Terminal differentiation and loss of tumorigenicity of human cancers via pluripotency based reprogramming.

Zhang et. al.

SUPPLEMENTAL SECTION:

Supplemental Figure Legends:

Figure 1: Percent frequency and time to iPSCs morphology formation following infection of reprogramming transcription factors as described in main text. Reprogrammed cells were grown on MEF feeder layers (as described in the Materials and Methods) and observed for colony formation characteristic of reprogrammed cells (as in Figure 1Aii). Earliest day of definitive colony formation following reprogramming as visually determined is shown in Figure 1. Colonies were stained with alkaline phosphatase (as shown in Supplemental Figure 3) to ensure that morphologically appearing colonies were reprogrammed. To determine the frequency of reprogramming, cells were harvested at the "earliest day of definitive colony formation following reprogramming as visually determined" as shown above, and stained individually with antibodies for Nanog, Oct4, SSEA4, and Tra-1-81 and with fluorescent conjugated secondary antibodies (FITC, PE, APC, and Cy5.5 - from Abcam: Figure 1). The stained cells were then washed twice in PBS+0.1%BSA, re-suspended in 500ul FACS buffer (PBS+3%BSA). The samples were analyzed using a flow cytometer immediately. Percentage frequency of cells expressing the four markers is shown.

Figure 2: Gene expression analysis of the reprogramming transcription factors based on Affymetrix U133Plus 2.0 arrays comparing parental sarcoma cells to ESCs and MSCs. Error bars show standard deviation. Each analysis was comprised of three replicates (of the indicated samples) and further averaged over the following probe sets (Affymetrix) U133 Plus2.0: per indicated gene: myc 202431_s_at; klf4 220266_s_at, 221841_s_at; sox2 213721_at, 213722_at, 214178_s_at, 228038_at; nanog 220184_at; lin28 219823_at; oct4=Pou51 208286_x_at, 214532_x_at, 210265_x_at, 210905_x_at.

Figure 3: (A) Expression of six reprogramming transcription factors in parental and reprogrammed sarcomas via RT-PCR using primers specific for to distinguish between the endogenous or the transgene. As could be seen in the figure, upon reprogramming the transgene is not present in either the parental or the reprogrammed sarcomas. The silencing of the transenge is a well established phenomenon of nuclear reprogramming. Additionally, as could be seen in the figure, some parental cell lines express some of the reprogramming factors (endogeneoulsy) at baseline (as shown in Figure 1), however upon reprogramming ALL reprogrammed sarcoma cell lines express ALL reprogramming factors; and specifically do so via activation of the endogenous versions (as the transgenes are not expressed).

Figure 4: (A) Alkaline Phosphatase staining of ESC/iPSC-like colonies following infection of reprogramming transcription factors as described in main text after the indicated number of days in Supplemental Figure 1. Scale bar=500um. Alkaline Phosphatase (AP) staining was performed on colonies grown on MEFs using the StemTAG™ Alkaline Phosphatase Staining Kit. Scale bar=500uM. (B) Alkaline phosphatase quantification as a function of time following infection with reprogramming transcription factors. Alkaline phosphatase quantification was determined using the StemTAG™ Alkaline Phosphatase Activity Assay Kit (Fluorometric). Specifically, MEF plates onto which infected sarcoma cell lines were plated were grown for the indicated number of days. Total cellular lysates were collected weekly and alkaline phosphatase quantification was performed using the StemTAG™ AP Fluorometric Substrate and fluorescence measured with a fluorescence plate reader at 480 nm/520 nm. Values were compared to a predetermined standardization curve. As controls MEF plates without infected cells were collected at each time point and alkaline phosphatase activity was similarly measured. Data is presented as Relative Fluorescence Units in which the value determined in MEF/sarcoma plates is shown relative to that in the corresponding control MEF plates. “C” represents the baseline value set at 1 of MEF cells determined prior to infection and plating of sarcoma cells. Each data represents the average of two independent experiments (i.e., infections and platings) in which three plates from each experiment were examined. Error bars=standard deviation.

Endogenous specific primers: [1-3]

NANOG[1]: CAGCCCCGATTCTTCCACCAGTCCC; CGGAAGATTCCCAGTCGGGTTCACC

SOX2[2]: GGGAAATGGGAGGGGTGCAAAAGAGG; TTGCGTGAGTGTGGATGGGATTGGTG

OCT3/4[2]: GACAGGGGGAGGGGAGGAGCTAGG; CTTCCCTCCAACCAGTTGCCCAAAC

c-MYC[2]: GCGTCCTGGGAAGGGAGATCCGGAGC: TTGAGGGGCATCGTCGCGGGAGGCTG

Klf4[2]: TGATTGTAGTGCTTTCTGGCTGGGCTCC; ACGATCGTGGCCCCGGAAAAGGACC

Lin28[3]: AGCCAAGCCACTACATTC; AGATACGTCATTCGCACA

Transgene specific primers:[4, 5]

Oct4[4]: CCTCACTTCACTGCACTGTA; CCTTGAGGTACCAGAGATCT

SOX2[4]: CCCAGCAGACTTCACATGT; CCTTGAGGTACCAGAGATCT

Klf4[4]: GATGAACTGACCAGGCACTA; CCTTGAGGTACCAGAGATCT

Myc[4]: TGCCTCAAATTGGACTTTGG; CGCTCGAGGTTAACGAATT

Lin28[5]: GCTCTGACTGACCGCGTTAC; CTGCTCCTCGAAACTTCCTG

NANOG[5]: GCTCTGACTGACCGCGTTAC; TCACACGTCTTCAGGTTGCAT

(B) Expression of Rex via RT-PCR in parental and reprogrammed sarcomas. Primers for REX1[2]:CAGATCCTAAACAGCTCGCAGAAT; GCGTACGCAAATTAAAGTCCAGA.

Figure 5: Pictographs of Hematoxylin, Ki67, and vimentin stained xenografts from parental and reprogrammed sarcomas. Specifically, for each xenograft, approximately 10^6 iPS cells were manually harvested, washed and resuspended in a 1.5 ml tube containing 300 µl iPS medium and then injected intramuscularly (tibialis anterior) into NOD SCID Gamma mice (Jackson Labs). Specifically, sarcoma cells and the reprogrammed sarcomas were injected into the left and right anterior tibialis muscle of NOD-SCID-gamma (NSG) mice, respectively, and observed. Eight mice were used for each cell line. The experiment was repeated independently two times. Any visible tumors 4–8 weeks post-transplantation were dissected and fixed overnight with 4% paraformaldehyde/PBS solution. The tissues were then paraffin embedded, sectioned, stained with hematoxylin and eosin, and examined for the presence of tissue representatives of all three germ layers. Scale bar=100um. No evidence of teratoma formation was observed and no morphological evidence of any non-sarcomatous element could be identified. All NOD-SCID-gamma mice were treated in accordance with IACUC Guidelines and Columbia veterinary policy as previously described[6]. IRB protocol AAAD-1669; PI Matushansky. NOTE: as stated in the

main text, both parental and reprogrammed sarcomas were negative for S100, pankeratin, neurophysin and desmin (data not shown).

