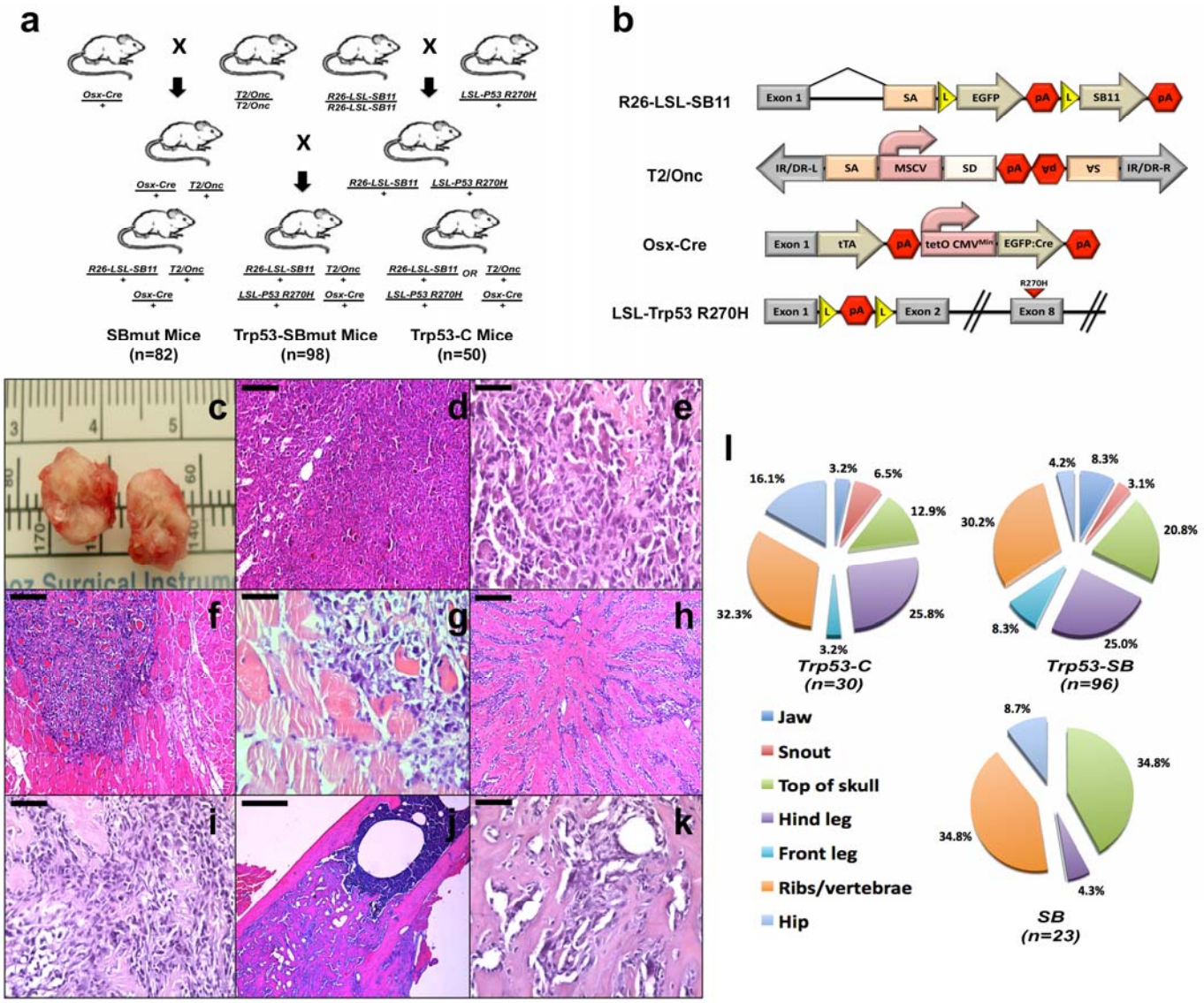
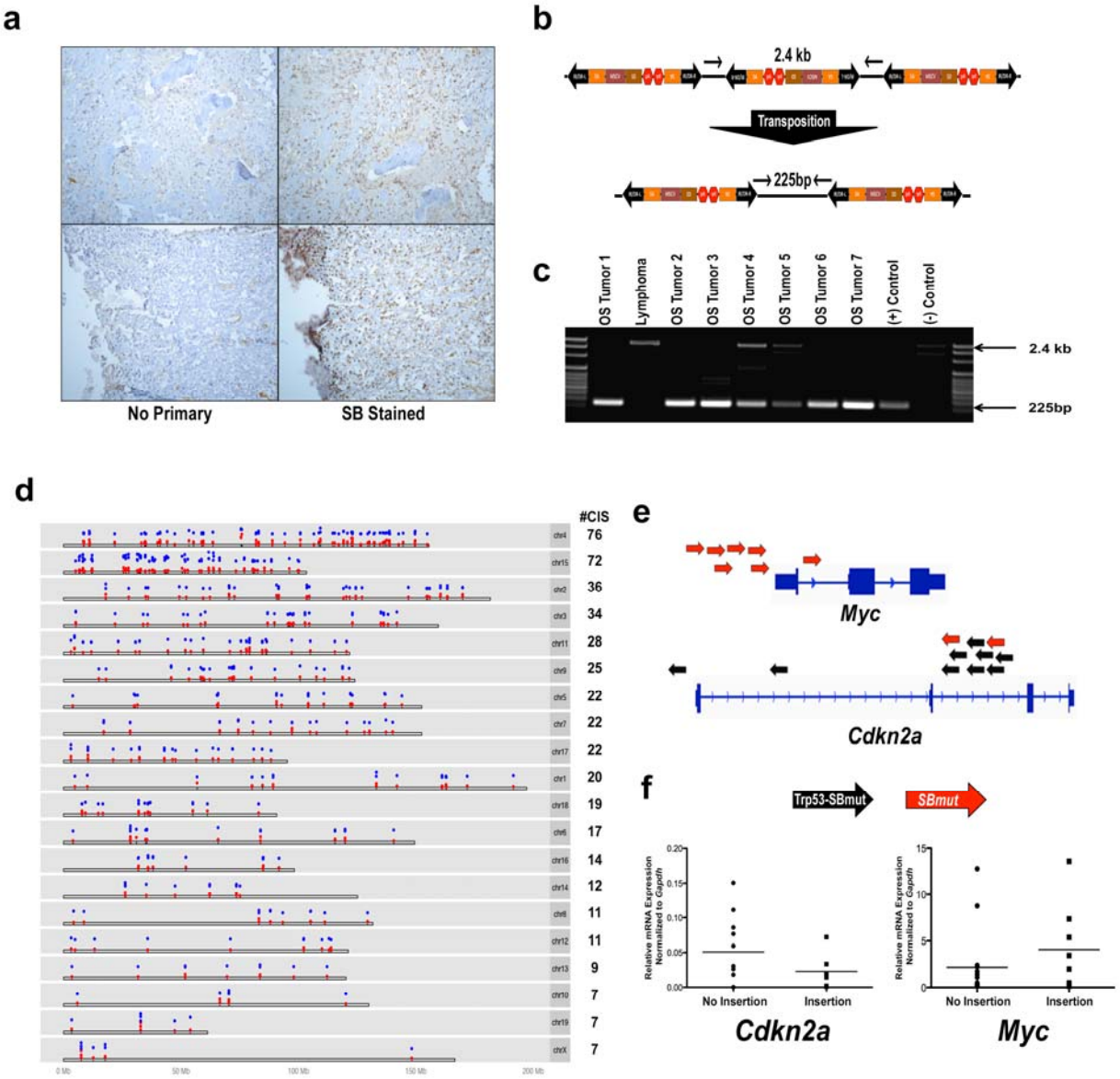


Supplementary Figure 1

Breeding scheme, transgenes, histological analysis and site distribution of SB-mutagenized osteosarcoma.

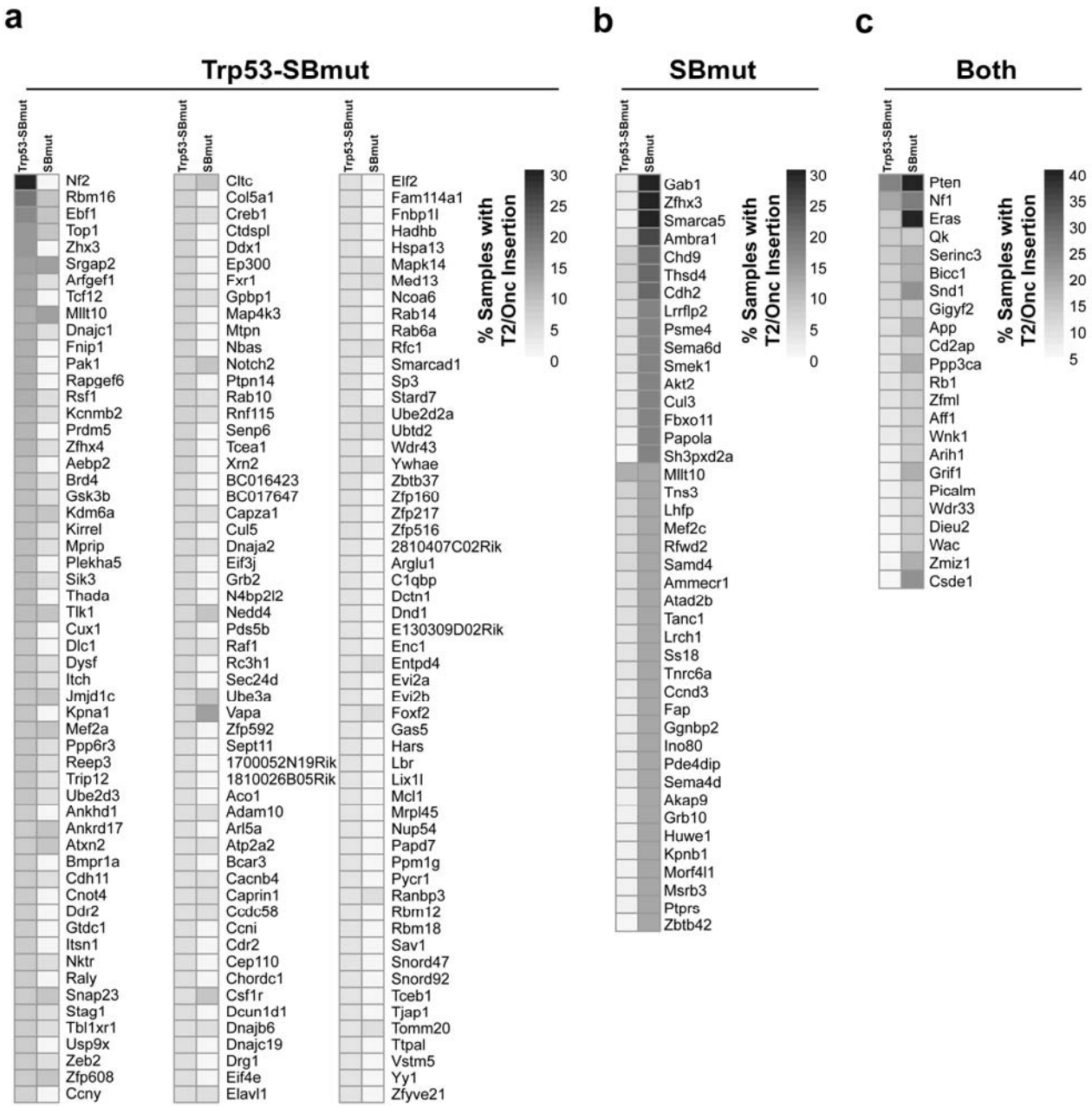
(**a**) Breeding scheme. *R26-LSL-SB11* homozygous mice were bred to *Trp53*$^{LSL\text{-}R270H/+}$ mice to generate doubly transgenic mice. Concurrently, *Osx-cre* mice were bred to T2/Onc mice. These doubly transgenic mice were then intercrossed to obtain experimental and control animals. (**b**) Transgene architecture of the *Rosa26-LSL-SB11*, T2/Onc, *Osx-cre* and *Trp53*$^{LSL\text{-}R270H/+}$ alleles used in this study. T2/Onc is engineered with splice acceptors (SA) and polyadenylation (pA) signals in both orientations for gene inactivation and the strong murine stem cell virus (MSCV) 5′ LTR promoter followed by a splice donor (SD) sequence for overexpressing genes. (**c**–**k**) Hematoxylin and eosin (H&E)-stained sections showing the representative tumor morphology of SB-mutagenized osteosarcoma with areas of high cellularity (**d**,**e**), invasive growth into surrounding tissue (**f**,**g**) and large areas of osteoid deposits (**h**–**k**). Scale bars, 200 μm for **c**,**e**,**g**,**i** and 50 μm for **d**,**f**,**h**,**k**. (**n**) Graphs displaying the percentage of tumors that developed at each site in Trp53-C (*n* = 30), Trp53-SBmut (*n* = 96) and SBmut (*n* = 23) animals.



Supplementary Figure 2

Validation of SB mutagenesis, local hopping and *Myc*/*Cdkn2a*.

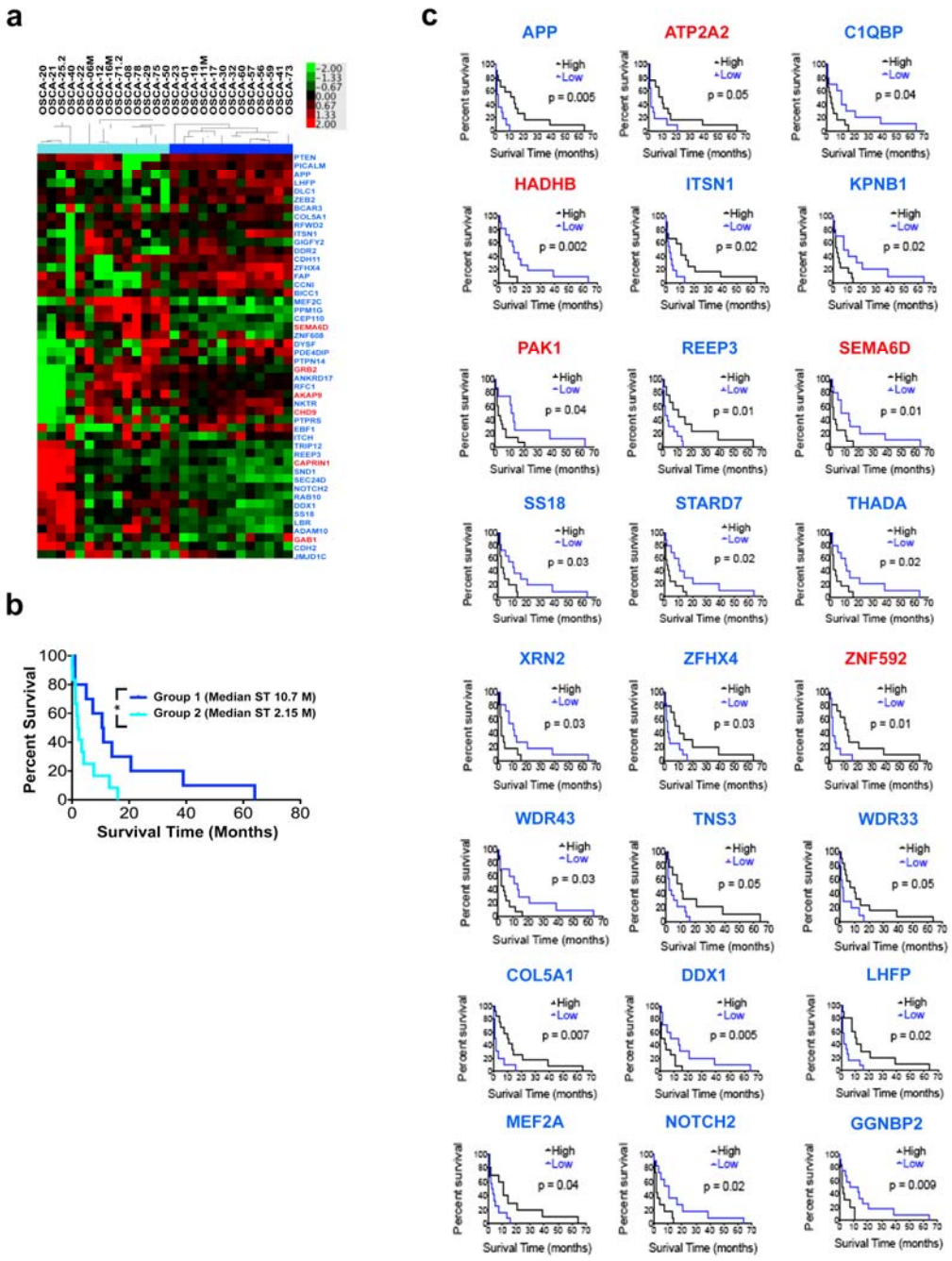
(**a**) Representative photomicrographs of positive IHC staining for SB protein and an appropriate control with no primary antibody, performed on SB-mutagenized osteosarcomas. (**b**,**c**) PCR-based transposon excision assay demonstrating that transposition is occurring in all SB-mutagenized osteosarcoma tumors (225-bp amplicon) and absent in background tumors (2.4-kb amplicon). Appropriate positive and negative excision controls are shown. (**d**) Diagram depicting the CISs identified on all chromosomes from animals with T2/Onc concatemers on chromosomes 4 (red symbols) and 15 (blue symbols). The height of each symbol indicates the frequency of the CIS, with symbols stacked for chromosomes 4 (0–40) and 15 (40–80). (**e**) Diagram depicting T2/Onc insertion sites driving overexpression of *Myc* and causing loss of function of *Cdkn2a*. Black and red arrows represent T2/Onc insertions identified in tumors from Trp53-SBmut and SBmut animals, respectively. The direction of the arrows is representative of the direction of the mouse stem cell virus (MSCV) LTR and splice donor (SD) sequence of T2/Onc. (**f**) Relative levels of mRNA for *Myc* and *Cdkn2a* isolated from tumors with or without T2/Onc insertions in the respective genes analyzed by quantitative PCR. Error bars, s.d.



Supplementary Figure 3

The CIS-associated genes identified from the Trp53-SBmut and SBmut tumor cohorts are mutated below significance in both cohorts.

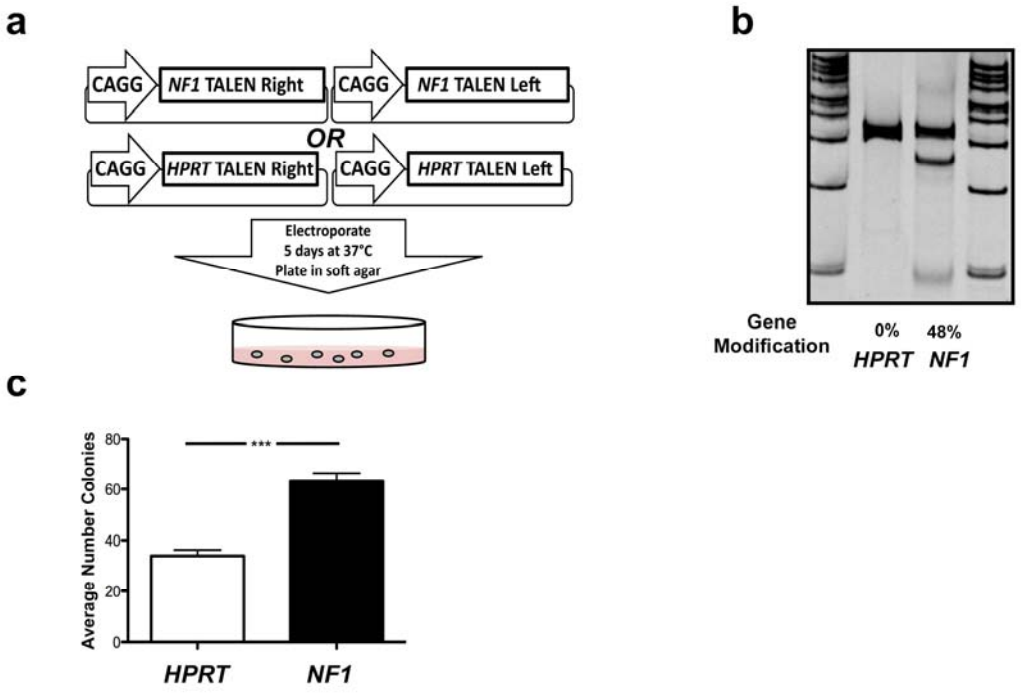
(**a**–**c**) Heat maps depicting the percentage of tumors in the Trp53-SBmut and SBmut cohorts that harbor T2/Onc insertions for the CIS-associated genes identified in only Trp53-SBmut (**a**) or SBmut (**b**) tumors or in both Trp53-SBmut and SBmut tumors (**c**).



Supplementary Figure 4

A subset of CIS-associated genes are predictive of outcome in canine osteosarcoma.

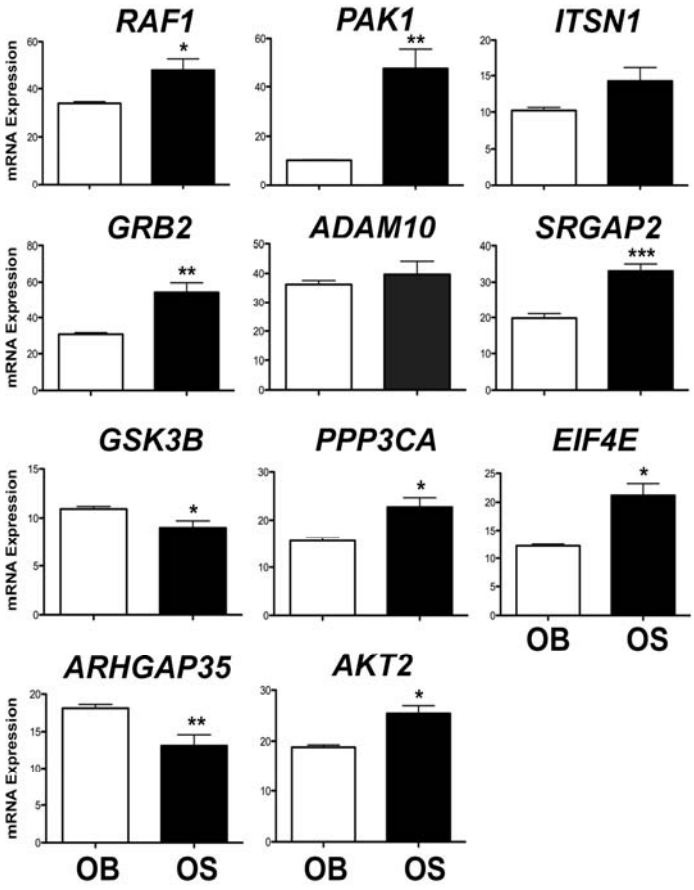
(**a**) Unsupervised hierarchical clustering of all CIS genes with appreciable expression variation (*n* = 48 genes) resulting in the clustering of afflicted dogs into two groups with significantly different survival times. Gene color denotes SB prediction as an oncogene (red) or TSG (blue). (**b**) Kaplan-Meier survival curve of canine osteosarcoma using the 48 CIS-associated genes with appreciable expression variation. Kaplan-Meier survival curve of dogs for CIS-associated genes whose expression significantly correlates with outcome. (**c**) CIS-associated genes were evaluated against survival in a set of 27 canine osteosarcoma cell lines established from patients with known outcome (overall survival time). Gene color denotes SB prediction as an oncogene (red) or TSG (blue). Impact on survival was analyzed by median-centered (high, low) expression. Log-rank Mantel-Cox *P* values of <0.05 were considered significant.



Supplementary Figure 5

***NF1* loss transforms iOB cells**.

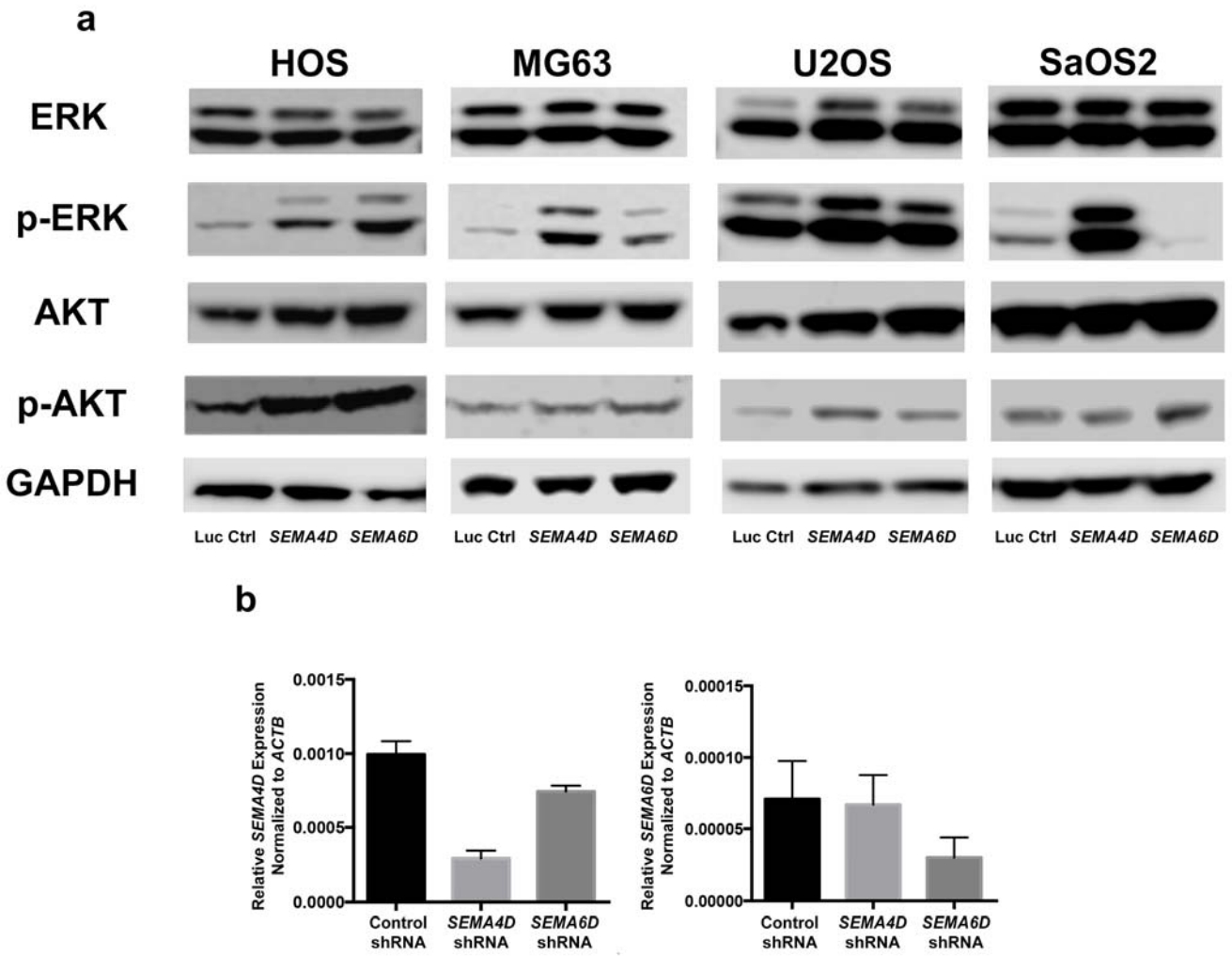
(**a**) Diagram of the experimental procedure used to knock out *NF1* with transcription activator-like effector nucleases (TALENs) in immortalized osteoblast (iOB) cells. (**b**) Results of a surveyor assay performed on DNA extracted from iOB cells 5 d after transfection with TALENs targeting *NF1* or *HPRT*. (**c**) Average number of colonies formed in soft agar by iOB cells treated with *NF1* or *HPRT* TALENs. Data are the means ± SE of five independent experiments; ***$P < 0.0001$, Student's *t* test. Error bars, s.d.



Supplementary Figure 6

Axon guidance genes are misexpressed in human osteosarcoma.

Relative mRNA levels of CIS genes involved in axon guidance in normal human osteoblast (OB) samples and human osteosarcoma tumors analyzed by RNA sequencing (n = 12 human osteosarcoma and 3 normal osteoblast samples). *P < 0.05, **P < 0.001, ***P < 0.0001, Student's t test. Error bars, s.d.



Supplementary Figure 7

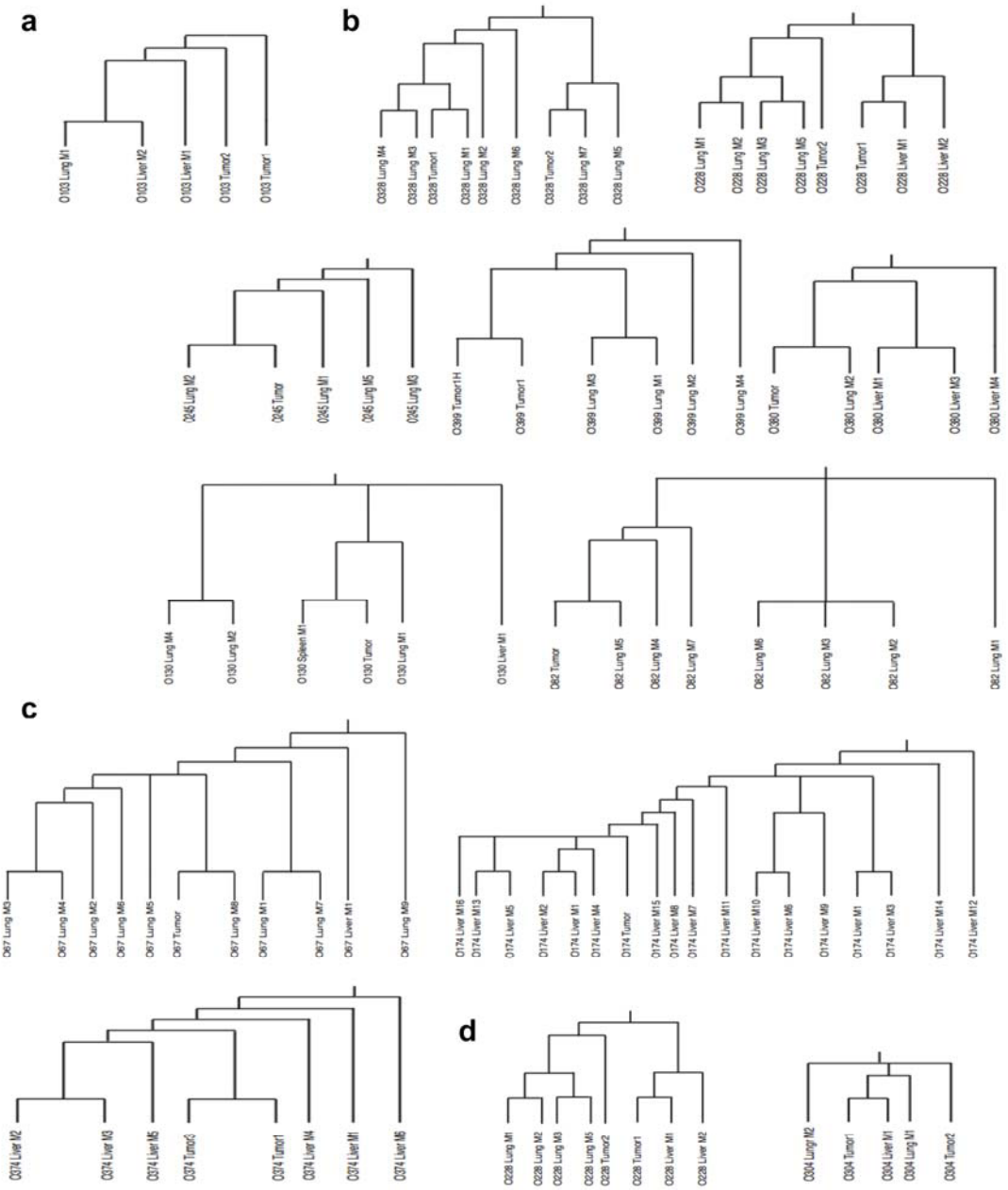
Immunoblot analysis of proteins involved in the *SEMA4D* and *SEMA6D* signaling pathways in human osteosarcoma cell lines and shRNA knockdown validation.

(**a**) Immunoblot analysis for the indicated proteins in lysates from HOS, MG63, U2OS and SaOS2 cells overexpressing luciferase control, *SEMA4D* or *SEMA6D* cDNA. (**b**) RT-PCR analysis of *SEMA4D* and *SEMA6D* transcripts in shRNA-expressing HOS cells normalized to *Actb* levels. Error bars, s.d.

Supplementary Figure 8

Hierarchical clustering analysis of T2/Onc insertion in animals that developed metastatic osteosarcoma.

(**a**) Group 1 animals in which primary tumors represent the most ancestral state in the phylogenic tree owing to the metastases sharing few insertions with the primary tumor. (**b**) Group 2 animals in which the primary tumor had the greatest separation from the ancestral state and therefore shared many insertions with the metastases. (**c**) Group 3 animals in which the primary tumor did not correspond to the most recent common ancestor and did not share as many insertion sites with the metastases as seen in group 2. (**d**) Animals containing multiple tumors that developed independent sets of metastases that were categorized into more than one of the three groups.



Supplementary Figure 9

Parsimony analysis of T2/Onc insertions in animals that developed metastatic osteosarcoma.

(**a**) Group 1 animals in which primary tumors corresponded to the most ancestral state in the phylogenic tree owing to the metastases sharing few insertions with the primary tumor. (**b**) Group 2 animals in which the primary tumor had the greatest separation from the ancestral state and therefore shared many insertions with the metastases. (**c**) Group 3 animals in which the primary tumor did not correspond to the most recent common ancestor and did not share as many insertion sites with the metastases as seen in group 2. (**d**) Animals containing multiple tumors that developed independent sets of metastases that were categorized into more than one of the three groups.