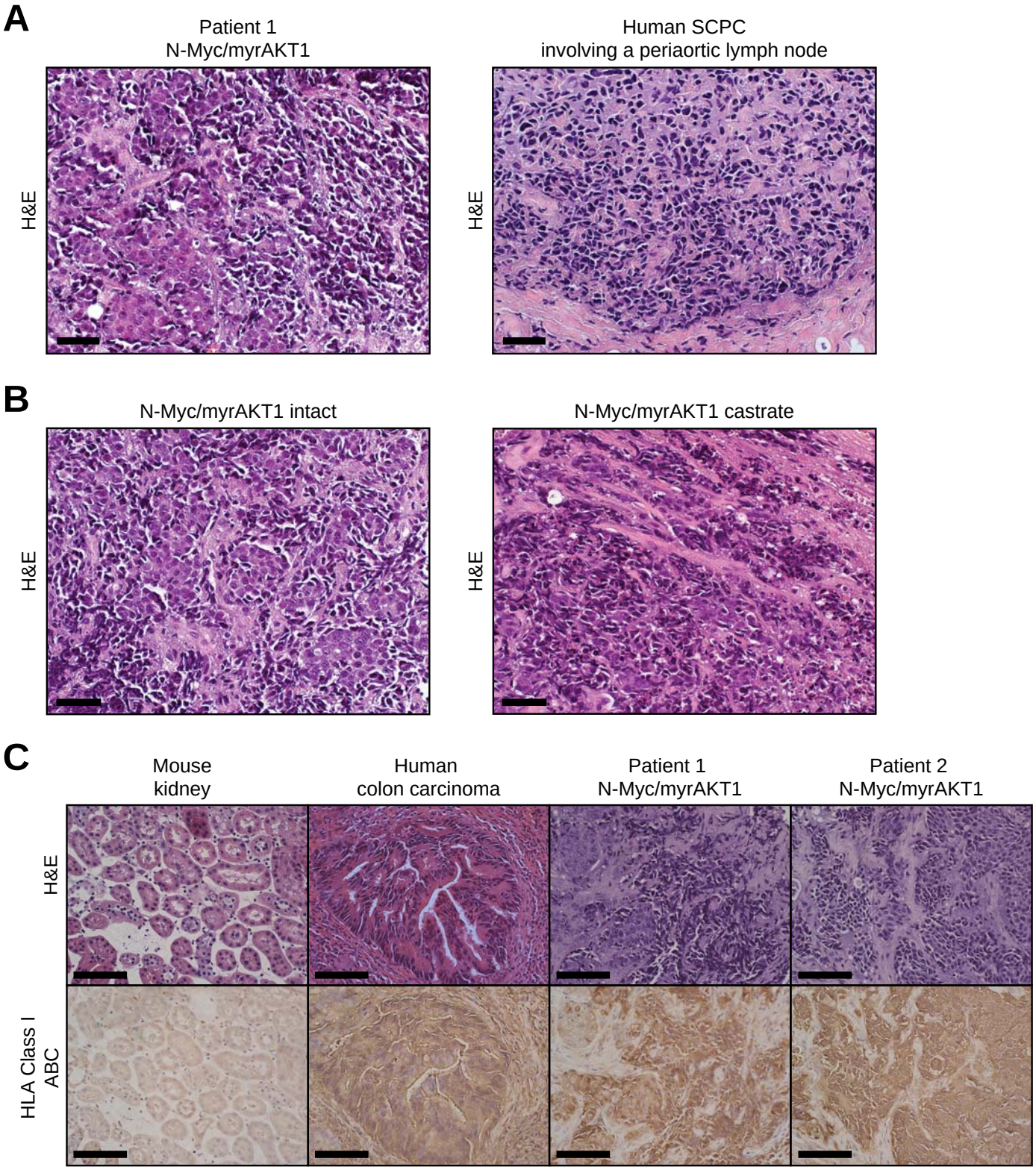
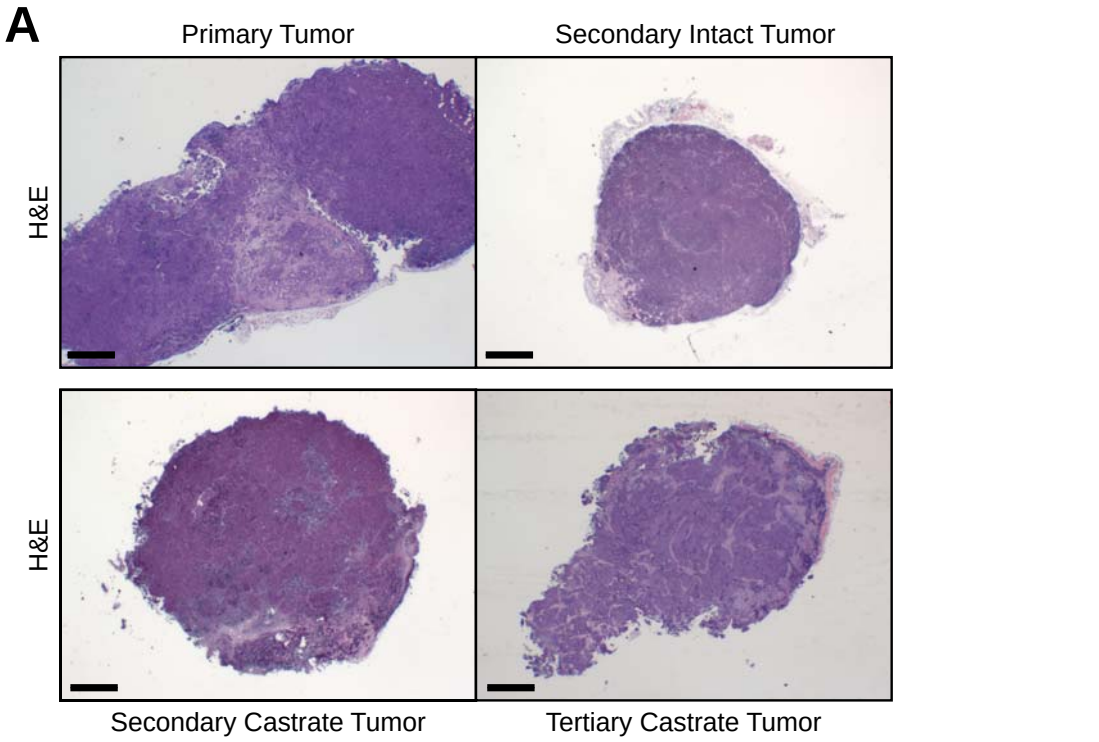
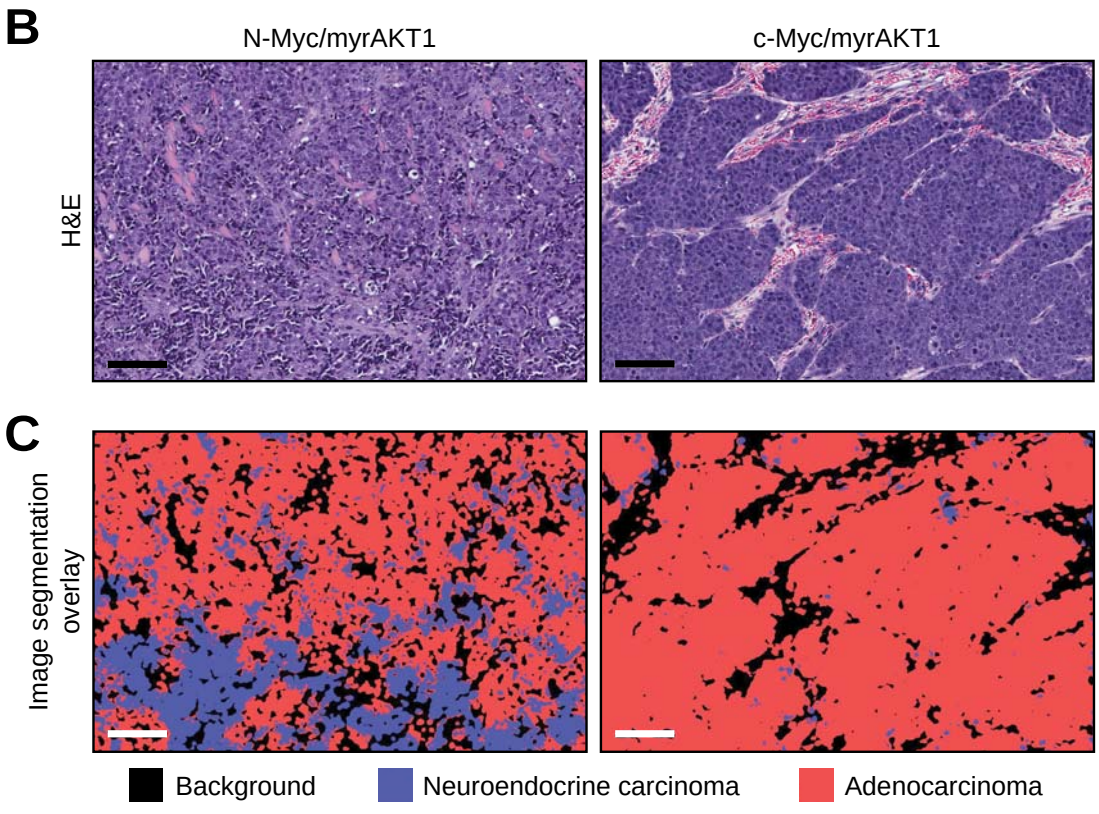