Figure 6: Tumors as described in Supplemental Figure 5 were weighed and measured (greatest dimension). Average weights (A) and measurements (B) are shown. Error bars=standard deviation.

Figure 7: Proliferation analysis of parental sarcomas and their reprogrammed counterparts undergoing adipogenic and osteogenic differentiation. Specifically, BOTH parental sarcomas and their reprogrammed counterparts were plated at a density of $1x10^5$ per 10 cm plate on matrigel. Cells were grown overnight in either the corresponsing normal sarcoma cell line medium (please see Materials and Methods) or ESC medium for parental and reprogrammed respectively. Then on "Day 1" cytokines constituting either adipocytic or osteogenic differentiation medium (please see Materials and Methods) were added to the corresponding medium and cultured in differentiation medium from that point on. Cell counting was performed daily using the Countess Cell Counter (Invitrogen). Each cell line was tested concurrently in triplicate. Two independent experiments set up a week apart were performed. Average counts on each subsequent day and standard deviation are shown. ”F” and ”B” following each cell line in the Figure Legend indicates that the cell line was differentiated into fat (F) or bone (B) respectively. The ”R” preceding each cell line in the Figure Legend refers to the reprogrammed counterpart of each sarcoma.

Figure 8: Schedule, inoculation, days of growth, size of greatest dimension, and weight of xenografts from sarcomas and reprogrammed sarcomas following in vitro adipogenic and osteogenic differentiation. Specifically, sarcomas and reprogrammed sarcomas were differentiated as described in either Supplemental Figure 11 or the Main Text - and subsequently injected into NOD SCID Gamma mice as described in Supplemental Figure 9 or the Main Text. Days to tumor formation and sacrifice, length of greatest dimension, and weight of tumors are shown: (A) Before and after differentiation along the osteogenic lineage. (B) Before and after differentiation along the adipogenic lineage. (C) To ensure that tumor formation was indeed eliminated and not delayed; reprogrammed sarcomas were solely injected into a seperate cohort of mice and followed for up to four months. Days to tumor formation and sacrifice, length of greatest dimension, and weight of tumors are shown.

Figure 9: : Immunoflourescence of CD71 and CD235(GPA) in reprogrammed SW872 cells after 14 and 28 days of erythroid differentiation. Left= goat anti-rabbit Oregon Green 488 to primary marker; middle panel= DAPI; right panel=merge. green circle=CD235a+/DAPI+; blue circle=CD235-/DAPI+; yellow circle=CD235a+/DAPI-. Scale bars=10uM.

Figure 10: Schedule, inoculation, days of growth, size of greatest dimension, and weight of xenografts from reprogrammed sarcomas following in vitro erythroid differentiaton as described in the Main Text.

Figure 11: Quantitative RT-PCR of myc RNA in whole cell populations of either the parental or the reprogrammed sarcoma cell lines before (control) and after 21 days of exposure in either adipocytic (fat) or osteogenic (bone) differentiation medium. Values are normalized to GADPH and each cell line is baselined to the value of “1” of the control of the parental sarcoma cell line. Each measurement was repeated three times. The entire experiment was repeated independently twice as a result of two independent differentiation inductions. The average of six samples is shown. Error bars show standard deivations. Specifically, Quantitative PCR was performed with Platinum SYBR Green qPCR Supermix UDG (Invitrogen) and analyzed with the 7300 real-time PCR system (Applied Biosystems). Primers c-Myc-Human: FOR-CCACAGCAAACCTCCTCACAG, REV-GCAGGATAGTCCTTCCGAGTG; GAPDH-Human: FOR-GGTCGTATTGGGCGCCTGGTCACC, REV-CACACCCATGACGAACATGGGGGC.

Figure 12: (A) Average value of differentially methylated promoter CpG islands ($p<0.05$) in sarcomas and reprogrammed sarcomas. (B) Average gene expression values (from Affymetrix Arrays) corresponsing to differentially methylated promoter CpG islands (from A) in sarcomas and reprogrammed sarcomas for that subset of corresponding genes that show differential expression between sarcomas and reprogrammed sarcomas (please see Main Text for details). Data was analyzed via GeneSpring Software and normalized as described in the Materials and Methods.

Figure 18: Quantitative RT-PCR of 26 of 29 genes (of the 205 gene promoters showing significant changes in CpG methylation) which show significant changes ($p<0.05$, paired T-Test) at the gene expression level via Affymetrix (please see Main Text for details). Primer sets for ACSBG2, FBLIM1, KIAA0141 could not be identified. Quantitative RT-PCR was performed as described in Supplemental Figure 15. Values were secondarily baselined ("1") to those for the parental sarcoma cell line and values relative to that baseline is shown for both sarcomas and their reprogrammed counterparts. Experiments were performed in triplicate. Average values are shown. Error bars=standard deviation. Primer sets used are shown in Supplemental Table 4 (sheet 2).

Supplemental Table Legends:

Supplemental Table 1:
Summary of genetic and karyotypic abnormalities in the five cell lines used.

Supplemental Table 2:
Global DNA-methylation changes in sarcomas and their reprogrammed counterparts using Infinium HumanMethylation27 arrays.

Supplemental Table 3:
216 CpG islands (corresponding to 205 gene promoters) demonstrating a 15% change in DNA promoter CpG methylation ($p<0.05$, paired T-Test).

Supplemental Table 4:
29 unique genes (37 Affymetrix IDs) showing significant changes ($p<0.05$, paired T-Test) at the gene expression level of the 205 gene promoters showing significant changes in CpG methylation.

Supplemental Table 5:
Oncogene list.

Supplemental Table 6:
Tumor suppressor gene list.

Supplemental Table 7:
1241 promoters whose degree of methylation was significantly changed by direct reprogramming ($p<0.05$, paired T-Test; no restriction on fold change. In contrast to Supplemental Table 3 which was restricted to a 15% fold change.

Supplemental Table 8:
Overlap of Supplemental Table 7 and Supplemental Table 6 (oncogene list).

Supplemental Table 9:
Overlap of Supplemental Table 7 and Supplemental Table 5 (tumor suppressor gene list).

Supplemental Table 10:
Discriminatory 50 gene set[7], previously used to differentiate between mouse ESCs, mouse embryonic fibroblasts, and reprogrammed fibroblasts.

Supplemental Table 11:
Analysis of variance (ANOVA) identifying a subset of genes discriminating among ESCs, MSCs, fibroblasts), reprogrammed fibroblasts, partially reprogrammed fibroblasts and reprogrammed-sarcomas.

Supplemental Table 12:
Paired T-Test (p 0.05; 1.5 fold change) comparing sarcomas to their reprogrammed counterparts identifying 125 differentially expressed unique genes.

