Supplementary Online Content

Abegglen LM, Caulin AF, Chan A, et al. Potential mechanisms for cancer resistance in elephants and comparative cellular response to DNA damage in humans. *JAMA*. doi:10.1001/jama.2015.13134

This supplementary material has been provided by the authors to give readers additional information about their work.

eAppendix. Supplemental Methods and Results

Methods

Analysis of Cancer Incidence of Captive Animals

Necropsy data was compiled from the collection by the San Diego Zoo[1] to estimate cancer incidence in mammalian species. The analysis was limited to species with a minimum sample size of ten, giving a total of 832 necropsies across 36 mammalian species and 37 reported cases of cancer and/or lethal tumors (see eTable 1). Tumors that did not have full pathology done, as well as cases noted as hyperplasia, were counted towards cancer incidence so as to not drastically underestimate the values. 95% confidence intervals (CI) for the cancer mortality estimates in captive zoo animals were calculated using the standard error with one pseudo-count added to both the positive and negative observation counts. Lower bounds of the confidence interval were set to zero if they were calculated to be negative after subtracting 1.96* (standard error) because cancer rates cannot be negative.

Adult body mass, maximum lifespan and mass specific basal metabolic rate (msBMR) data were collected from the AnAge database[2] and are provided in eTable 1. Maximum lifespan was used because the average lifespan of many species is unknown. A logistic regression was performed to determine if body mass, lifespan, msBMR or a combination of these variables was a good predictor of cancer incidence within a species (eFigure 1). In the regressions, the influence of each species was weighted by the number of samples for that species. Analyses with or without that weighting led to the same conclusion that Peto's paradox is real. The regression model was done on the log scale as well as linear, none of which were significant. The log-scale was plotted for easy visualization of the large range of masses. Due to the limited amount of annotated mortality data available for non-human animals, the cancer mortality rates in this analysis should be considered rough estimates.

Data on 644 captive elephants kept in zoos and sanctuaries (both African *[Loxodonta Africana]* and Asian *[Elephas maximus]*) were obtained from the online Elephant Encyclopedia database.[3] Causes of death were divided into seven categories: cancer, euthanized because of cancer, non-cancer disease, euthanized for a reason other than cancer, unspecified disease, euthanized for an unspecified reason, and exogenous cause of mortality. Exogenous causes of mortality include accidents (e.g. falling in the enclosure) and animal fights that cause fatal injury. Inferred cancer rates were calculated by assuming the same percentage of deaths with an unknown cause would be due to cancer as deaths with known causes. For example, if cancer makes up *x* percent of deaths with a known cause, then it was assumed that cancer was also responsible for *x* percent of the deaths with an unspecified cause (i.e. "disease unspecified" and "euthanized unspecified").

Specifically, the fraction of cancers reported in deaths with a specified disease was f_{dk} and the fraction of elephant euthanizations attributed to cancer is f_{ek}, where the subscript *k* represented 'known' and the *d* and *e* represented 'disease' and 'euthanized' respectively. The number of deaths from unspecified diseases that can be inferred to be cancer was equal to $f_{dk} \times N_{du}$, where N_{du} was the number of deaths caused by an unspecified disease. Similarly the number of unspecified euthanizations that were estimated due to cancer was equal to $f_{ek} \times N_{eu}$, where N_{eu} was the number of euthanizations with no specified reason. The ceiling integer for each of these values was taken as a conservative measure of the cancer incidence. The inferred cancer rate was equal to $\frac{(f_{dk} \times N_{du}) + (f_{ek} \times N_{eu}) + C_{dk} + C_{ek}}{N}$, where C_{dk} and C_{ek} were the number of cancer cases in the known disease population and the known euthanized population respectively and *N* was the total number of elephant deaths. 95% confidence intervals were calculated with the standard error (observed cancer % ± 1.96 x SE, were SE is the standard error). One pseudo-count was added to both the tumor positive and tumor negative observations for each species and the standard error was calculated as $SE = \sqrt{p(1-p)/n}$ where n is the number of necropsies (including the 2 pseudo-counts) and p is the proportion of necropsies with a tumor (including pseudo-counts).

In this study, a simplified estimation of cancer mortality rates was used due to the lack of detailed information in elephant populations. If the standard age population structure of elephants was better understood and one could obtain well documented cancer mortality rates per year (requiring a much larger sample size), then the estimates would be more comparable to those calculated for humans by GLOBOCAN and the American Cancer Society (ACS). To make as realistic as possible estimates given the available data for elephants, the inferred mortality rates were calculated to account for the unknown causes

of death as described above. Additionally, the age-adjusted lifetime risk was calculated by weighting each age-specific cancer death rate by the proportion of the population in the given age group and summing these for the cumulative risk. This is based on the methods used for human data[4,5]; however, there was not a standard population structure available to use for elephants so the distribution of ages found in this population of 644 deaths was applied. The age-adjusted lifetime risk was not statistically different from the crude calculation of lifetime risk of cancer mortality in elephants (4.81% vs 4.82%). As more data becomes available, these cancer mortality rates can be refined to enable a more direct comparison with human cancer mortality rates.

Sample Collection

Whole blood samples from an individual male African elephant supplied by the Oakland Zoo (Oakland, California) were used for DNA analyses including cloning and re-sequencing experiments. For the ionizing irradiation and doxorubicin experiments, RNA collection, and DNA whole genome sequencing, elephant whole blood samples were obtained from two female African elephants at Utah's Hogle Zoo (Salt Lake City, Utah). Processing of blood for live cell experiments began within 1 hour of the blood being drawn. Asian elephant blood samples were provided by the Ringling Bros. Center for Elephant Conservation (Polk City, Florida) and immediately flown to Utah for processing to begin live cell experiments within 10 hours after the blood draw. Each organization's respective ethical and scientific review boards approved all the experiments that included collection and use of elephant blood. Human whole blood samples were obtained from IRB-consented patients with Li-Fraumeni Syndrome (LFS) and healthy controls enrolled under the Institutional Review Board (IRB)-approved Cancer Genetics Study (IRB# 00041211) at the University of Utah.

