Supplemental Data

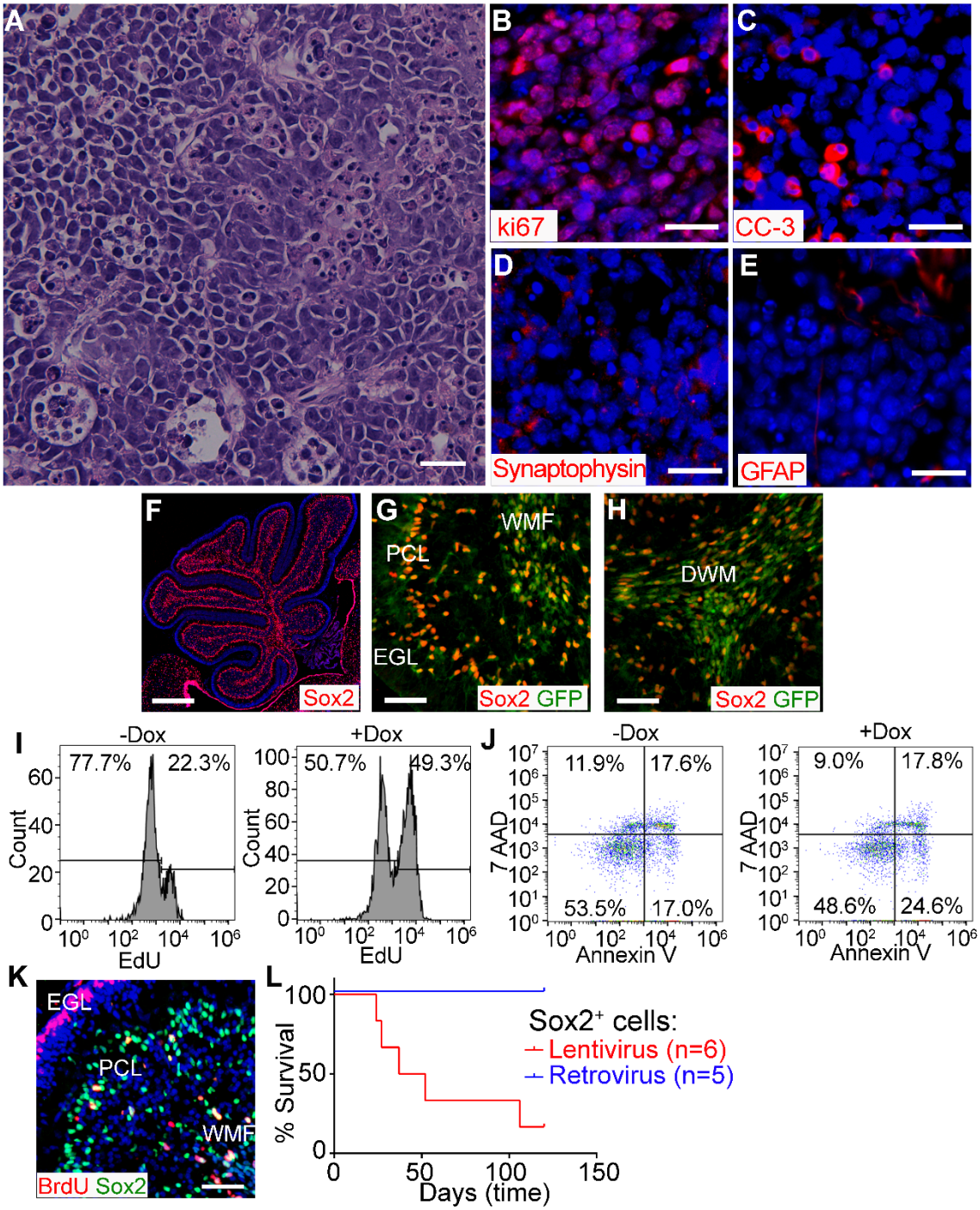

Figure S1. *MYC* overexpression alone drives tumorigenesis from Sox2⁺ cells in the postnatal cerebellum. (A) H&E staining in section of tumor generated from total cerebellar cells. (B-E) IHC staining in section of tumor generated from total cerebellar cells. Nuclei were counter-stained with DAPI (blue). Scale bars = 25 μm. (F) IHC staining of cerebellar sections from wild type mice at P5. Scale bars = 500 μm. (G) and (H) IHC staining of cerebellar sections from Sox2^{-GFP} mice at P5. GFP was shown as green. EGL, External Granular Layer; PCL, Purkinje Cell Layer; WMF, White Matter Folium; DWM, Deep White Matter. Scale bars = 100 μm. (I) and (J) Proliferation and apoptosis of Sox2⁺ cells transduced by *MYC*. (K) BrdU and Sox2 staining of cerebellar sections from wild type mice at P5 Scale bars = 50 μm. (L) Survival curve of mice transplanted with P5 Sox2⁺ cells infected with *MYC* lentivirus or retrovirus.

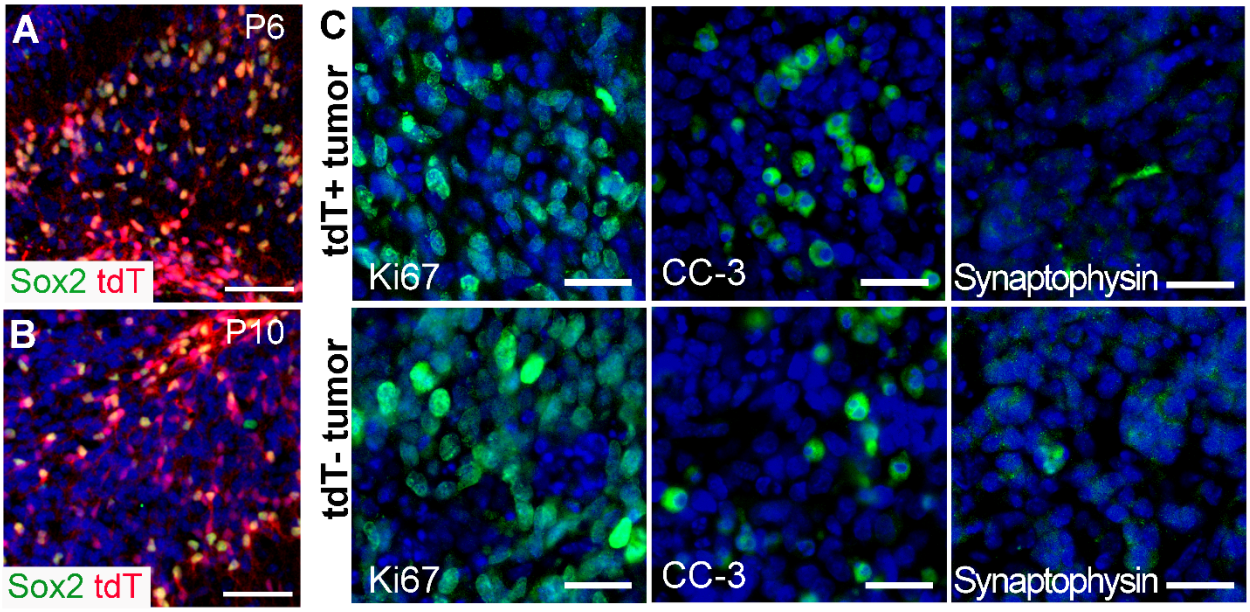

Figure S2. Tracing tumor cell-of-origin with tdT labeled Sox2$^+$ cells. (A) and (B) tdT expression permanently labels Sox2$^+$ cells and their progenies. Sox2-CreER^{T2}/Rosa-CAG-LSL-tdTomato mice were given tamoxifen at P5 and brains were collected at P6 (A) or P10 (B). Frozen sections were stained with antibody specific for Sox2 (green). tdT$^+$ is shown as red. Scale bars = 50 μm. (C) Characterization of tdT$^+$ and tdT$^-$ tumors. Frozen sections of tumor generated from Sox2-CreER^{T2}/Rosa-CAG-LSL-tdTomato cells treated with or without 4-OHT (tdT$^+$ tumor or tdT$^-$ tumor, respectively) were stained with antibodies (green) specific for Ki67, CC-3 or Synaptophysin. Scale bars = 25 μm. Nuclei were counter-stained with DAPI (blue).