**Figure S1, related to Figure 1. N-Myc/myrAKT1 prostate tumors demonstrate histologic features of human small cell prostate carcinoma and are of a human cellular origin.** (A) H&E-stained sections obtained from an N-Myc/myrAKT1 tumor and an extensive stage human small cell prostate carcinoma specimen from a periaortic lymph node (scale bar=100 μm). (B) H&E-stained sections of N-Myc/myrAKT1 transplanted tumors from intact and castrate conditions (scale bar=100 μm). (C) H&E and HLA Class I ABC stains of mouse kidney, human colon carcinoma, and N-Myc/myrAKT1 tumors generated from two different patient samples (scale bar=100 μm).

**Table S1, related to Figure 1. Characteristics of human prostate specimens.**

| Patient | Age Range | Race | Diagnosis | Gleason Score (Tumor Size) | Pre-operative PSA |
|---|---|---|---|---|---|
| 1 | 65-70 | White | Prostatic adenocarcinoma | 3+3=6 (N/A) | 7.4 |
| 2 | 75-80 | White | Prostatic adenocarcinoma | Focus 1: 4+5=9 (1.2 cm)<br>Focus 2: 4+3=7 (1.1 cm)<br>Focus 3: 3+3=6 (0.1 cm)<br>Focus 4: 3+3=6 (0.1 cm) | <0.01 |
| 3 | 60-65 | White | Prostatic adenocarcinoma | Focus 1: 3+4=7 with tertiary 5 (3.2 cm)<br>Focus 2: 3+3=6 (0.25 cm) | 9.4 |
| 4 | 60-65 | White | Benign prostate gland | N/A (N/A) | N/A |
| 5 | 65-70 | White | Prostatic adenocarcinoma | 3+4=7 (1.5 cm) | N/A |
| 6 | 60-65 | White | Prostatic adenocarcinoma | 3+4=7 (N/A) | 3.4 |


A
Primary Tumor
Secondary Intact Tumor
H&E
H&E
Secondary Castrate Tumor
Tertiary Castrate Tumor


B
N-Myc/myrAKT1
c-Myc/myrAKT1
H&E
C
Image segmentation overlay
Background
Neuroendocrine carcinoma
Adenocarcinoma