The Cancer Genetics Study (CGS) is a comprehensive registry at Huntsman Cancer Institute (HCI) at the University of Utah that collects medical, family history, and genetic testing data as well as biological samples (peripheral blood samples, tumor samples, cell lines, etc.). Participation in the registry is offered to all patients seen for clinical genetics evaluation through HCI's Family Cancer Assessment Clinic. The CGS includes individuals with a personal history of cancer as well as at-risk relatives and unaffected individuals with a known molecular predisposition to cancer. Unaffected spouses and family members who test negative for known familial cancer predisposition mutations are included as controls. Unrelated controls also can be recruited into the CGS if they are healthy individuals without a personal or family history of cancer. All CGS participants are contacted annually to update information on cancers diagnoses, screening and changes in family history. Currently, the CGS has 2992 living participants.

DNA Isolation

The DNA was purified with the DNeasy® Blood & Tissue Kit (QIAGEN) and concentrated by precipitating with 1/10 volume 3M sodium acetate (pH 5.2) and one volume of isopropanol followed by a 75% ethanol wash. The pellet was re-suspended in an appropriate volume of 1X TE buffer, based on the concentration determined by nanodrop, and stored at -20°C.

Cloning and Capillary Sequencing of TP53 Processed Paralogs

Primers were designed to amplify the *TP53* processed copies simultaneously in the African elephant to be used in downstream cloning. 15 of the 19 published processed copies have sequences that perfectly match the primers and the others only differ by a couple of bases. All of the sequenced clones could be captured with the one primer set (Forward 5'-GTCAGGTCACCTAGTTTCTGAATTG-3', Reverse 5'-GTCAATCCATCAACCAACAGG-3'). A second reverse primer (5'-GTCAATCCATCAAAAAACAGG-3') was used with the same forward primer to match other copies more specifically, but the same loci were picked up with each set. Each 50 uL PCR reaction contained 1X final concentration of PCR Buffer (Invitrogen), a 200 uM concentration of each dNTP, 1.5 mM concentration of MgCl2, 0.2 uM of each primer (IDT), 1 unit of Platinum® *Taq* DNA Polymerase (Invitrogen), and ~50 ng of template DNA. The following PCR program was used: 94°C for 2 min; 40 cycles of 94°C for 30 sec, 55°C for 30 sec, and 72°C for 2 min; and ending with 72°C for 5 min. Product sizes were verified by running the samples on a 1.5% agarose gel and visualizing bands with UV light.

The TOPO® TA Cloning® Kit for Sequencing (Invitrogen) was used to clone the PCR products into the PCR®4-TOPO® vector and transform chemically competent TOP10 E. coli cells. The kit protocol was followed for 3 cloning and transformation reactions: 2 replicates for the PCR products designed to capture the multiple copies of TP53 and one control transformation using the pUC19 plasmid. Each

reaction was plated onto LB 100ug ampicillin agar at 20 µL, 40 µL and 100 µL. Competent cells transformed with pUC19 were used as a positive transformation control while untransformed cells plated on ampicillin were used as a negative control.

To PCR the cloned product and verify the insert, 25 µL reactions were prepared so that each reaction contained 2.5 µL 10X PCR Buffer with MgCl2, 0.5uL dNTPs, 0.25 µL T3 primer, 0.25 µL T7 primer, 21.25 µL water and 0.25uL Taq. A single isolated colony was scraped from a plate with a pipette tip and placed into a well containing the 25 µL of PCR master mix. Two 96 well plates of colonies were prepared for a total of 192 reactions. The PCR program was run as follows: 94°C for 5 min; 30 cycles of 94°C for 30 sec, 55°C for 30 sec, and 72°C for 2.5 min followed by a 10 minute final extension at 72°C.

PCR reactions were diluted by adding 20uL water to 5uL of the PCR product and cleaned by adding 1uL SAP (1u/ µL), 1 uL Exonuclease I (10u/ µL) and 2uL SAP reaction buffer (USB® Products, Affymetrix, Inc). The samples (15pmol primer and ~150ng purified PCR product in a total volume of 6 µL) were submitted to the Genome Core Facility at the University of California San Francisco for sequencing using ABI BigDye v3.1 dye terminator sequencing chemistry and the new generation ABI PRISM 3730xl capillary DNA analyzer.

Asian elephant clones followed the same protocol for sequencing as described above, with the exception of the ExoI/SAP cleanup before sequencing. Instead the plasmids were isolated from the clones with the Qiagen Plasmid Prep kit and were then sequenced at the University of Utah HSC Sequencing Core Facility with the Applied Biosystems 3730xl capillary sequencer.

Four sequencing primers were used for each clone to fully cover the fragment with overlap: the T3 promoter 5'-AATTAACCCTCACTAAAGGG-3', the T7 promoter 5'-TAATACGACTCACTATAGGG-3', 5'-CCTGAGAAGCTGGTTCTGTCC-3', and 5'-CCAGACGTCAGCATATGATGGA-3'. Sequences traces were examined trimmed and assembled using CLC Genomics Workbench.

The presence of the ancestral *TP53* genomic sequence was verified by a PCR reaction using forward primer 5'-AGGAAGTCGGGTGGGGAGCC-3', reverse primer 5'-GGCAGGGTGGGGACAGCAAC-3' and the same master mix described previously for the initial retrogene PCR. The following PCR program was used: 94°C for 2 min; 35 cycles of 94°C for 30 sec, 62°C for 30 sec, and 72°C for 1 min; and ending with 72°C for 5 min. The PCR reactions were run on a 1.5% agarose gel and visualized by UV light. The bands were extracted and purified using the PureLink™ Quick Gel Extraction Kit (Invitrogen). This was followed by capillary sequencing as described above.

Sequence analysis

The African elephant genome assembly *LoxAfr3* was used for sequence analyses. *TP53* gene locations and sequences from the African elephant genome were obtained from the Ensembl database (release 72 www.ensembl.org) and NCBI (www.ncbi.nlm.nih.gov/gene) (eTable2). From these two databases, a total of 20 genomic positions were found to be homologous to *TP53*. The UCSC Genome Browser was used to view these 20 regions and verify that *TP53* transcripts of other species mapped to each of these locations. The start and stop of these locations was manually re-annotated based on these alignments to update some annotations that had truncated parts of the region homologous to the ancestral *TP53* coding sequence.