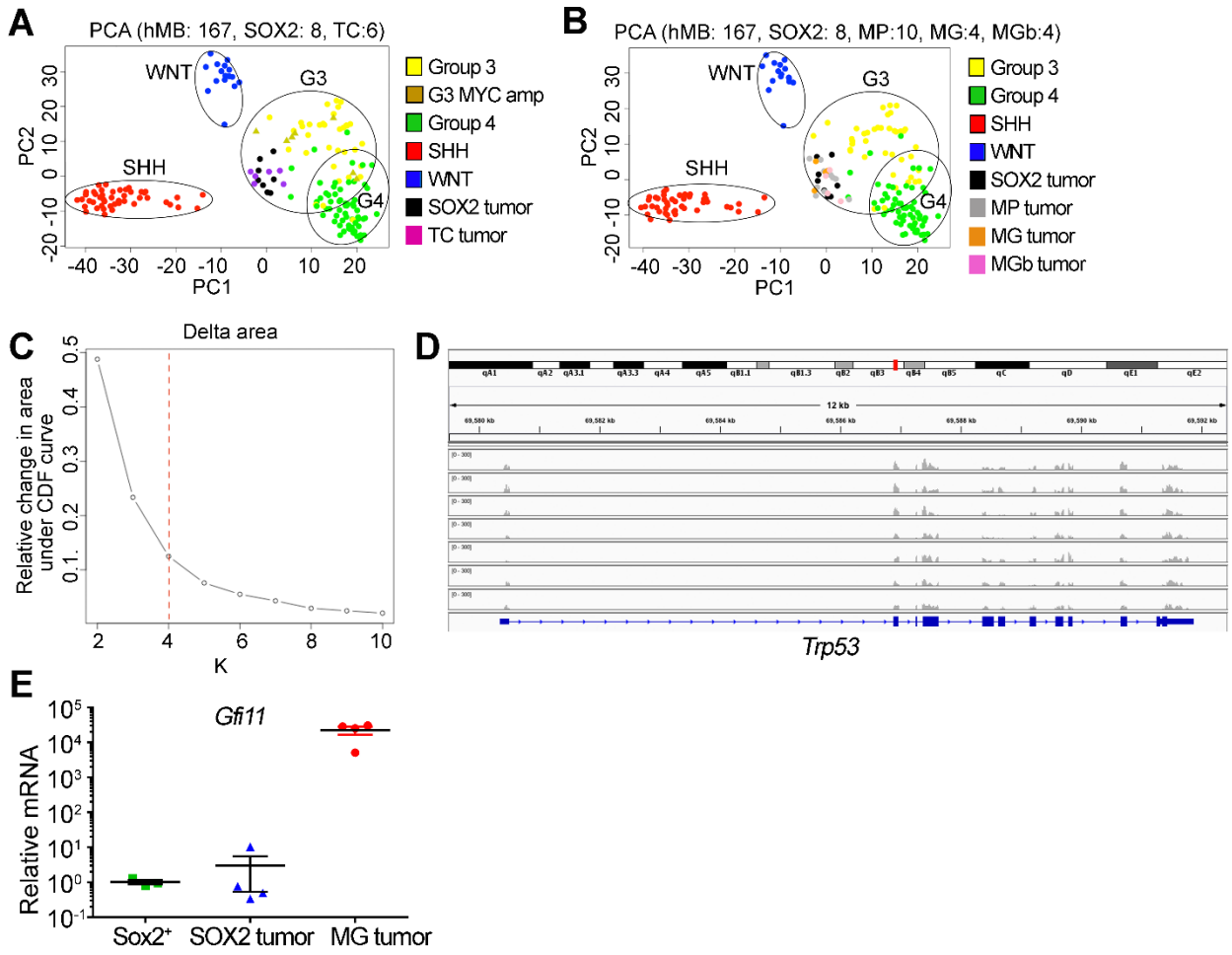

Figure S3. *MYC*-driven tumors resemble human Group 3 medulloblastoma at the molecular level. (A) and (B) PCA for gene expression (RPKM) RNA-seq data from human medulloblastoma, SOX2 tumors (n=8) and TC tumors (n=6) combined with Affy tumor data from MP (n=10), MG (n=4) and MGb (n=4) models based on batch effect adjustment for top 500 most variable genes. The axes represent two linear combinations of genes that account for most of the variance in the original data set. SOX2 tumors were generated from Sox2$^+$ cells with *MYC* overexpression. TC tumor were generated from total cerebellar cells with *MYC* overexpression. MP tumor , MG tumor or MGb tumor were generated from CD133$^+$ cells with *MYC* + *DNp53*, *MYC* + *Gfi1* or *MYC* + *Gfi1b* overexpression respectively. (C) Cumulative cluster distribution analysis for combined mice and human batch effect adjusted RNA-seq data demonstrates that ~80% of relative change is covered by assignment of k=4 clusters, representing the MB tumor subgroups. (D) Visualization of the *Trp53* gene sequence acquired from RNA-seq using IGV. Peaks with gray color indicate sequence without mutations. (E) Expression of *Gfi1* in Sox2$^+$ cerebellar cells, SOX2 tumors or MG tumors assessed by qRT-PCR. mRNA levels of *Gfi1* in indicated tumors were normalized to the mean mRNA levels of *Gfi1* in Sox2$^+$ cerebellar cells (value=1).

