SUPPLEMENTARY MATERIAL for

Subclone eradication analysis identifies targets for enhanced cancer therapy and reveals L1 retrotransposition as a dynamic source of cancer heterogeneity

Kirsi Ketola[1], Heidi Kaljunen[1†], Sinja Taavitsainen[2†], Roosa Kaarijärvi[1], Emmi Järvelä[1], Bernardo Rodriguez Martin[3,4], Kerstin Haase[5], Dan J. Woodcock[6], Jose Tubio[3,4], David C. Wedge[7], Matti Nykter[2], and G. Steven Bova[2]

[1]Institute of Biomedicine, University of Eastern Finland, FI-70210, Kuopio, Finland

[2]Faculty of Medicine and Health Technology, Tampere University and Tays Cancer Center, Tampere, FI-33014, Finland

[3]Genomes and Disease, Centre for Research in Molecular Medicine and Chronic Diseases (CIMUS), Universidade de Santiago de Compostela, Santiago de Compostela, Spain

[4]Department of Zoology, Genetics and Physical Anthropology, Universidade de Santiago de Compostela, Santiago de Compostela, Spain

[5]Experimental and Clinical Research Center, Max Delbrück Center and Charité-Universitätsmedizin Berlin, Berlin, Germany

[6]Big Data Institute, University of Oxford, Old Road Campus, Headington, Oxford, UK.

[7]Manchester Cancer Research Centre, University of Manchester, Manchester, UK

[†]These authors contributed equally

Corresponding Authors:

Kirsi Ketola, Institute of Biomedicine, University of Eastern Finland, Kuopio FI-70210, Finland phone: +358503299984 email: kirsi.ketola@uef.fi

G. Steven Bova, Faculty of Medicine and Health Technology, Tampere University and Tays Cancer Center, Tampere, FI-33014, Finland. Phone: +3582945211 email: steve.bova@tuni.fi

The authors declare no potential conflicts of interest.

Running title: Cancer subclone eradication and L1 retrotransposon dynamics

Supplementary Material Index

1. Methods

A34 Samples and DNA data

Tissue and blood samples from patient A34 were collected as part of the PELICAN integrated clinical-molecular autopsy study of lethal prostate cancer (Table 1). The patient gave informed written consent to participate in the John Hopkins Medicine IRB-approved study. Detailed specimen isolation and analysis methods are contained in Woodcock et al(1). In brief, 11 DNA samples studied included five high molecular weight (HMW) samples from microdissected frozen tissue, three laser-microdissected paraffin-embedded tissue samples, and three samples derived from blood (serum and plasma). WGS (average 40x) and targeted deep sequencing (average 785x) data were generated using library creation methods and Illumina sequencing technology. Samples were collected both during life and at autopsy and metastatic samples (Table 1). Potential functional effects of variants were explored algorithmically using MutationTaster(2) and Provean(3). Allele-specific copy number analysis was performed using FACETS v0.5.0(4)**.**

Genome References

WGS, targeted deep sequencing, and methylation DNA data from the tissue and blood samples of A34 were mapped to the hg19 reference genome for comparison with the results of TraFiC-mem(5), which currently requires use of the hg19 reference genome. RNA data was not compared to TraFiC-mem and was therefore mapped to the hg38 reference genome.

Identification of L1 insertion sites

Whole genome sequencing reads aligned to hg19 from A34 metastatic samples were analyzed for somatic L1 insertions (solo-L1 insertions or L1-mediated transductions) using TraFiC-mem v1.1(5) (Supplementary Table 1). Differences in L1 insertion calls made in the current study and those previously reported by Tubio *et al* (5) are attributable to the use of a more recent version of TraFiC-mem and changes to the algorithm (such as the use of BWA-mem rather than RepeatMasker in the search of retrotransposon-like sequences in reads) (Supplementary Tables 1 and 5). All putative L1 insertions were manually validated via examination of their supporting reads in IGV(6) as previously described(7). Ideogram visualizations of L1 transduction events were generated using Circos v0.69-8(8).

Generation and analysis of CpG methylation data

From A34 metastatic and autopsy blood DNA samples, paired-end reads generated from Illumina TruSeq Methyl Capture EPIC libraries were assessed for quality using FastQC v0.11.9 (https://github.com/s-andrews/FastQC) and trimmed using TrimGalore v0.6.5 (https://github.com/FelixKrueger/TrimGalore) (base quality < 20). Trimmed reads were aligned against hg19 using Bismark v0.22.3(9) and Bowtie2 v2.3.4.1(10) (using arguments *--fastq --score-min L,0,-0.2 --no-mixed --no-discordant --dovetail --maxins 500 --ignore-quals*). The methylation information for every C base analyzed was extracted using the Bismark methylation extractor script with arguments *--paired-end --bedGraph --remove_spaces --cytosine_report --ignore_r2 2*. Differential methylation status was assessed using methylKit v1.12.0(10) package in R version 3.6.2. CpG sites with at least 20% absolute methylation percentage change between compared groups and q-value < 0.01 were considered to be differentially methylated.

Cell culture

LNCaP and 22Rv1 cells were cultured in a humidified CO2-incubator at 37°C in Gibco™ RPMI 1640 (1X) media (Thermo Fisher Scientific) supplemented with 10% FBS (Gibco standard FBS, Thermo Fisher Scientific), 2 mM L-Glutamine (Gibco®, Thermo Fisher Scientific), and combination of 100 U/ml Penicillin and 100 µg/mL Streptomycin (Gibco® Pen Strep, Thermo Fisher Scientific). For VCaP cells, Dulbecco's Modified Eagle Medium (DMEM, Gibco, Thermo Fisher Scientific) supplemented with 10% FBS (GE Healthcare™ HyClone™) and for PC-3 F-12 K medium supplemented with 10% FBS, and antibiotics as for LNCaP growth medium was used. Cell lines were authenticated using Promega STR (Short Tandem Repeat) systems referenced to ATCC STR database by the the Johns Hopkins Genetic Resources Core Facility (JHGRCF) (PC3, LNCaP, 22Rv1) and by the FIMM HiLife Unit (VCaP).. PC3, LNCaP, 22Rv1 cell lines were obtained from ATCC in 2011 and were authenticated at p0, and p1 cells were used for the experiments. Mycoplasma testing of PC3, LNCaP, 22Rv1 cell lines was done by JHGRCF in p0 cells using MycoDtect™ (Greiner Bio-One). VCaP cells were obtained from ATCC, were authenticated in the passage just prior to passage of the cells used in the experiments, and were Mycoplasma tested just prior to the experiments using Venor®GeM Classic (Sigma).