Multiple alignments of the resequenced retrogenes were created with MUSCLE (Version 3.8)[6] and hand edited using Seaview (Version 4.5.2). PhyML (Version 3.0) was used to create the maximum likelihood phylogeny for which gap sites were ignored. Default settings for PhyML using DNA input were used. The alignments of the clustered clones were used to get a consensus sequence of the 18 retrogenes with Seaview. The percent identities of the pairwise alignments of each processed retrogene copy to the coding sequence (CDS) of the ancestral *TP53* copy range from 85-88%. The NCBI sequence (gene ID:100663725; accession ID: **XM_003416902.2**) was used over the Ensembl prediction for the ancestral copy (**ENSLAFG00000007483**) because it showed higher homology to other mammalian p53 amino acid sequences and had the expected exon structures. Within the processed copies, the percent identity shared ranged from 86-99%, based on all pairwise alignments. The copies clustered into two groups of closely related paralogs: 6 in Group A and 12 in Group B) (Figure 2). Between groups A and B, the average percent identity was approximately 88% and within each group the percent identity was greater than 95% on average. The African elephant *TP53* retrogene consensus sequences from this study are available through GenBank (KF7185855-KF715872).

DNA was extracted from whole blood of a female African elephant (Hogle Zoo) using the DNeasy Blood and Tissue Kit (Qiagen). DNA was treated with RNase A (Thermo Scientific) as described in

DNeasy Blood and Tissue kit protocol. DNA was eluted in 10mM Tris, pH8.0. Sequencing libraries were made using the Illumina TruSeq DNA PCR-free sample prep kit with a 200bp mean insert size. Sequencing was done using HiSeq 125 cycle paired-end sequencing with version 4 chemistry to an average sequencing depth of 30x. Short reads were aligned to the elephant reference genome loxAfr3[7] using BWA mem.[8] Variants were called using the GATK pipeline with Unified Genotyper.[9,10]

Cell Culture

Peripheral blood mononuclear cells (PBMCs) were isolated by Ficoll-Paque density-gradient followed by three washes in large volumes of PBS to remove platelets. Where indicated, cells were washed with PBS supplemented with 2% filtered autologous plasma. Cells were cultured in RPMI 1640 (Life Technologies) supplemented with 20% filtered (0.22 μm filter) autologous plasma, penicillin/streptomycin, and L-glutamine. To enrich for lymphocytes, PBMCs were incubated at 37°C for 30 minutes in tissue culture treated dishes to allow monocytes to adhere. Suspension cells (lymphocytes) were harvested and counted. Cells were resuspended at $2x10^6$ cells/ml and either left untreated, exposed ionizing irradiation (IR) in a RS-2000 X-ray Biological Research Irradiator, or cultured continuously with the indicated concentrations of doxorubicin (5μM or 15μM). Cells were plated in 96 well plates for flow cytometry assays (100μl/well) and 6 well plates for protein and RNA extraction (2ml/well). Cells were then incubated at 37°C for the indicated amounts of time. Treated (ionizing irradiation or doxorubicin) and untreated (NT) lymphocytes were stained with annexin V and propidium iodide (PI) and analyzed by flow cytometry at each time point (5hr, 10hr, 18hr and 24hr).

Flow Cytometry

Untreated, irradiated, or doxorubicin treated peripheral blood lymphocytes were seeded into 96 well plates at $2x10^5$ cells/well. After incubating for the indicated time points, cells were washed with room temperature PBS or PBS+2% autologous plasma and stained with APC Annexin V and propidium iodide (PI) (Apoptosis Detection Kit II, BD Biosciences). Fluorescence was analyzed using a BD FACSCanto II (BD Biosciences). Cell gates were set using forward- and side-scatter profiles to exclude small debri from analysis. Data was collected for 10,000 events/sample and analyzed using FlowJo software (Tree Star, Inc.). Flow cytometery experiments were repeated two (Asian elephant) to three (African elephant) times and each experiment was performed with three to four biological replicates. Data was graphed and 95% confidence intervals were calculated using GraphPad Prism (GraphPad Software, Inc) software for Figure 3, Figure 7, eFigure 6, and eFigure 8. Significance ($P < 0.05$) was determined using a two-sided *t*-test with no assumption of consistent variance at different time points. For Figure 4, to measure statistical significance, R was used to calculate exact p values. First, *F* tests were run for each of the three comparisons (LFS to healthy human, LFS to elephant, and healthy human to elephant) with $\alpha=0.05$ to determine equal versus unequal variance. *P* values for the F tests reveal no significant differences in variance (LFS to healthy human $P=0.16$, LFS to elephant $P=0.41$, and healthy human to elephant $P=0.97$). Two-sided *t*-tests were then run assuming equal variances for each of the three comparisons. *P* values for the equal variance, two-sided *t*-test are listed with results.

Western Blots

Cells were lysed in Cell Lysis Buffer (Cell Signaling Technology) supplemented with Protease/Phosphatase Inhibitor Cocktail (Cell Signaling Technology). 25-75 μg of total protein lysate was run on 12% Bolt Bis-Tris Plus Gels (Life Technologies) and transferred to PVDF membranes. The following antibodies were used for immunodetection: Monoclonal anti-GAPDH (Clone GAPDH-71.1, Sigma Aldrich), Purified Mouse anti-p21 (556430, BD Biosciences), and anti-TP53 antibody (AAS55020E, Antibody Verify), anti-p53 (clone DO-1, Santa Cruz), anti-phospho-p53(Ser15) (9284, Cell Signaling), anti-MDM2 (clone 2A10, ab16895, Abcam, Cambridge, MA), anti-c-Myc (clone A-14, Santa Cruz). The p53 antibody used for the lymphocyte western blot in Figure 6, Figure 7, and eFigure 6 was the Antibody Verify antibody, which was raised using an immunogen with 87% homology between human and elephant p53. The epitope contains phosphorylation sites, which prevent the antibody from recognizing activated p53. The human epitope recognized by the p21 antibody is an exact match for elephant p21. Secondary antibodies were HRP-linked mouse IgG and HRP-linked rabbit IgG (GE Healthcare Life Sciences). Substrates used for detection were Western Lightning Ultra (PerkinElmer) and SuperSignal West Dura Chemiluminescent Substrate. Signal was detected with a Bio-Rad ChemiDoc Gel Imager or developed on Amersham Hyperfilm ECL (GE Healthcare Life Sciences).