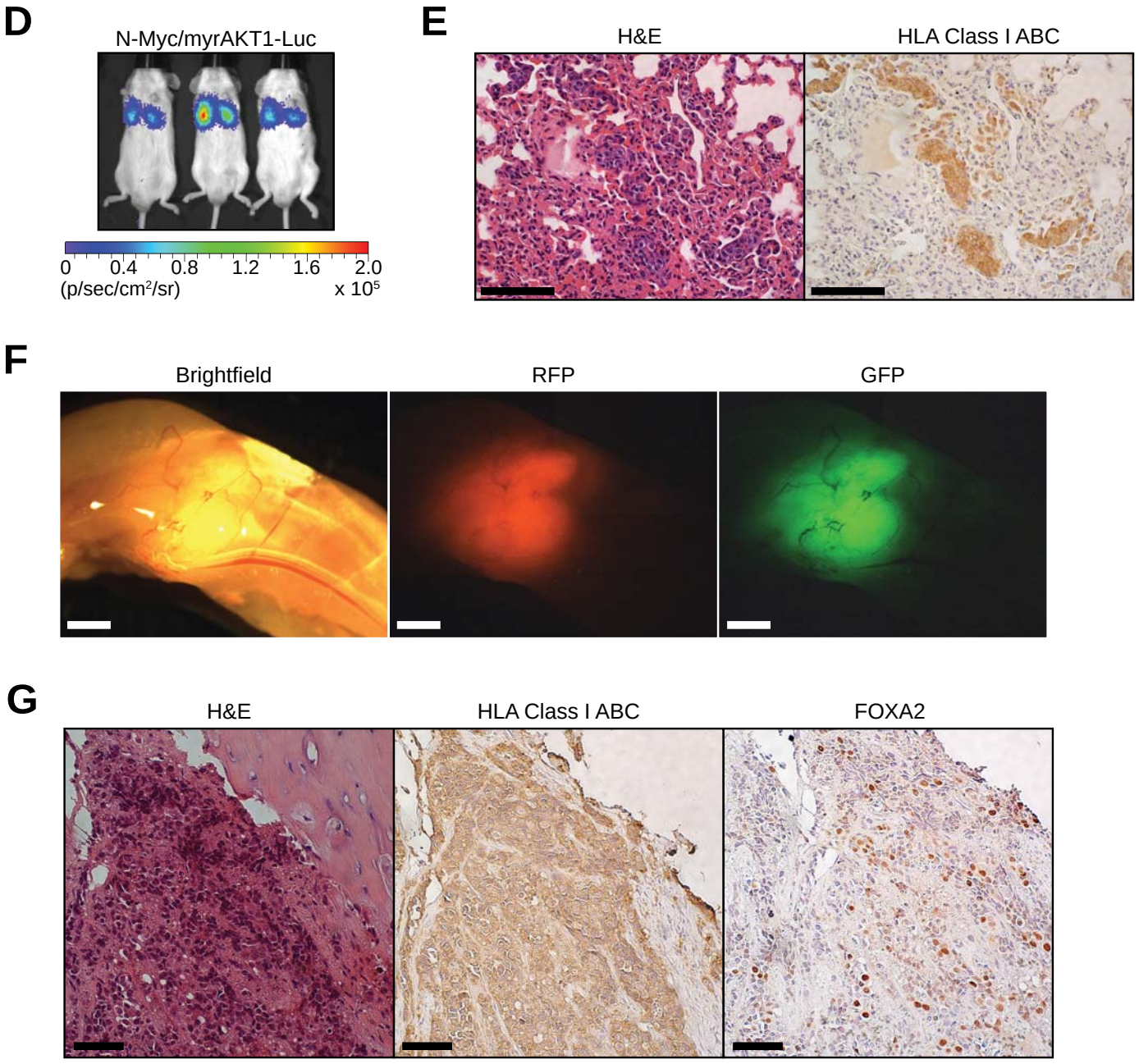



**Figure S2, related to Figure 3. N-Myc/myrAKT1 prostate tumors are enriched for NEPC after castration and demonstrate widespread metastases marked by HLA Class I ABC and neuroendocrine marker FOXA2 expression.** (A) H&E-stained sections of primary, secondary, and tertiary N-Myc/myrAKT1 prostate tumors at low magnification (scale bar=800 μm). (B) Representative H&E-stained sections of N-Myc/myrAKT1 and c-Myc/myrAKT1 tumors (scale bar=100 μm). (C) Image segmentation overlay of the photomicrographs in (B) with features characterized as background (black), neuroendocrine carcinoma (blue), and adenocarcinoma (red) (scale bar=100 μm). (D) Bioluminescent imaging of mice immediately post-tail vein injection with the N-Myc/myrAKT1-Luc cell line (signal intensity is represented by radiance, p/sec/cm2/sr). (E) H&E and HLA Class I ABC immunostaining of lung sections from mice 21 days after tail vein injection with the N-Myc/myrAKT1-Luc cell line (scale bar=100 μm). (F) Brightfield, red fluorescent, and green fluorescent images of a hind limb from an N-Myc/myrAKT1-Luc mouse bearing a tumor deposit (scale bar=2 mm). (G) H&E, HLA Class I ABC, and FOXA2 immunostaining of tissue sections from metastatic N-Myc/myrAKT1-Luc tumors involving the femur (scale bar=100 μm).

**Table S2, related to Figure 4. Weighted 50-gene NEPC signature.**

| Gene | Weight |
|---|---|
| CPLX2 | 0.054237 |
| ACTL6B | 0.0397921 |
| CA9 | 0.0390316 |
| MYT1 | 0.0355834 |
| UNC13A | 0.0326315 |
| TNNT1 | 0.0318484 |
| DPYSL5 | 0.0293946 |
| LRTM2 | 0.0273999 |
| SYT4 | 0.0254472 |
| XKR7 | 0.0234302 |
| CHRNB2 | 0.0211502 |
| SRRM4 | 0.0206361 |
| ATP1A3 | 0.0193364 |
| CDK5R2 | 0.0151874 |
| FAM123C | 0.0147568 |
| MARCH4 | 0.0104644 |
| RTBDN | 0.0091143 |
| RUNDC3A | 0.0084861 |
| GRM4 | 0.008138 |
| VGF | 0.0080349 |
| NKX2-2 | 0.0065812 |
| SEZ6 | 0.0065713 |
| ASCL1 | 0.0060928 |
| SCRT1 | 0.005412 |
| HMP19 | 0.003171 |
| LHX2 | 0.0025981 |
| MAST1 | 0.0025208 |
| SYT5 | 7.45E-05 |

| Gene | Weight |
|---|---|
| POTEH | -0.001174 |
| ANO7 | -0.001527 |
| AZGP1 | -0.00187 |
| OR51E2 | -0.002243 |
| TARP | -0.004231 |
| SRD5A2 | -0.004461 |
| ALOX15B | -0.004484 |
| POTEG | -0.005139 |
| ACPP | -0.005406 |
| GDEP | -0.006899 |
| PAGE4 | -0.00748 |
| P704P | -0.007655 |
| ACSM1 | -0.009399 |
| DES | -0.009427 |
| KLK2 | -0.014421 |
| C15orf21 | -0.015193 |
| MSMB | -0.016603 |
| PCGEM1 | -0.018376 |
| NPY | -0.024701 |
| PCA3 | -0.027402 |
| KLK3 | -0.034047 |
| MYBPC1 | -0.036665 |