The culturing of LNCaP, VCaP, PC-3 and 22Rv1 prostate cancer cells was performed on 6-well plates for the RNA analysis with quantitative real-time PCR (RT-PCR). The seeding density of LNCaP, PC-3 and 22Rv1 cells was $3.5x10^5$ cells per well and the corresponding value for VCaP cells was $4.0x10^5$ cells per well.

Exposure of cell lines to carboplatin/enzalutamide alone

Carboplatin or enzalutamide dissolved in DMSO and culture medium without FBS was added after a 24-hour initial incubation period. To control cells DMSO diluted in culture medium was added to a final concentration of 0.02%. LNCaP prostate cancer cells were transfected with L1 plasmid or positive and negative control plasmids and treated with carboplatin (5 µM) or ENZ (10 µM) the day after transfection, and cells were monitored for five days.

Exposure of cell lines to carboplatin/enzalutamide and AZT

First, the potential cell toxicity of AZT alone was determined and no effects of AZT on cell viability were seen up to 50 µM tested in LNCaP or VCaP cells (Supplementary Fig. S1). Next, AZT was introduced to both LNCaP and VCaP cells alone and in combination with carboplatin or ENZ.

L1-EGFP Retrotransposition assay

The setup of the retrotransposition assay and the creation of the plasmids containing an L1RP(11) element tagged with Green Fluorescent Protein (GFP) were as described by Ostertag et al.(12) and illustrated in Faulkner et al (13) with the following modifications. The plasmids used in our assays were purchased from Addgene and contained a puromycin resistance gene instead of hygromycin with derivatization by Farkash et al.(14). The control plasmids used included two negative controls, one with disabling mutations at ORF1 (plasmid ID EF05J, L1Neg-EGFP thereafter) and other with deactivated endonuclease domain (plasmid ID EF13E, L1Mut-EGFP thereafter), and a positive control plasmid (EF12J, L1Pos-EGFP thereafter) where L1RP element had been swapped with less active L1.3 element. The transfections of LNCaP cells were performed on 96-well plates using a reverse transfection approach and Lipofectamine 3000 reagent (Thermo Fisher Scientific) according to the

manufacturer's guidelines. LNCaP cells were seeded at a density of 5000 cells per well. The green fluorescence of activated L1 in cells was monitored using IncuCyte S3 live-cell imaging (Sartorius) equipped with a green fluorescence channel (Fig. 3A). L1 positive EGFP+ counts per mm^2 were adjusted for overall cell confluency. Briefly, the cell confluence was determined by automatic counting using IncuCyte S3 image analysis tools; the cell confluence (confluence mask) was calculated based on phase-contrast imaging with a minimum area filter of 150 μm^2, segmentation adjustment of 1.0, 100 μm^2 cleanup and size adjust of −2 pixels. Cell green fluorescence (green emission wavelength: 524 nm, excitation wavelength: 460 nm) was determined using green fluorescence object count and top-hat image processing method (Sartorius) including background subtraction segmentation radius of 10 µm, 1 μm^2 cleanup, and one pixel size adjust as well as filters for area (100 μm^2), cell eccentricity (0.1), and mean intensity (0.1)(15).

siRNA silencing

LNCaP or PC-3 cells were reverse transfected with 25 nM siRNAs against FANCI (Dharmacon, ON-TARGETplus SMARTpool, L-022320-01-0005) or Scr control (Dharmacon, non-targeting pool), using OPTI-MEM and Lipofectamine 2000 transfection reagent (Invitrogen) in three biological replicates on 12-well plates. Samples were collected after 72 hours for mRNA isolation (Supplementary Table 2).

Cell proliferation assay

For cell proliferation assays, cells were reverse transfected with 25 nM siRNAs against FANCI (Dharmacon, ON-TARGETplus SMARTpool, L-022320-01-0005) or Scr control (Dharmacon, non-targeting pool), using OPTI-MEM and Lipofectamine 2000 transfection reagent (Invitrogen) in four biological replicates in a 384-well plate. LNCaP cells were plated on the wells (1000 cells/well) in antibiotic free regular growth medium. After 48h, vehicle control (DMSO) or Carboplatin dilutions were added into wells in appropriate concentrations diluted in FBS free medium and the cell confluence was monitored for five days using live-cell imaging (IncuCyte S3, Sartorius). Cell confluence was determined by automatic counting using IncuCyte S3 Image analysis tools. Cell confluence (confluence mask) on each time point was calculated based on phase-contrast imaging with a minimum area filter of 150 μm^2, segmentation adjustment of 1.0, 100 μm^2 cleanup and size adjust of −2 pixels. Confluence curves were compared for statistical differences between treatment conditions at each timepoint using t-test.

Quantitative Real-time PCR

The isolation of RNA from cell lines was conducted using TriPure Isolation Reagent (Roche) following the manufacturer's protocol. The concentration of the RNA samples was measured using NanoDrop™ One/OneC Microvolume UV-Vis Spectrophotometer, followed by dilution of the samples to 1 µg/µl. The conversion of 1 µg of RNA to cDNA was done using Transcriptor First Strand cDNA Synthesis Kit (Roche) according to the manufacturer's instructions. The RT-PCR run was performed using LightCycler™ 480 SYBR Green I Master (Roche) and The LightCycler® 480 Real-Time PCR System (Roche) with 96-multiwell format. To allow separate analysis of ORF1 and ORF2, two sets of primer pairs were designed to target the L1RP-ORF1 (ORF1 mRNA thereafter), L1RP-ORF2 (ORF2 mRNA thereafter) and FANCI (Supplementary Table 3). Each run included two biological and two technical replicates per treatment with fold change calculated based on the obtained Ct-values. Normalization was done against GAPDH values measured (Supplementary Table 3).

Analysis of LuCaP xenograft samples was done using the same approach with Trizol (Invitrogen) used for RNA isolation, Qubit 4 fluorometer (Thermo Scientific) used to measure concentration, and Maxima Reverse Transcriptase (Thermo Scientific) used for cDNA synthesis. For measurements, Thermo Scientific Maxima SYBR Green/ROX qPCR Master Mix (2x) was used on Bio-Rad CFX96 Real-Time PCR Detection System (96-well format).