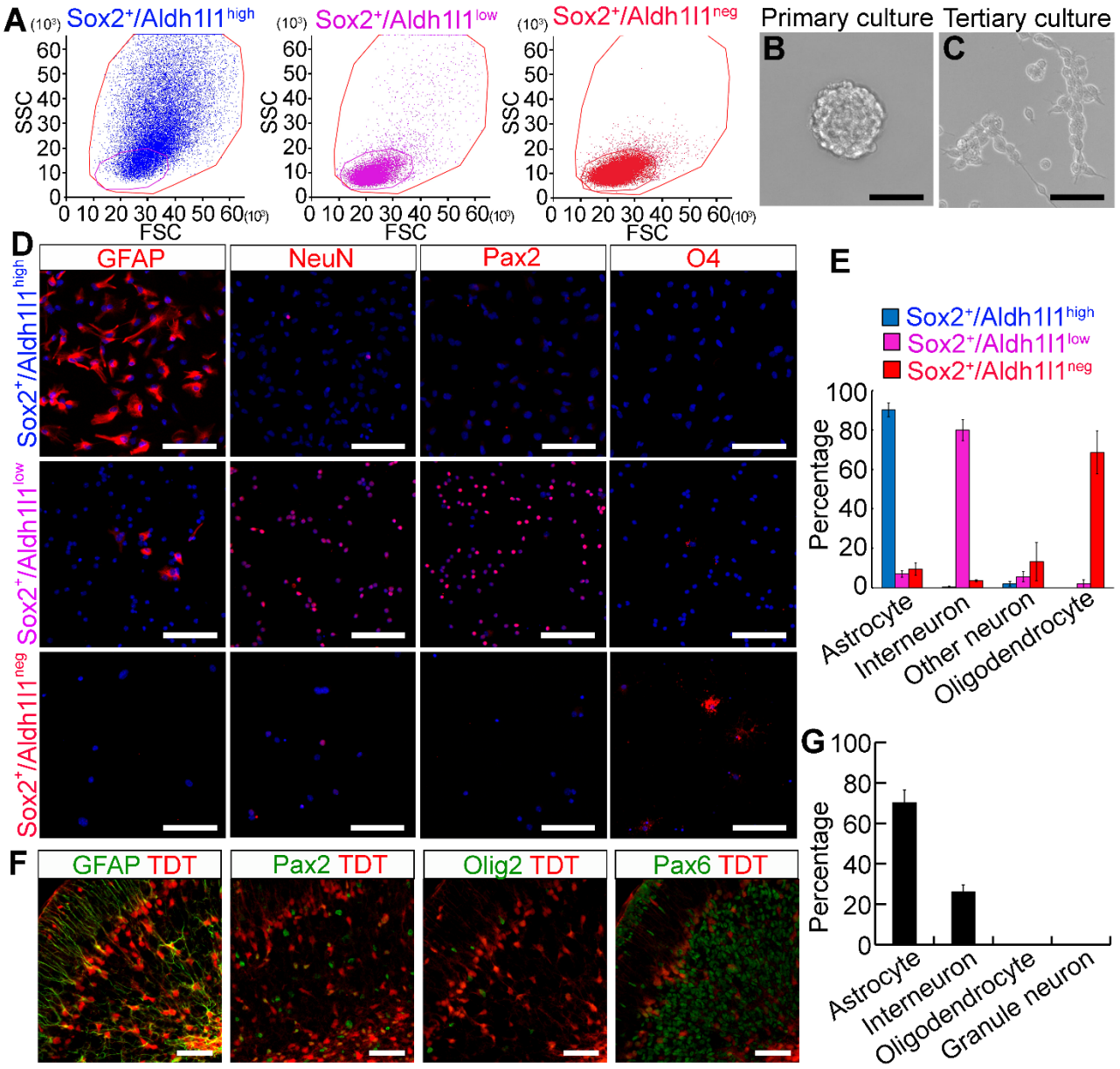

Figure S4. Characterization of Sox2$^+$/Aldh1L1^{high}, Sox2$^+$/Aldh1L1^{low} and Sox2$^+$/Aldh1L1^{neg} cells. (A) FACS-sorting of Sox2$^+$/Aldh1L1^{high}, Sox2$^+$/Aldh1L1^{low}, and Sox2$^+$/Aldh1L1^{neg} cells. Aldh1L1^{high} cells are larger in size and exhibit higher granularity than Sox2$^+$/Aldh1L1^{low} and Sox2$^+$/Aldh1L1^{neg} cells. Sox2$^+$/Aldh1L1^{low} and Sox2$^+$/Aldh1L1^{neg} cells were further gated by SSC and FSC (circles in the middle and right panels) to collect small cells with low granularity. (B) Primary neurospheres derived from Sox2$^+$/Aldh1L1^{high} cells. Scale bars = 500 μm. (C) Sox2$^+$/Aldh1L1^{high} cells form astrocytic morphology when cultured for more than 3 passages. Scale bars = 500 μm. (D) and (E) Differentiation of freshly isolated Sox2$^+$/Aldh1L1^{high}, Sox2$^+$/Aldh1L1^{low} and Sox2$^+$/Aldh1L1^{neg} cells. Cells were stained for indicated antibodies (red). Nuclei were counter-stained with DAPI (blue). Scale bars = 100 μm. (E) Quantification of differentiation in (D). Percentage of astrocytes (GFAP$^+$), interneurons (Pax2$^+$), other neurons (NeuN$^+$/Pax2$^-$) or oligodendrocytes (O4$^+$) was calculated, Error bars represent the mean ± SD. (F) and (G) *In vivo* lineage tracing of Aldh1L1$^+$ cells. P5 Aldh1L1-CreER^{T2}/Rosa-CAG-LSL-tdTomato mice were treated with tamoxifen and cerebella were collected at P10. Frozen sections were stained with indicated antibodies (green). tdT was shown as red. Scale bars = 50 μm. (G) Quantification of cells in (F). Percentage of astrocytes (GFAP$^+$), interneurons (Pax2$^+$), oligodendrocytes (Olig2$^+$) or granule neurons (Pax6$^+$) in tdT$^+$ cells was calculated, Error bars represent the mean ± SD.