Supplemental Table 13:
Analysis of variance (ANOVA) identifying a subset of genes discriminating between ESCs, MSCs, and fibroblasts.

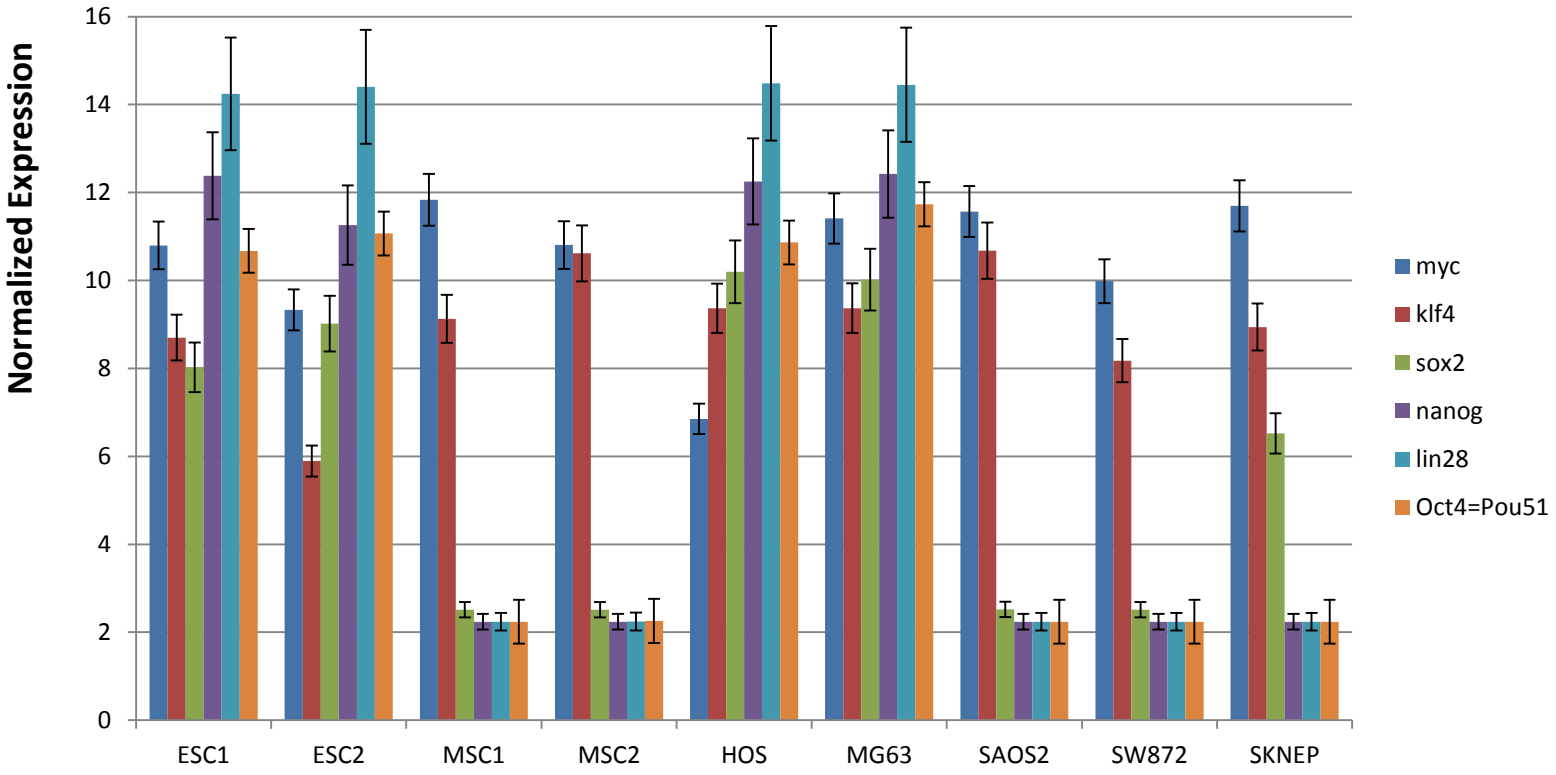
REFERENCES (Supplemental Section)

1. Wang Y, Chen J, Hu JL, et al. Reprogramming of mouse and human somatic cells by high-performance engineered factors. EMBO Rep. 2011;12:373-378.
2. Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell. 2006;126:663-676.
3. Narsinh KH, Jia F, Robbins RC, et al. Generation of adult human induced pluripotent stem cells using nonviral minicircle DNA vectors. Nat Protoc. 2011;6:78-88.
4. Park IH, Zhao R, West JA, et al. Reprogramming of human somatic cells to pluripotency with defined factors. Nature. 2008;451:141-146.
5. Darr H, Benvenisty N. Genetic analysis of the role of the reprogramming gene LIN-28 in human embryonic stem cells. Stem Cells. 2009;27:352-362.
6. Mills J, Hricik T, Siddiqi S, et al. Chromatin structure predicts epigenetic therapy responsiveness in sarcoma. Mol Cancer Ther. 2011.
7. Mikkelsen TS, Hanna J, Zhang X, et al. Dissecting direct reprogramming through integrative genomic analysis. Nature. 2008;454:49-55.