Western Blot

Whole cell lysates were prepared using SDS sample buffer (66 mM Tris-HCl pH 6.8, 13% Glycerol, 2.1% SDS and 0,01% Bromophenol Blue) with protease inhibitor added (cOmplete™ Protease Inhibitor Cocktail, Roche). Samples were sonicated prior to addition of β-mercaptoethanol (2% v/v; Biorad, cat. no. 1610710) and boiled for 5 min at 95°C before loading into the gel. Proteins were separated by SDS-PAGE and wet transferred in methanol-based transfer buffer (20% v/v methanol, 25 mM Tris, 192 mM Glycine) onto 0.45 µm nitrocellulose membrane (Thermo Scientific; cat. no. 88018) using constant 250 mA current at 4°C for 60 min. Blocking was conducted by incubating membranes for 1 h at room temperature in 5% nonfat dry milk in 1xTBS-Tween (20mM Tris-Cl, 137mM NaCl, 0.1% Tween 20) after which primary antibodies were applied overnight at 4°C. Primary antibody solutions were prepared in blocking buffer. Anti-ORF1p (Millipore Sigma, cat. no. MABC1152) was used at 0.4 µg/ml and anti-GAPDH (Santa Cruz Biotechnology, cat. no. sc-25778) at 0.2 µg/ml. Secondary antibody incubation was performed at room temperature for 45 min. Goat anti-rabbit IgG horseradish peroxidase conjugate (Invitrogen, cat. no. G21234) was used as the secondary antibody at 0.1 µg/ml diluted in 1xTBS-tween. For detection Pierce™ ECL Western Blotting Substrate (Thermo Scientific; cat. no. 32106) was used and imaging was done using the Chemidoc Imaging system (Biorad).

Estimation of L1 mRNA levels from RNA-seq

RNA sequencing reads were downloaded for two prostate cancer patient-derived pre- and post-castration resistance LuCaP xenografts (77 and 105)(16) and from head and neck squamous cell carcinoma cell line SCC25 samples treated in a time-series with either PBS or cetuximab for 11 weeks(17). Reads were processed with TrimGalore v0.6.5, a tool that runs FastQC v0.11.9 to assess the quality of the reads and Cutadapt v1.18(18) to filter and trim the reads for poor quality bases and adapters. STAR v2.7.8(19) was used to align the reads to the hg38 reference genome using gene annotations from Gencode Release 33. featureCounts v2.0.2(20) was used to quantify reads (including fractional counts for multi-mapping reads) within 146 putatively retrotransposition-active human L1 elements with intact ORF1 and ORF2 as annotated in the L1Base2 database(21). Counts were normalized in each sample to represent the number of reads mapping to the putatively active L1 elements per million aligned reads in the sample.

2. Supplementary Text: Retrotransposons Background

Retrotransposons represent a class of mobile genetic elements that contribute to modification of the human genome during evolution and play a role in regulation of important biological processes at the genetic or epigenetic level(22–27). There are three categories of retrotransposons that have remained active: Long interspersed nuclear element-1 (often termed LINE-1 or L1), Alu, and SVA elements. The activity of the latter two are dependent on L1 retrotransposons.

L1 retrotransposons account for about 17% of human DNA with approximately 500,000 copies found in the human genome, although only a subset of these are retrotransposition competent(28,29). The majority of known L1s have lost their functionality due to truncations, inverted rearrangements, or point mutations occurring during reverse transcription or subsequent chromosomal replication of the inserted element(30). On average the human genome contains about 80-100 retrotransposition-competent L1s of which 5-10% are highly active (also termed “hot” L1s) and thus responsible for most detected L1 activity(29).

An active, retrotransposition-competent human L1 contains a 5´ untranslated region (UTR), two open reading frames (ORF1 and ORF2), and a 3´UTR with a poly(A) tail and is about 6 kilobases in length(31,32). Human L1s have two internal promoters, sense and antisense, included in the 5´UTR(33–35). The binding site of RNA polymerase II is located at the sense promoter, which is also responsible for the initiation of the transcription process from the 5´ end to the 3´ end. The antisense promoter works in the opposite direction and produces chimeric transcripts through transcriptional activation of adjacent genes(35). The proteins encoded by ORF1 and ORF2 are required for retrotransposition to occur. ORF1 translates into a 40 kDa protein (referred as ORF1p) with a nucleic acid chaperone activity and RNA-recognition motif(36), while ORF2 protein (ORF2p) is a 150 kDa protein that has both endonuclease and reverse transcriptase domains(31,37).

L1 retrotransposons belong to the category of non-LTR retrotransposons as they lack the long terminal repeats (LTR) found in human endogenous retroviruses(38). The propagation mechanism of L1s also differs from that employed by retroviruses. Binding of RNA polymerase II initiates L1 transcription from the 5´ to 3´ end, after which the L1 mRNA is transported from the nucleus to the cytoplasm. In the cytoplasm, the translation of the L1 mRNA into two L1-encoded proteins, ORF1p and ORF2p, is followed by their binding to their encoding RNA to form an L1 ribonucleoprotein particle(12,39,40). This particle is responsible for the insertion of the L1 retrotransposon at a new genomic location through a process called target primed reverse transcription (TPRT)(12,39). In TPRT, the endonuclease of ORF2p creates a nick the target site, producing a free 3´-hydroxyl, which serves as a primer for the reverse transcription by the reverse transcriptase (RT) also found in ORF2p. The retrotransposon RNA that anneals at the endonuclease cleavage site functions as a template for the reverse transcription. The final steps of the TPRT include cleavage of the second DNA strand, integration of the newly synthesized L1 cDNA and completion of the DNA synthesis resulting in L1 insert flanked by target site duplications (TSDs)

In normal differentiated tissues the expression and thus the retrotransposition activity of L1s is kept repressed by methylation of its 5’ promoter(41–43). The upregulation of L1 RNA and

protein results in increased genomic instability in the form of DNA lesions, such as mutations, deletions or large genomic rearrangements produced by L1-mediated insertional events(12,44–48). L1 activation has been observed as a consequence of oxidative stress(49) and radiation(14)f. L1 activation has a known role in tumorigenesis and elevated levels of both L1-encoded proteins, ORF1p and ORF2p, have been detected in various cancer tissues and cell lines, including prostate cancer(50–52).

3. Supplementary References

1. Woodcock DJ, Riabchenko E, Taavitsainen S, Kankainen M, Gundem G, Brewer DS, et al. Prostate cancer evolution from multilineage primary to single lineage metastases with implications for liquid biopsy. Nat Commun. Nature Publishing Group; 2020;11:5070.

2. Schwarz JM, Cooper DN, Schuelke M, Seelow D. MutationTaster2: mutation prediction for the deep-sequencing age. Nat Methods. Nature Publishing Group; 2014;11:361–2.

3. Choi Y, Sims GE, Murphy S, Miller JR, Chan AP. Predicting the Functional Effect of Amino Acid Substitutions and Indels. PLOS ONE. Public Library of Science; 2012;7:e46688.

4. Shen R, Seshan VE. FACETS: allele-specific copy number and clonal heterogeneity analysis tool for high-throughput DNA sequencing. Nucleic Acids Res. 2016;44:e131.

5. Tubio JMC, Li Y, Ju YS, Martincorena I, Cooke SL, Tojo M, et al. Mobile DNA in cancer. Extensive transduction of nonrepetitive DNA mediated by L1 retrotransposition in cancer genomes. Science. 2014;345:1251343.

6. Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, et al. Integrative Genomics Viewer. Nat Biotechnol. 2011;29:24–6.

7. Rodriguez-Martin B, Alvarez EG, Baez-Ortega A, Zamora J, Supek F, Demeulemeester J, et al. Pan-cancer analysis of whole genomes identifies driver rearrangements promoted by LINE-1 retrotransposition. Nat Genet. 2020;1–14.

8. Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, et al. Circos: an information aesthetic for comparative genomics. Genome Res. 2009;19:1639–45.

9. Krueger F, Andrews SR. Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. Bioinforma Oxf Engl. 2011;27:1571–2.

10. Akalin A, Kormaksson M, Li S, Garrett-Bakelman FE, Figueroa ME, Melnick A, et al. methylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles. Genome Biol. 2012;13:R87.

11. Kimberland ML, Divoky V, Prchal J, Schwahn U, Berger W, Kazazian HH. Full-length human L1 insertions retain the capacity for high frequency retrotransposition in cultured cells. Hum Mol Genet. 1999;8:1557–60.

12. Ostertag EM, Kazazian HH. Twin Priming: A Proposed Mechanism for the Creation of Inversions in L1 Retrotransposition. Genome Res. 2001;11:2059–65.

13. Faulkner GJ, Billon V. L1 retrotransposition in the soma: a field jumping ahead. Mob DNA. 2018;9:22.

14. Farkash EA, Kao GD, Horman SR, Prak ETL. Gamma radiation increases endonuclease-dependent L1 retrotransposition in a cultured cell assay. Nucleic Acids Res. Oxford Academic; 2006;34:1196–204.

15. Gelles JD, Mohammed JN, Santos LC, Legarda D, Ting AT, Chipuk JE. Single-Cell and Population-Level Analyses Using Real-Time Kinetic Labeling Couples Proliferation and Cell Death Mechanisms. Dev Cell. 2019;51:277-291.e4.

16. Nguyen HM, Vessella RL, Morrissey C, Brown LG, Coleman IM, Higano CS, et al. LuCaP Prostate Cancer Patient-Derived Xenografts Reflect the Molecular Heterogeneity of Advanced Disease an--d Serve as Models for Evaluating Cancer Therapeutics. The Prostate. 2017;77:654–71.

17. Stein-O'Brien G, Kagohara LT, Li S, Thakar M, Ranaweera R, Ozawa H, et al. Integrated time course omics analysis distinguishes immediate therapeutic response from acquired resistance. Genome Med. 2018;10:37.

18. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal. 2011;17:10–2.

19. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinforma Oxf Engl. 2013;29:15–21.

20. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinforma Oxf Engl. 2014;30:923–30.

21. Penzkofer T, Jäger M, Figlerowicz M, Badge R, Mundlos S, Robinson PN, et al. L1Base 2: more retrotransposition-active LINE-1s, more mammalian genomes. Nucleic Acids Res. 2017;45:D68–73.

22. Burns KH. Transposable elements in cancer. Nat Rev Cancer. 2017;17:415–24.

23. Apostolou P, Toloudi M, Chatziioannou M, Kourtidou E, Mimikakou G, Vlachou I, et al. Involvement of retrotransposon L1 in stemness and cellular plasticity. Cell Commun Adhes. Taylor & Francis; 2015;22:1–7.

24. Beck CR, Garcia-Perez JL, Badge RM, Moran JV. LINE-1 Elements in Structural Variation and Disease. Annu Rev Genomics Hum Genet. Annual Reviews; 2011;12:187–215.

25. Rodić N, Burns KH. Long Interspersed Element–1 (LINE-1): Passenger or Driver in Human Neoplasms? PLOS Genet. Public Library of Science; 2013;9:e1003402.

26. Belancio VP, Roy-Engel AM, Deininger PL. All y'all need to know 'bout retroelements in cancer. Semin Cancer Biol. 2010;20:200–10.

27. Sciamanna I, Landriscina M, Pittoggi C, Quirino M, Mearelli C, Beraldi R, et al. Inhibition of endogenous reverse transcriptase antagonizes human tumor growth. Oncogene. Nature Publishing Group; 2005;24:3923–31.

28. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, et al. Initial sequencing and analysis of the human genome. Nature. Nature Publishing Group; 2001;409:860–921.

29. Brouha B, Schustak J, Badge RM, Lutz-Prigge S, Farley AH, Moran JV, et al. Hot L1s account for the bulk of retrotransposition in the human population. Proc Natl Acad Sci. 2003;100:5280–5.

30. Szak ST, Pickeral OK, Makalowski W, Boguski MS, Landsman D, Boeke JD. Molecular archeology of L1 insertions in the human genome. Genome Biol. 2002;3:research0052.1.

31. Mathias SL, Scott AF, Kazazian HH, Boeke JD, Gabriel A. Reverse transcriptase encoded by a human transposable element. Science. American Association for the Advancement of Science; 1991;254:1808–10.

32. Burns KH, Boeke JD. Human Transposon Tectonics. Cell. 2012;149:740–52.

33. Mizrokhi LJ, Georgieva SG, Ilyin YV. jockey, a mobile drosophila element similar to mammalian LINEs, is transcribed from the internal promoter by RNA polymerase II. Cell. 1988;54:685–91.

34. Swergold GD. Identification, characterization, and cell specificity of a human LINE-1 promoter. Mol Cell Biol. American Society for Microbiology Journals; 1990;10:6718–29.

35. Speek M. Antisense Promoter of Human L1 Retrotransposon Drives Transcription of Adjacent Cellular Genes. Mol Cell Biol. American Society for Microbiology Journals; 2001;21:1973–85.

36. Martin SL. Nucleic acid chaperone properties of ORF1p from the non-LTR retrotransposon, LINE-1. RNA Biol. Taylor & Francis; 2010;7:706–11.

37. Feng Q, Moran JV, Kazazian HH, Boeke JD. Human L1 Retrotransposon Encodes a Conserved Endonuclease Required for Retrotransposition. Cell. 1996;87:905–16.

38. Malik HS, Burke WD, Eickbush TH. The age and evolution of non-LTR retrotransposable elements. Mol Biol Evol. Oxford Academic; 1999;16:793–805.

39. Luan DD, Korman MH, Jakubczak JL, Eickbush TH. Reverse transcription of R2Bm RNA is primed by a nick at the chromosomal target site: A mechanism for non-LTR retrotransposition. Cell. 1993;72:595–605.

40. Kolosha VO, Martin SL. In vitro properties of the first ORF protein from mouse LINE-1 support its role in ribonucleoprotein particle formation during retrotransposition. Proc Natl Acad Sci. National Academy of Sciences; 1997;94:10155–60.