For the p21 experiments, PBLs and fibroblasts were exposed to IR (0.5, 2 and 6 GY) and cultured for either 5 hours (PBL) or 2 hours (fibroblasts). Protein expression of p21 from western blots was quantified using ImageJ's gel analysis feature. Images were converted to 8-bit, and then bands were selected. Area was calculated for p21 and the loading control (GAPDH). Relative density for each was calculated by dividing the percent of the treated band by the untreated band for each individual. Normalized relative density was then calculated by dividing the relative density of p21 bands to the relative density of the GAPDH bands. Data from two independent PBL experiments, one performed in triplicate and the other performed in quadruplicate, were combined. Data was graphed and 95% confidence intervals were calculated using GraphPad Prism (GraphPad Software, Inc) software and significance (α=0.05) was determined using two-sided, unpaired *t*-tests with no assumption of consistent variance.

TP53 Retrogene Transcript Expression

RNA was collected from the treated (2Gy) and untreated elephant cells after 5 hours. The samples were stabilized with RNAlater and purified with the Qiagen RNeasey Kit with on column DNase I treatment. Samples were reverse transcribed (RT) with the TaqMan® Reverse Transcription kit (Applied Biosystems) using poly-T and random hexamer oligos in separate reactions. The poly-T samples were used in downstream analyses.

Primers to distinguish the ancestral transcript from possible processed *TP53* transcripts were designed using Primer3 and their specificity to these regions was verified using BLAST against all predicted transcripts in the African elephant genome. The ancestral p53 cDNA was amplified using forward and reverse primers (5'-CCTCCTGGACCCTGTCATCTT-3' and 5'-AAGCCCAGACGGAAACCATA-3') and the processed p53 cDNA copies were amplified with forward and reverse primers (5'-CCTGAGAAGCTGGTTCTGTCC-3' and 5'-GCAGTAGGTCTTCTGGGAAGG-3'). The primers allow for degenerate binding and do not match all of the retrogene *TP53* copies exactly, but Group A and Group B of the retrogenes would be predicted to bind and form products of 201 and 185 base pairs long, respectively. If the primers bound the ancestral *TP53* copy, though this would not be expected due to lower sequence identity, the product would be 214bp. The forward primer is located in the homologous region spanning exon two and three and the reverse primer is found in exon three thereby eliminating the possibility of the multiple products being from different isoforms of the ancestral *TP53* transcript.

Each 25 μL PCR reaction contained a final concentration of 1X PCR Buffer (Invitrogen), a 200 μM of each dNTP, 1.5 mM of MgCl2, 0.2 μM of each primer (Operon), 1 unit of Platinum® *Taq* DNA Polymerase (Invitrogen), and 1uL template cDNA directly from the RT reaction. The following PCR program was used: 94°C for 2 min; 35 cycles of 94°C for 30 sec, 55°C for 30 sec, and 72°C for 1 min; and ending with 72°C for 5 min. Product sizes were verified by running the samples on a 1.5% agarose gel and visualizing bands with UV light. Samples were also run on a 10% polyacrylamide gel to get better separation of the bands. Fragments at 201bp and 185bp were excised from the gel (eFigures 3 and 4), purified with the Qiax® II Gel Extraction Kit (Qiagen) and sequenced with capillary sequencing at the UCSF Genome Core Facility.

DNA Repair Assay

PBLs were fixed with 4% PFA and 200-500x10^5 cells were cytospun onto slides. Immunostaining was performed using the primary monoclonal anti-phospho-Histone H2A.X (Ser139) antibody (Millipore Corp)[11] and the secondary Alexa Fluor® 488 goat anti-mouse IgG (Invitrogen). Slides were coated with mounting medium containing Dapi (Vectashield; Vector Laboratories, Burlingame, CA) and cover slips applied. Cells were visualized using the Axioscope 2 Plus with Axiovision software (Zeiss; Carl Zeiss Microscopy, LLC, Thornwood, NY). Foci from 100 cells were counted from each slide and allocated into one of four categories (0-5, 6-10, 11-15, 16-20+). A total of three slides were counted from two separate experiments, and the average counts are presented in Table 2 of the main text.

Dapi staining for apoptosis

Human and African elephant PBLs (washed in PBS+2% autologous serum after ficoll gradient) were Dapi stained for a visual representation of apoptotic cells. 18 hours after culture without treatment (NT) or after 0.5GY IR, cells were fixed and nuclei were stained as described. 100,000 cells were cytospun onto slides. Cells were permeablized in 0.5% triton in PBS for 1 minute, washed three times in PBS for 1 minute each, and blocked in 5% BSA in PBS for 30 minutes. Cell nuclei were stained with Vectashield

mounting medium containing Dapi (Vector Laboratories). Cells were visualized and images were captured using an EVOS FL Imaging System (Life Technologies).

qPCR of* TP53 *Retrogenes

PBLs and fibroblasts were processed and treated as described. At each time point, cells were harvested, centrifuged to remove media, and cell pellets were frozen at -80C. RNA was extracted using PureLink RNA mini kit (Life Technologies) and Qiashredder (Qiagen) for homogenization. RNA was treated with Turbo DNase (Life Technologies) to remove genomic DNA. RNA concentrations were measured with Ribogreen (Life Technologies). To measure African elephant TP53 retrogene expression in PBLs, 300ng total RNA from each sample was used for cDNA synthesis using iScript cDNA kit (Bio-Rad). Each sample was ran including a no reverse transcriptase control to rule out genomic DNA contamination. No template controls were also included to rule out contamination. 6µl of cDNA from the 20µl iScript reaction was used to pre-amplify elephant TP53 retrogenes using TaqMan PreAmplification kit (Life Technologies) with 10 cycles (total reaction volume for each sample was 25ul). qPCR was carried out using TaqMan gene expression kit (Life Technologies). 5µl from the preamplification reaction was added to each reaction (final volume 25µl) along with the indicated TaqMan Gene expression primer/probe sets.

To measure *TP53* expression in human and elephant PBLs, one step qPCR reactions were performed using Qiagen's QuantiTect Probe RT-PCR Kit. Each sample was run in a 25µl reaction with 40ng RNA. Elephant *TP53* expression was also measured from cDNA after preamplification of retrogenes using the method described above. Both methods yielded the same results. PBL qPCR reactions were run on a Bio-Rad CFX96 Real-Time PCR detection system. GAPDH was used as the reference gene, and fold change was calculated using the delta delta Cq method. Each experiment was performed at least three times with three biological replicates.

To measure *TP53* retrogene expression in African elephant fibroblasts, 50ng of total RNA was used for one-step qPCR with Express One-Step SuperScript qRT-PCR SuperMix (Life Technologies). GAPDH was used as the reference gene. Reactions were run on a StepOnePlus Real Time PCR system (Life Technologies).