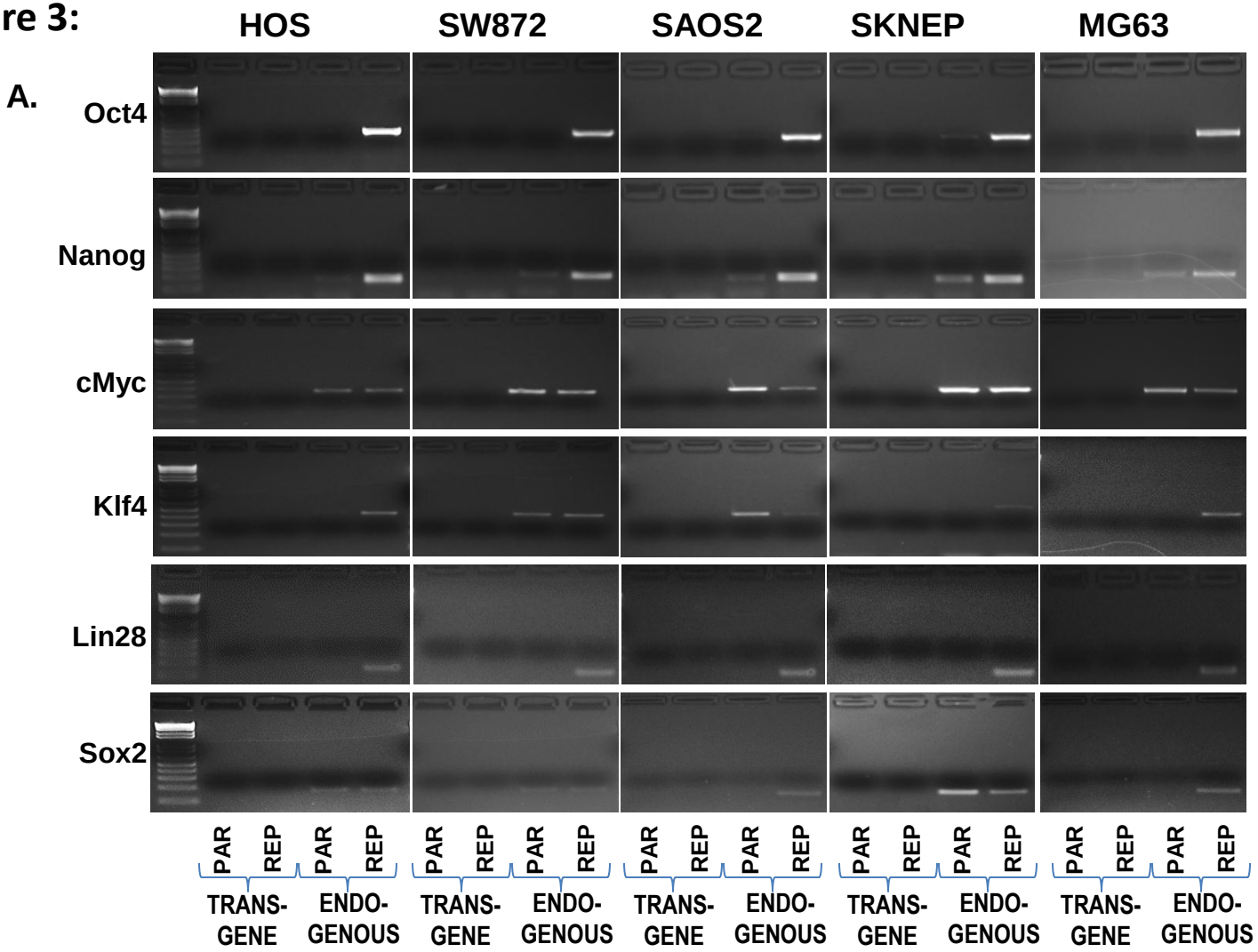
Supplemental Figure 1:

Rep-Sarcoma	% Frequency	Time (days)
SAOS2	2	25
HOS	5	27
MG63	3	42
SW872	3	26
SKNEP	4	18

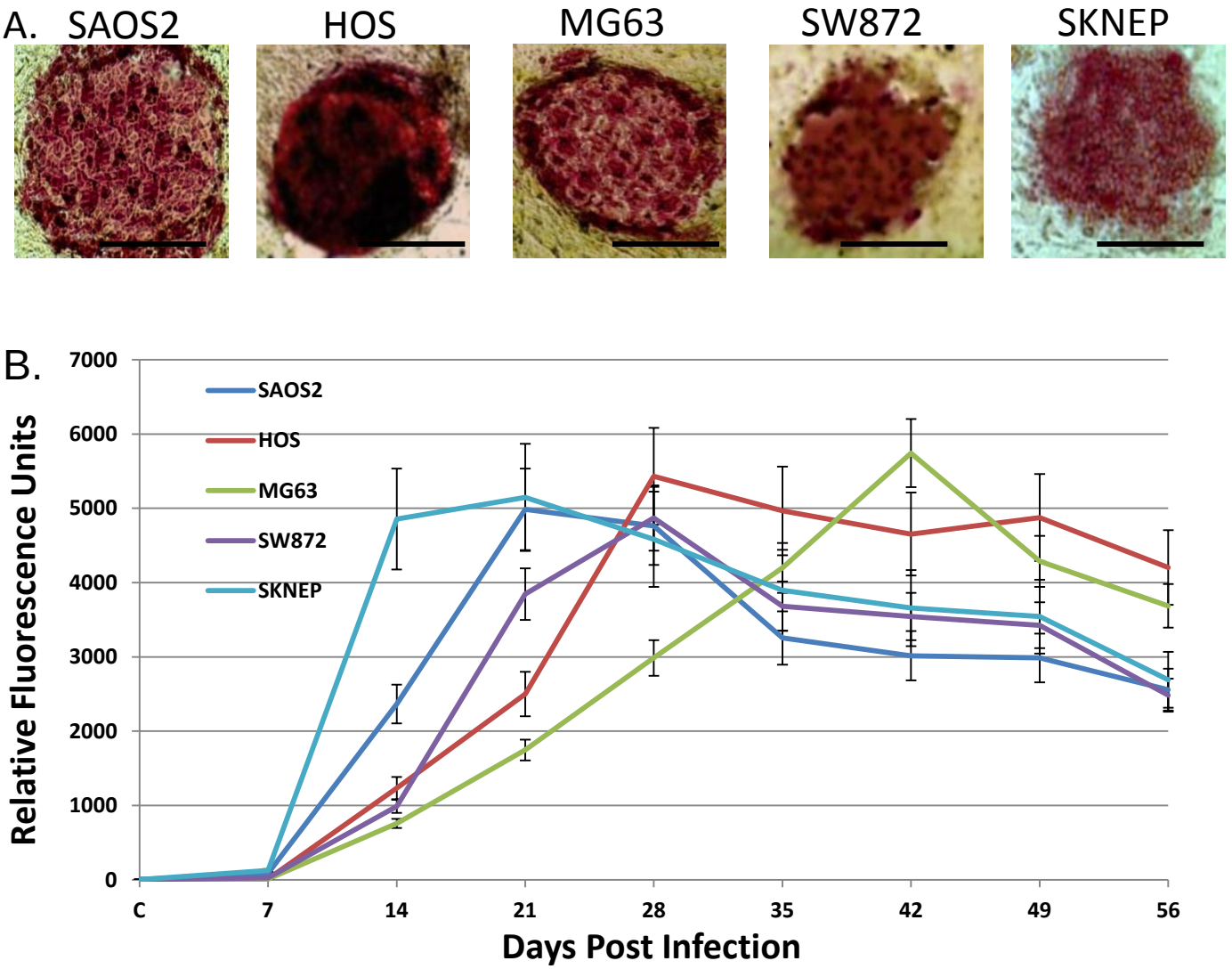
Supplemental Figure 2:



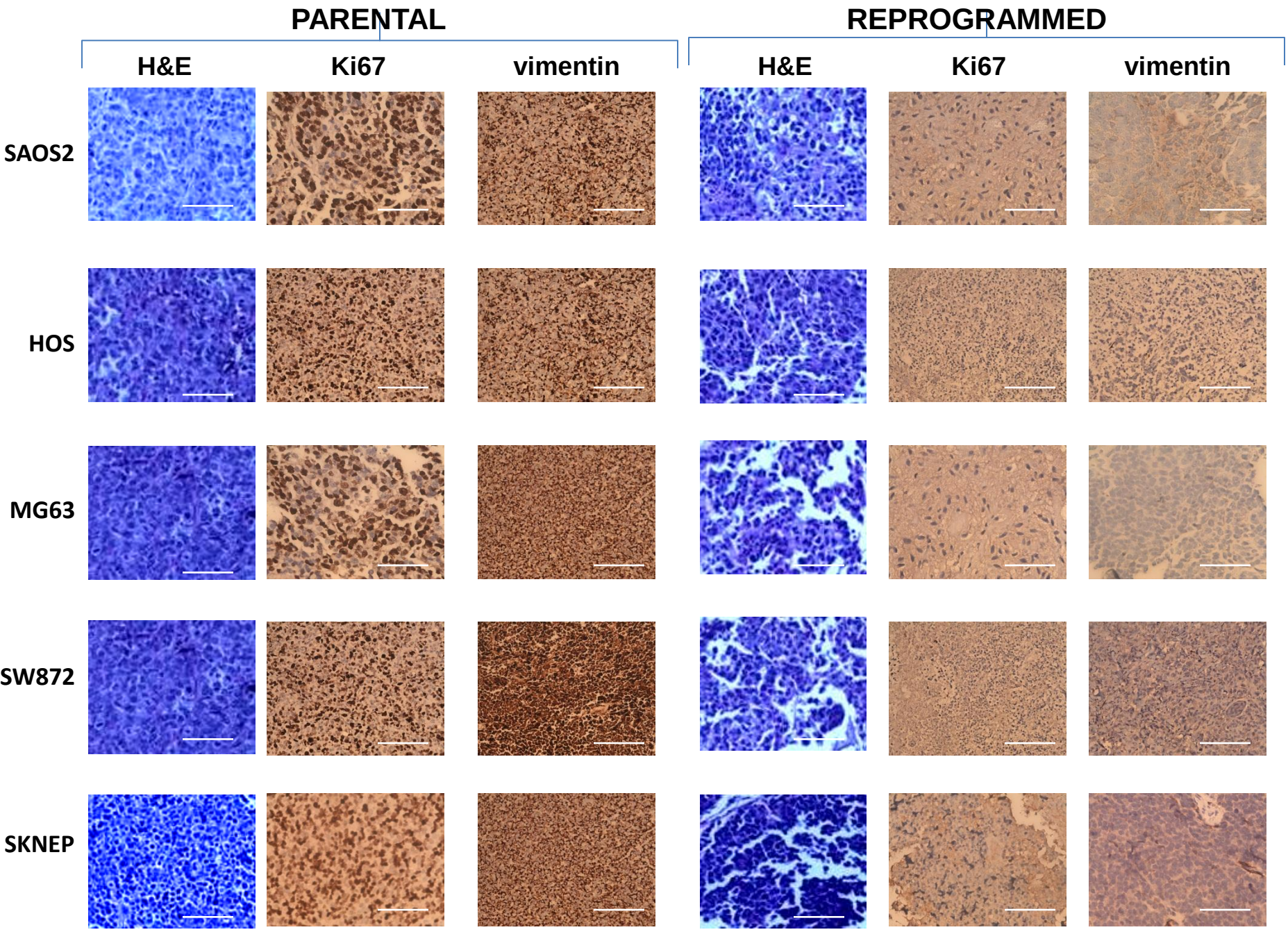
Supplemental Figure 3:



Supplemental Figure 4:

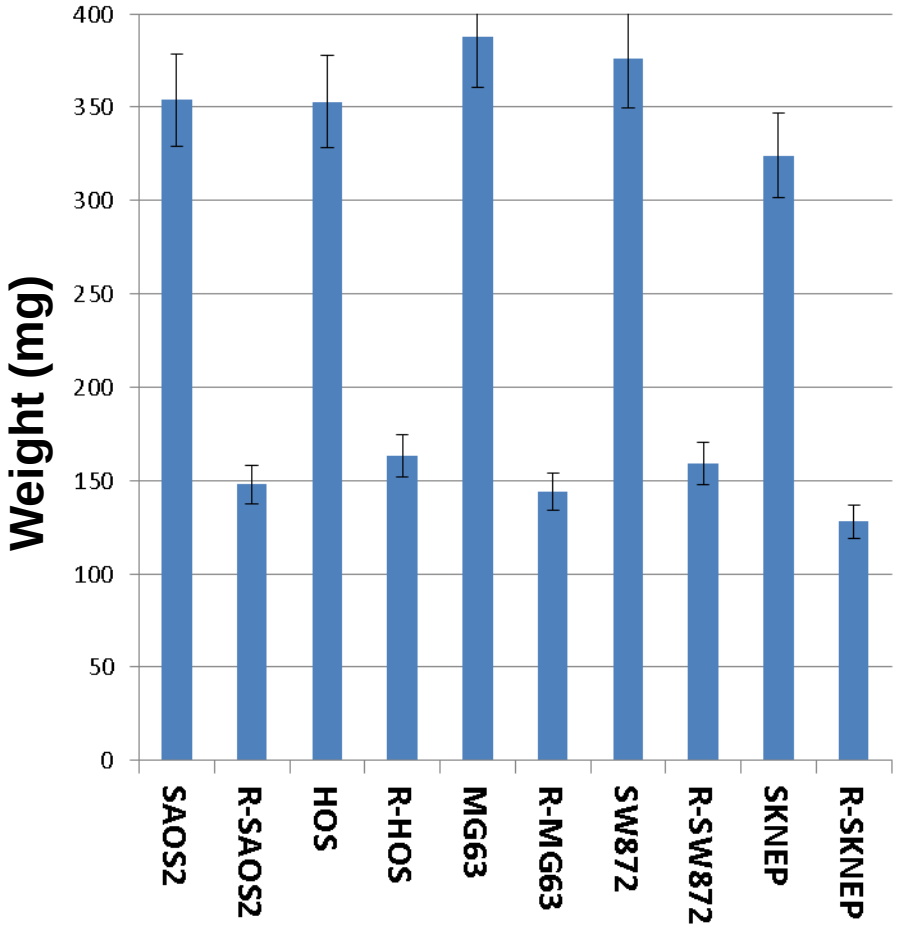


Supplemental Figure 5:

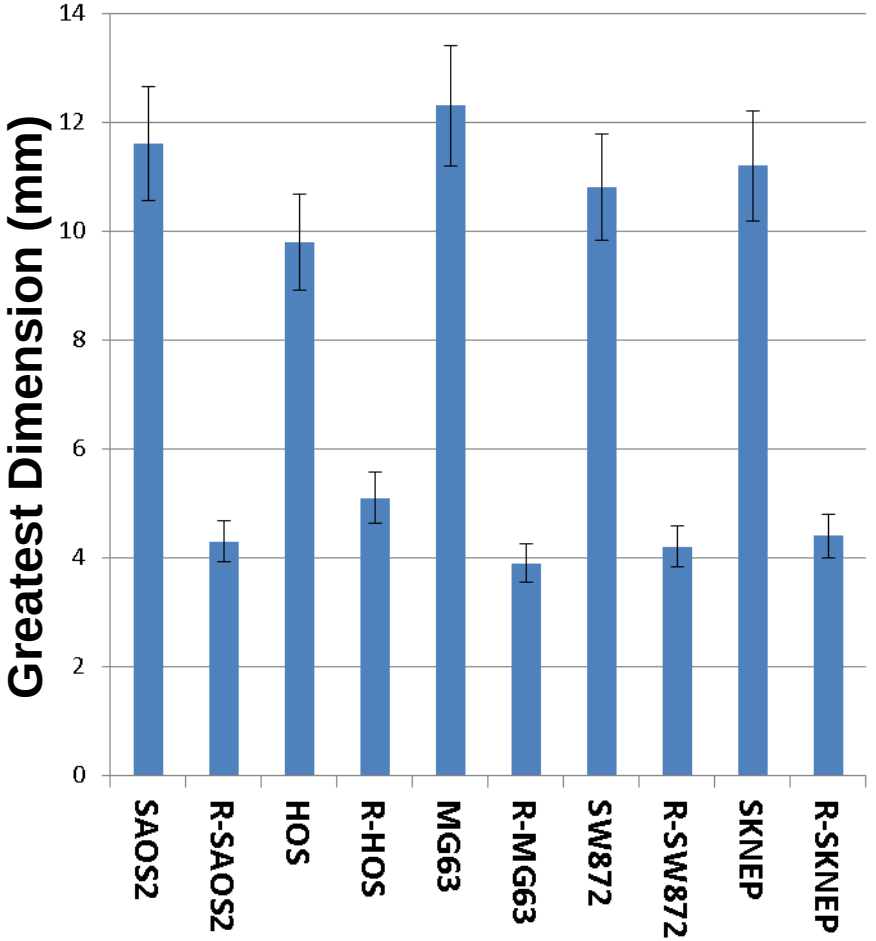


Supplemental Figure 6:

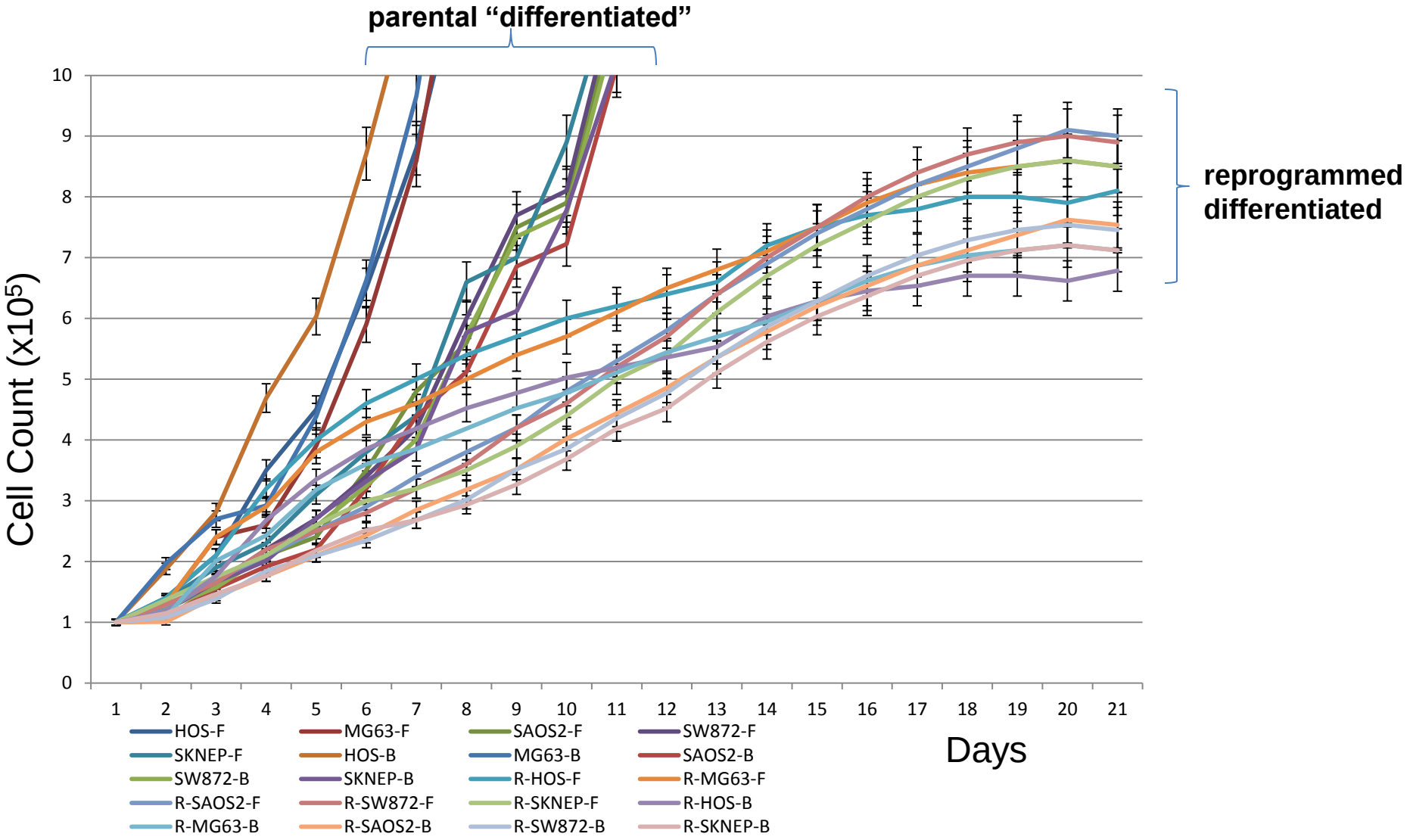
A.



B.



Supplemental Figure 7:



Supplemental Figure 8:

A.