41. Schulz WA, Steinhoff C, Florl AR. Methylation of Endogenous Human Retroelements in Health and Disease. In: Doerfler W, Böhm P, editors. DNA Methylation Dev Genet Dis Cancer [Internet]. Berlin, Heidelberg: Springer; 2006 [cited 2020 May 21]. page 211–50. Available from: https://doi.org/10.1007/3-540-31181-5_11

42. Harris CR, Normart R, Yang Q, Stevenson E, Haffty BG, Ganesan S, et al. Association of Nuclear Localization of a Long Interspersed Nuclear Element-1 Protein in Breast Tumors with Poor Prognostic Outcomes. Genes Cancer. SAGE Publications; 2010;1:115–24.

43. Su Y, Davies S, Davis M, Lu H, Giller R, Krailo M, et al. Expression of LINE-1 p40 protein in pediatric malignant germ cell tumors and its association with clinicopathological parameters: A report from the Children's Oncology Group. Cancer Lett. 2007;247:204–12.

44. Kazazian HH. Mobile elements and disease. Curr Opin Genet Dev. 1998;8:343–50.

45. Symer DE, Connelly C, Szak ST, Caputo EM, Cost GJ, Parmigiani G, et al. Human L1 Retrotransposition Is Associated with Genetic Instability In Vivo. Cell. 2002;110:327–38.

46. Gilbert N, Lutz-Prigge S, Moran JV. Genomic Deletions Created upon LINE-1 Retrotransposition. Cell. 2002;110:315–25.

47. Han K, Sen SK, Wang J, Callinan PA, Lee J, Cordaux R, et al. Genomic rearrangements by LINE-1 insertion-mediated deletion in the human and chimpanzee lineages. Nucleic Acids Res. Oxford Academic; 2005;33:4040–52.

48. Wallace NA, Belancio VP, Deininger PL. L1 mobile element expression causes multiple types of toxicity. Gene. 2008;419:75–81.

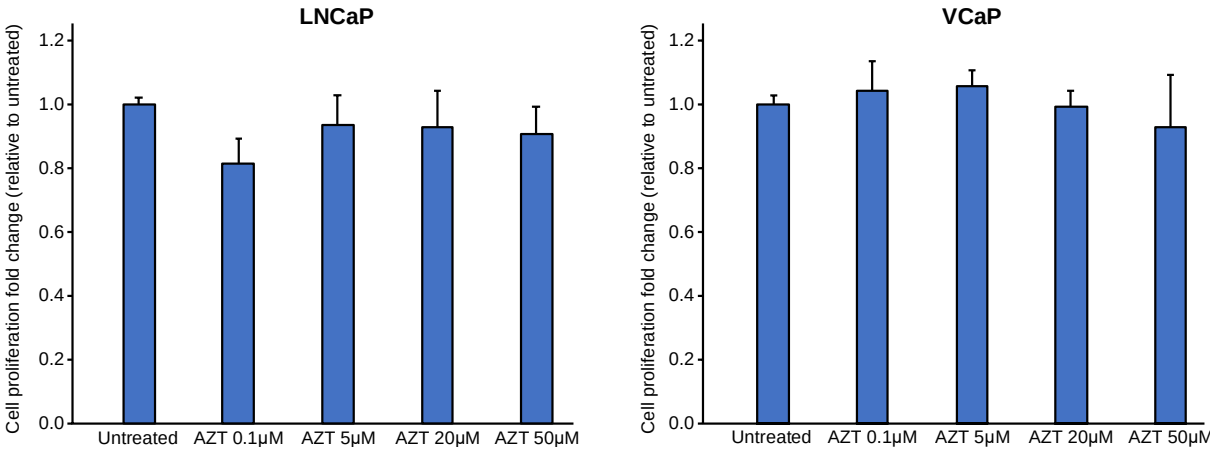
49. Giorgi G, Marcantonio P, Del Re B. LINE-1 retrotransposition in human neuroblastoma cells is affected by oxidative stress. Cell Tissue Res. 2011;346:383–91.

50. Cho N-Y, Kim B-H, Choi M, Yoo EJ, Moon KC, Cho Y-M, et al. Hypermethylation of CpG island loci and hypomethylation of LINE-1 and Alu repeats in prostate adenocarcinoma and their relationship to clinicopathological features. J Pathol. 2007;211:269–77.

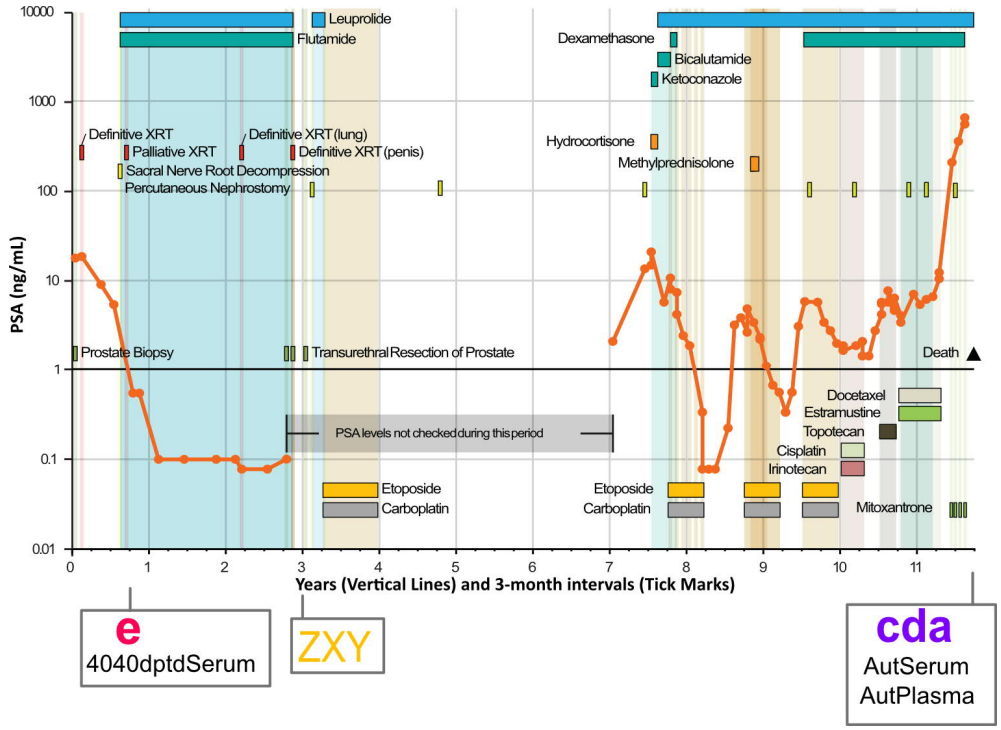
51. Luca CD, Guadagni F, Sinibaldi-Vallebona P, Sentinelli S, Gallucci M, Hoffmann A, et al. Enhanced expression of LINE-1-encoded ORF2 protein in early stages of colon and prostate transformation. Oncotarget. Impact Journals; 2015;7:4048–61.

52. Briggs EM, Ha S, Mita P, Brittingham G, Sciamanna I, Spadafora C, et al. Long interspersed nuclear element-1 expression and retrotransposition in prostate cancer cells. Mob DNA. 2018;9:1.

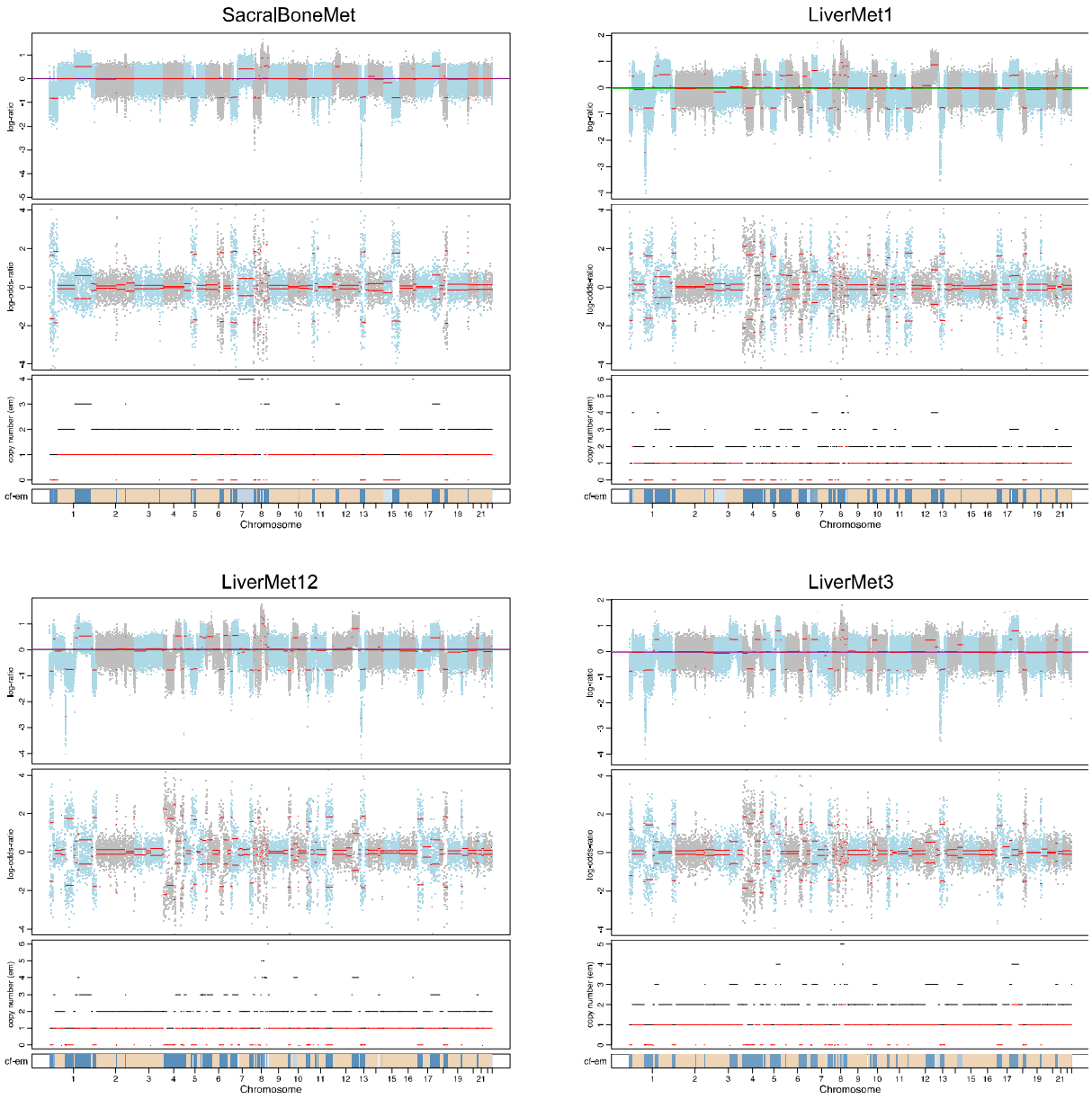
53. Behan FM, Iorio F, Picco G, Gonçalves E, Beaver CM, Migliardi G, et al. Prioritization of cancer therapeutic targets using CRISPR-Cas9 screens. Nature. 2019;568:511–6.



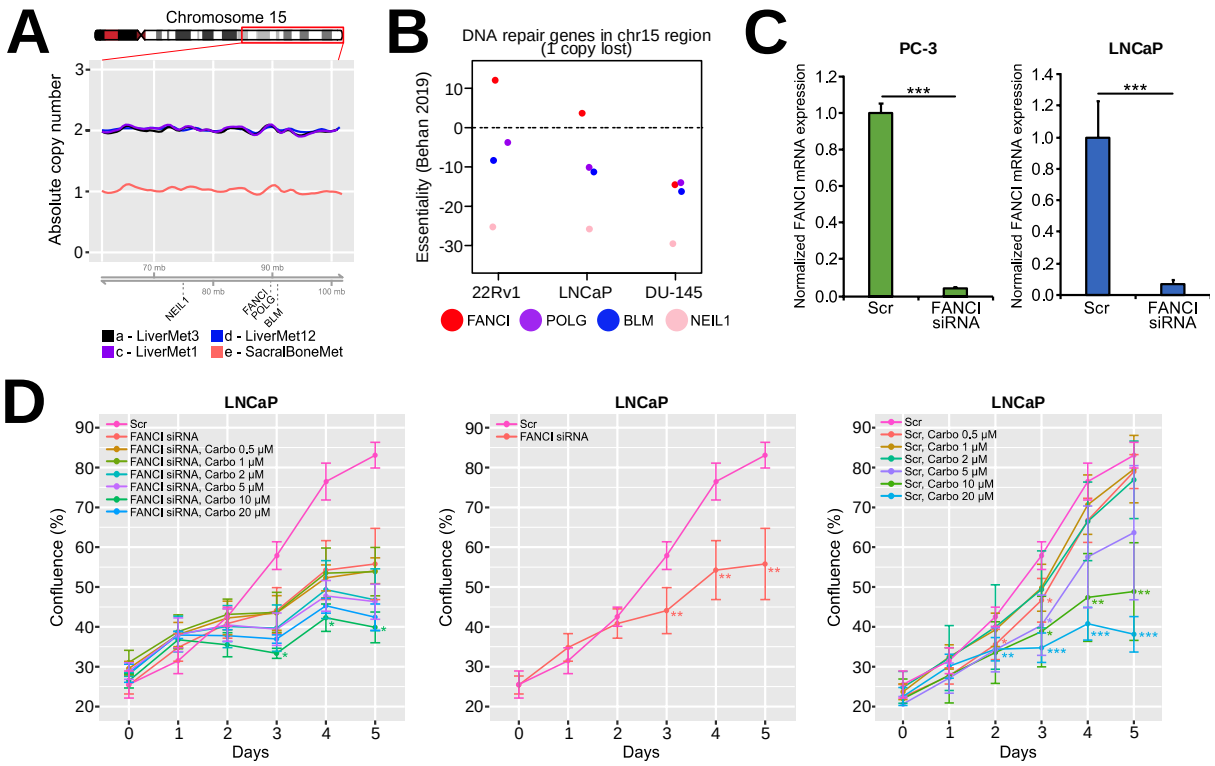
Supplementary Fig. S1. AZT treatment of LNCaP and VCaP cell lines. As a preliminary study, we explored whether AZT has any cytotoxic effect in LNCaP or VCaP cells. The cells were exposed to various concentrations of AZT (0.1, 5, 20, 50 uM) for 5 days and cell confluence was determined using IncuCyte S3. The effect of AZT on cell proliferation was calculated relative to the untreated control. No significant cytotoxic effect was seen up to 50uM AZT treatment.



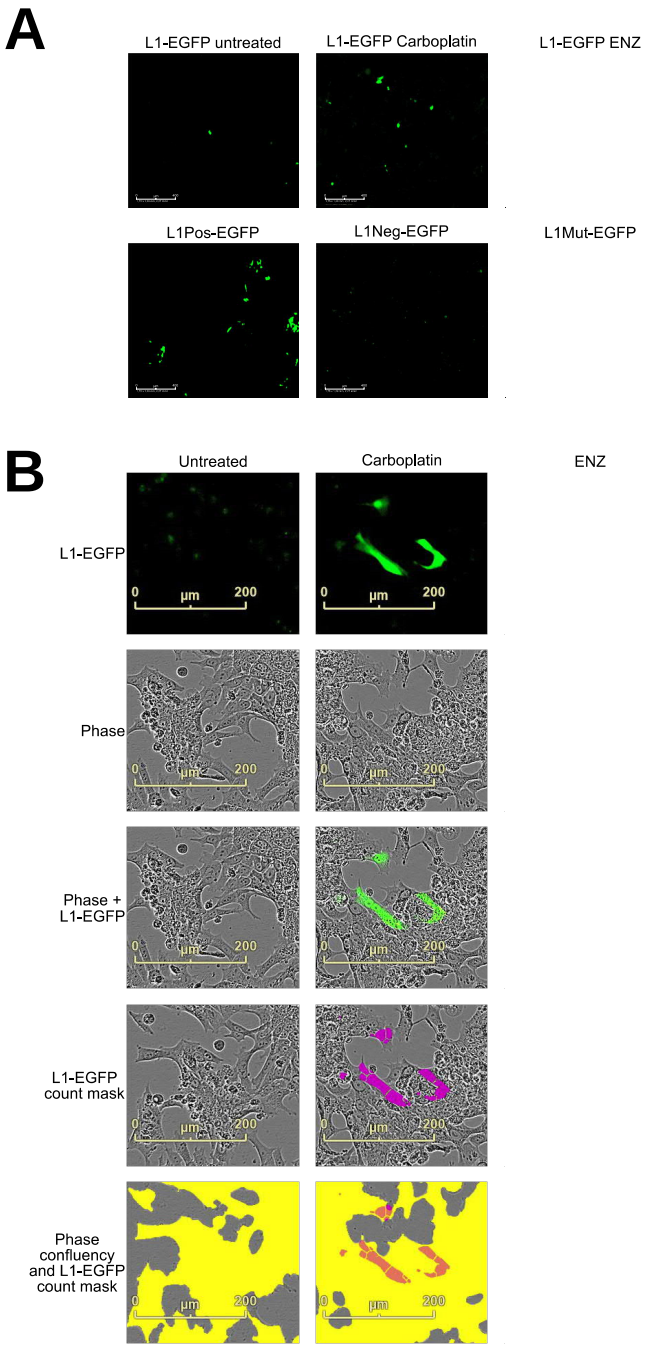
Supplementary Fig. S2. A34 clinical timeline and samples analyzed. At age 54 (age changed between -3 and 3 years according to study consent) patient A34 developed a single episode of painless hematuria, triggering a visit to his physician. On digital rectal exam he was found to have an indurated right lobe of the prostate. Prostate biopsy showed Gleason score 2+3=5 prostate adenocarcinoma. Serum PSA at diagnosis was 16.2 ng/mL and technetium (99mTc MDP) bone scan and prostate acid phosphatase were normal. He then received 6400 cGy external beam radiation to his prostate with intent to cure. Seven months later he developed leg pain and perineal numbness, and underwent S1-S3 decompressive laminectomy, when a sample of the metastatic tumor tissue ("e" in red) and serum sample (used for cell-free circulating DNA isolation) were obtained. A34 commenced leuprolide/flutamide androgen deprivation at the time of laminectomy, and received 4500 cGy palliative radiation to the sacrum starting 1 month after laminectomy. A routine chest radiogram 18 months after laminectomy revealed a left hilar mass, bronchoscopic biopsy showed poorly differentiated carcinoma with glandular features, and he received 6000 cGy radiation to the left lung hilum. Eight months later he was found to have a prostate cancer penile skin metastasis confirmed by biopsy. He also underwent transurethral resection of the prostate for urinary obstructive symptoms at the same time when prostate carcinoma tissue Z, X, and Y were obtained as shown. He subsequently received 4000 cGy radiation to the penis. Bilateral percutaneous nephrostomy tubes were placed when urinary obstruction symptoms persisted. Six months later he received a first course of Carboplatin and Etoposide chemotherapy. The patient felt dramatically better after chemotherapy, and refused to see doctors except for nephrostomy tube change for the ensuing three years. He then presented with severe lethargy, and over the ensuing four years underwent three additional courses of carboplatin/etoposide and other chemotherapy as shown. Peak serum PSA close to death was 603.1 ng/mL. At autopsy, large liver metastases nearly consumed the liver, and metastases to lumbosacral bone and perirenal area were present. No lung metastases were present. Metastatic cancer DNA and RNA samples from three separate liver metastases ("cda" in purple), and cell-free circulating DNA were obtained at autopsy for the analysis.



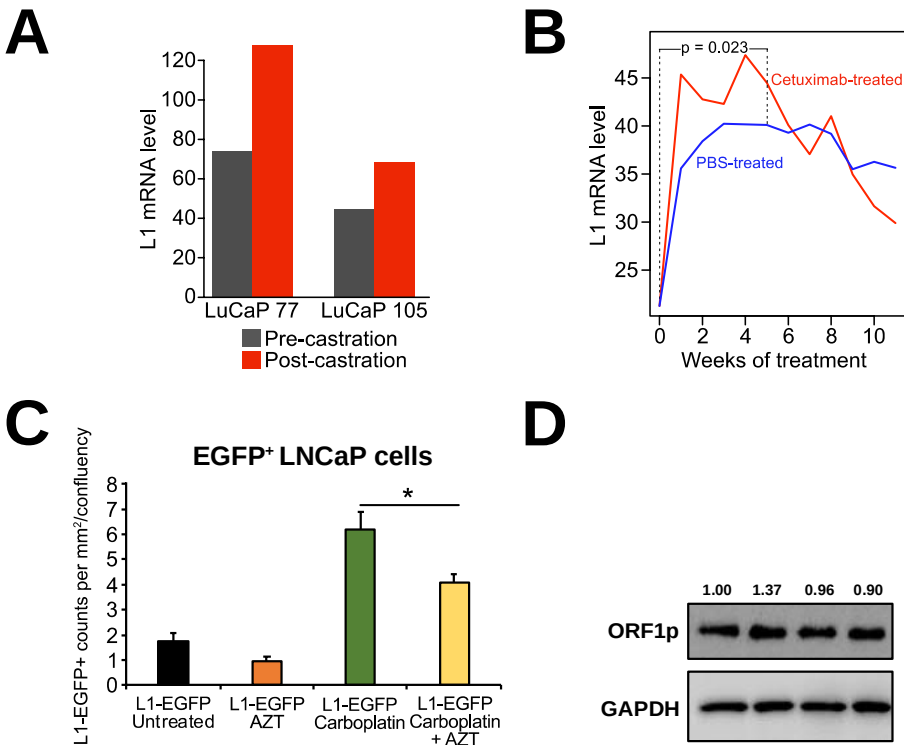
Supplementary Fig. S3. Allele-specific copy number profiles of A34 metastatic samples. A figure is included for each sample, with chromosomes shown as alternating light blue and grey colors. Mean values for segments are shown as red lines. The top panel of each figure shows the log-ratio of read depth in the metastatic sample compared to the normal sample. Green and purple lines indicate the sample's median log-ratio and the diploid log-ratio, respectively. The second panel shows the log-odds ratio of variant allele counts in the metastatic sample compared to the normal sample. The third panel shows total (black lines) and minor (red lines) copy number. The chromosome bar at the bottom of each figure depicts estimated cellular fraction (cf), where the dark blue color indicates high cellular fraction, light blue indicates low cellular fraction, and beige indicates a diploid segment where total copy number is 2 and minor copy number is 1(4).



Supplementary Fig. S4. Additional data supporting FANCI as target identified by DSER analysis. (A) Line plots of absolute copy number showing one copy loss of a ~40 Mbp region of chromosome 15 specific to the SacralBoneMet sample harboring the eradicated subclones in A34. We confirmed the absence of 15q copy neutral loss-of-heterozygosity in the LiverMet samples (Supplementary Fig. S3). The lost region of 15q contains 295 protein-coding genes, of which four are DNA repair-implicated: NEIL1, FANCI, POLG, and BLM. These genes are marked along the genomic axis at the bottom of the plot. (B) PICKLES database essentiality scores based on data from Behan et al(53) for the four DNA repair-related genes (*FANCI*, *POLG*, *BLM*, and *NEIL1*) located in the chr15q region with loss of heterozygosity in the sacral bone metastasis sample. *FANCI* has a positive essentiality score in 22Rv1 and LNCaP cells, suggesting it may be necessary for cell survival. The essentiality score is a quantile normalized Bayes factor that represents the level of confidence that the gene is essential. (C) Barplots confirming FANCI siRNA knockdown in PC-3 and LNCaP cells based on significantly reduced FANCI mRNA expression level compared to scrambled (Scr). mRNA expression values were normalized against measured GAPDH expression values. (D) LNCaP cell confluency curves when exposed to FANCI siRNA and varying concentrations of carboplatin. Scrambled (Scr) siRNA is shown as a control. t-test was used to determine statistical significance of sample conditions at each timepoint compared to FANCI siRNA (left) and Scr (middle and right) (*, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$).



Supplementary Fig. S5. LNCaP L1 response to carboplatin and enzalutamide exposure including image processing details. (A) The effects of carboplatin (5 µM) or enzalutamide (ENZ, 10 µM) on L1 activity in LNCaP prostate cancer cells analyzed using a retrotransposition L1-EGFP reporter assay and IncuCyte S3 imaging system equipped with a green fluorescence channel. The control plasmids used included two negative controls (L1Neg-EGFP) and (L1Mut-EGFP) and a positive control plasmid (L1Pos-EGFP). (B) Representative IncuCyte cell images and visualized image processing calculations used in L1 activity assays in LNCaP prostate cancer cells. Separate images are shown for untreated, carboplatin, and ENZ exposed LNCaP cells with green channel only (L1-EGFP), phase contrast only, and phase contrast + L1-EGFP, as well as for the masking of the individual green cells "L1-EGFP count mask" (purple) and "Phase confluency and L1-EGFP count mask" (yellow). See Methods.



Supplementary Fig. S6. Further validation of L1 activation and suppression. (A) L1 mRNA levels from pre- and post-castration samples from LuCaP 77 and LuCaP 105(16). The y-axis shows the number of reads mapping to the 146 putatively retrotransposition active L1 elements with intact ORF1 and ORF2 from L1Base2, divided by million total aligned reads to the human genome in each sample. (B) L1 mRNA levels from head and neck squamous cell carcinoma cell line SCC25 samples treated with PBS or cetuximab in a time-series for 11 weeks(17). The y-axis shows the number of reads mapping to the 146 putatively retrotransposition active L1 elements with intact ORF1 and ORF2 from L1Base2, divided by million total aligned reads to the human genome in each sample. L1 mRNA levels are significantly different between cetuximab- and PBS-treated samples during the first five weeks of treatment ($p = 0.023$, paired t-test). (C) Azidothymidine (AZT) blocks L1 transposon activity in the L1-EGFP reporter assay performed on LNCaP cells. (D) Western blot of L1 ORF1 protein in LNCaP cells when treated with carboplatin, azidothymidine (AZT), or both. Asterisks indicate significant differences between treatment conditions based on t-test (*, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$).