Human TaqMan Gene Expression assays included – *TP53* (Hs01034249_m1, FAM labeled) and GAPDH Endogenous Control (VIC/TAMRA probe, catalog number 431088E) from Life Technologies.

Elephant primer/probe sets were custom synthesized TaqMan Gene Expression assays from Life Technologies. Primer and probe sequences - African elephant ancestral TP53 – forward and reverse primers (5'-TGGGAACTCCTTCCTGAGAATC-3', 5'-TTCTGAGAGTAGCAGATCGTCCAT-3') and FAM labeled probe (5'-TCCCCCACACTACCCCCGGC-3') - African elephant GAPDH – forward and reverse primers (5'-TCAGATCCACCACTGACACGTT-3', 5'-CCTGAGCTGAATGGGAAGCT-3') and VIC labeled probe (5'-ACTGGCATGGCCTTCCGTGTCC-3') – African elephant group A TP53 retrogenes – forward and reverse primers (5'-CATGTGCTGACATCTGGTAGATGG-3', 5'-CTGTCCCTTCCCAGAAGACC-3') and FAM labeled probe (5'-ACTGCAGCAACTGTGGTTTCCATCTTGGC-3') – African elephant group B TP53 retrogenes – forward and reverse primers (5'-CATGTGCTGACATCTGGTAGATGG-3', 5'-CTCTGTCCCTTCTCAGAAGACC-3') and FAM labeled probe (5'-CAGCACCTATCATTTCTGTCTGGGCTTCTTGC-3').

Elephant primer/probe sets were manually designed based on regions of heterogeneity between ancestral and processed p53 transcripts. To verify specificity of primer/probes, the amplified product was run on a 1% agarose gel. Bands within the predicted template size were excised from the gel and purified with a Qiagen Gel Purification Kit. cDNA was quantified using NanoDrop. 40ng PCR product and 5pmol primer were submitted for Sangar sequencing. An additional confirmation of specificity was carried out using BioRad's SYBR Green and primers only. The resulting melt curve was analyzed for non-specific binding.

Specificity of the TP53 retrogene primer probe sets was confirmed by testing their ability to amplify a product from each of the 18 cloned retrogene cDNAs in separate reactions. The retrogene group A TaqMan assay only amplifies retrogene group A clones (1-6), and the retrogene group B TaqMan assay only amplifies retrogene group B clones (7-18). Significance ($P < 0.05$) of expression changes from the NT samples was determined using a two-sided *t*-test with no assumption of consistent variance at different time points. GraphPad Prism (GraphPad Software, Inc) was used to plot the fold changes in expression and 95% confidence intervals.

Fibroblast Cells

Primary Normal Human Dermal Fibroblasts (Lonza), Primary African Elephant Dermal Fibroblasts (kindly provided by George Church) and HEK293 (human embryonic kidney cells, kindly provided by Michael Deininger) were cultured in DMEM supplemented with glutamine, sodium pyruvate, HEPES, and 10% fetal bovine serum. One-step qPCR was performed using TaqMan gene expression assays for Group A *TP53* retrogenes, Group B *TP53* retrogenes, and GAPDH as described above. African elephant and human fibroblasts were treated with various doses of ionizing irradiation and doxorubicin. ApoTox-Glo (Promega) reagent was used to measure cell viability, cytotoxicity and apoptosis (caspase 3/7 cleavage). Elephant fibroblasts and human fibroblasts were cultured for 72 hours after exposure to the indicated doses of IR and for 18 hours with the indicated doses of doxorubicin.

Transfection with Retro9

HEK293 were transfected with either pcDNA3.1+ empty vector or pcDNA3.1+N-Myc_retro9. Elephant retrogene 9 (GenBank ID: KF715863) was synthesized and cloned into the pcDNA3.1+N-Myc plasmid by GenScript (GenScript USA Inc.). The construct contained the entire coding sequence of retrogene 9 that aligns with the coding sequence for ancTP53 (cDNA sequence: 100663725). Codon optimization for expression in human cells was performed by GenScript, and the optimized sequence was cloned into the pcDNA3.1+N-Myc vector by GenScript.

The plasmids were transfected into HEK293 cells using Mirus TransIT-LT1 transfection reagent (Mirus Bio LLC). T25 flasks were seeded with $0.75x10^6$ cells and incubated for 6 hours prior to transfection. Transfection mixture included 5µg of plasmid DNA, 30µl of transfection reagent, and 150µl of Opti-MEM. 24 hours post-transfection, media containing transfection mixture was replaced with fresh media containing 500µg/ml G418. Selection media was replaced daily for 7 days until all of the untransfected control cells died. The cells remained under selection in media containing 500µg/ml G418 throughout culture and during all assays.

The Myc-tagged (amino terminal) Retrogene 9 and empty vector transfected cells were treated with either proteasome inhibitor (10µM MG132) or the indicated doses of doxorubicin (0.1, 1, and 10µM) and IR (2, 5, and 10GY). After 5 hours in culture, cells were harvested, lysed, and expression of retrogene protein was measured by immunoblotting with antibody to Myc. GAPDH expression was measured to assess equal loading. To determine if retrogene 9 interacted with Mdm2 in HEK293 cells, co-immunoprecipations of lysates from transfected HEK293 cells with Myc antibody (clone 9E10, kindly provided by Michael Engel) were performed. Cells were either not treated (NT) or treated with 6 GY IR and incubated for 2 hours, followed by immunoprecipitation (see below). The co-immunoprecipitated proteins were loaded equally on duplicate protein gels and transferred to membranes for western blots. One membrane was probed with a MYC antibody to detect retrogene 9. The other membrane was probed with an antibody to phospho-p53 (Ser15) (9284, Cell Signaling) to determine if retrogene 9 is phosphorylated at this residue. This blot was then striped and reprobed with antibody to MDM2 (clone 2A10, ab16895, Abcam, Cambridge, MA). Whole cells lysates (pre-immunoprecipitation) were probed for GAPDH to determine if equal amounts of protein were used for co-immunoprecipitation.