MOUSE ID	CELL TYPE	CELL LINE-DIFF 21 DAYS **ODM**	# CELLs INNOCULATEd	FLANK (L or R)	DAYS TO SAC	Greatest Dimension (mm)	WEIGHT (mg)
111710A1	osteosarcoma	MG63	one million	L	72	11	388
111710A2	osteosarcoma	rep-MG63	one million	R	72	0	0
111710B1	osteosarcoma	MG63	one million	L	74	10	345
111710B2	osteosarcoma	rep-MG63	one million	R	74	0	0
111710G1	liposarcoma	SW872	one million	L	76	12	421
111710G2	liposarcoma	rep-SW872	one million	R	76	0	0
111710H1	liposarcoma	SW872	one million	L	73	11	398
111710H2	liposarcoma	rep-SW872	one million	R	73	0	0
111710I1	Ewings	SKNEP	one million	L	56	13	388
111710I2	Ewings	rep-SKNEP	one million	R	56	0	0
111710J1	Ewings	SKNEP	one million	L	58	12	376
111710J2	Ewings	rep-SKNEP	one million	R	58	0	0
111710A1	osteosarcoma	MG63	two million	L	68	11	356
111710A2	osteosarcoma	rep-MG63	two million	R	68	0	0
111710B1	osteosarcoma	MG63	two million	L	69	13	382
111710B2	osteosarcoma	rep-MG63	two million	R	69	0	0
111710G1	liposarcoma	SW872	two million	L	62	11	321
111710G2	liposarcoma	rep-SW872	two million	R	62	0	0
111710H1	liposarcoma	SW872	two million	L	60	10	308
111710H2	liposarcoma	rep-SW872	two million	R	60	0	0
111710I1	Ewings	SKNEP	two million	L	54	11	369
111710I2	Ewings	rep-SKNEP	two million	R	54	0	0
111710J1	Ewings	SKNEP	two million	L	58	12	389
111710J2	Ewings	rep-SKNEP	two million	R	58	0	0

B.

MOUSE ID	CELL TYPE	CELL LINE-DIFF 21 DAYS **ADM**	# CELLs INNOCULATEd	FLANK (L or R)	DAYS TO SAC	Greatest Dimension (mm)	WEIGHT (mg)
112110A1	osteosarcoma	MG63	one million	L	71	12	377
112110A2	osteosarcoma	rep-MG63	one million	R	71	0	0
112110B1	osteosarcoma	MG63	one million	L	75	11	365
112110B2	osteosarcoma	rep-MG63	one million	R	75	0	0
112110G1	liposarcoma	SW872	one million	L	74	11	398
112110G2	liposarcoma	rep-SW872	one million	R	74	0	0
112110H1	liposarcoma	SW872	one million	L	71	12	349
112110H2	liposarcoma	rep-SW872	one million	R	71	0	0
112110I1	Ewings	SKNEP	one million	L	54	13	376
112110I2	Ewings	rep-SKNEP	one million	R	54	0	0
112110J1	Ewings	SKNEP	one million	L	55	10	359
112110J2	Ewings	rep-SKNEP	one million	R	55	0	0
112110A1	osteosarcoma	MG63	two million	L	66	12	352
112110A2	osteosarcoma	rep-MG63	two million	R	66	0	0
112110B1	osteosarcoma	MG63	two million	L	68	11	326
112110B2	osteosarcoma	rep-MG63	two million	R	68	0	0
112110G1	liposarcoma	SW872	two million	L	64	12	346
112110G2	liposarcoma	rep-SW872	two million	R	64	0	0
112110H1	liposarcoma	SW872	two million	L	65	11	358
112110H2	liposarcoma	rep-SW872	two million	R	65	0	0
112110I1	Ewings	SKNEP	two million	L	55	12	329
112110I2	Ewings	rep-SKNEP	two million	R	55	0	0
112110J1	Ewings	SKNEP	two million	L	57	11	349
112110J2	Ewings	rep-SKNEP	two million	R	57	0	0

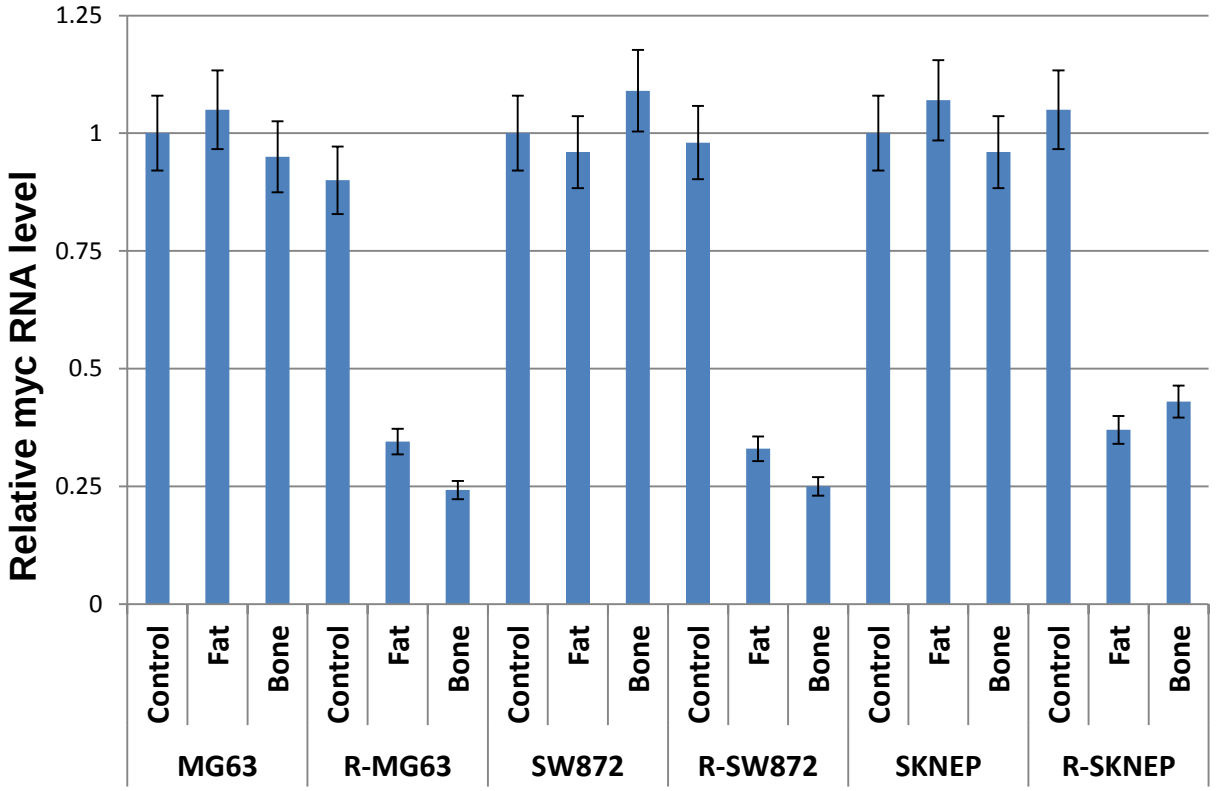
C.