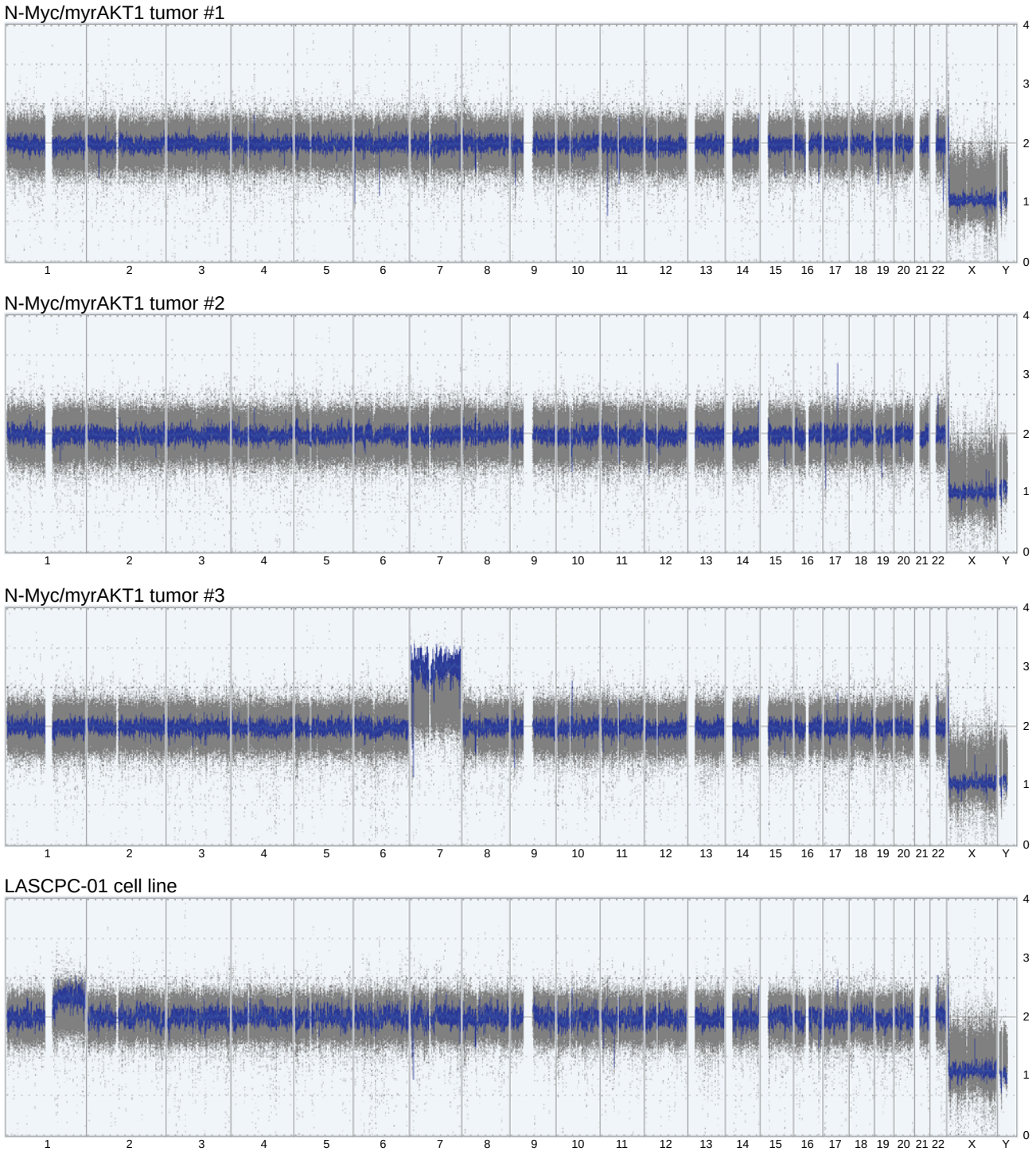



**Figure S3, related to Figure 4. N-Myc/myrAKT1 tumors and the LASCPC-01 cell line exhibit few chromosomal abnormalities.** Whole genome array CGH plots showing chromosomal location and copy number of three N-Myc/myrAKT1 tumors and the LASCPC-01 cell line analyzed using the high-resolution Affymetrix CytoScan HD platform (blue line represents the smoothed copy number profile).

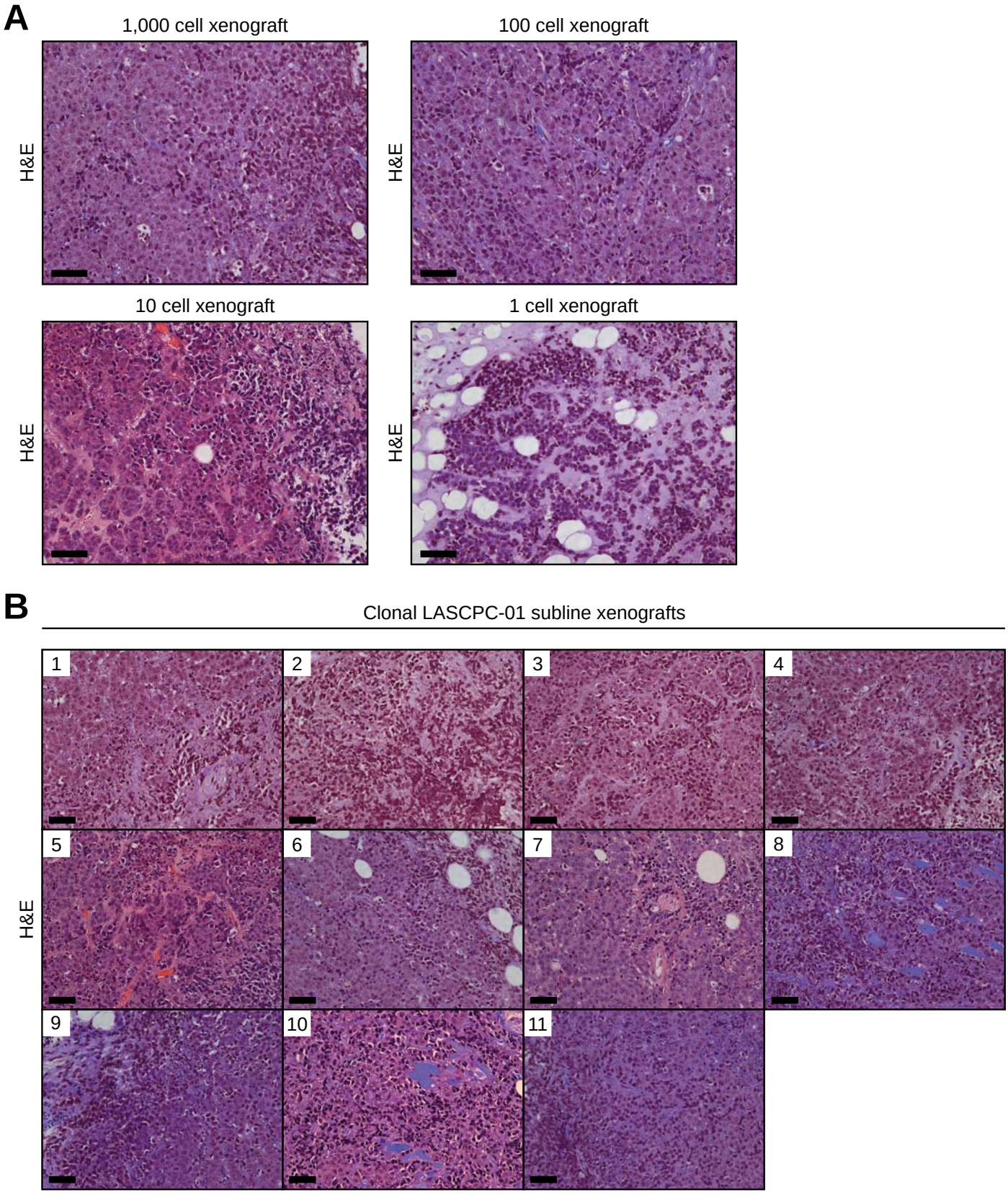
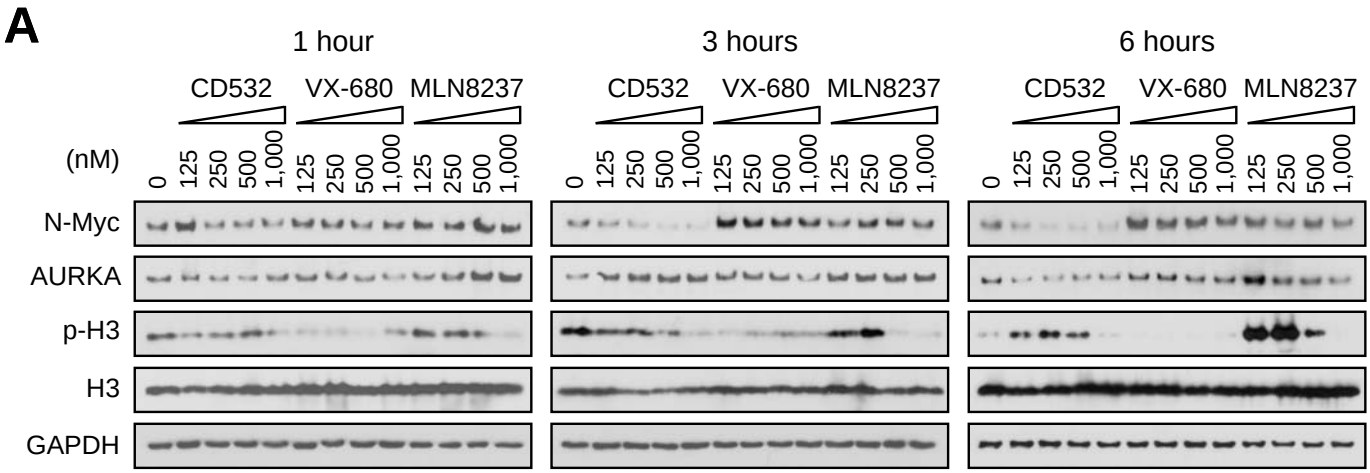
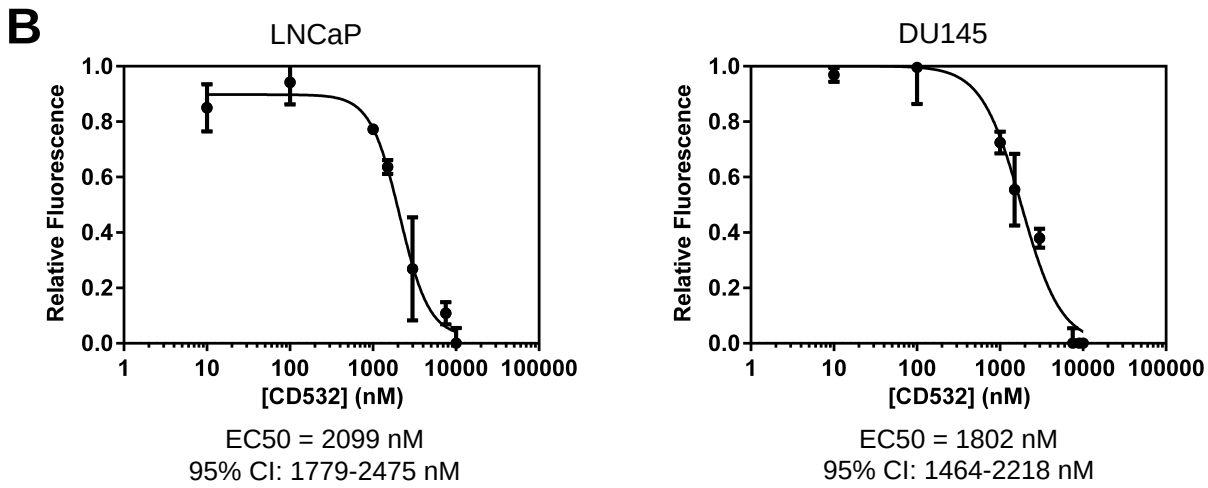