ApoTox-Glo Assay

Primary human and elephant fibroblasts were harvested, counted and treated with ionizing irradiation prior to being seeded into black 96 well clear bottom plates at a density of 2000 cells/well. Cells were incubated at 37C for the indicated amount of time. ApoTox-Glo (Promega) reagent was used to measure viability, cytotoxicity, and apoptosis (caspase activation). Fluorescence and luminescence were measured using a Wallac Envision (PerkinElmer) plate reader. After background subtraction, data was normalized to no treatment controls for irradiation and DMSO controls for doxorubicin treated samples. Assays measuring response to IR were run twice, including 4 replicates, and assays measuring response to doxorubicin were run twice with 3 replicates. Data was graphed and 95% confidence intervals were calculated using GraphPad Prism (GraphPad Software, Inc) software and significance (α=0.05) was determined using two-sided, unpaired *t*-tests with no assumption of consistent variance.

Immunoprecipitation

HEK293 Retro 9 transfected cells were harvested. $1x10^6$ cells were either exposed to 6 Gray ionizing irraditation or left untreated. The cells were then transferred into a 24 well plate and incubated at 37C and 5% CO2 for 2 hours. After incubation, the plate was placed on ice, and cells were washed twice

with 2.5 mL of ice cold 1X PBS. Each wash was aspirated and cells were scraped into 1 mL of lysis buffer. The lysates were then transferred into pre-chilled microcentrifuge tubes. Each tube was sonicated (on ice) and samples were rotated at 4C for 30 min. Debris was pelleted, and a small portion of the supernatant was reserved as a pre-IP control. The remaining lysate was transferred to pre-chilled tubes. Supernatant was pre-cleared by adding protein-G sepharose slurry and incubating on a nutator for 1 hour at 4C. Beads were pelleted and supernatants transferred to fresh, pre-chilled tubes. Three μg of anti-MYC antibody (clone 9E10, kindly provided by Michael Engel) was added to each tube and incubated at 4C for 90 min. Protein-G sepharose slurry was added to each tube and incubated at 4C for 60 min with rotation. Beads were pelleted and the supernatant aspirated and discarded. Beads were washed 4 times with lysis buffer, and after the final wash the bead pellet was resuspended in protein loading buffer and incubated at 95C for 10 min. Samples were stored at -20C until loaded on protein gels for western blot.

Lymphocyte wash conditions

To determine if wash conditions affect the viability of lymphocytes, after ficoll gradient (in the absence of culture), African elephant PBLs and human PBLs were plated in 96 well plates, pelleted, and washed in either PBS or PBS+2% autologous plasma. To determine if an increase in annexinV+/PI+ cells could be eliminated by washing in PBS+2% autologous plasma, all wash steps for lymphocyte processing used this buffer. Cells were treated with IR (0.5 and 2 GY), cultured and stained with annexinV and PI as described above. Data was graphed and 95% confidence intervals were calculated using GraphPad Prism (GraphPad Software, Inc) software and significance (α=0.05) was determined using two-sided, unpaired *t*-tests with no assumption of consistent variance.

Results

Tumor incidence in zoo animals

Logistic regression models for tumor incidence in zoo animals did not show significant correlation between tumor incidence and body size, lifespan, mass specific basal metabolic rate (msBMR) or any combination of those variables (eFigure 1). Even when the upper limit of the wide confidence intervals were used to avoid underestimating cancer incidence, there was still no significant correlation between body mass, lifespan and frequency of cancer deaths across species. All examples showed a decrease in actual cancer incidence to what would be expected without the evolution of cancer protective mechanisms, but this was not significant. The possibility of a minor correlation between body size, lifespan and cancer cannot be ruled out completely due to the lack of data and small sample sizes. As highlighted in the main text, increasing the sample size for this necropsy data with more data points will improve the accuracy of future results of cross-species cancer estimates.

Asian elephant TP53 retrogenes

The Asian elephant genome also contained 15-20 *TP53* retrogenes (eFigure 2). The capillary sequenced clones from the Asian elephant, African elephant, and published sequences from GenBank were aligned and clustered using a maximum likelihood phylogeny. There was evidence for retrogenes from both Group A and Group B seen originally in the African elephant genome.

***TP53* retrogene transcription**

RNA isolated from irradiated and non-irradiated elephant PBMCs showed evidence for retrogene transcription (eFigure 3). PCR band products specific for the *TP53* retrogenes were seen in all time points and in both treated and untreated samples for the elephant cells, confirming the transcription of the *TP53* retrogenes. Capillary sequencing of excised gel bands from PCR of this elephant cDNA confirmed transcription (eFigure 4). Retrogene sequences can been found in GenBank (KF715855-KF715872).

Elephant fibroblasts and TP53 retrogenes

African elephant fibroblasts express *TP53* retrogenes and respond more to DNA damage than human fibroblasts (eFigure 11). Specifically, the Group B TP53 retrogenes were found to be expressed in African elephant fibroblasts, and the expression level did not change 4 hours after 2GY IR. Expression of group A retrogenes was not detected in this experiment, but the possibility of low level expression was not ruled out. A significant decrease in viability after IR was observed for elephant fibroblasts compared to human fibroblasts. No increase in cytotoxicity or caspase activity was observed for either cell type,

indicating that these cells were undergoing cell cycle arrest. Treatment of these cells with doxorubicin led to a greater increase in caspase activity in elephant cells compared to human. In addition, elephant fibroblast cytotoxicity increased and viability decreased, consistent with apoptosis. Human fibroblasts showed a pattern more consistent with cell cycle arrest, including loss of viability with little increase in caspase and no increase in cytoxicity.

TP53 retrogene translation

The elephant p53 retrogene 9 was expressed as a protein in transfected HEK293 and responded to DNA damage similar to p53 in human fibroblasts (eFigure 5). The Myc-tagged protein expressed in HEK293 cells ran around 28 KDa, consistent with retrogene 9 expressed from the start sequence before Myc through to the first in frame stop codon at residue 181 of this retrogene (predicted size of 20.88 KDa). Expression of retrogene 9 in HEK293 cells increased with proteasome inhibition and doxorubicin treatment, as well as ionizing irradiation similar to the increase in activated p53 (clone DO-1) in human fibroblasts treated with IR. A band of the same size as the band detected with the Myc antibody was detected (around 28 KDa) with the phospho-p53 (Ser15) antibody, indicating that retrogene 9 is phosphorylated at serine 15. In addition, a 90 KDa band was detected with the MDM2 antibody, indicating that retrogene 9 binds to MDM2. This data indicates that elephant retrogene 9 expression is regulated post-translationally, similar to p53. Unfortunately, HEK293 cells are inappropriate for further characterization of the functional consequence of retrogene 9 expression because these cells are transformed with adenovirus proteins that inhibit the function of p53.[12,13]

Doxorubicin response

African elephant lymphocytes showed increased sensitivity to treatment with doxorubicin (5 μM and 15 μM) compared to human cells (eFigure 6, 5 μM: 24.77%, 95% CI [23.0%-26.53%] vs. 8.1%, 95% CI [6.55%-9.66%] late apoptotic after 18h, $P<0.001$ and 15 μM: 38.7%, 95% CI [28.97%-48.43%] vs. 16.73%, 95% CI [15.16%-18.31%] $P<0.001$). The western blot showed induction of p21 at 5 hours in African elephant and human lymphocytes with doxorubicin treatment indicating the expected p53 pathway activation. The p53 antibody used on the elephant cells cannot detect activated p53, only non-phosphorylated p53 was measured.

Dapi staining for apoptosis.

Few apoptotic cells were seen in the no treatment condition for both elephant and human cells. The 0.5GY exposure visually displayed more apoptotic elephant cells compared to human cells (eFigure 7).

Lymphocyte wash conditions

See eFigure 8. A large increase in annexinV+/PI+ cells was observed for both human (15.65%, 95% CI [11.21%-20.09%]) and elephant PBLs (72.13%, 95% CI [50.79%-93.46%]) washed in PBS, but the elephant PBLs were significantly more sensitive to the wash conditions ($P<0.001$). The addition of protein (autologous plasma) to the wash buffer protected the cells from death. Since the cells were viable prior to being washed (as measured by staining with acridine orange/PI and counting on a Nexcelom Cellometer Auto 2000), this rapid increase in cell death likely represented necrosis rather than apoptosis. Elephant cells washed in the presence of protein showed a significant increase in annexinV+/PI+ even in the no treatment (NT) group compared to human cells. A greater increase in p21 expression in the untreated elephant PBLs compared to human PBLs cultured for 5 hours (see eFigure 10) indicated that this increase in apoptosis of cultured elephant PBLs was potentially mediated by p53.

P21 protein expression

The expression of p21 protein was significantly increased in elephant cells compared to human cells as measured by immunoblot and ImageJ analysis (eFigure 10). More p21 protein expression was observed in elephant PBLs (20.13 fold increase, 95% CI [8.72-31.5] relative to 0h NT with 0.5GY IR) and fibroblasts (1.9 fold increase relative to 0h NT with 2GY IR) at 5 and 2 hours, respectively, following IR exposure compared to human cells (3.5 fold increase, 95% CI [1.7-5.31] for PBLs with 0.5GY IR and no increase for fibroblasts with 2GY IR) when normalized to a loading control ($P=0.0042$ for PBL).

eTable 1. Tumor Incidence, Mass, Lifespan, and Metabolic Rate of Zoo Mammals

Common Name	# Necropsies	# Tumors	Adult mass (Kg)	Maximum lifespan (yrs)	msBMR (W/g)	% Tumors	Lower limit 95% CI	Upper limit 95% CI
Striped Grass Mouse	13	0	0.05	4.5	0.00452	0.00%	0.00%	12.62%
Philippine Tarsier	13	0	0.12	16	0.00381	0.00%	0.00%	12.62%
Pygmy Marmoset	15	0	0.12	18.6	0.00541	0.00%	0.00%	11.19%
Treeshrew	13	1	0.2	12.4	0.00424	7.69%	0.00%	24.90%
Marmoset	18	3	0.26	16.5	0.00446	16.67%	0.00%	34.20%
Squirrel Monkey	17	0	0.93	30.2	0.00529	0.00%	0.00%	10.04%
Prairie Dog	16	3	1.13	11	0.00212	18.75%	0.00%	37.96%
Fennec Fox	10	0	1.25	16.3	0.00232	0.00%	0.00%	15.64%
Viriginia Opossum	11	0	3	6.6	0.00186	0.00%	0.00%	14.49%
Rock Hyrax	76	1	3.6	14.8	0.00202	1.32%	0.00%	4.82%
Parma Wallaby	34	1	4.25	15.9	unknown	2.94%	0.00%	10.42%
Armadillo	74	2	5.5	22.3	0.00136	2.70%	0.00%	7.08%
Racoon	15	0	6	21	0.00215	0.00%	0.00%	11.19%
Tasmanian devil	18	9	6.5	13	0.00141	50.00%	28.09 %	71.91%
Darma Wallaby	14	0	6.5	15.1	0.00162	0.00%	0.00%	11.86%
Tree Kangaroo	18	0	7.2	26.9	0.00114	0.00%	0.00%	9.55%
Blue Monkey	27	2	9	37.8	0.00223	7.41%	0.00%	18.49%
Koala	26	1	9.3	22.1	0.00121	3.85%	0.00%	13.39%
Black-backed Jackal	18	1	10.25	16.7	0.00279	5.56%	0.00%	18.70%
Hamadryas Baboon	18	0	18	37.5	0.00166	0.00%	0.00%	9.55%
Collared peccary	16	0	20.2	31.5	0.00162	0.00%	0.00%	10.58%
African Wild Dog	40	3	26.5	17	0.00377	7.50%	0.00%	16.38%
Eastern Wallaroo	40	0	30	22	0.00111	0.00%	0.00%	4.61%
Pronghorn	25	1	46.1	15.5	0.00147	4.00%	0.00%	13.88%
Cheetah	13	3	53.5	20.5	0.00161	23.08%	0.70%	45.46%
Red Kangaroo	15	0	55	25	0.0011	0.00%	0.00%	11.19%

eTable 1. Tumor Incidence, Mass, Lifespan, and Metabolic Rate of Zoo Mammals (continued)

Common Name	# Necropsies	# Tumors	Adult mass (Kg)	Maximum lifespan (yrs)	msBMR (W/g)	% Tumors	Lower limit 95% CI	Upper limit 95% CI
Capybara	13	0	55	15.1	0.00139	0.00%	0.00%	12.62%
Cougar	11	0	63	23.8	0.00133	0.00%	0.00%	14.49%
Reindeer	16	0	101.25	21.7	0.00141	0.00%	0.00%	10.58%
Harbor Seal	35	2	115	47.6	0.00267	5.71%	0.00%	14.51%
Tiger	17	2	119.7	26.3	0.00097	11.76%	0.00%	28.16%
Llama	18	0	140	28.9	0.0013	0.00%	0.00%	9.55%
Blue Wildebeest	25	0	164.5	24.3	0.00117	0.00%	0.00%	7.12%
Donkey	19	1	165	47	0.00093	5.26%	0.00%	17.82%
Lion	52	1	175	27	0.00097	1.92%	0.00%	6.96%
Moose	13	0	386	22	0.00088	0.00%	0.00%	12.62%
Elephant	644	20	4800	65	unknown	3.11%	1.74%	4.47%

