Title

Transformation of Human Mesenchymal Stem Cells into High-Grade Sarcomas by YAP1 and K-RAS Reflects the Undifferentiated Pleomorphic Sarcoma-Myxofibrosarcoma Disease Spectrum

Authors

Maria A. Munoz[1,2], Jack Freeland[1,3], Edmond O'Donnell III[4], Justin Langerman[5], Morgan Darrow[6], Steven Thorpe[4], Robert Randall[4], Kathrin Plath[5], Kermit Carraway[7], Owen Witte[8], Thomas Graeber[3], Janai R. Carr-Ascher*[,2,4]

[1]These authors contributed equally to this work

Affiliations

[2]Department of Internal Medicine, Division of Hematology/Oncology, University of California, Davis; Sacramento, CA, 95817, USA.
[3]Department of Molecular and Medical Pharmacology, University of California, Los Angeles; Los Angeles, CA, 90095, USA
[4]Department of Orthopedic Surgery, University of California, Davis; Sacramento, CA, 95817, USA
[5]Department of Biological Chemistry, University of California, Los Angeles; Los Angeles, CA, 90095, USA
[6]Department of Pathology and Laboratory Medicine, University of California, Davis; Sacramento, CA, 95817, USA
[7]Department of Biochemistry and Molecular Medicine, University of California, Davis; Sacramento, CA, 95817, USA
[8]Department of Microbiology, Immunology, and Molecular Genetics, University of California, Los Angeles; Los Angeles, CA, 90095, USA

Lead Contact

*Correspondence: jrcarr@ucdavis.edu

Declaration of Interests

The authors declare no competing interests

Abstract

High-grade complex karyotype sarcomas are a heterogeneous group of more than seventy tumors that vary in histology, clinical course, and patient demographics. Despite these differences, sarcomas are treated similarly with varying efficacy. This is in part due to the rarity of the disease and the lack of preclinical models for the study of individual sarcoma subtypes. Here, we describe the development of a model of high-grade sarcoma formation from human mesenchymal stem cells by introducing genetic changes that are seen in complex karyotype sarcomas. RB1 and P53 are tumor suppressors that are often mutated or functionally inactivated in sarcomas and these were targeted using CRISPR-Cas9 technology which, alone did not result in sarcoma formation. To identify drivers, a lentiviral library screen was performed and the results were validated in a secondary screen. From this, YAP1 and K-RAS were found to robustly drive the formation of undifferentiated pleomorphic sarcoma and myxofibrosarcoma which exist on a disease spectrum. Other drivers were identified that resulted in high grade sarcomas. The majority of these tumors were undifferentiated pleomorphic sarcoma although pleomorphic leiomyosarcoma and osteosarcoma were present in a subset of tumors driven by CDK4 and PIK3CA. This tool allows for the study of human sarcoma subtype development from mesenchymal stem cells and provides a platform for further mechanistic studies, investigations of metastasis, and drug development for this aggressive and heterogeneous disease.

Introduction

Sarcomas are a heterogeneous group of tumors with a 65% survival rate at 5 years despite multimodality treatment with surgery, radiation, and chemotherapy. These aggressive tumors carry a poor prognosis due to high rates of metastasis after initial treatment(Fletcher et al., 2013; NCI). Current therapies are unable to control the disease and new treatments are needed. Targeted therapies have led to a significant improvement in patient survival across tumor types, yet, within sarcomas, few targets have been identified and therapies directed towards these have been met with varied success(Goutsouliak et al., 2020; Llovet et al., 2018; Tap et al., 2020).

There are more than 70 sarcoma subtypes that vary in their histology, genetics, and patient demographics (Fletcher et al., 2013). These can be genetically subcategorized into two distinct groups. First, those that have a simple karyotype where the biology is driven by a known translocation. The most common in adults is synovial sarcoma, which contains a SYT-SSX gene fusion leading to chromatin remodeling and subsequent tumor formation (Nacev et al., 2020). Second, there are complex karyotype sarcomas that display an aneuploidy phenotype with numerous chromosomal gains and losses (Helman and Meltzer, 2003). The majority of sarcomas are of complex karyotype (85% in adults) including undifferentiated pleomorphic sarcoma (UPS), a highly aggressive tumor characterized by a lack of specific markers of differentiation. While sarcomas containing a translocation have a clearly defined driver, pathways and factors leading to the development of complex karyotype sarcomas are not known. As a result, clinically, these complex karyotype sarcomas are treated similarly with varying degrees of success. This is in part due to the rarity of sarcomas as they constitute less than 1% of all tumors leading to limited biological material for study and small sample sizes for clinical trials resulting in a generalization of treatments for a heterogeneous disease(Lee et al., 2019). Identification of critical drivers of individual sarcoma subtypes could result in an improved understanding of sarcoma biology and identify new subtype specific therapeutic targets for this devastating disease.

To gain insight into the biology of complex karyotype soft tissue sarcomas, The Cancer Genome Atlas (TCGA) evaluated the most common complex karyotype sarcomas. These tumors were found to have a low mutation burden and a high rate of aneuploidy(Abeshouse et al., 2017). The most common mutations and chromosomal alterations identified led to loss of *RB1* or *P53*, known tumor suppressors that have established roles in sarcoma biology (Helman and Meltzer, 2003; Kirsch et al., 2007; Thoenen et al., 2019). Incorporating TCGA data as well as other large datasets, *P53* was found to be altered in 47% of sarcomas. *RB1* mutation or deletion was seen in 22% of tumors with the majority of these also carrying an alteration in P53(Bui et al., 2019). These studies indicate that RB1 and P53 play important roles in sarcoma biology and this has been shown in mouse models where targeting of P53 drives sarcoma formation (Kirsch et al., 2007; Lin et al., 2009).

Mesenchymal stem cells (MSCs) are capable of giving rise to connective tissues of the body and therefore, are the presumed cell of origin for bone and soft tissue sarcomas (Pittenger et al., 1999). This is exemplified by studies in mice where loss of P53 in mouse mesenchymal stem cells was sufficient for transformation to sarcoma(Rubio et al., 2010). In contrast, it is clear that loss of P53 alone is insufficient to induce sarcoma formation from human mesenchymal stem cells (Rodriguez et al., 2012). Prior models of human sarcoma have relied primarily on viral proteins or mutated oncogenes to drive transformation, and these have resulted in the formation of a single sarcoma subtype (Funes et al., 2007; Li et al., 2009; Matushansky et al., 2007; Serakinci et al., 2004; Shima et al., 2007). While these models have provided valuable insights, the ability to make comparisons among human sarcoma subtypes has been limited.

Here, we describe the development and characterization of a new pre-clinical model of human sarcoma formation from mesenchymal stem cells by introducing genetic alterations observed across human sarcomas. This model provides the opportunity to understand the development of human complex karyotype sarcomas and identify drivers of sarcoma development which can inform new subtype specific targeted therapies.

Results

Establishment of a Pre-Clinical Sarcoma Model

To generate an i*n vivo* model of sarcoma formation, we began with human MSCs. While mouse MSCs can be transformed to sarcomas by targeting of single genes such as *RB1* and *P53*, human MSCs require multiple hits (Rodriguez et al., 2012). A putative mechanism for this variation between species is due to differences in telomere maintenance. This is exemplified by the successful transformation of human MSCs when *hTERT* has been added along with viral proteins or mutated *H-Ras* (Funes et al., 2007; Serakinci et al., 2004; Shima et al., 2007). Therefore, as a basis for these studies, we utilized a mesenchymal stem cell line that has been immortalized by *hTERT*, ASC52telo(Wolbank et al., 2011). This cell line expresses surface markers indicative of mesenchymal stem cells including CD44, CD73, and CD90 in addition to staining negative for CD19, CD34, and CD45. These cells are capable of undergoing in vitro differentiation in adipogenic, osteogenic, and chondrogenic lineages (Supplemental Figure 1A).

RB1 and P53 are commonly inactivated or mutated in human soft tissue sarcomas (Helman and Meltzer, 2003). Data from The Cancer Genome Atlas (TCGA) demonstrated not only the loss of *RB1* and/or *P53* by deletion at the chromosomal level or through somatic mutation but, there is also evidence of amplification of negative regulators of these proteins further emphasizing their importance in sarcoma biology (Abeshouse et al., 2017). *RB1* and *P53* were targeted in ASC52telo using a CRISPR-Cas9 system (Doudna and Charpentier, 2014). A lentivirus containing inducible Cas9 was transduced into the ASC52telo cell line(Wang et al., 2014). Single cell clones were expanded and tested to ensure tight regulation of Cas9 (Supplemental Figure 1B). An inducible system was used to limit any potential in vitro or in vivo toxicity secondary to high levels of Cas9 (Yang et al., 2018). CRISPR guides targeting either *RB1* and *P53* or both were inserted into a lentiviral vector containing GFP (Shi et al., 2015)(Supplemental Figure 1D). Cas9 expressing ASC52telo cells were transduced with this lentivirus and single GFP+ cells were sorted, expanded, and gene knockout was verified (Supplemental Figure 1D-1E).

Implantation of *RB1-/-P53+/-* cell lines subcutaneously into mice did not yield tumors within 3 months therefore, targeting of *RB1* and *P53* alone in the background of *hTERT* immortalization is not sufficient to drive robust sarcoma formation in this model. We hypothesized that additional genetic alterations were needed for transformation of human MSCs to sarcoma therefore, we generated a lentiviral library that would allow for a genetic screen approach. Genes that have been implicated in sarcoma biology in the literature or genes that are upregulated or amplified in the TCGA soft tissue sarcoma data were included (Abeshouse et al., 2017). A total of 27 lentiviral constructs containing a gene of interest tagged with V5 and a barcode sequence were utilized in the screen (Supplemental Table 1). Control *RB1+/+P53+/+* cells and *RB1-/-P53+/-* cells were transduced with the lentiviral gene library and implanted subcutaneously into immunocompromised non-SCID gamma (NSG) mice (Figure 1A). After 4-8 weeks, 37/48 (77%) of RB1-/-P53+/- cells transduced with the library formed tumors while control cells did not, indicating that the lentiviral library alone is not capable of transforming MSCs to sarcomas.

Histologic evaluation of these tumors demonstrated highly pleomorphic cells with numerous mitotic figures and areas of necrosis consistent with high-grade sarcoma (Figure 1B-1C). Notably, there were two distinct histologies seen. The first and most abundant histology is consistent with undifferentiated pleomorphic sarcoma (UPS) while the second area contains features of aggressive sarcoma but also has abundant myxoid accumulation, consistent with high grade myxofibrosarcoma (MFS) (Figure 1D-1E). Immunohistochemistry of these tumors showed strong positivity for vimentin with alpha smooth muscle actin marking endothelial cells and infrequent tumor cells. Myogenin was consistently negative indicating the tumors did not demonstrate evidence of muscle differentiation that would be consistent with rhabdomyosarcoma (Figure 1F). Taken together, these tumors contain predominantly undifferentiated pleomorphic sarcoma histology with areas consistent with myxofibrosarcoma. Decreasing the cell number or amount of lentivirus led to a longer tumor latency but did not have an impact on histologic features (data not shown).

The tumors generated are histologically consistent with human high-grade sarcomas therefore, we wanted to test if they are functionally similar human tumors. To do this, tumors were injected into the gastrocnemius muscle of immunocompromised mice. In humans the majority of sarcomas are found in the extremities and metastasize almost exclusively to the lungs(Billingsley et al., 1999). Consistent with this, implantation of *RB-/-P53+/-* cells with the lentiviral library into the gastrocnemius gave rise to lung metastases in 5/12 (42%) animals without metastases to lymph nodes, liver, or spleen (Figure 1G). Using this defined and highly efficient system, mesenchymal stem cells can be transformed into high grade sarcomas that histologically and functionally represent the human disease.

Loss of Either RB1 or P53 but not Both is Required for Transformation

In mouse models, loss of P53 is sufficient for tumor formation but in human MSC transformation to high-grade sarcoma, the requirements for RB1 and P53 are not known (Donehower et al., 1991; Jacks et al., 1994). To test this, ASC52telo cells that have loss of RB1 alone (*RB1-/-P53+/+*), P53 alone (*RB+/+P53+/-*) or both (*RB1-/-P53+/-* cells) were generated in addition to control cells (*RB1+/+P53+/+*) (Supplemental Figures 1D, 1F). Interestingly, screening of numerous clones all showed heterozygous loss of *P53*, homozygous allele loss was not observed and presumably not tolerated by the cells. Functionally, P53 is inactive as evidenced by the lack of induction in response to doxorubicin treatment (Figure 2A).

To determine the requirement for RB1 and P53 in sarcoma development, we transduced each of these four cell types with the lentiviral library and implanted these cells into immunocompromised mice. As expected, control cells with RB1 and P53 intact did not generate tumor outgrowths within 12 weeks. *RB1-/-P53+/-, RB1+/+P53+/-, and RB1-/-P53-/-* with addition of the lentiviral library gave rise to high grade sarcomas. Evaluation of these tumors histologically did not show any differences between those generated by loss of RB1, P53, or both (Figure 2B). This indicates that loss of one tumor suppressor is necessary for transformation but targeting both RB1 and P53 are not required.

YAP1 and K-RAS Driven Tumors Histologically Reflect the Undifferentiated Pleomorphic Sarcoma-Myxofibrosarcoma Spectrum

To identify which genes in the lentiviral library are capable of driving the transformation of MSCs into high-grade sarcomas, the unique barcodes were analyzed using next generation sequencing. Of note, all lentiviral library constructs were expressed although at varying levels prior to implantation and there was a minimal change in the relative levels over five days of cell culture, indicating significant *in vitro* selection did not occur (Figure 3A). From barcode sequencing of tumor outgrowths, YAP1, DDIT3, and K-RAS were enriched for and expressed at high levels across multiple tumors (Figure 3B).

Yes associated protein (YAP1) is a key downstream mediator of Hippo signaling and has been shown to be overexpressed or amplified in sarcomas as well as other cancer types (Eisinger-Mathason et al., 2015; Moroishi et al., 2015; Zanconato et al., 2016). The FUS-DDIT3 translocation leads to the formation of myxoid/round cell liposarcoma and amplification of DDIT3 has been shown to be associated with a myxoid liposarcoma like histology (Mantilla et al., 2019). Rare mutations in oncogenic H-RAS and K-RAS have been noted in genomic analysis of sarcomas and implicated in transformation, but a role for wildtype K-RAS has not been established (Shima et al., 2007; Yoo et al., 1999). While these factors have been implicated in sarcoma biology, they have not been shown to be drivers of human soft tissue sarcoma formation.

To determine if these factors are capable of transforming MSCs to sarcomas as single genes, a secondary screen was performed. *RB1-/-P53+/-* ASC52telo cells were transduced with a negative control, each lentivirus alone (YAP1, K-RAS, and DDIT3) or the combination as a positive control. As expected, the combination of all three genes was sufficient to drive transformation. Interestingly, YAP1 and K-RAS alone led to tumor formation within 8-12 weeks while addition of *DDIT3* alone did not after 26 weeks (Figure 3C). These tumors can be passaged in vivo with second passage cells giving rise to tumors within 2 weeks (data not shown). Of note, the negative control led to 1/8 (12.5%) implants forming tumors after 26 weeks that histologically represented undifferentiated pleomorphic sarcoma indicating that spontaneous transformation can occur in these cells at a long latency.

TCGA data demonstrated that undifferentiated pleomorphic sarcoma and myxofibrosarcoma exist on a disease spectrum as the molecular analysis could not differentiate the two by unsupervised clustering (Abeshouse et al., 2017). YAP1 tumors were histologically consistent with undifferentiated pleomorphic sarcoma as evidenced by a lack of expression of markers of differentiated cell types (Figure 3D). K-RAS tumors had the distinct appearance of myxofibrosarcoma histologically (Figure 3E). While DDIT3 alone did not give rise to tumors within 26 weeks (0/8 implants), the addition of DDIT3 to K-RAS tumors led to a more aggressive phenotype with increased pleomorphism (Figure 3F) indicating this is likely a modifier of sarcoma development. Taken together, this data demonstrates that YAP1 and K-RAS driven tumors histologically recreate the human undifferentiated pleomorphic sarcoma/myxofibrosarcoma spectrum.

Identification of Additional Drivers of High-Grade Sarcoma Formation

It is possible that additional factors in the lentiviral library could lead to transformation and this was masked by a proliferative advantage conferred by K-RAS, YAP1, or DDIT3. Therefore, the primary screen was repeated removing K-RAS, YAP1, and DDIT3 from the lentiviral library. Evaluation of these tumors showed histology consistent with undifferentiated pleomorphic sarcoma. Barcode sequencing of these tumors identified several potential drivers. Genes whose barcodes comprised >10% of represented barcodes in one tumor or >5% in two outgrowths were tested individually in a secondary screen. This included SOX2, PIK3CA, JUN, DKK1, CDK4, CCND1, and BCL2 (Figure 4A).

PIK3CA encodes the catalytic subunit of PIK3 which activates Akt signaling. Rarely, mutations of PIK3CA have been noted in soft tissue sarcomas and these are associated with decreased survival (Barretina et al., 2010). Addition of PI3KCA to RB1-/-P53+/- cells led to formation of tumors although with a significantly lower efficacy of approximately 25% compared to 75% of implants forming tumors when expressing YAP1 and 87.5% with K-RAS. Additionally, a longer latency was noted as compared to YAP1 and K-RAS driven tumors (average 115.3 and 62 days respectively). Interestingly, the histology of 20% of these tumors was consistent with osteosarcoma and this was confirmed by strong staining for SATB2 (Figure 4B) while 80% represented undifferentiated pleomorphic sarcoma.

Cyclin dependent kinase 4 (CDK4) binds in a complex with cyclin D to phosphorylate RB1 and alleviate repression of E2F to promote cell cycle progression(Hsu et al., 2022). Alterations in this pathway have been previously noted in large sarcoma sequencing datasets(Jardim et al., 2021). In TCGA data, CDK4 was shown to be amplified in up to 18.5% of samples(Abeshouse et al., 2017). CDK4 and MDM2 are both located on chromosome 12q and amplification of this region is diagnostic for liposarcomas and CDK4/6 inhibitors have shown some efficacy against this sarcoma subtype(Dickson et al., 2016). Similar to PIK3CA, these tumors had a significantly longer latency with an average of 142 days until tumors reach 1cm and a decreased penetrance with 37.5% of implants giving rise to tumors. Of those, 50% of tumors displayed strong staining of SMA and had a more spindled appearance consistent with pleomorphic leiomyosarcoma while the remaining tumors were undifferentiated pleomorophic sarcoma (Figure 4C).

JUN robustly drove the formation of high-grade sarcoma with an average latency of 71 days and 100% (8/8) penetrance. Tumors showed areas of both undifferentiated pleomorphic sarcoma and pleomorphic leiomyosarcoma as evidenced by positive SMA staining and a more spindled morphology (Supplemental Figure 2A). BCL2 and DKK1 gave rise to undifferentiated pleomorphic sarcoma with an increase in giant bizarre cells with a latency of 31 and 36 weeks (Supplemental Figure 2B). Penetrance of these tumors was 50% and 37.5% respectively.

Discussion

Here we describe the development a preclinical model that robustly drives the formation of high-grade sarcoma subtypes from human mesenchymal stem cells. The step-wise requirement for transformation by first targeting tumor suppressor, then adding an oncogenic driver provides an unprecedented opportunity to study the biology of sarcoma development. This model allows for simple in vitro genetic manipulations of the roles of other tumor suppressors such as PTEN and ATRX that have been implicated in sarcoma biology(Abeshouse et al., 2017). This also provides a platform for the unbiased identification of other drivers of sarcoma formation using additional lentiviral libraries.

Sarcomas are a highly heterogenous group of tumors and an understanding of the relationships between these sarcomas is limited. Here, we have generated four sarcoma subtypes from a single cell of origin, undifferentiated pleomorphic sarcoma, myxofibrosarcoma, osteosarcoma, and leiomyosarcoma which indicates that these may have similar underlying biology. In particular, areas consistent with undifferentiated pleomorphic sarcoma were seen in tumors that had osteosarcoma or leiomyosarcoma which indicates these could also exist on a spectrum. The starting material for these tumors is a cell line with loss of hTERT and this may have some inherent genomic instability and could reflect a heterogenous population raising the possibility that this could contribute to sarcoma subtypes formed. This warrants further investigation as these studies could provide important insights into the initial events leading to sarcoma formation.

Undifferentiated pleomorphic sarcoma and myxofibrosarcoma exist on a spectrum and clinically, this is seen as well. These two tumors are treated similarly although they have different typical characteristics with myxofibrosarcomas having higher rates of local relapse and lower grade at diagnosis and undifferentiated pleomorphic sarcoma having higher rates of distant metastasis(Abeshouse et al., 2017). This model will allow for a better understanding of the biology of these tumors and shared therapeutic vulnerabilities which could inform clinical trial design and drug development. This system provides a much needed tool for the study of high grade sarcoma subtype development and provides a platform for further investigations into the biology of this heterogeneous and devastating disease.

Methods

Cloning of CRISPR/Cas9 Constructs

Guide RNA sequences targeting RB1, P53, or a control sequence (GTAATCCTAGCACTTTTAGG) were individually cloned into the LRG vector as previously described(Shi et al., 2015). Sequences were previously used in a genome wide screen by Wang et al(Wang et al., 2014). In order to combine these guides into one vector, the Gibson cloning method was used. The U6 promoter, guide sequence, and scaffold were amplified with custom primers (IDT) that contained a tag overlapping with the LRG sequence flanking the EcoRI site. LRG vectors containing the control sequence or RB1 targeting sequence were cut with EcoRI and then ligated with the amplified sequence using the NEB HiFi Assembly kit (Cat#E5520).

Lentivirus Production

To generate lentivirus, 293T cells (ATCC) were transfected with pMDL, pREV, and pVSVg(Tiscornia et al., 2006), 3rd generation lentivirus packaging plasmids and a lentivirus construct using calcium chloride transfection. Media containing virus was collected 48 hours after transfection and filtered through a 0.22µm filter. Virus was subsequently concentrated by centrifugation using Amicon 100kDa centrifugal filter units (Millipore Cat # UFC9100). Virus was titered using the Lenti-X p24 Rapid Titer Kit (Takara Cat #632200).

Cell Line Creation

To generate the ASC52telo cell lines with RB1 and P53 gene knockout or controls, ASC52 cells were cultured in MSC basal media (Lonza Cat #PT-3001). Lentivirus was generated from pCW-Cas9, Doxycycline inducible Cas9 as described above and this was added to cells using polybrene and spinfection(Wang et al., 2014). Cells were selected in puromycin and then single cells were seeded into 96 well plates. The clones were expanded and tested by western blot for Cas9 in the presence or absence of doxycycline. The clone with the tightest regulation of Cas9 was used for cell line generation. Lentivirus constructs containing CRISPR guides were transduced into cells using spinfection and polybrene. Cells were expanded and then sorted for green fluorescent protein (GFP) positivity. Individual GFP positive cells were sorted into 96 well plates to create single cell clones. Clones were expanded and tested by western blot to determine which had loss of RB1 and/or P53. The DNA level mutation leading to the knockout was determined by PCR of the region containing the CRISPR target sequence and subsequent TOPO cloning and sanger sequencing (Invitrogen K2800J10).

In Vitro Differentiation

To differentiate mesenchymal stem cells in vitro, the Lonza protocol was used. Briefly, adipogenic, osteogenic, and chondrogenic medias were prepared by addition of supplements (Cat #PT-3002, PT-3003, PT-3004). ASC52 cells were plated at $1x10^4$ per cm^2 for adipogenic differentiation and $1.5x10^3$ per cm^2 for osteogenic differentiation. Cell numbers used were less than recommended by the manufacturer because these cells lack contact inhibition. For chondrogenic differentiation, cells were resuspended at a concentration of $1.6x10^7$ and 5ul was used per pellet. Differentiation medias were added and changed every three days for 14-21 days. Cells were fixed in 10% PBS buffered formalin then stained with oil red o, alizarian red, and alcian blue as previously described for visualization(Fierro et al., 2011).

Lentivirus Library Production

LentiORF clones were purchased from Sigma containing genes of interest in TCR3 backbone (Millipore Sigma). All constructs are V5 tagged except for K-Ras. Each construct was verified by western blot using a V5 antibody. Plasmids were maxiprepped, quantified, then pooled together for lentivirus production. Lentivirus was produced using the 3rd generation system as described above. In brief, 293T cells were transfected, media containing virus was collected after 48 hours and filtered. The virus was concentrated by centrifugation and tittered for addition to cells.

Primary Screen

Cells lines with RB1 and P53 targeted or a control line were transduced with the lentiviral library at MOI of 200 using spinfection and polybrene. Cells were expanded in culture for 3-5 days and then trypsinized, counted, and suspended in 60ul of a 1:1 mix of HBSS and Matrigel (Corning). These cells were injected subcutaneously into 6-8 week old NSG mice (Jackson Labs Cat#005557). Tumors were monitored by palpation and once tumors reached 1cm, were harvested and both frozen and formalin fixed samples were collected.

Secondary Screen
To test the ability of individual oncogenes to drive tumor transformation, genes were cloned into a lentiviral construct that contains red fluorescent protein (RFP). Lentivirus was produced as described above and tittered. Cells were transduced with lentivirus at MOI 200. After 3-5 days, cells were sorted for RFP positivity. Cells were grown in culture and expanded then trypsizined, counted, and suspended in 60ul of a 1:1 mix of HBSS and Matrigel then injected subcutaneously into NSG mice.

Barcode Sequencing
For analysis of barcodes in tumor samples, DNA was extracted from paraffin embedded samples. Tumors were sectioned and stained by H&E to ensure the presence of viable tissue. The subsequent section was removed from the slide and DNA was extracted (Qiagen, Cat#56404). Barcodes were amplified using flanking primers (CTTGAAAGTATTTCGATTTCTTGGC and TCCAGAGGTTGATTGTCGAC) and KAPA HiFi HotStart ReadyMix (Cat#K2602) to form a 104bp product. Adapters for next generation sequencing were ligated to the PCR products using the KAPA Hyper Plus Kit and KAPA dual indexed adapter kit (Cat#K8727 and Cat#KK8514). Libraries were evaluated using a bioanalyzer and samples were submitted for next generation sequencing to UC Berkeley QB3 genomics and sequenced using Illumina technology. Barcodes were extracted from sequenced library fastQs by correcting for read direction using grep to sort, then taking the first 24bp of each read, and then determining the top 100 unique most frequent barcodes in each library using uniq in unix. Sequences were checked for matches within 1 levenshtein distance by building a dictionary of possible sequences in R software and then searching for matches to known transgene barcodes. The total of each transgene barcode detected per library was tabulated from this data.

Histology
Paraffin embedding, sectioning, and H&E staining was performed by the UC Davis Genomic Pathology Lab. SATB2, myogenin, and alpha SMA staining was performed by the UCLA TPCL facility. Vimentin staining was done according to the following protocol. Immunohistochemistry (IHC) staining was performed on formalin-fixed paraffin embedded (FFPE) tissue sections. Sections were de-paraffinized in Xylene, then re-hydrated. Antigen Unmasking Solution (Vector Laboratories, Cat# H-3300), citrate-based pH 6.0, was used for antigen retrieval. Endogenous peroxidase activity was blocked by incubating sections with BLOXALL (Vector, Cat# SP-6000) for 10 minutes. The M.O.M Mouse IgG Blocking Reagent from the M.O.M ImmPRESS Polymer Kit (Vector, Cat# MP-2400) was used to block endogenous mouse IgG staining. The Vimentin primary antibody (Agilent, Cat#M072529-2) was diluted 1:75 in 2.5% horse serum. Secondary antibody staining was done with the M.O.M ImmPress Horse Anti-Mouse IgG for 10 minutes. The ImmPACT DAB EqV (Vector, Cat# SK-4103) reagent was used for visualization. Sections were counterstained with hematoxylin and mounted with VectaMount AQ Mounting Medium (Vector, CAT# H-5501).

Acknowledgements
This work was supported in part by the UC Davis Paul Calabresi Career Development Award for Clinical Oncology as funded by the National Cancer Institute/National Institutes of Health through grant #5K12-CA138464 (Carr-Ascher).

The authors thank the Qian Chen and the UC Davis Center for Genomic Pathology Lab and the UCLA Translational Pathology Core Laboratory (TPCL) for histology and immunohistochemistry support. In addition, Jonathan Van Dyke and the UC Davis Flow Cytometry Shared Resource for FACS and cell sorting support. The authors thank the UC Berkeley QB3 Genomics core for next generation sequencing support.

pCW-Cas9 was a gift from Eric Lander & David Sabatini (Addgene plasmid # 50661 ; http://n2t.net/addgene:50661 ; RRID:Addgene_50661). LRG (Lenti_sgRNA_EFS_GFP) was a gift from Christopher Vakoc (Addgene plasmid # 65656 ; http://n2t.net/addgene:65656 ; RRID:Addgene_65656).

Citations
Abeshouse, A., Adebamowo, C., Adebamowo, S.N., Akbani, R., Akeredolu, T., Ally, A., Anderson, M.L., Anur, P., Appelbaum, E.L., Armenia, J., et al. (2017). Comprehensive and Integrated Genomic Characterization of Adult Soft Tissue Sarcomas. Cell *171*, 950-965.e28. https://doi.org/10.1016/j.cell.2017.10.014.

Barretina, J., Taylor, B.S., Banerji, S., Ramos, A.H., Lagos-Quintana, M., Decarolis, P.L., Shah, K., Socci, N.D., Weir, B.A., Ho, A., et al. (2010). Subtype-specific genomic alterations define new targets for soft-tissue sarcoma therapy. Nature Genetics *42*, 715–721. https://doi.org/10.1038/ng.619.

Billingsley, K.G., Burt, M.E., Jara, E., Ginsberg, R.J., Woodruff, J.M., Leung, D.H.Y., and Brennan, M.F. (1999). Pulmonary Metastases From Soft Tissue Sarcoma. Ann Surg *229*, 602. .

Bui, N.Q., Przybyl, J., Trabucco, S.E., Frampton, G., Hastie, T., van de Rijn, M., and Ganjoo, K.N. (2019). A clinico-genomic analysis of soft tissue sarcoma patients reveals CDKN2A deletion as a biomarker for poor prognosis. Clinical Sarcoma Research *9*, 1–11. https://doi.org/10.1186/s13569-019-0122-5.

Dickson, M.A., Schwartz, G.K., Keohan, M.L., D'Angelo, S.P., Gounder, M.M., Chi, P., Antonescu, C.R., Landa, J., Qin, L.-X., Crago, A.M., et al. (2016). Progression-Free Survival Among Patients With Well-Differentiated or Dedifferentiated Liposarcoma Treated With CDK4 Inhibitor Palbociclib: A Phase 2 Clinical Trial. JAMA Oncology *2*, 937–940. https://doi.org/10.1001/jamaoncol.2016.0264.

Donehower, L.A., Harvey, M., Slagle, B.L., Mcarthurt, M.J., Jrt, C.A.M., Butel, J.S., and Bradley, A. (1991). Mice deficient for p53 are developmentally normal but susceptible to spontaneous tumours. 215–221. .

Doudna, J.A., and Charpentier, E. (2014). The new frontier of genome engineering with CRISPR-Cas9. Science *346*. https://doi.org/10.1126/science.1258096.

Eisinger-Mathason, T.S.K., Mucaj, V., Biju, K.M., Nakazawa, M.S., Gohil, M., Cash, T.P., Yoon, S.S., Skuli, N., Park, K.M., Gerecht, S., et al. (2015). Deregulation of the Hippo pathway in soft-tissue sarcoma promotes FOXM1 expression and tumorigenesis. Proceedings of the National Academy of Sciences *112*, E3402–E3411. https://doi.org/10.1073/PNAS.1420005112.

Fierro, F.A., Kalomoiris, S., Sondergaard, C.S., and Nolta, J.A. (2011). Effects on Proliferation and Differentiation of Multipotent Bone Marrow Stromal Cells Engineered to Express Growth Factors for Combined Cell and Gene Therapy. Stem Cells *29*, 1727–1737. https://doi.org/10.1002/stem.720.

Fletcher, C., Bridge, J., Hogendoorn, P., and Mertens, F. (2013). WHO Classification of Tumours of Soft Tissue and Bone (IARC: Lyon).

Funes, J.M., Quintero, M., Henderson, S., Martinez, D., Qureshi, U., Westwood, C., Clements, M.O., Bourboulia, D., Pedley, R.B., Moncada, S., et al. (2007). Transformation of human mesenchymal stem cells increases their dependency on oxidative phosphorylation for energy production. Proceedings of the National Academy of Sciences of the United States of America *104*, 6223–6228. https://doi.org/10.1073/pnas.0700690104.

Goutsouliak, K., Veeraraghavan, J., Sethunath, V., De Angelis, C., Osborne, C.K., Rimawi, M.F., and Schiff, R. (2020). Towards personalized treatment for early stage HER2-positive breast cancer. Nature Reviews Clinical Oncology *17*, 233–250. https://doi.org/10.1038/s41571-019-0299-9.

Helman, L.J., and Meltzer, P. (2003). Mechanisms of sarcoma development. Nature Reviews Cancer *3*, 685–694. https://doi.org/10.1038/nrc1168.

Hsu, J.Y., Seligson, N.D., Hays, J.L., Miles, W.O., and Chen, J.L. (2022). Clinical Utility of CDK4/6 Inhibitors in Sarcoma: Successes and Future Challenges. JCO Precision Oncology e2100211. https://doi.org/10.1200/PO.21.00211.

Jacks, T., Remington, L., Williams, B.O., Schmitt, E.M., Halachmi, S., Bronson, R.T., and Weinberg, R.A. (1994). Tumor spectrum analysis in p53-mutant mice. Current Biology *4*, 1–7. https://doi.org/10.1016/S0960-9822(00)00002-6.

Jardim, D.L., Millis, S.Z., Ross, J.S., Woo, M.S., Ali, S.M., and Kurzrock, R. (2021). Cyclin Pathway Genomic Alterations Across 190,247 Solid Tumors: Leveraging Large-Scale Data to Inform Therapeutic Directions. The Oncologist *26*, e78–e89. https://doi.org/10.1634/theoncologist.2020-0509.

Kirsch, D.G., Dinulescu, D.M., Miller, J.B., Grimm, J., Santiago, P.M., Young, N.P., Nielsen, G.P., Quade, B.J., Chaber, C.J., Schultz, C.P., et al. (2007). A spatially and temporally restricted mouse model of soft tissue sarcoma. Nature Medicine *13*, 992–997. https://doi.org/10.1038/nm1602.

Lee, D.Y., Staddon, A.P., Shabason, J.E., and Sebro, R. (2019). Phase I and phase II clinical trials in sarcoma: Implications for drug discovery and development. Cancer Medicine *8*, 585–592. https://doi.org/10.1002/cam4.1958.

Li, N., Yang, R., Zhang, W., Dorfman, H., Rao, P., and Gorlick, R. (2009). Genetically transforming human mesenchymal stem cells to sarcomas: Changes in cellular phenotype and multilineage differentiation potential. Cancer *115*, 4795–4806. https://doi.org/10.1002/cncr.24519.

Lin, P.P., Pandey, M.K., Jin, F., Raymond, A.K., Akiyama, H., and Lozano, G. (2009). Targeted mutation of p53 and Rb in mesenchymal cells of the limb bud produces sarcomas in mice. Carcinogenesis *30*, 1789–1795. https://doi.org/10.1093/carcin/bgp180.

Llovet, J.M., Montal, R., Sia, D., and Finn, R.S. (2018). Molecular therapies and precision medicine for hepatocellular carcinoma. Nature Reviews Clinical Oncology *15*, 599–616. https://doi.org/10.1038/s41571-018-0073-4.

Mantilla, J.G., Ricciotti, R.W., Chen, E.Y., Liu, Y.J., and Hoch, B.L. (2019). Amplification of DNA damage-inducible transcript 3 (DDIT3) is associated with myxoid liposarcoma-like morphology and homologous lipoblastic differentiation in dedifferentiated liposarcoma. Modern Pathology *32*, 585–592. https://doi.org/10.1038/s41379-018-0171-y.

Matushansky, I., Hernando, E., Socci, N.D., Mills, J.E., Matos, T.A., Edgar, M.A., Singer, S., Maki, R.G., and Cordon-Cardo, C. (2007). Derivation of sarcomas from mesenchymal stem cells via inactivation of the Wnt pathway. Journal of Clinical Investigation *117*, 3248–3257. https://doi.org/10.1172/JCI31377.

Moroishi, T., Hansen, C.G., and Guan, K.L. (2015). The emerging roles of YAP and TAZ in cancer. Nature Reviews Cancer *15*, 73–79. https://doi.org/10.1038/nrc3876.

Nacev, B.A., Jones, K.B., Intlekofer, A.M., Yu, J.S.E., Allis, C.D., Tap, W.D., Ladanyi, M., and Nielsen, T.O. (2020). The epigenomics of sarcoma. Nature Reviews Cancer *20*, 608–623. https://doi.org/10.1038/s41568-020-0288-4.

NCI Surveillance, Epidemiology, and End Results (SEER).

Pittenger, M.F., Mackay, A.M., Beck, S.C., Jaiswal, R.K., Douglas, R., Mosca, J.D., Moorman, M.A., Simonetti, D.W., Craig, S., and Marshak, D.R. (1999). Multilineage potential of adult human mesenchymal stem cells. Science *284*, 143–147. https://doi.org/10.1126/science.284.5411.143.

Rodriguez, R., Rubio, R., and Menendez, P. (2012). Modeling sarcomagenesis using multipotent mesenchymal stem cells. Cell Research *22*, 62–77. https://doi.org/10.1038/cr.2011.157.

Rubio, R., García-Castro, J., Gutiérrez-Aranda, I., Paramio, J., Santos, M., Catalina, P., Leone, P.E., Menendez, P., and Rodríguez, R. (2010). Deficiency in p53 but not Retinoblastoma Induces the Transformation of Mesenchymal Stem Cells In vitro and Initiates Leiomyosarcoma In vivo. Cancer Research *70*, 4185–4194. https://doi.org/10.1158/0008-5472.CAN-09-4640.

Serakinci, N., Guldberg, P., Burns, J.S., Abdallah, B., Schrødder, H., Jensen, T., and Kassem, M. (2004). Adult human mesenchymal stem cell as a target for neoplastic transformation. Oncogene *23*, 5095–5098. https://doi.org/10.1038/sj.onc.1207651.

Shi, J., Wang, E., Milazzo, J.P., Wang, Z., Kinney, J.B., and Vakoc, C.R. (2015). Discovery of cancer drug targets by CRISPR-Cas9 screening of protein domains. Nature Biotechnology *33*, 661–667. https://doi.org/10.1038/nbt.3235.

Shima, Y., Okamoto, T., Aoyama, T., Yasura, K., Ishibe, T., Nishijo, K., Shibata, K.R., Kohno, Y., Fukiage, K., Otsuka, S., et al. (2007). In vitro transformation of mesenchymal stem cells by oncogenic H-rasVal12. Biochemical and Biophysical Research Communications *353*, 60–66. https://doi.org/10.1016/j.bbrc.2006.11.137.

Tap, W.D., Wagner, A.J., Schöffski, P., Martin-Broto, J., Krarup-Hansen, A., Ganjoo, K.N., Yen, C.C., Abdul Razak, A.R., Spira, A., Kawai, A., et al. (2020). Effect of Doxorubicin Plus Olaratumab vs Doxorubicin Plus Placebo on Survival in Patients with Advanced Soft Tissue Sarcomas: The ANNOUNCE Randomized Clinical Trial. JAMA - Journal of the American Medical Association *323*, 1266–1276. https://doi.org/10.1001/jama.2020.1707.

Thoenen, E., Curl, A., and Iwakuma, T. (2019). TP53 in bone and soft tissue sarcomas. Pharmacology and Therapeutics *202*, 149–164. https://doi.org/10.1016/j.pharmthera.2019.06.010.

Tiscornia, G., Singer, O., and Verma, I.M. (2006). Production and purification of lentiviral vectors. Nat Protoc *1*, 241–245. https://doi.org/10.1038/nprot.2006.37.

Wang, T., Wei, J.J., Sabatini, D.M., and Lander, E.S. (2014). Genetic screens in human cells using the CRISPR-Cas9 system. Science *343*, 80–84. https://doi.org/10.1126/science.1246981.

Wolbank, S., Ph, D., Stadler, G., Peterbauer, A., Sc, M., Katinger, H., Griensven, M. Van, and Redl, H. (2011). Telomerase Immortalized Human Amnion- and Adipose-Derived Mesenchymal Stem Cells: Maintenance of Differentiation and Immunomodulatory Characteristics. Tissue Engineering Part A *15*, 1843–1854. https://doi.org/10.1089/ten.tea.2008.0205.Telomerase.

Yang, S., Li, S., and Li, X.J. (2018). Shortening the Half-Life of Cas9 Maintains Its Gene Editing Ability and Reduces Neuronal Toxicity. Cell Reports *25*, 2653-2659.e3. https://doi.org/10.1016/j.celrep.2018.11.019.

Yoo, J., Robinson, R.A., and Lee, J.Y. (1999). H-ras and K-ras gene mutations in primary human soft tissue sarcoma: concomitant mutations of the ras genes. Modern Pathology : An Official Journal of the United States and Canadian Academy of Pathology, Inc *12*, 775–780. .

Zanconato, F., Cordenonsi, M., and Piccolo, S. (2016). YAP/TAZ at the Roots of Cancer. Cancer Cell *29*, 783–803. https://doi.org/10.1016/j.ccell.2016.05.005.

Figure 1.

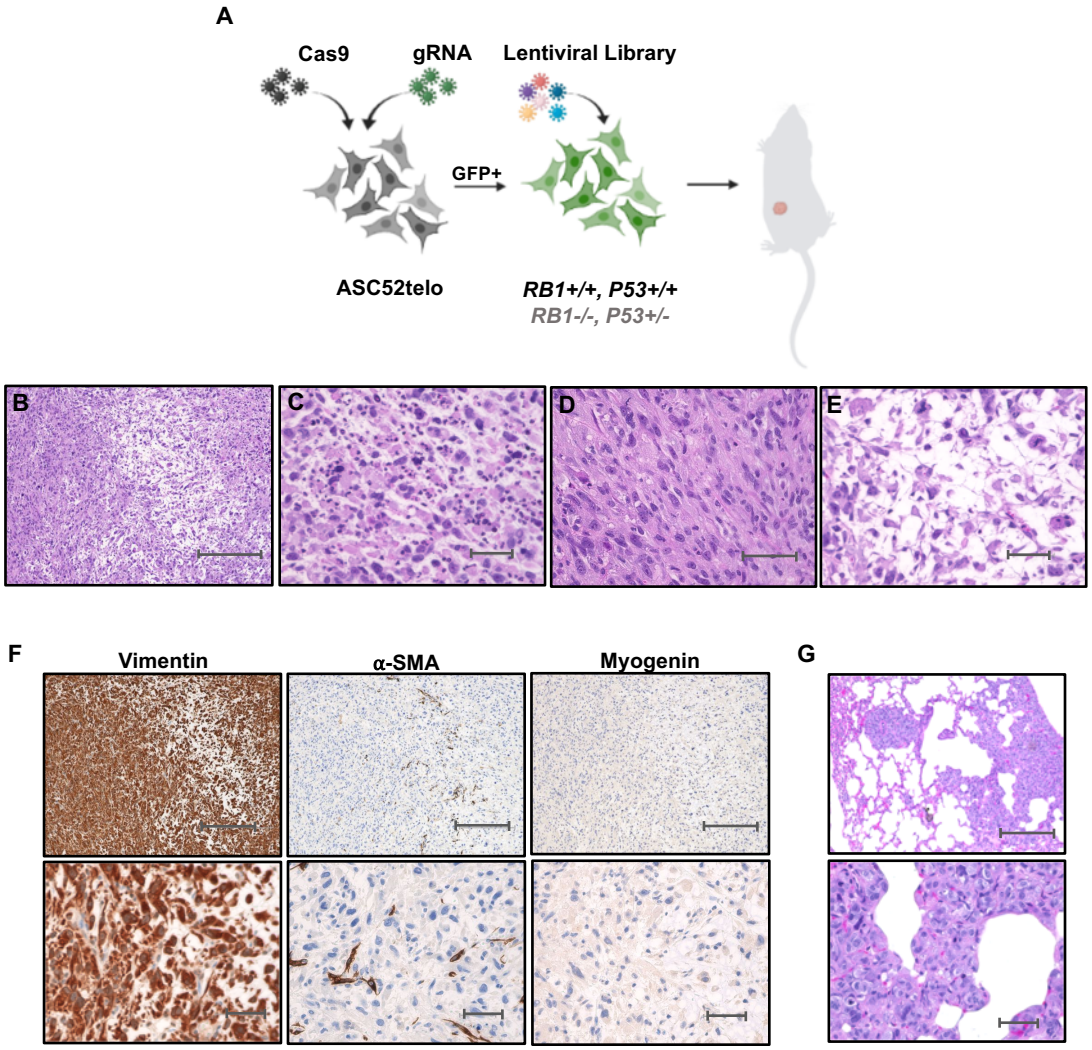

Figure 1. Establishment of a Pre-Clinical Sarcoma Model (A) Schematic diagram of forward transformation model of sarcoma development. (B) Low power view of H&E staining of tumors formed that demonstrated two distinct histologies. Scale bar represents 300um. (C) Distinct areas of necrosis were seen as evidenced by an immune infiltrate, scale bar represents 50um. High power view of the two distinct histologies demonstrate (D) Undifferentiated pleomorphic sarcoma and (E) Myxofibrosarcoma. (F) Immunohistochemistry of tumors formed show strong positivity for vimentin, primarily endothelial cell staining with infrequent tumor cells showing positivity, and negative staining for myogenin. Top row shows low power views with scale bars representing 300um and bottom row shows high power views with 50um scale bars. (G) Representative lung sections show the presence of metastatic sarcoma cells with low (top) and high (bottom) power views and scale bars representing 300um and 50um respectively.

Figure 2.

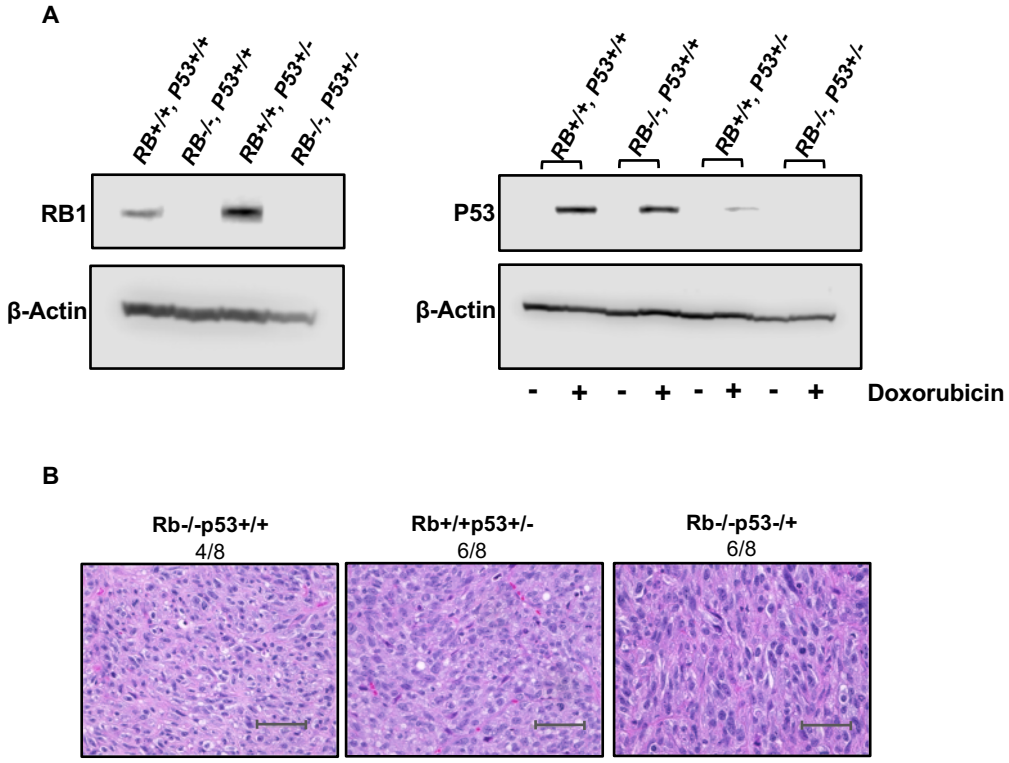

Figure 2. Loss of Either RB1 or P53 but not Both is Required for Transformation (A) Western blot showing expression of RB1, P53, or a loading control across each of the four cell lines. In the panel on the right, P53 expression is measured 48 hours after culture in the presence of absence of 20uM of doxorubicin. (B) Representative tumor sections from outgrowths with either loss of RB1 alone, P53 alone, or both. The number of total tumors formed is listed under each cell genotype. Scale bars represent 50um.

Figure 3.

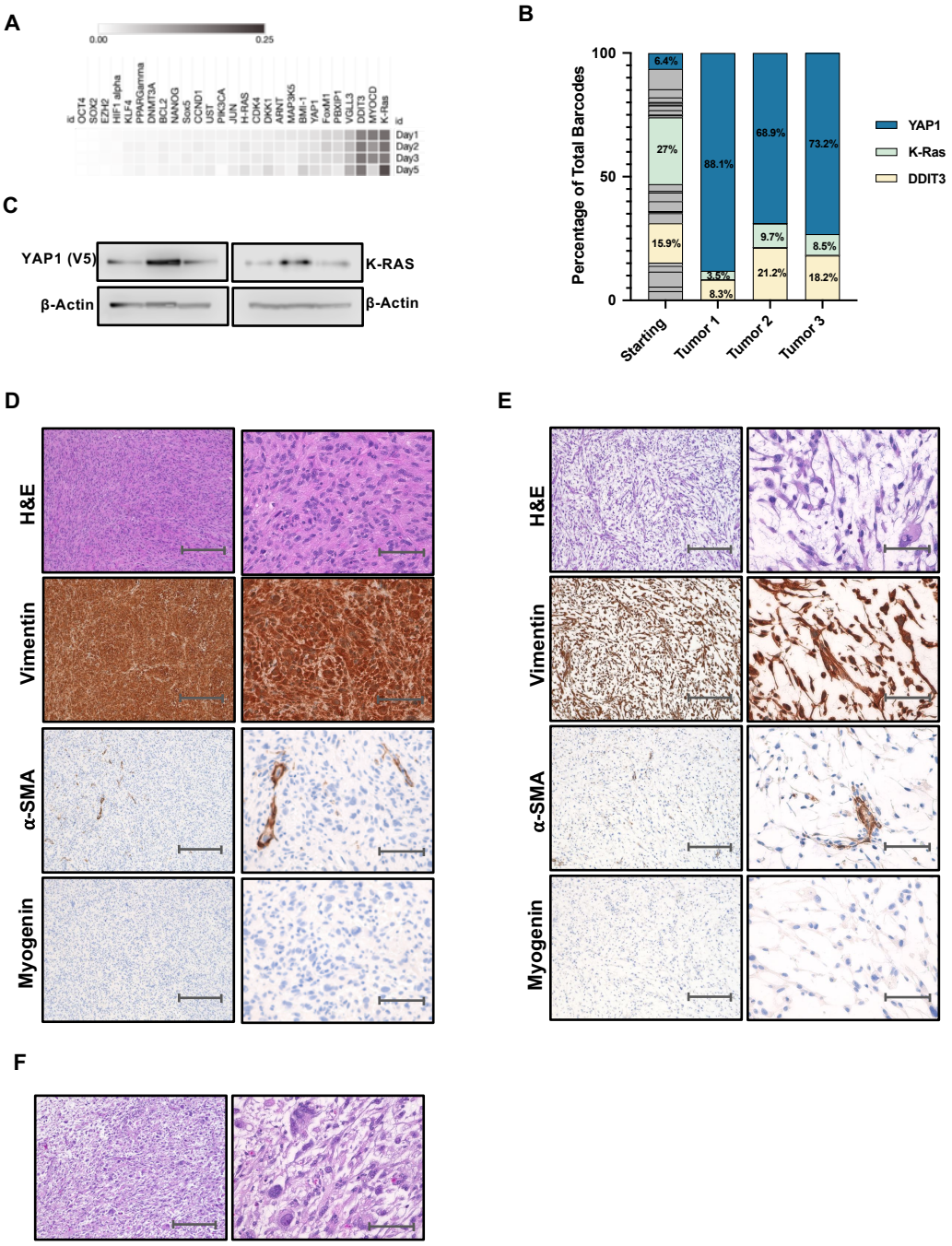

Figure 3. YAP1 and K-RAS Driven Tumors Histologically Reflect the Undifferentiated Pleomorphic Sarcoma-Myxofibrosarcoma Spectrum (A) Barcode sequencing of cell lines transduced with lentiviral library to show the relative distribution of barcodes over a five day period. Scale bar represents relative percentage of barcodes amplified. (B) Bar graph showing the percent of barcodes in the cell line prior to implantation (starting) and the tumor outgrowths. (C) Western blot showing YAP1 and KRAS expression in tumor outgrowths. (D) Histology from YAP1 driven tumors are shown with lower power on the left column, scale bars 300um and high power on the right, scale bars 50um. Tumors histologically represent undifferentiated pleomorphic sarcoma with strong staining for vimentin and negative for SMA and myogenin. (E) KRAS driven tumors are shown in this panel with histology consistent with myxofibrosarcoma. Low power (scale bars 300um) are shown on the left with high power (scale bars 50um) on the right. (F) Histology of combination KRAS and DDIT3 tumors that show increased pleomorphism in both low (left, scale bars 300um) and high power (right, scale bars 50um) views.

Figure 4.

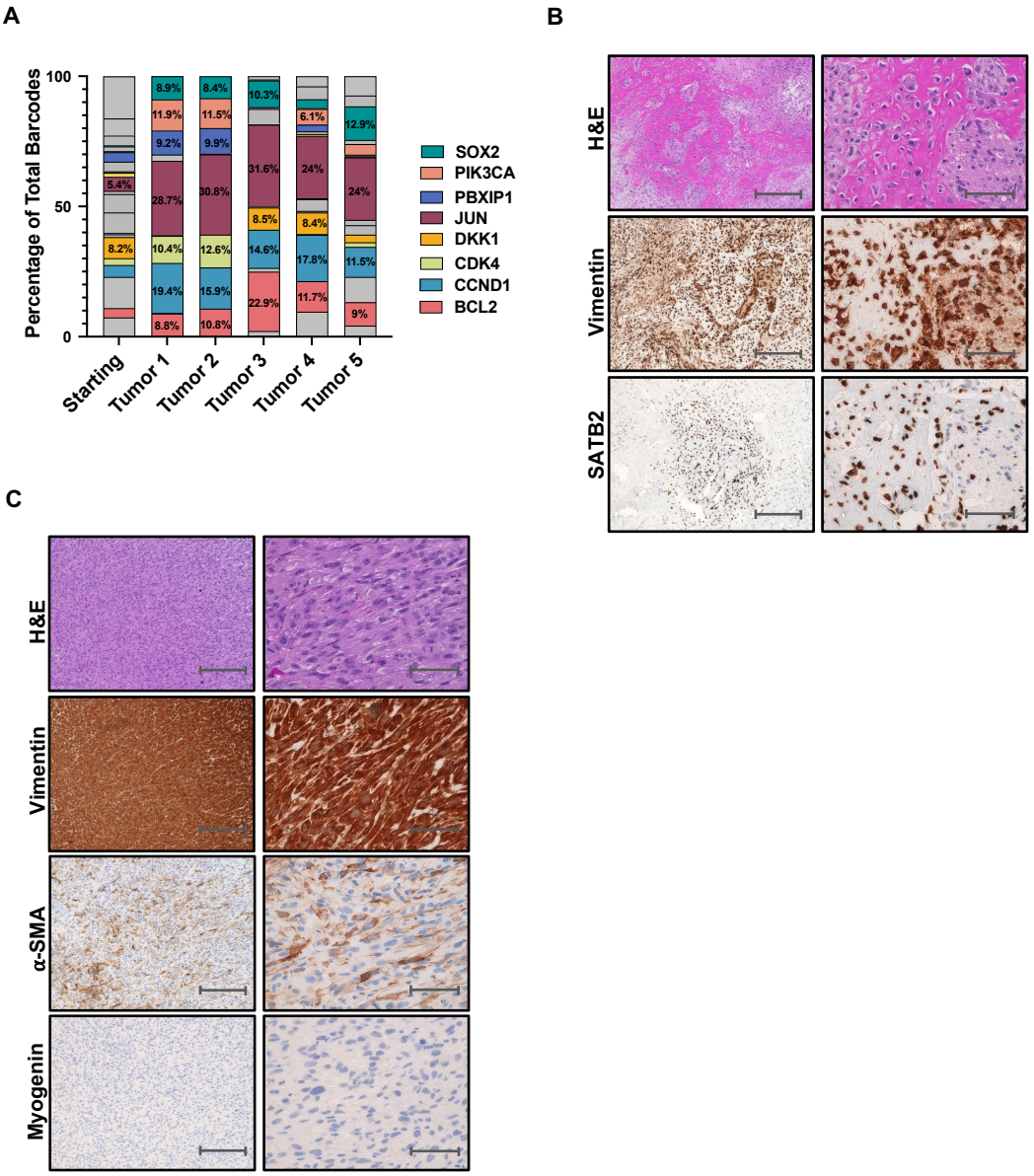

Figure 4. Identification of Additional Drivers of High-Grade Sarcoma Formation. (A) Barcode sequencing of cell lines transduced with lentiviral library without YAP1, KRAS and DDIT3. The graph shows the percent of barcodes in the cell line prior to implantation (starting) and the tumor outgrowths. (B) Histology from PI3K driven tumors are shown with lower power on the left column, scale bars 300um and high power on the right, scale bars 50um. Tumors histologically represent osteosarcoma with strong nuclear staining for SATB2. (C) Histology from CDK4 driven tumors are shown with lower power on the left column, scale bars 300um and high power on the right, scale bars 50um. Tumors histologically represent leiomyosarcoma with strong staining for SMA.