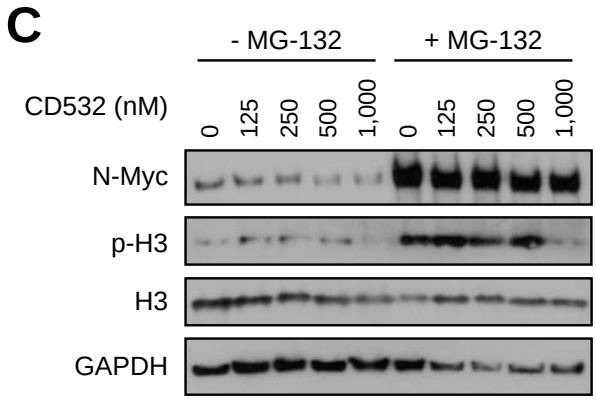

**Figure S4, related to Figure 5. N-Myc/myrAKT1 tumor cells are highly tumorigenic and demonstrate plasticity.** (A) H&E-stained sections of xenograft tumors derived from the implantation of 1,000, 100, 10, or 1 LASPC-01 tumor cell. (B) H&E-stained sections of xenograft tumors from each of eleven clonal LASCPC-01 sublines demonstrating mixed neuroendocrine carcinoma and adenocarcinoma.




EC50 = 2099 nM
95% CI: 1779-2475 nM

EC50 = 1802 nM
95% CI: 1464-2218 nM



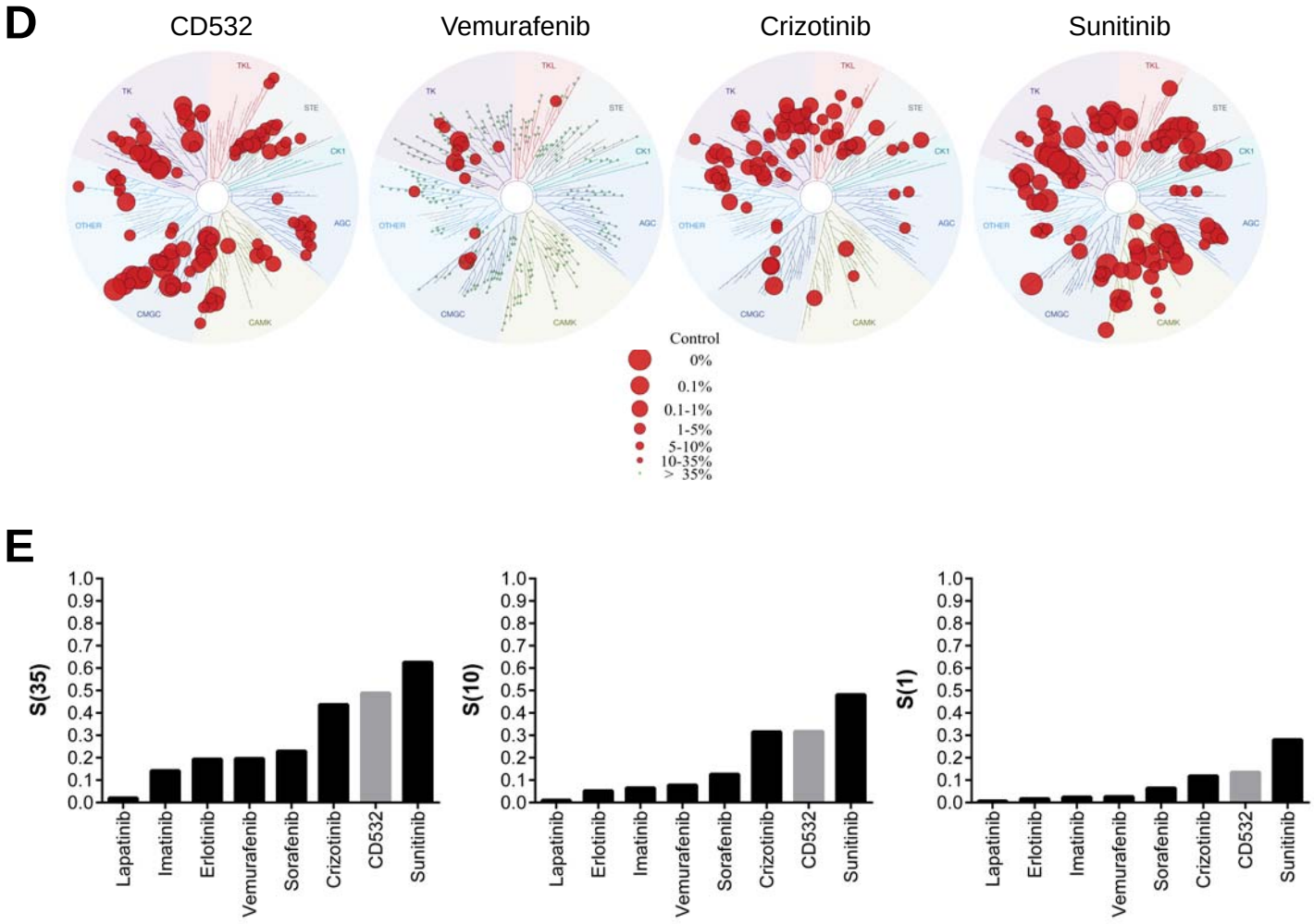



**Figure S5, related to Figure 7. On-target effects of CD532 on N-Myc destabilization and Aurora A kinase inhibition and kinase selectivity of CD532.** (A) Immunoblot analysis of LASCPC-01 cells treated with the indicated inhibitors and doses for the indicated periods of time with antibodies against N-Myc, AURKA, phosphorylated histone H3 (p-H3), histone H3, and GAPDH as a loading control. (B) Dose response of CD532 +/- SD (normalized to DMSO treatment only) at 48 hours using the CellTiter-Glo cell viability assay in LNCaP cells and DU145 cells. (C) Immunoblot analysis of LASCPC-01 cells treated with increasing concentrations of CD532 for 3 hours in the presence or absence of 10 μM MG-132 pre-treated for 3 hours. (D) Small molecule kinase interaction map for CD532 after screening a panel of >400 human kinases at a dose of 10 μM using the DiscoveRx KINOME*scan* platform. Also shown are maps for vemurafenib, crizotinib, and sunitinib generated from Karaman *et al.* (Nature Biotechnology, 2008) and publically available data from the Harvard Medical School LINCS Center. Red circles represent small molecule-kinase interactions and the size of the circles indicates relative binding, represented as percent of control [(test compound signal - positive control signal)/(negative control signal - positive control signal) x 100]. (E) Selectivity scores of kinase inhibitors for binding interactions with $S_{35}$ or $S_{10}$ or $S_1$ [(number of non-mutated kinases with percent of control <35 or <10 or <1)/ (number of non-mutated kinases tested)].

## SUPPLEMENTAL EXPERIMENTAL PROCEDURES

### Lentiviral Constructs

The myrAKT1 lentiviral vector has been described previously (Xin et al., 2005). The N-Myc lentiviral vector was cloned by PCR amplification of pMXs-hu-N-Myc (Plasmid 50772, Addgene) with the forward primer 5'-AGTTCTAGAACCATGCCGAGCTGCTCCACG-3' and reverse primer 5'-AGTGA ATTCTTAGCAAGTCCGAGCGTGTTC-3'. The PCR product was cloned into pCR2.1-TOPO using the TOPO TA Cloning Kit (Life Technologies). The *MYCN* insert was then released by digestion with XbaI and EcoRI and cloned into the XbaI and EcoRI sites of the FU-CRW lentiviral backbone (Memarzadeh et al., 2007). The final plasmid was named FU-MYCN-CRW. The FU-IYLW lentiviral vector was used for ectopic expression of firefly luciferase. The inducible N-Myc lentiviral vector was cloned by inserting *MYCN* into the BamHI site of the PSTV lentiviral backbone. The final plasmid was named PSTV-MYCN-CGW. Lentiviruses were prepared and titered as described (Xin et al., 2005).

### Cell Lines

DU145, LNCaP, and PC-3 (ATCC) were grown in RPMI with 10% FBS. LAPC4 (gift from Robert Reiter, UCLA) was grown in Iscove's with 20% FBS. NCI-H660 (ATCC) was grown in HITES media containing RPMI, 5% FBS, 10 nM hydrocortisone, 10 nM beta-estradiol (Sigma), insulin-transferrin-selenium, and Glutamax (Life Technologies). LASCPC-01 was established by dissociating an N-Myc/myrAKT1 tumor as previously described (Stoyanova et al., 2013) and plating cells in HITES media. Clonal sublines of LASCPC-01 were established by singly sorting the LASCPC-01 cells on a BD FACS ARIA II (BD Biosciences) into individual wells of a 96-

well plate each containing 100 µl of HITES media. Single cell deposition was verified by direct microscopic visualization and cultures were monitored daily for colony formation.

**Histology, Immunohistochemistry, and Immunoblotting**

Tumor samples were formalin-fixed, paraffin-embedded, sectioned to 4 µm thickness, and mounted on glass slides. For each tumor, sections were stained with a standard H&E protocol. For immunohistochemistry, unstained sections were deparaffinized, hydrated, and subjected to heat-induced antigen retrieval using 40 mM sodium citrate buffer pH 6. Primary and secondary antibodies for immunohistochemistry are listed below. For immunoblot analysis, tumor tissues and prostate cancer cell lines were homogenized and lysed in either urea lysis buffer (20 mM HEPES pH 8.0, 9 M urea, 2.5 mM sodium pyrophosphate, 1.0 M beta-glycerophosphate, 1% N-octyl glycoside, 2 mM sodium orthovanadate) or TNN lysis buffer (50 nM Tris pH 8, 120 nM NaC1, 0.5% NP-40, 1 mM dithiothreitol, and protease inhibitors). Primary and secondary antibodies for immunoblotting are listed below. Films were scanned on an HP Scanjet G4050 for analysis. Densitometry was performed using Adobe Photoshop CS5 and corrected for background intensity.

**Antibodies for Immunohistochemistry and Immunoblotting**

Primary antibodies used for immunohistochemistry include CK8 (1:1,000, Covance MMS-162P), p63 (1:250, Santa Cruz sc-8431), AR (1:250, Santa Cruz sc-816), CHGA (1:600, Dako M0869), NCAM1 (1:100, Novocastra NCL-CD56-504), SYP (1:200, Novocastra NCL-SYNAP-299), TTF-1 (1:5,000, Upstate #07-601), FOXA2 (1:200, Abcam ab108422), and HLA Class I ABC (1:200, Abcam ab70328). Secondary antibodies used were ImmPRESS Anti-Rabbit Ig

Peroxidase and Anti-Mouse Ig Peroxidase (Vector Labs). Liquid DAB+ substrate reagent (Dako) was used to perform direct chromogenic visualization. Primary antibodies used for immunoblotting include N-Myc (1:1,000, Santa Cruz B8.4.B), AURKA (1:1,000, Abcam ab1287), NSE (1:1,000, Abcam ab139749), ASCL1 (1:1,000, BD Pharmingen 24B72D11.1), AKT (1:1,000, Cell Signaling #4691), phospho-AKT Ser473 (1:1,000, Cell Signaling #9271), phospho-histone H3 Ser10 (1:1,000, Cell Signaling #3642), histone H3 (1:5,000, Cell Signaling #4499), cleaved PARP (1:1,000, Cell Signaling #5625), cleaved caspase-3 (1:1,000, Cell Signaling #9661), GAPDH (1:5,000, GeneTex GT239) and p84 (1:2,000, GeneTex 5E10). Secondary antibodies used were Goat Anti-Rabbit-HRP Conjugate and Goat Anti-Mouse-HRP Conjugate (BioRad) and Rabbit anti-Chicken Secondary Antibody (Thermo Scientific).

**Castration-Resistance Assay**

Dissociated N-Myc/myrAKT1 tumor cells were incubated with CD49f-APC antibody (eBiosciences) at 4ºC for 15 minutes. Cells were sorted on a BD FACS ARIA II to isolate the $GFP^+$, $RFP^+$, and $CD49f^{low}$ population. $10^5$ $GFP^+$, $RFP^+$, and $CD49f^{low}$ tumor cells or LNCaP cells were resuspended in 50 µl of cold Matrigel (BD Biosciences) and implanted subcutaneously in intact or surgically castrated NSG mice. Surgical castration (orchiectomy) of mice was performed in accordance with a protocol approved by the Animal Research Committee at UCLA.

**Metastasis Assays**

Dissociated N-Myc/myrAKT1 tumor cells were incubated in HITES media and propagated for two weeks in culture prior to infection with lentivirus expressing firefly luciferase. $10^6$ N-Myc/myrAKT1 or LNCaP cells expressing luciferase were washed twice with PBS, resuspended

in 100 µl of Hank's buffered saline solution (HBSS), and injected into the tail veins of NSG mice. $10^4$ N-Myc/myrAKT1 or LNCaP cells expressing luciferase were also washed in PBS, resuspended in 50 µl of HBSS, and injected into the left anterior lobe of the prostate of NSG mice. Mice were injected with 150 µl of 15 mg/ml luciferin intraperitoneally five minutes prior to live bioluminescent imaging with an IVIS Lumina II (Caliper Life Sciences) under inhaled isoflurane anesthesia. Metastasis assays and live bioluminescent imaging of mice were performed in accordance with a protocol approved by the Animal Research Committee at UCLA.

**Image Segmentation Analysis**

Primary, secondary, and tertiary N-Myc/myrAKT1 tumors passaged in intact and castrate mice were harvested and fixed in 10% buffered formalin for 16 hours. Two tumors for each condition were embedded in paraffin, sectioned and mounted on glass slides, and stained with H&E. High-resolution scans of whole stained sections were obtained using an Aperio ScanScope AT (Leica Biosystems). For each tumor section, ten randomly selected fields of 20X magnification were exported in TIFF format. For image segmentation, the smart segmentation feature of Image-Pro Premier (Media Cybernetics) was trained to classify two classes of cell objects (neuroendocrine carcinoma or adenocarcinoma) and background in H&E-stained sections of human prostate adenocarcinoma and small cell prostate carcinoma using the supervised class assignment of at least 50 individual cellular objects. TIFF images of each H&E-stained N-Myc/myrAKT1 tumor section were processed and the total pixel area of each class type was exported for quantification.

**Laser Capture Microdissection and RNA Isolation**

Tumor tissues were embedded and frozen in O.C.T. freezing compound. 10 μm sections of tumor were cryosectioned and mounted on laser capture microdissection slides. The slides were stained using the Arcturus Histogene LCM Frozen Section Staining Kit (Life Technologies). Laser capture microdissection was performed on a Leica Microsystems LMD7000 with visualization at 5X and 10X magnification. Microdissected specimens were collected into sterile PCR tubes and RNA isolation was performed with the Arcturus PicoPure RNA Isolation Kit (Life Technologies). RNA was analyzed using an RNA Bioanalyzer Kit (Agilent Technologies).

**Whole Transcriptome Sequencing Analysis**

cDNA libraries were prepared from isolated RNA using the TruSeq RNA Sample Prep Kit v2 (Illumina). High-throughput sequencing with 75 bp paired-end reads was performed using an Illumina HiSeq 2500 in rapid run mode. Reads were mapped to human genome reference HG19 using MapSplice (Wang et al., 2010b). Gene expression was quantified using RSEM (Li and Dewey, 2011) and quantile normalized. Gene set enrichment analysis was performed using GSEA software from Broad Institute (Subramanian et al., 2005) with a pre-ranked list of genes differentially expressed (>4-fold) between neuroendocrine carcinoma carcinoma and adenocarcinoma in each tumor sample.

**Neuroendocrine Gene Signature**

A computational model was designed to discriminate the seven NEPC and 30 prostate adenocarcinoma samples from the Beltran *et al.* RNA-seq gene expression dataset. Using the dichotomy based on the clinical diagnoses, we trained a logistic regression model with elastic net regularization (Friedman et al., 2010). We characterized the elastic net regularization with

parameters for the ridge regression term and LASSO term. The ridge regression term was fixed at 1.0 given the absence of preceding information about its importance. A LASSO term was selected to generate a gene expression signature with 50 non-zero weights. The model evaluated *in silico* through leave-pair-out cross validation. This scheme evaluates all possible pairs of one neuroendocrine prostate cancer and one adenocarcinoma sample and withholds each pair from training. The model trains on all other samples and then is applied back to the withheld pair. The final weighted neuroendocrine gene signature score was able to accurately classify all pairs of NEPC and prostate adenocarcinoma samples. Gene ontology was performed with DAVID 6.7 (Dennis et al., 2003) using official gene symbols.

**Copy Number Variation and Whole Exome Sequencing Analysis**

Snap frozen tissues were manually homogenized with a Dounce homogenizer and DNA was extracted using the DNeasy Blood and Tissue Kit (QIAGEN). Copy number analysis was performed on an Affymetrix Cytoscan HD Array and data was processed using the Affymetrix Chromosome Analysis Suite v3.1.0.15 with the NA33 reference model and using the High-Resolution setting. Exomes were isolated with the SeqCap EZ Human Exome Library v3.0 (Roche) and libraries were prepared using the Low-Throughput Library Preparation Kit with standard PCR amplification module (KAPA Biosystems). High-throughput sequencing with 150 bp paired-end reads was performed using an Illumina HiSeq 3000. Sequence data were aligned to the GRCh37 human reference genome using BWA v0.7.7-r411 (Li and Durbin, 2009). PCR duplicates were marked using the MarkDuplicates program in the Picard-Tools-1.115 tool set. GATK v3.2-2 was used for INDEL realignment and base quality recalibration (McKenna et al., 2010). Exome coverage was calculated using BEDTools (Quinlan and Hall, 2010). SAMtools

was used to call the single nucleotide variants and small INDELs (Li et al., 2009). All variants were annotated using ANNOVAR (Wang et al., 2010a). Data was subsequently filtered for genotype quality score, read depth (>15), and single nucleotide polymorphisms using the dbSNP and 1000 Genomes databases (Sherry et al., 1999).

**Serial Dilution and Clonal Subline Xenografts**

For serial dilution experiments, 1, 10, 100, or 1,000 LASCPC-01 cells were resuspended in 20 µl of cold Matrigel and implanted subcutaneously in NSG mice. For clonal subline xenografts, $10^6$ LASCPC-01 cells were resuspended in 50 µl of cold Matrigel and injected subcutaneously in NSG mice. All xenograft studies were performed in accordance with a protocol approved by the Animal Research Committee at UCLA.

**N-Myc Tumor Dependence Studies**

For inducible N-Myc/myrAKT1 experiments, 2 x $10^6$ inducible N-Myc/myrAKT1 tumor cells were resuspended in 50 µl of cold Matrigel and implanted subcutaneously in NSG mice that were fed doxycycline food pellets (Bio-Serv) for induction and/or fed regular food pellets with daily changes in cages and bedding for doxycycline withdrawal. Tumor dimensions were measured by calipers and tumor volumes were calculated using the following equation, V=(L*W*H)/2.

***In vitro* Testing of Aurora A Kinase Inhibitors**

CD532 (EMD Millipore), MLN8237, VX-680, cabazitaxel, and MG-132 (Selleck Chemicals) were dissolved in DMSO. 5 x $10^6$ LASCPC-01 cells were plated in a 10 cm dish in HITES media and incubated with the indicated dose of drug at 37º C prior to collection. For cell cycle analysis,

cells were washed with PBS and fixed in 70% ethanol in PBS overnight at -20ºC. Fixed cells were then stained with propidium iodide (50 µg/ml) and treated with RNase A (100 ng/ml) for 30 minutes at 37ºC. Cell cycle analysis was performed on a BD FACSCanto (BD Biosciences). Cell viability studies were performed by seeding $10^4$ LASCP-01 cells in 100 µl of HITES media in each well of a 96-well white wall optical plate. Cell viability after drug treatment was analyzed relative to DMSO treatment using the CellTiter-Glo Luminescent Cell Viability Assay (Promega). For proteasome inhibition studies, 2 x $10^6$ LASCPC-01 cells were plated in a 10 cm dish in HITES media and pre-treated with 10 µM MG-132 for three hours then treated with the indicated dose of CD532 for an additional three hours.

**Kinase Selectivity Analysis**

Kinase selectivity analysis of CD532 was performed using the KINOME*scan* screening and profiling service with the *scan*MAX panel (DiscoveRx) at a fixed dose of 10 µM (Davis et al., 2011). Small molecule kinase interaction maps were generated using TREE*spot* Compound Profile Visualization Tool and images were reprinted with permission from DiscoveRx Corporation. Quantitative selectivity scores were calculated by enumerating the small molecule kinase interactions at specific thresholds divided by the total number of kinases evaluated.

***In vivo* CD532 Studies**

For short-term CD532 treatment studies, $10^6$ dissociated cells from two independent N-Myc/myrAKT1 tumors were resuspended in 50 µl of cold Matrigel and subcutaneously xenografted in NSG mice. Once tumors achieved a volume of 200 $mm^3$, mice were injected with CD532 60 mg/kg or vehicle (95% PEG 300 and 5% DMSO) intraperitoneally daily for two

doses. For CD532 tumor challenge experiments, $10^6$ LASCPC-01 cells were resuspended in 50 µl of cold Matrigel and implanted subcutaneously in NSG mice. Once tumors achieved a volume of 35 $mm^3$, mice were injected with CD532 25 mg/kg or vehicle (95% PEG 300 and 5% DMSO) intraperitoneally twice per week. Tumor dimensions were measured by calipers and tumor volumes were calculated using the following equation, V=(L*W*H)/2. *In vivo* administration of CD532 and treatment monitoring were performed in accordance with a protocol approved by the Animal Research Committee at UCLA.

**SUPPLEMENTAL REFERENCES**


Davis, M.I., Hunt, J.P., Herrgard, S., Ciceri, P., Wodicka, L.M., Pallares, G., Hocker, M., Treiber, D.K., and Zarrinkar, P.P. (2011). Comprehensive analysis of kinase inhibitor selectivity. Nat Biotech *29*, 1046-1051.

Dennis, G., Jr., Sherman, B.T., Hosack, D.A., Yang, J., Gao, W., Lane, H.C., and Lempicki, R.A. (2003). DAVID: Database for Annotation, Visualization, and Integrated Discovery. Genome Biol *4*, P3.

Friedman, J., Hastie, T., and Tibshirani, R. (2010). Regularization Paths for Generalized Linear Models via Coordinate Descent. Journal of statistical software *33*, 1-22.

Li, B., and Dewey, C.N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC bioinformatics *12*, 323.

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics (Oxford, England) *25*, 1754-1760.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics (Oxford, England) *25*, 2078-2079.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M.*, et al.* (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome research *20*, 1297-1303.

Memarzadeh, S., Xin, L., Mulholland, D.J., Mansukhani, A., Wu, H., Teitell, M.A., and Witte, O.N. (2007). Enhanced paracrine FGF10 expression promotes formation of multifocal prostate adenocarcinoma and an increase in epithelial androgen receptor. Cancer cell *12*, 572-585.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics (Oxford, England) *26*, 841-842.

Sherry, S.T., Ward, M., and Sirotkin, K. (1999). dbSNP-database for single nucleotide polymorphisms and other classes of minor genetic variation. Genome research *9*, 677-679.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S.*, et al.* (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles.

Proceedings of the National Academy of Sciences of the United States of America *102*, 15545-15550.

Wang, K., Li, M., and Hakonarson, H. (2010a). ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic acids research *38*, e164.

Wang, K., Singh, D., Zeng, Z., Coleman, S.J., Huang, Y., Savich, G.L., He, X., Mieczkowski, P., Grimm, S.A., Perou, C.M.*, et al.* (2010b). MapSplice: accurate mapping of RNA-seq reads for splice junction discovery. Nucleic acids research *38*, e178.

Xin, L., Lawson, D.A., and Witte, O.N. (2005). The Sca-1 cell surface marker enriches for a prostate-regenerating cell subpopulation that can initiate prostate tumorigenesis. Proceedings of the National Academy of Sciences of the United States of America *102*, 6942-6947.