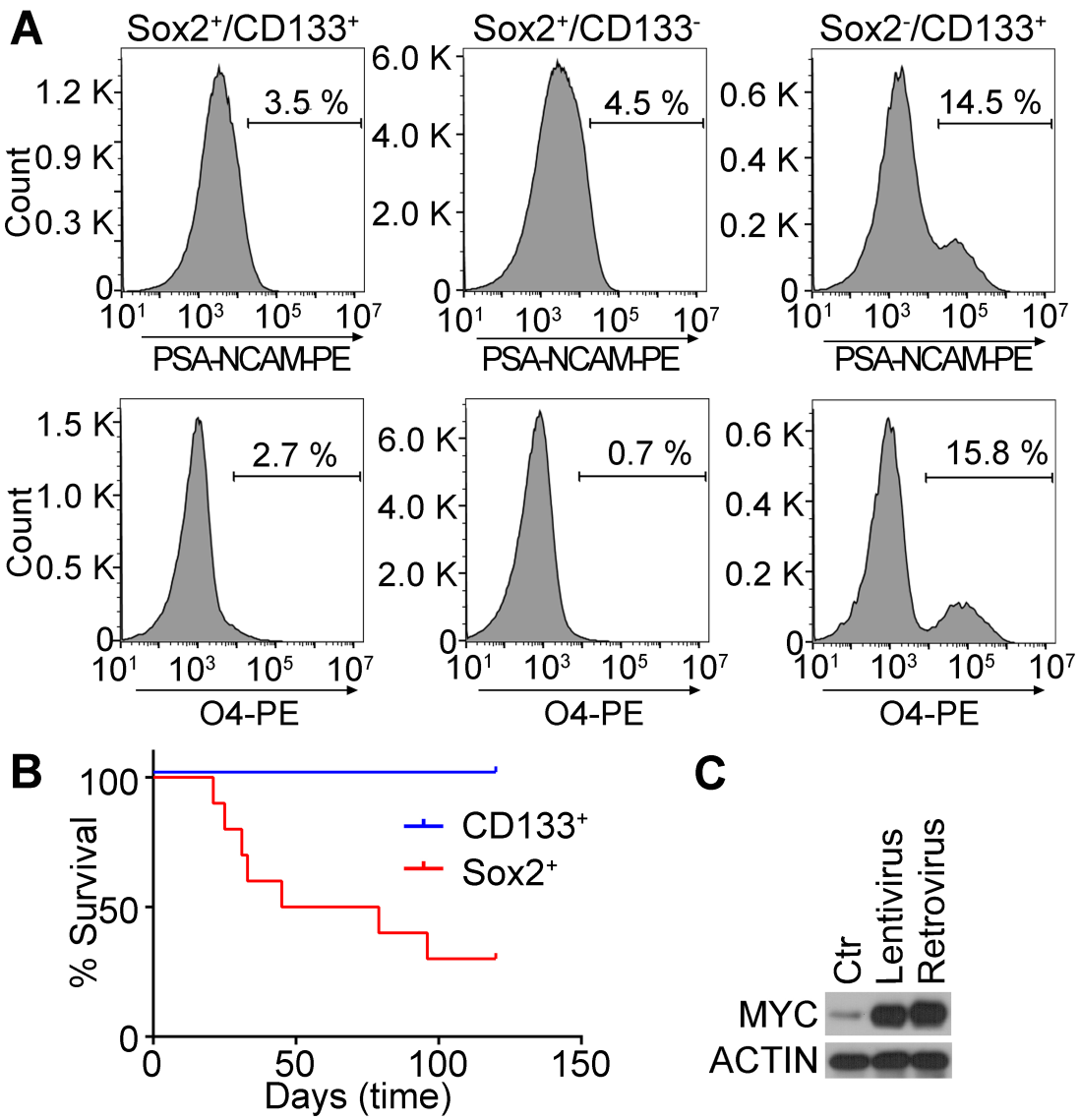

Figure S5. Characterization of Sox2⁺ vs. CD133⁺ cells in the early postnatal cerebellum. (A) Expression of PSA-NCAM or O4 in Sox2⁺/CD133⁺, Sox2⁺/CD133⁻ or Sox2⁻/CD133⁺ cells examined by flow cytometry. (B) Survival curve of mice transplanted with P5 Sox2⁺ cells infected with *MYC* lentivirus (n=10) or CD133⁺ cells infected with *MYC* retrovirus (n=6). (C) *MYC* expression in 293T cells infected with *MYC* lentivirus or retrovirus examined by western blot.

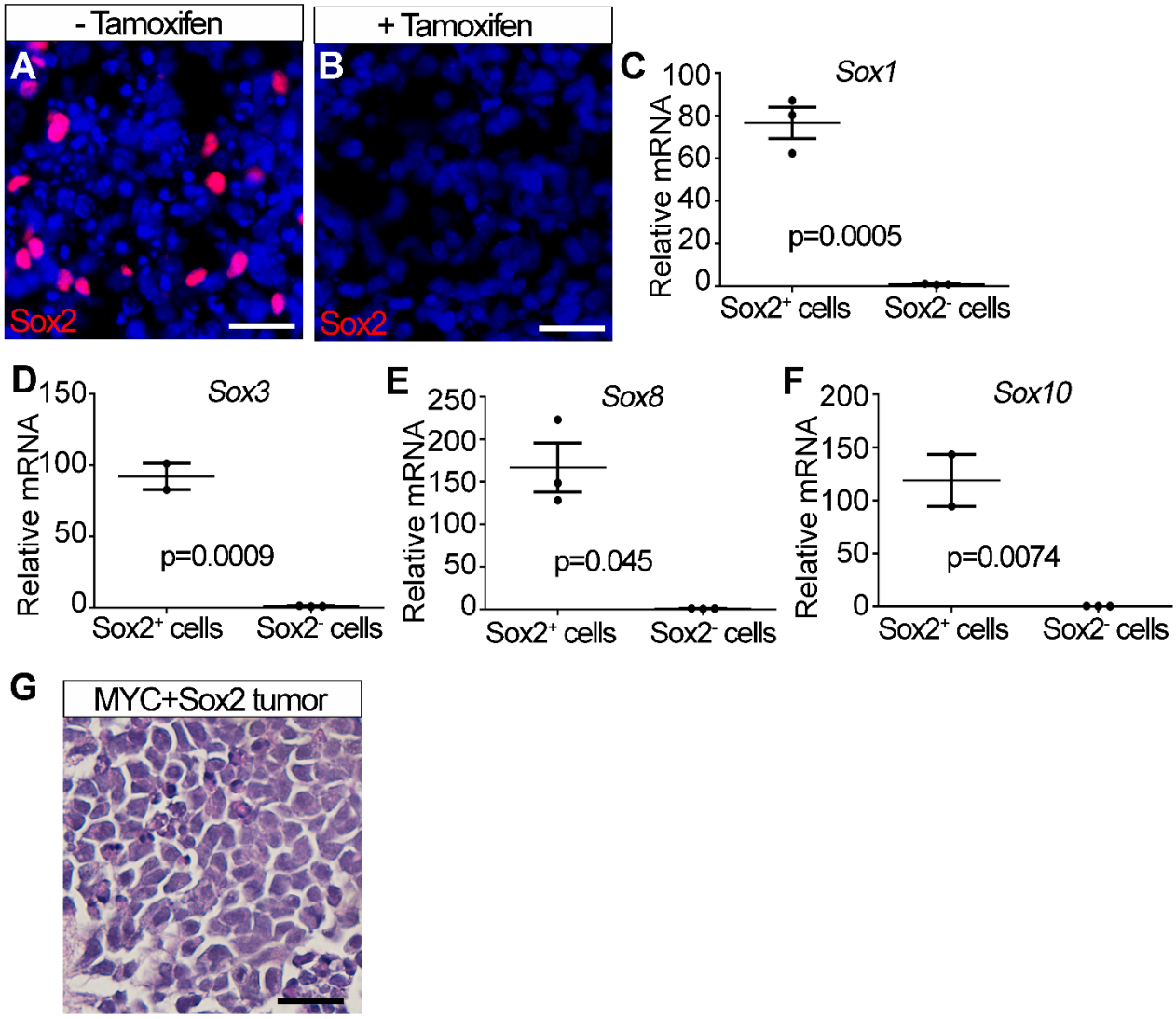

Figure S6. *Sox2* promotes *MYC*-induced tumorigenesis. (A) and (B) IHC staining of Sox2 (red) in tumor generated from mice transplanted with P5 Sox2-CreER^{T2}/Sox2^{loxp} cells with or without tamoxifen treatment. Scale bars = 25 μm. (C-F) Expression of *Sox1, Sox3, Sox8* and *Sox10* in P5 Sox2$^+$ cells and Sox2$^-$ cells assessed by qRT-PCR. mRNA levels of each gene in Sox2$^+$ cells were normalized to the mean mRNA levels of gene in Sox2$^-$ cells (value=1). Error bars represent the mean ± SEM. (G) H&E staining of tumor generated from Sox2$^-$ cells with *MYC*+*Sox2* overexpression. Scale bars = 25 μm.