MOUSE ID	CELL TYPE	CELL LINE-DIFF 21 DAYS	Differentiated	# CELLs INNOCULATEd	FLANK (L or R)	DAYS TO SAC	Greatest Dimension (mm)	WEIGHT (mg)
112810A1	osteosarcoma	rep-MG63	ODM	one million	L	125	0	0
112810A2	osteosarcoma	rep-MG63	ADM	one million	R	125	0	0
112810B1	osteosarcoma	rep-MG63	ODM	one million	L	125	0	0
112810B2	osteosarcoma	rep-MG63	ADM	one million	R	125	0	0
112810G1	liposarcoma	rep-SW872	ODM	one million	L	125	0	0
112810G2	liposarcoma	rep-SW872	ADM	one million	R	125	0	0
112810H1	liposarcoma	rep-SW872	ODM	one million	L	125	0	0
112810H2	liposarcoma	rep-SW872	ADM	one million	R	125	0	0
112810I1	Ewings	rep-SKNEP	ODM	one million	L	125	0	0
112810I2	Ewings	rep-SKNEP	ADM	one million	R	125	0	0
112810J1	Ewings	rep-SKNEP	ODM	one million	L	125	0	0
112810J2	Ewings	rep-SKNEP	ADM	one million	R	125	0	0
112810A1	osteosarcoma	rep-MG63	ODM	two million	L	125	0	0
112810A2	osteosarcoma	rep-MG63	ADM	two million	R	125	0	0
112810B1	osteosarcoma	rep-MG63	ODM	two million	L	125	0	0
112810B2	osteosarcoma	rep-MG63	ADM	two million	R	125	0	0
112810G1	liposarcoma	rep-SW872	ODM	two million	L	125	0	0
112810G2	liposarcoma	rep-SW872	ADM	two million	R	125	0	0
112810H1	liposarcoma	rep-SW872	ODM	two million	L	125	0	0
112810H2	liposarcoma	rep-SW872	ADM	two million	R	125	0	0
112810I1	Ewings	rep-SKNEP	ODM	two million	L	125	0	0
112810I2	Ewings	rep-SKNEP	ADM	two million	R	125	0	0
112810J1	Ewings	rep-SKNEP	ODM	two million	L	125	0	0
112810J2	Ewings	rep-SKNEP	ADM	two million	R	125	0	0

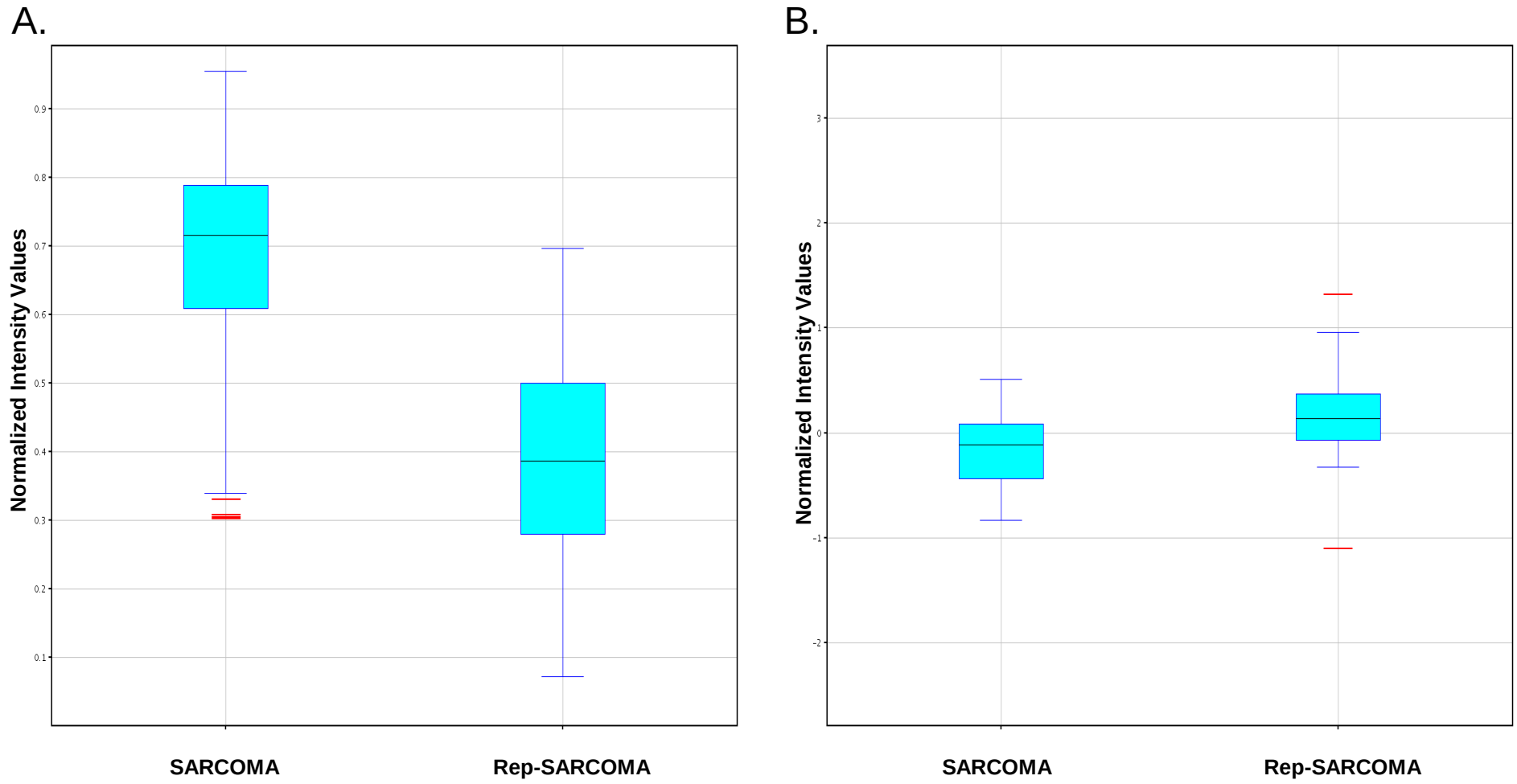
Supplemental Figure 9:

MOUSE ID	CELL TYPE	CELL LINE-DIFF 21 DAYS	Differentiated	# CELLs INNOCULATed	FLANK (L or R)	DAYS TO SAC	Greatest Dimension (mm)	WEIGHT (mg)
120110A1	osteosarcoma	rep-MG63	erythroid	one million	L	122	0	0
120110A2	osteosarcoma	rep-MG63	erythroid	one million	R	122	0	0
120110B1	liposarcoma	rep-SW872	erythroid	one million	L	122	0	0
120110B2	liposarcoma	rep-SW872	erythroid	one million	R	122	0	0
120110G1	Ewings	rep-SKNEP	erythroid	one million	L	122	0	0
120110G2	Ewings	rep-SKNEP	erythroid	one million	R	122	0	0
120110H1	osteosarcoma	rep-MG63	erythroid	two million	L	122	0	0
120110H2	osteosarcoma	rep-MG63	erythroid	two million	R	122	0	0
120110I1	liposarcoma	rep-SW872	erythroid	two million	L	122	0	0
120110I2	liposarcoma	rep-SW872	erythroid	two million	R	122	0	0
120110J1	Ewings	rep-SKNEP	erythroid	two million	L	122	0	0
120110J2	Ewings	rep-SKNEP	erythroid	two million	R	122	0	0

Supplemental Figure 10:



Supplemental Figure 11:



Supplemental Figure 12: