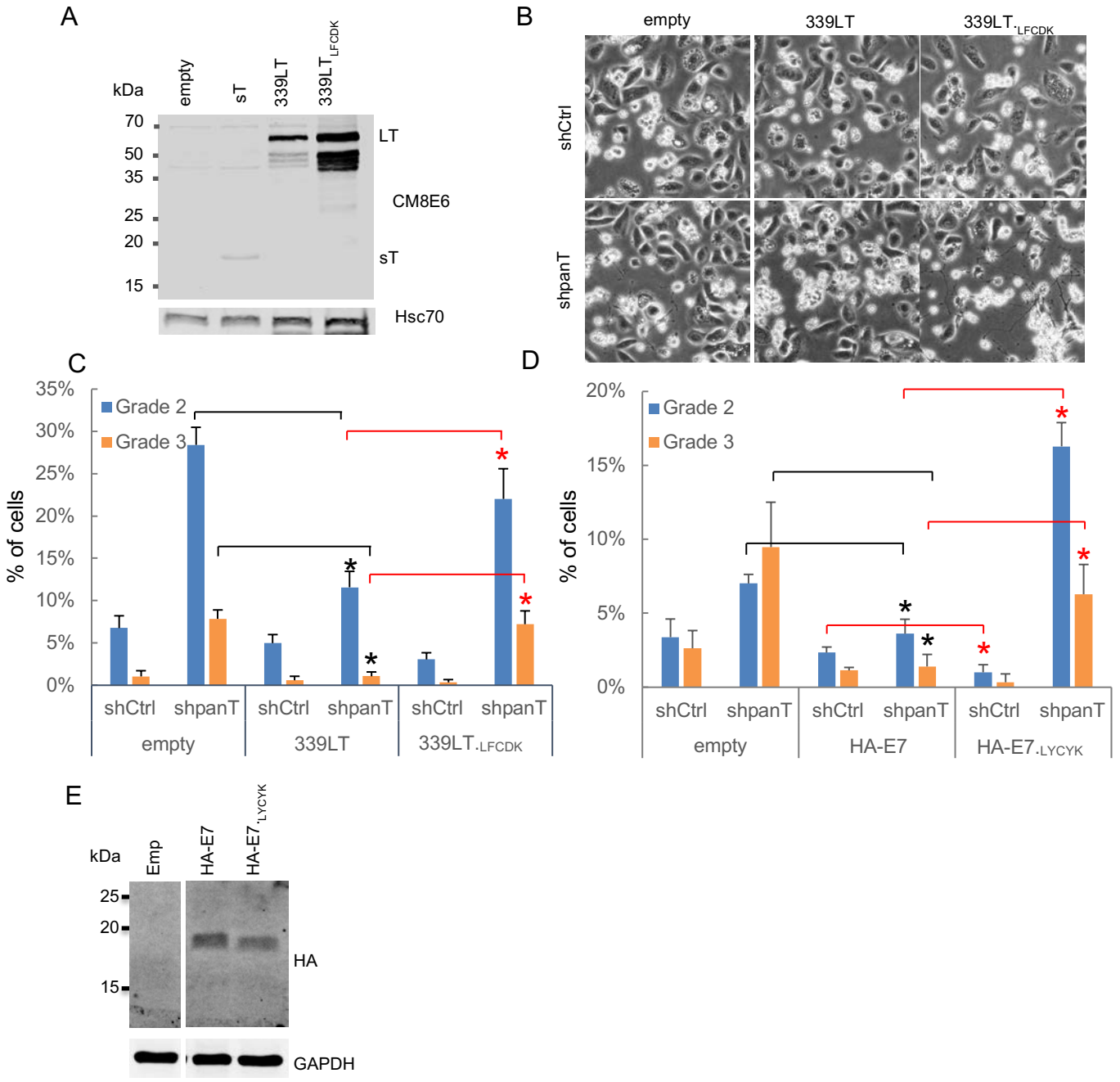


Fig. S5. Tumor-derived LT inhibits MS-1 cell differentiation induced by T antigen knockdown through its Rb-targeting domain. (*A*) Expression of codon-optimized sT, truncated 339LT, and the Rb-binding mutant 339LT.$_{LFCDK}$ proteins in KRT used for qRT-PCR experiments in Fig. 5*A*. MCV sT and LT protein expression was detected by immunoblot using CM8E6 antibody that recognizes a common N-terminal epitope. (*B*) MS-1 cells transduced with shCtrl-neo or shpanT-neo were selected with medium containing 300 μg/mL G418 for 5 days and then cocultured with KRT for 5 days. Expression of the shpanT-resistant codon-optimized truncated LT (339LT) inhibits T antigen knockdown-induced differentiation of MS-1 whereas the expression of Rb-binding mutant (339LT.$_{LFCDK}$) or transduction with empty vectors do not. (Magnification: 10x) (*C*) MS-1 differentiation was evaluated according to the grading criteria in Fig. S1*E*. Differentiation of MS-1 cells displaying grade 2 and grade 3 morphologies are significantly inhibited by 339LT expression. Error bars indicate standard errors. Significance was determined for the comparison between empty and 339LT (black bars) or between 339LT and 339LT.$_{LFCDK}$ (red bars). Data are shown as average ± SE *$p < 0.05$ (Student *t* test). (*D*) Expression of HA-E7 inhibits T antigen knockdown-induced neurite formation in MS-1 whereas the expression of Rb-binding mutant (E7.$_{LYCYK}$) cancels its effect. (*E*) HA-E7 and HA-E7.$_{LYCYK}$ expression was confirmed by anti-HA immunoblots. GAPDH was used as a loading control.