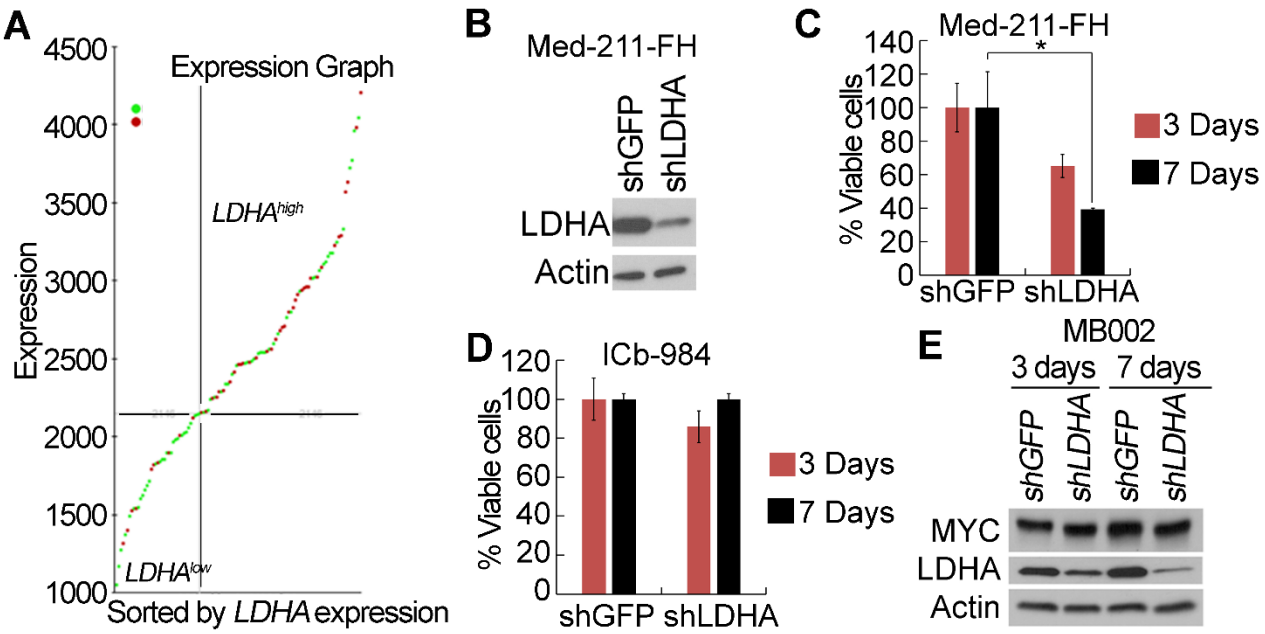

Figure S7. Inhibition of LDHA inhibits growth of *MYC*-driven medulloblastoma. (A) *LDHA* expression in human Group 3 medulloblastoma. (B) Knockdown of *LDHA* in Med-211-FH cells (Group 3 medulloblastoma PDX). Freshly isolated Med-211-FH cells were infected with retroviruses encoding *LDHA* shRNA or corresponding control shRNA for 24hr and then cultured for additional 6 days. Cells were collected and *LDHA* expression was examined by western blot. (C) Knockdown of *LDHA* inhibited cell growth of Group 3 medulloblastoma PDX *in vitro*. Freshly isolated Med-211-FH cells were infected with retroviruses encoding *LDHA* shRNA or corresponding control shRNA for 24hr and then cultured for totally 3 or 7 days. Cell viability was assessed by CellTiter-Glo® Assay. Viability of cells infected with *shGFP* virus was set to 100%. Error bars represent the mean ± SD. *p<0.05. (D) Knockdown of *LDHA* did not inhibit cell growth of SHH medulloblastoma PDX cells *in vitro*. Cell viability assay were similarly done as in (C) with ICb-984 cells. (E) *LDHA* knockdown didn't affect *MYC* expression in Goup 3 medulloblastoma PDX cells. Freshly isolated MB002 cells were infected with retroviruses encoding *LDHA* shRNA or corresponding control shRNA for 24hr and then cultured for totally 3 or 7 days. Western blot was used to determine *MYC* or *LDHA* expression in these samples.

Table S1. Top 100 selected genes in specific subgroups of human medulloblastoma.
Table S2. Top 500 variable genes selected for PCA applied in Fig. 2M.
Table S3. Top 500 variable genes selected for PCA applied in Fig. S3A.
Table S4. Top 500 variable genes selected for PCA applied in Fig. S3B.
Table S5. Expression of marker genes characterizing medulloblastoma molecular subgroups.
Table S6. Genes differentially expressed in SOX2 tumors compared to Sox2+ normal cerebellar cells.
Table S7. Pathways significantly enriched in SOX2 tumors compared to Sox2+ normal cerebellar cells.
Table S8. Genes involved in glycolysis differentially expressed in SOX2 tumors compared to Sox2+ normal cerebellar cells.
Table S9. Gene expression of glucose metabolism enzymes and *MYC* in human medulloblastoma and healthy cerebellum.

All Supplementary Tables are provided as excel files.

Supplemental Experimental Procedures

In Vivo Bioluminescent Imaging

Tumor-bearing animals were subjected weekly to bioluminescence imaging. Briefly, mice were given intraperitoneal injections of 150 ug/g D-Luciferine (Caliper Life Sciences, cat#12279) and anesthetized with 2.5% isoflurane. 10 min post-injection, the animals were imaged using the IVIS® Lumina III (IVIS-200) imaging system.

FACS Sorting

$Sox2^+$, $Sox2^-$ or $Math1^+$ cells were FACS-sorted from the cerebella of P5-6 $Sox2^{-GFP}$ or $Math1^{-GFP}$ mice as previously described [1, 2]. To isolate $Sox2^+/Aldh1L1^{high}$, $Sox2^+/Aldh1L1^{low}$, and $Sox2^+/Aldh1L1^{neg}$ cells, we crossed $Aldh1L1^{-GFP}$, Sox2-$CreER^{T2}$ with Rosa-CAG-LSL-tdTomato mice, treated P5 progeny with 0.2mg/g body weight tamoxifen (Sigma) for 24 hr, collected cerebella at P6 and FACS-sorted the cells by GFP and tdT. To remove possible contamination of astrocytes (which are usually bigger in size), $Sox2^+/Aldh1L1^{low}$ and $Sox2^+/Aldh1L1^{neg}$ subpopulations were further gated based on FSC and SSC to select small cells with lower granularity. To isolate $Sox2^+/CD133^+$, $Sox2^+/CD133^-$ and $Sox2^-/CD133^+$ cells, total cerebellar cells prepped from P5 $Sox2^{-GFP}$ mice were stained with PE or APC conjugated anti-CD133 antibody and FACS-sorted by GFP and PE or APC. To isolate $CD133^+$ cells, total cerebellar cells prepped from P5 C57BL/6J mice were stained with PE or APC conjugated anti-CD133 antibody and FACS-sorted by PE or APC. To isolate $Sox2^+$ cells or Sox2 KO cells by tdT expression, Sox2-$CreER^{T2}$/Rosa-CAG-LSL-tdTomato or Sox2-$CreER^{T2}$/$Sox2^{loxp}$/Rosa-CAG-LSL-tdTomato mice were administered tamoxifen (Sigma-Aldrich) at a single dose of 0.2 mg/g body weight by oral gavage at p5. Cerebellar were collected at P6 and tdT^+ cells were sorted. A BD Influx™ cell sorter was used in all experiments.

Flow cytometry analysis

To examine percentage of tdT^+ cells in tumor generated from *MYC* overexpressing Sox2-$CreER^{T2}$/Rosa-CAG-LSL-tdTomato cells with or without 4-hydroxytamoxifen treatment, cells were isolated from these tumors and analyzed by flow cytometry. To analyze Sox2 and CD133 expression, total cerebellar cells prepped from P5 C57BL/6J or $Sox2^{-GFP}$ mice were stained with APC conjugated isotype or anti-CD133 antibody and subjected to flow cytometry. To examine progenitors in $Sox2^+/CD133^+$, $Sox2^+/CD133^-$ and $Sox2^-/CD133^+$ population, sorted cells were stained PE conjugated isotype, anti-PSA-NCAM or anti-O4 antibody and subjected to flow cytometry.

Retrovirus and Lentivirus Plasmids

Lentivirus plasmids used include pICUE-*MYCT58A*-IRES-Luciferase (Doxycycline inducible *MYC*), pWPI-*MYCT58A*-IRES-GFP, pWPI-*MYCT58A*-IRES-Luciferase and pICUE-*Sox2*-IRES-Luciferase (Doxycycline inducible *Sox2*). Retroviruses used include pMKO.1 puro-*shGFP* [3] (targeting sequence: GCT ACG TCC AGG AGC GCA CCC TCG AGG GTG CGC TCC TGG ACG TAG C, a gift from William Hahn, Addgene plasmid # 10675), pMKO.1 puro-*shLDHA* [4] (targeting sequence: GCT ACA CAT CCT GGG CTA TTG CTC GAG CAA TAG CCC AGG ATG TGT AGC, a gift from Huadong Pei at The George Washington University), MSCV-*MYCT58A*-IRES-Luciferase , MSCV-*MycT58A*-IRES-Luciferase, MSCV-*MycT58A*-IRES-GFP, MSCV-IRES-GFP [5], MSCV-IRES-Luciferase (a gift from Scott Lowe, Addgene plasmid # 18760) and pMXs-*Sox2* (a gift from Shinya Yamanaka, Addgene plasmid # 13367) [6] . pICUE-*MYCT58A*-IRES-Luciferase was used to generate tumor from total cerebellar cells, $Sox2^+$ cells, subpopulations of $Sox2^+$ cells with different *Aldh1L1* level, and subpopulations of cells with differential *CD133* and *Sox2* expression. pICUE-*MYCT58A*-IRES-Luciferase, pICUE-*MYCT58A*-IRES-Luciferase + pMXs-*Sox2*, pWPI-*MYCT58A*-IRES-GFP, and pWPI-*MYCT58A*-IRES-GFP + pICUE-*Sox2*-IRES-Luciferase were used to generate tumor from $Sox2^-$ cells. To generate tumor for drug treatment, pWPI-*MYCT58A*-IRES-GFP + MSCV-IRES-Luciferase or pWPI-*MYCT58A*-IRES-Luciferase was used. MSCV-*MycT58A*-IRES-GFP + MSCV-IRES-Luciferase or pWPI-*MYCT58A*-IRES-GFP were used to generate preneoplasms from $CD133^+$ or $Sox2^+$ cells.

To create pICUE-*MYCT58A*-IRES-Luciferase, a fragment containing rtTA-IRES was PCR amplified from pICUE-*MYCT58A*-IRES-GFP [5] by using primers: Forward GGT CAT ATG TGG CCT GGA GAA ACA GCT and Reverse TAT GTC GAC GGT GGA CCG GTA AGC TTA. Luciferase was PCR amplified from MSCV-IRES-Luciferase by using primers: Forward AGC GTC GAC ATG GAA GAC GCC AAA AAC and Reverse ATA CGG TCC GTT ACA CGG CGA TCT TTC C. To replace GFP with Luciferase, rtTA-IRES was cut with NdeI and SalI. Luciferase was cut with SalI and CpoI. pICUE-*MYCT58A*-IRES-GFP vector was cut with NdeI and CpoI. rtTA-IRES, luciferase and pICUE-*MYCT58A*-IRES-GFP vector were ligated in one reaction with T4 ligase (New England Biolabs). To generate pWPI-*MYCT58A*-IRES-GFP, *MYCT58A* was PCR amplified from pICUE-*MYCT58A*-IRES-GFP by using primers: Forward GCA TTA ATT AAG CAG GCA CCA TGC CCC TC and Reverse GCC GTT TAA ACC TAC GCA CAA GAG TTC CG. *MYCT58A* was then digested with PacI and PmeI, and ligated with pWPI (a gift from Didier Trono, Addgene plasmid # 12254) digested with the same enzymes. To create pWPI-*MYCT58A*-IRES-Luciferase, IRES was PCR amplified from pWPI by using primers: Forward GTC GTT TAA ACT ACG GGC TGC AGG AAT TC and Reverse GGC GGA TCC GGT ATT ATC ATC GTG TTT TTC. Luciferase was PCR amplified from pICUE-*MYCT58A*-IRES-Luciferase by using primers: Forward GGC GGA TCC ATG GAA GAC GCC AAA AAC and Reverse GTC ATC GAT CAT ATG TTA CAC GGC GAT CTT TCC. To replace GFP with Luciferase, IRES was cut with PmeI and BamHI. Luciferase was cut with BamHI and ClaI. pWPI-*MYCT58A*-IRES-GFP vector was cut with PmeI and BstBI. IRES, luciferase and pWPI-*MYCT58A*-IRES-GFP vector were ligated in one reaction. Sequence of inserts were verified by DNA sequencing.

Immunohistochemistry

For immunohistochemical staining of frozen sections, mice were perfused with 4% paraformaldehyde (PFA) and then brains were removed and fixed overnight in 4% PFA. The tissues were cryoprotected in 25% sucrose overnight and then were frozen in Tissue Tek-OCT (Sakura). 12 μM cryosections were permeabilized for 20 min with PBS containing 0.1% Triton X-100, blocked for 1 hr with PBS containing 0.1% Triton X-100 and 10% normal donkey serum, and incubated overnight with primary antibodies at 4 °C. The next day, sections were stained with corresponding secondary antibodies for 1 hour at room temperature and counterstained with DAPI. Antibodies used for immunostaining include: Ki67 (1:200, Abcam), cleaved Caspase-3 (1:1000, Cell Signaling Technologies), Nestin (1:100, EMD Millipore), Olig2 (1:1000, EMD Millipore), Sox2 (1:200, Santa Cruz), Sox9 (1:250, Abcam), Synaptophysin (1:500, Abcam), GFAP (1:1000, Dako), Pax2 (1:100, Biolegend), Pax6 (1:100, Biolegend).

To analyze active proliferation by BrdU, P5 C57BL/6J mice were administered 10 mg/mL BrdU by intraperitoneal injection at a dose of 1mg/kg. Two hours after injection, mice were perfused with 4% paraformaldehyde and brains were collected and treated as above. Frozen sections were then permeabilized with 0.3% Triton X for 20 min at room temperature, incubated with 2N HCl for 30 min at room temperature, then neutralized in 0.1M sodium borate solution for 20 min at room temperature. Following HCl treatment, slides were washed with PBS, then blocked for 1 hr with PBS containing 0.1% Triton X-100 and 10% normal donkey serum, and incubated overnight with primary antibodies against BrdU (1:1000, Abcam) and Sox2 at 4 °C. The next day, sections were stained with corresponding secondary antibodies for 1 hour at room temperature and counterstained with DAPI.

For immunostaining of cultured cells, cells were cultured in 8 well PDL-coated chamber slides for 5 days. Slides were fixed in 4% PFA for 20 min, then blocked and incubated with antibodies as described above. Antibodies used for cultured cells include GFAP (1:2000, Dako), NeuN (1:200, EMD Millipore), Pax2 (1:200, Biolegend) and O4 (1:200, EMD Millipore).

For immunohistochemical staining of paraffin sections, tissues were fixed overnight in 4% PFA, transferred to 70% ethanol, paraffin embedded, and sectioned at a thickness of 5 μm. Sections were stained with hematoxylin and eosin (Sigma).

To test whether tdT faithfully labels $Sox2^{+}$ cells and their progeny during development, Sox2-$CreER^{T2}$ mice were crossed with Rosa-CAG-LSL-tdTomato mice. Offspring were administered tamoxifen (Sigma-Aldrich) at a single dose of 0.2 mg/g body weight by oral gavage at p5. Animals were then perfused with 4% PFA at P6 or P10 and cerebella were collected for immunohistochemistry staining.

To examine Sox9 and CC-3 expression after *Sox2* knockout, Sox2-CreER^{T2}/Sox2^{-loxp} mice were administered tamoxifen (Sigma-Aldrich) at a single dose of 0.2 mg/g body weight by oral gavage or mock treated at p5. Animals were then perfused with 4% PFA at P10 and cerebella were collected for immunohistochemistry staining.

qRT-PCR

For qRT-PCR, RNA was extracted from tumor tissue using the QIAGEN RNAeasy prep kit and reverse transcribed into cDNA using the Bio-Rad Reverse Transcriptase kit. Samples were then prepared with Powerup SYBR green mix (Life Technologies) and primers, and evaluated using the ABI 7500 Real Time PCR System. Primers were listed below:

Gene	Species	Forward (5' - 3')	Reverse (5' - 3')
Sox1	Mouse	AGA CTT CGA GCC GAC AAG AG	TCA CTC AGG GCT GAA CTG TG
Sox3	Mouse	CTC GAG AGA ACG CAT CAG GTG	GCC ACC GTG AAA AGG CCG
Sox8	Mouse	AGC GAG AAG AGG CCG TTT G	CGC CTT GGC TGG TAT TTG TAA T
Sox10	Mouse	AGG TTG CTG AAC GAA AGT GAC	CCG AGG TTG GTA CTT GTA GTC C
Sox2	mouse	TAG AGC TAG ACT CCG GGC GAT GA	TTG CCT TAA ACA AGA CCA CGA AA
Ldha	Mouse	TGC GTG CTG GAG CCA CT	GCG AGG AGA AGC AGC GTG
Pkm2	Mouse	GTC TGG AGA AAC AGC CAA GG	CGG AGT TCC TCG AAT AGC TG
Hk2	Mouse	TGA TCG CCT GCT TAT TCA CGG	AAC CGC CTA GAA ATC TCC AGA
Pdk1	Mouse	GGA CTT CGG GTC AGT GAA TGC	TCC TGA GAA GAT TGT CGG GGA
Gapdh	Mouse	TTC CAG TAT GAC TCC ACT CAC GG	TGA AGA CAC CAG TAG ACT CCA CGA C
Gfi1	Mouse	TGC TCC GAG TTC GAG GAC TT	CAG AGA GCG GCA CAG TGA CTT

Western Blot

To examine *MYC* expression mediated by MYC lentivirus or retrovirus, 293T cells were infected with or without indicated viruses and cultured for 48 hrs. Cells were then collected for western blot.

To examine CC-3 expression in cultured Sox2^{+} cells or Sox2 KO cells with or without *MYC* overexpression, sorted cells were infected with control or *MYC* lentivirus for overnight and cultured for totally 4 days. Cells were then collected for western blot.

Cells were lysed with Pierce RIPA buffer (Thermo Fisher Scientific) supplemented with proteinase and phosphatase inhibitor cocktail (Thermo Fisher Scientific). Protein concentration was measured by BCA assay (Bio-Rad). Lysates were separated on 8-16% Mini-Protean TGX gels (Bio-Rad) by electrophoresis, then transferred to Immobilon-P membranes (EMD Millipore). Membranes were blocked in 5% non-fat dry milk (Bio-Rad) for one hour at room temperature, and incubated with primary antibodies to human LDHA (1:1000, Cell Signaling Technologies), mouse LDHA (1:4000, Abcam), PKM2 (1:1000, Cell Signaling Technologies), HK2 (1:1000, Cell Signaling Technologies), PDK1 (1:1000, Cell Signaling Technologies), MYC (1:3000, Abcam), CC-3 (1:1000, Cell Signaling Technologies) or beta-actin (1:3000, Abcam) at 4°C for overnight. Membranes were then incubated with HRP-conjugated secondary antibodies (1:10,000, GE Life Sciences) for one hour at room temperature. Signals were developed using Amersham ECL (GE Healthcare).

Neurosphere and Differentiation Assays

For neurosphere assay, cells were plated at low density (5000 cells per 2mL per well) in 6 well plates with NeuroCult™ medium supplemented with bFGF and EGF and cultured for 7 days. Neurospheres were counted manually using an inverted fluorescent microscope. To evaluate self-renewal, neurospheres were dissociated into single cells with 0.05% Trypsin-EDTA (Life Technologies) and cultured in the same conditions as stated above. For differentiation assay, neurospheres or freshly isolated cells (50000 cells) were plated in PDL-coated Nunc 8 well chamber cells (Lab-Tek) and cultured in differentiation medium (Stemcell Technologies) with 8% FBS (Gibco) for 5 days. Cells were then fixed in 4% PFA followed by immunohistochemistry staining.

Proliferation and Apoptosis Assays

To determine the effects of MYC on cell proliferation and apoptosis of $Sox2^{+}$ cells, $Sox2^{+}$ cells were FACS-sorted from the cerebella of P5 $Sox2^{-GFP}$ mice. Cells were infected with inducible *MYC* lentivirus overnight and cultured in NeuroCult™ media supplemented with bFGF and EGF in the presence or absence of doxycycline for totally two days for apoptosis assay or five days for proliferation assay. To test for apoptosis, cells were then collected, digested to single cell suspension with papain or 0.05% Trypsin with EDTA (Life Technologies), and stained for Annexin V (Thermo Fisher Scientific) and 7AAD (BD Biosciences). To test for proliferation, cells were incubated for 1 hr with 10 μM Edu (Thermo Fisher Scientific), then fixed and stained for Edu using the Click-iT Plus Edu Alexa Fluor 647 kit (Thermo Fisher Scientific). Samples were analyzed by flow cytometry for early and late stage apoptosis or proliferation using the BD Accuri flow cytometer.

Lineage Tracing

To trace the lineage of $Aldh1L1^{+}$ cells during development, Aldh1L1-$CreER^{T2}$ mice were crossed with Rosa-CAG-LSL-tdTomato mice. Offspring were administered tamoxifen (Sigma-Aldrich) at a single dose of 0.1 mg/g body weight by oral gavage at P5. Animals were then perfused with 4% PFA at P10 and cerebella were collected for immunohistochemistry staining.

Lactate Assay

Freshly isolated MB002 cells were cultured for 3 days in stem cell medium and then pre-treated with indicated concentrations of GSK for 3.5 hr to inhibit LDHA activity. Cells were then resuspended and cultured with fresh media containing GSK. Culture media were collected at 2 hr and 4 hr after resuspending cells with new media and filtered through 10 kDa cut-off spin filter (EMD Millipore) to remove LDHA in the media. Cell viability at each time point was also determined by CellTiter-Glo® assay. Lactate level were measured by a lactate oxidase based Lactate Assay Kit (MAK064, Sigma) and were normalized to cell viability for each sample.

RNA-Seq Data Analysis

The single-end reads from mouse MYC-driven medulloblastoma samples were aligned to mouse genome mm10 using STAR v2.4.1 tool [7]. Gene expression counts were computed using featureCounts module of the Subread package v1.4.6 with Ensembl GRCm38 v72 anntoation based only on uniquley mapped reads [8] and normalized as Reads Per Kilobase per Million mapped reads (RPKM). Alignment and counts data quality control was performed using Qualimap2 [9].

The human data (170 RNA-seq samples) presented in the manuscript by P.Northcott et al [10] was obtained from the European Genome-Phenome Archive (study ID: EGAS00001001953). General analysis (alignment, quality control, counts computation, differential gene expression analysis among tumor groups) was performed using the same tools as for mice data with GRCh37 1000G as reference genome and Gencode v.19 as transcriptome annotation. In the final analysis, 3 samples were discarded due to several quality control issues.

The confident list of orthologous genes for comparison between mice models and human tumors was selected with BiomaRt R package application [11]. Batch effect adjustment to compare and cluster the samples in combined gene counts was performed using ComBat tool from SVA package [12]. Principal component analysis and unsupervised hierarchical clustering visualization was performed using the top 500 most highly variable genes. Clustering stability was estimated with usage of ConsensusClusterPlus CDS method [13]. Additional verification was performed based on

the selection of top 100 group-specific differentially expressed genes detected from medulloblastoma RNA-seq data using DESeq2 tool [14] and sorted by adjusted p-value. This gene set was extracted from the combined orthologous gene counts dataset (resulting in present gene subsets SHH: 63 out of 100, WNT: 67 out of 100, G3: 59 out of 100, G4: 55 out of 100) and used as a reference set for unsupervised hierarchical clustering.

Additional independent group used different approaches to analyze the RNA-Seq data. Raw fastq files were analyzed using FastQCv0.11.5 and trimming was performed on raw reads using Trimmomatic v0.35 [15]: standard parameters for phred33 encoding were used. Reads mapping to the reference genome (GRCm38) was performed on quality-checked and trimmed reads using HISAT2 [16]. The reference annotation is Ensembl v87. The overlap of reads with annotation features found in the reference.gtf was calculated using HT-seq v0.6.1 [17]. The output computed for each sample (raw read counts) was then used as input for DESeq2 [14] analysis. Raw counts were normalized using DESeq2's function "fpkm" and the expression value of FPKM was used to perform and visualize principal component analysis (PCA) results (using DESeq2's "plotPCA" function).

For differential expression and gene enrichment analysis, we compared the expression values of genes in the tumor and normal tissues. A gene was considered differentially expressed when the absolute difference of FPKM value between compartments was greater than 5, the log2 fold-change was greater than 1 and the adjusted P-value (FDR) was less than 0.05. The genes that both expressed up-regulated compared to matched normal samples and the other tumor types were defined as the signature genes for each tumor types. Gene sets enrichment analysis was performed for differentially expressed genes that were observed by subgroups, using DAVID [18, 19].

To analyze the expression of glycolytic enzymes, RNA-seq data from publicly available human medulloblastoma tumors and normal cerebellums (Datasets: EGAD00001001899, EGAD00001002683, EGAD00001001210, EGAD00001001620, EGAD00001000328 and GSE67196) were mapped to human genome GRCh38/hg38 using STAR(v.2.5.1) [7]. Genes and isoform expression was quantified using the cufflinks software (v2.2.1) [20] and hg38 UCSC transcript annotation.

Visualization of the *Trp53* gene sequence acquired from RNA-seq were performed by using IGV [21, 22]. Peaks with gray color indicate sequence without mutations.

Affymetrics Data Analysis

The Affymeterics data for MP tumor models was obtained from GEO public database (GSE4126), while for MG and MGb directly from the corresponding manuscript authors. The general data processing and normalization was performed using Affy R package [23]]. Further, normalized log2 expression AffyId values from each sample were aggregated into matrix by computing mean values per gene and filtered based on mice/human ortholgos selection. Further combined analysis with RNA-seq data was performed based on ComBat method application.

***MYC-LDHA* correlation**

MYC and *LDHA* mRNA expression from publicly available human MB tumors were correlated using linear regression and Spearman correlation methods. The coefficient of determination (R^2) and Spearman coefficient were calculated in the R statistical environment.

Survival analysis based on *LDHA* level

Overall survival functions were estimated using the Kaplan-Meier method and p-values were calculated using the log-rank test. The statistical analysis was performed in the R statistical environment using the R package survival (v2.41-3) and survminer (v0.4.0). Gene expression and overall survival data from Group 3 medulloblastoma tumors were downloaded from a previously published dataset (GEO accession data: GSE85218[RS1]). Group 3 medulloblastoma tumors were separated according to LDHA expression.

Statistical Analyses

Statistical analysis was performed using Excel or GraphPad Prism software. All data are presented as means ± SD or means ± SEM as indicated. Comparisons between different groups were made using Student's t test or ANOVA as appropriate. The statistical significance of Kaplan-Meier survival curves was assessed using the log-rank (Mantel-Cox) test. p Values of 0.05 or lower were considered statistically significant for all experiments.

Supplemental References

1. Lee, A., et al., *Isolation of neural stem cells from the postnatal cerebellum.* Nat Neurosci, 2005. **8**(6): p. 723-9.
2. Ellis, P., et al., *SOX2, a persistent marker for multipotential neural stem cells derived from embryonic stem cells, the embryo or the adult.* Dev Neurosci, 2004. **26**(2-4): p. 148-65.
3. Masutomi, K., et al., *The telomerase reverse transcriptase regulates chromatin state and DNA damage responses.* Proc Natl Acad Sci U S A, 2005. **102**(23): p. 8222-7.
4. Zhao, D., et al., *Lysine-5 acetylation negatively regulates lactate dehydrogenase A and is decreased in pancreatic cancer.* Cancer Cell, 2013. **23**(4): p. 464-76.
5. Pei, Y., et al., *An animal model of MYC-driven medulloblastoma.* Cancer Cell, 2012. **21**(2): p. 155-67.
6. Takahashi, K. and S. Yamanaka, *Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors.* Cell, 2006. **126**(4): p. 663-76.
7. Dobin, A., et al., *STAR: ultrafast universal RNA-seq aligner.* Bioinformatics, 2013. **29**(1): p. 15-21.
8. Liao, Y., G.K. Smyth, and W. Shi, *featureCounts: an efficient general purpose program for assigning sequence reads to genomic features.* Bioinformatics, 2014. **30**(7): p. 923-30.
9. Okonechnikov, K., A. Conesa, and F. Garcia-Alcalde, *Qualimap 2: advanced multi-sample quality control for high-throughput sequencing data.* Bioinformatics, 2016. **32**(2): p. 292-4.
10. Northcott, P.A., et al., *The whole-genome landscape of medulloblastoma subtypes.* Nature, 2017. **547**(7663): p. 311-317.
11. Durinck, S., et al., *Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt.* Nat Protoc, 2009. **4**(8): p. 1184-91.
12. Leek, J.T., et al., *The sva package for removing batch effects and other unwanted variation in high-throughput experiments.* Bioinformatics, 2012. **28**(6): p. 882-3.
13. Wilkerson, M.D. and D.N. Hayes, *ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking.* Bioinformatics, 2010. **26**(12): p. 1572-3.
14. Love, M.I., W. Huber, and S. Anders, *Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2.* Genome Biol, 2014. **15**(12): p. 550.
15. Bolger, A.M., M. Lohse, and B. Usadel, *Trimmomatic: a flexible trimmer for Illumina sequence data.* Bioinformatics, 2014. **30**(15): p. 2114-20.
16. Kim, D., B. Langmead, and S.L. Salzberg, *HISAT: a fast spliced aligner with low memory requirements.* Nat Methods, 2015. **12**(4): p. 357-60.
17. Anders, S., P.T. Pyl, and W. Huber, *HTSeq--a Python framework to work with high-throughput sequencing data.* Bioinformatics, 2015. **31**(2): p. 166-9.
18. Huang da, W., B.T. Sherman, and R.A. Lempicki, *Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources.* Nat Protoc, 2009. **4**(1): p. 44-57.
19. Huang da, W., B.T. Sherman, and R.A. Lempicki, *Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists.* Nucleic Acids Res, 2009. **37**(1): p. 1-13.
20. Trapnell, C., et al., *Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks.* Nat Protoc, 2012. **7**(3): p. 562-78.
21. Robinson, J.T., et al., *Integrative genomics viewer.* Nat Biotechnol, 2011. **29**(1): p. 24-6.
22. Thorvaldsdottir, H., J.T. Robinson, and J.P. Mesirov, *Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration.* Brief Bioinform, 2013. **14**(2): p. 178-92.

23. Gautier, L., et al., *affy--analysis of Affymetrix GeneChip data at the probe level.* Bioinformatics, 2004. **20**(3): p. 307-15.