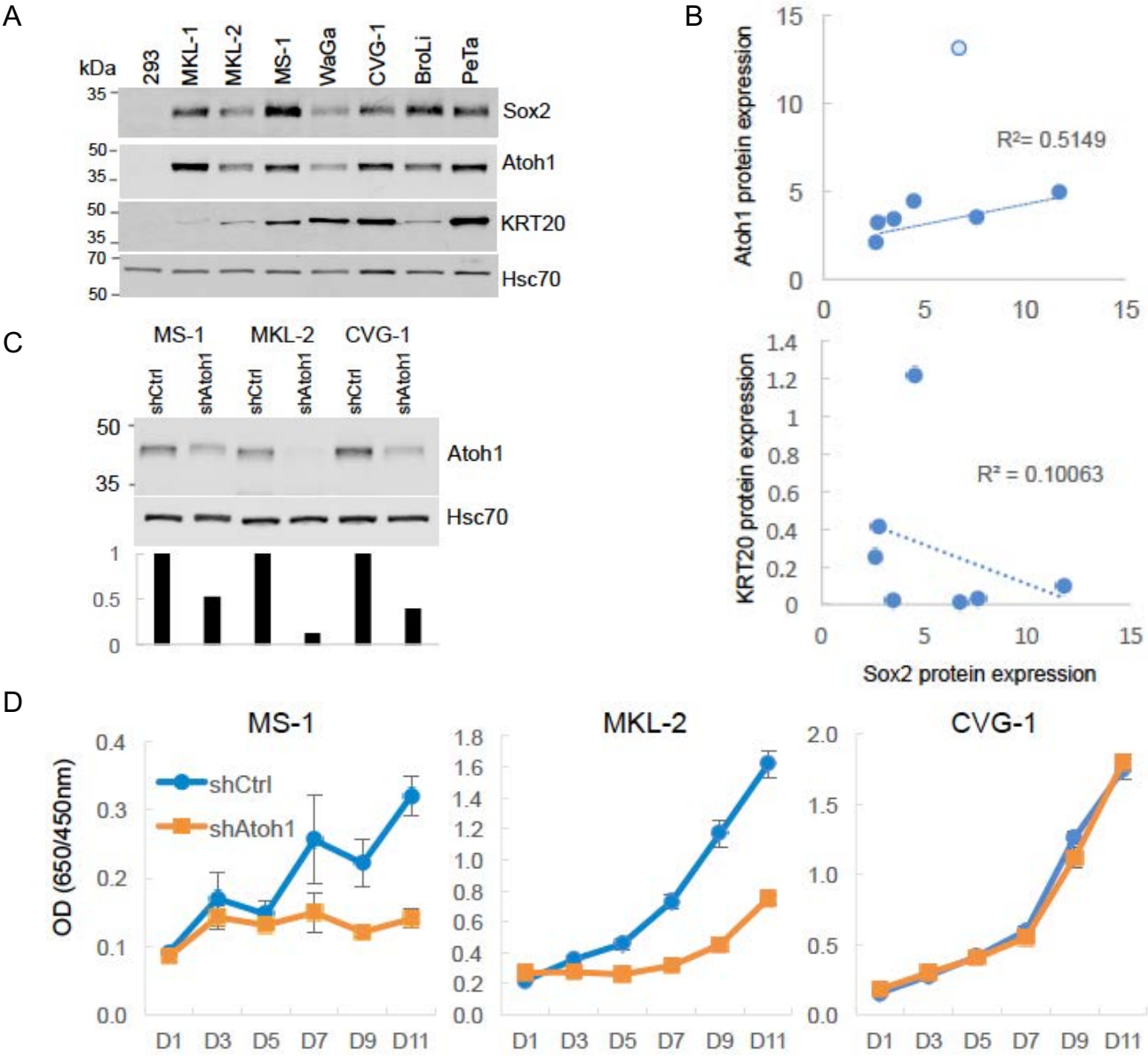


Fig. S6. (*A*) All MCV-positive MCC cell lines tested express Sox2, Atoh1, KRT20 proteins. Immunoblot experiments were performed in multiple MCV-positive MCC cells with 293 cells as a control. Hsc70 was used as a loading control. Images were captured by LI-COR Odyssey scanner. (*B*) Sox2 protein expression is correlated with the Atoh1 protein expression, but not with KRT20 protein expression. Respective protein expression was quantitated by LI-COR and normalized by Hsc70 expression. The calculated values were plotted for regression analysis. (*C*) shRNA knockdown of Atoh1 in MKL-2 and CVG-1. A LI-COR quantitative immunoblot experiment was performed to confirm knockdown of Atoh1 protein expression by shAtoh1. Graph indicates a quantitation result of above immunoblots normalized by Hsc70. (*D*) Effect of Atoh1 knockdown on cell proliferation. Atoh1 knockdown inhibits MS-1 and MKL-2 proliferation. The samples in C. were subjected to WST-8 cell proliferation assays.

SI Table 1: Primers used for this study are listed in the table below.

Primers for qRT-PCR	
CCK-F	CGTAGGCAGCTGAGGGTATC
CCK-R	CCGGTCACTTATCCTGTGGC
SCN1A-F	GTTCATGGGCAACCTGAGGA
SCN1A-R	GGACATTGGCCTGCATCAGA
SCN2A-F	TGTGACAGAGTTTGTGGACCT
SCN2A-R	GGTTGCCCATGAACAACTGC
SCN3A-F	CCGGAAAAAGGTGCTGTGAC
SCN3A-R	TGGCTTTGGTTTGTTCTCATCA
SCN8A-F	CAAGAAACCACCAAAGGCCG
SCN8A-R	AGGCCGTGGCACTAAATCTGA
SCN9A-F	CCGTTTCAATGCCACACCTG
SCN9A-R	ACTCTGAAAGTTCGAAGAGCTG
Sox2-F	GGGAAATGGGAGGGGTGCAAAAGAGG
Sox2-R	TTGCGTGAGTGTGGATGGGATTGGTG
Atoh1-F	GCAGGAGGAAAACAGCAAAA
Atoh1-R	ACTTGCCTCATCCGAGTCAC
panT-F	GCTCCTAATTGTTATGGCAACAT
panT-R	GCTCCAAAGGGTGTTCAATTC
18S-F	GGACACGGACAGGATTGACA
18S-R	ACCCACGGAATCGAGAAAGA
ACTB-F	GGACTTCGAGCAAGAGATGG
ACTB-R	AGCACTGTGTTGGCGTACAG
Oligonucleotide for shRNA construction	
shSox2.1-F	CCGGCGCTCATGAAGAAGGATAAGTTCAAGAGACTTATCCTTCTTCATGAGCGTTTTTG
shSox2.1-R	AATTCAAAAACGCTCATGAAGAAGGATAAGTCTCTTGAACTTATCCTTCTTCATGAGCG
shSox2.2-F	CCGGCTGCCGAGAATCCATGTATATTCAAGAGATATACATGGATTCTCGGCAGTTTTTG
shSox2.2-R	AATTCAAAAACTGCCGAGAATCCATGTATATCTCTTGAATATACATGGATTCTCGGCAG
shAroh1-F	CCGGTGAAAGTGCGGGAACAGCTGTTCAAGAGACAGCTGTTCCCGCACTTTCATTTTTTG
shAtoh1-R	AATTCAAAAAATGAAAGTGCGGGAACAGCTGTCTCTTGAACAGCTGTTCCCGCACTTTCA
Primers for ChIP	
ChIP-NC_F	TTGTGTAGCACCAGGGTCA
ChIP-NC_R	GCCTGGCCTGCAAATTATTA
ChIP-ACTB_F	TCGAGCCATAAAAGGCAACT
ChIP-ACTB_R	GAGGGGAGAGGGGGTAAAA
ChIP-Atoh1 A_F	GAGCTTCTGATTGGCTCTTT
ChIP-Atoh1 A_R	CCAATTCCCAGGTGAACG
ChIP-Atoh1 B_F	TTCTTTGTTCTCATGGATGGT
ChIP-Atoh1 B_R	TTTGGGCTCTGCTGAAATAA
ChIP-Atoh1 C_F	TGTTGAAACTCAAAGGAAAGTG
ChIP-Atoh1 C_R	AATGTGTAGGATGCCCAAAT
ChIP-Atoh1 D_F	GTCCTCTGCACACAAGAA
ChIP-Atoh1 D_R	CGTGTCTTCTCAGCCAAT

SI References

1. Shuda M*, et al.* (2008) T antigen mutations are a human tumor-specific signature for Merkel cell polyomavirus. *Proc Natl Acad Sci U S A* 105(42):16272-16277.
2. Velasquez C*, et al.* (2018) Characterization of a Merkel Cell Polyomavirus-Positive Merkel Cell Carcinoma Cell Line CVG-1. *Front Microbiol* 9:713.
3. Haraguchi T*, et al.* (2016) Dynamics and plasticity of the epithelial to mesenchymal transition induced by miR-200 family inhibition. *Sci Rep* 6:21117.
4. Houben R*, et al.* (2010) Merkel cell polyomavirus-infected Merkel cell carcinoma cells require expression of viral T antigens. *J Virol* 84(14):7064-7072.
5. Shuda M, Kwun HJ, Feng H, Chang Y, & Moore PS (2011) Human Merkel cell polyomavirus small T antigen is an oncoprotein targeting the 4E-BP1 translation regulator. *J Clin Invest* 121(9):3623-3634.
6. Shuda M*, et al.* (2009) Human Merkel cell polyomavirus infection I. MCV T antigen expression in Merkel cell carcinoma, lymphoid tissues and lymphoid tumors. *Int J Cancer* 125(6):1243-1249.
7. Kwun HJ*, et al.* (2009) The minimum replication origin of merkel cell polyomavirus has a unique large T-antigen loading architecture and requires small T-antigen expression for optimal replication. *J Virol* 83(23):12118-12128.
8. McCarthy DJ, Campbell KR, Lun AT, & Wills QF (2017) Scater: pre-processing, quality control, normalization and visualization of single-cell RNA-seq data in R. *Bioinformatics* 33(8):1179-1186.
9. Butler A, Hoffman P, Smibert P, Papalexi E, & Satija R (2018) Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol* 36(5):411-420.
10. Finak G*, et al.* (2015) MAST: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data. *Genome Biol* 16:278.
11. Maksimovic J, Phipson B, & Oshlack A (2016) A cross-package Bioconductor workflow for analysing methylation array data. *F1000Res* 5:1281.
12. Aryee MJ*, et al.* (2014) Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics* 30(10):1363-1369.
13. Ritchie ME*, et al.* (2015) limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 43(7):e47.
14. Fortin JP*, et al.* (2014) Functional normalization of 450k methylation array data improves replication in large cancer studies. *Genome Biol* 15(12):503.
15. Van Gele M*, et al.* (2004) Gene-expression profiling reveals distinct expression patterns for Classic versus Variant Merkel cell phenotypes and new classifier genes to distinguish Merkel cell from small-cell lung carcinoma. *Oncogene* 23(15):2732-2742.
16. Fischer N, Brandner J, Fuchs F, Moll I, & Grundhoff A (2010) Detection of Merkel cell polyomavirus (MCPyV) in Merkel cell carcinoma cell lines: cell morphology and growth phenotype do not reflect presence of the virus. *Int J Cancer* 126(9):2133-2142.