Necropsy data was collected by the San Diego Zoo over 14 years[1]. Body mass, maximum lifespan and mass specific basal metabolic rate (msBMR) were all obtained from the AnAge database.[2] 95% confidence intervals were calculated using the standard error (see Methods).

eTable 2. Genomic Locations of 20 *TP53* Gene Copies in the Published African Elephant Genome and the Local Whole Genome Sequence (WGS) Dataset

Gene type	scaffold	start	stop	strand	Ensembl Gene ID	NCBI Gene ID	WGS variants /bp	WGS mean coverage
R	175	436531	437673	+	None	100660069	0.0061	21.18
R	208	307663	308805	-	None	100670203	0.0009	23
R	217	57192	58321	-	None	100657221	0.0053	35.24
R	221	32722	33852	+	ENSLAFG00000027348	100669451	0.0221	109.47
R	221	320215	321342	+	ENSLAFG00000030555	100669732	0.0186	50.88
R	281	127150	128270	+	None	100670118	0.0223	54.81
R	294	64102	65230	+	ENSLAFG00000027820	100660838	0.0071	22.42
R	342	119172	120298	-	None	100673852	0.0151	94.31
R	378	23269	24394	+	ENSLAFG00000030880	100673452	0.0116	35.9
R	406	137208	138342	-	ENSLAFG00000027474	100666240	0.0185	78.87
R	458	14552	15678	+	ENSLAFG00000032042	100661323	0.0124	72.29
A	**47**	**11688313**	**11693871**	**-**	**ENSLAFG00000007483**	**100663725**	0.0036	33.16
R	498	44787	45912	+	None	100675551	0.0107	19.14
R	552	13399	14524	-	ENSLAFG00000028692	100668616	0.0044	58.28
R	627	40469	41597	+	None	100667946	0.0195	200.58
R	656	10157	11282	-	ENSLAFG00000027365	100673935	0.0133	74.24
R	76	9269289	9270442	+	None	100671320	0.026	27.22
R	786	1954	3080	+	ENSLAFG00000027669	100669552	0.0204	78.25
R	825	4052	5178	+	ENSLAFG00000028299	100673857	0.0053	38.82
R	928	6773	7899	+	ENSLAFG00000026238	100660953	0.0098	79.7

Genes with no high identity support from re-sequencing of cloned loci are indicated in gray and black entries have high sequence identity with our sequenced clones. The ancestral copy containing introns is highlighted in bold. Gene type 'R' indicates 'retrogene' and gene type 'A' indicates 'ancestral'. 'WGS mean coverage' indicates the average depth of WGS reads aligned to each gene sequence. Ensembl Gene IDs are from www.ensembl.org and NCBI Gene IDs are from http://www.ncbi.nlm.nih.gov/gene/.

eTable 3. Study Participants in the Cancer Genetics Study at the University of Utah

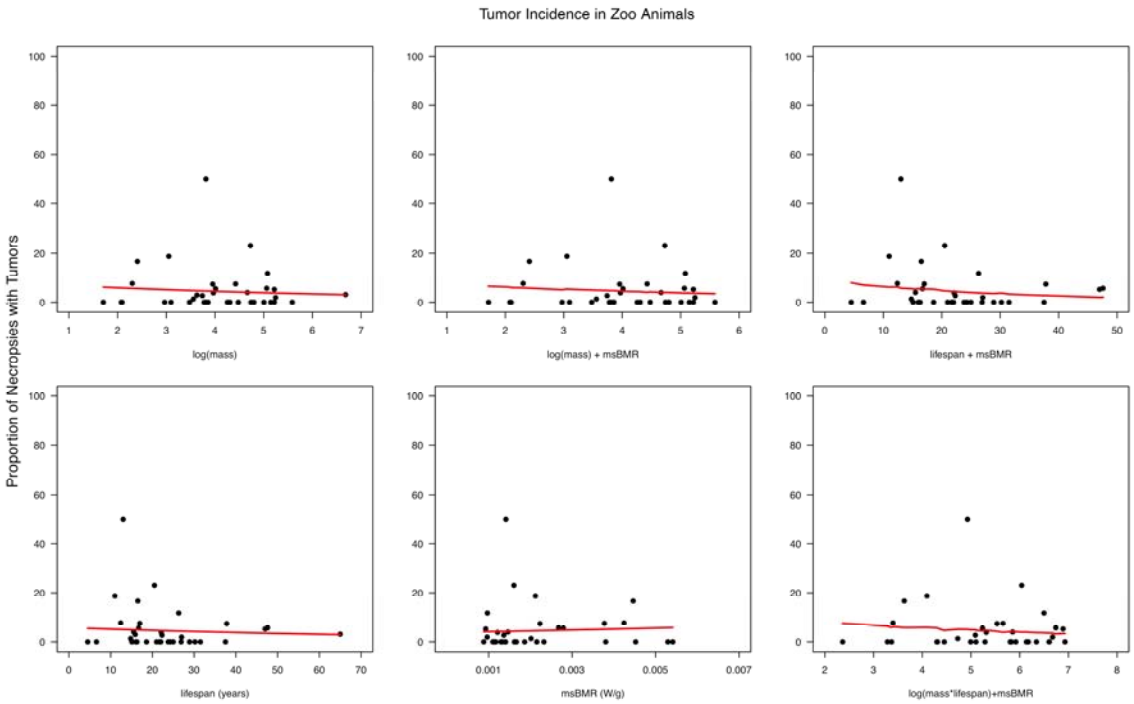
Study Participants Recruited from Cancer Genetics Study (University of Utah)				
LFS Patient Age	**LFS Patient Gender**	***TP53* Mutation**	**Cancer Type (Age at diagnosis)**	**Matched Non-LFS Healthy Control (Age / Gender)**
14 year old	Male	p.P177R	None	25 yo, male
18 year old	Female	p.P177R	Spinal chordoma (18 yo)	19 yo, female
20 year old	Female	deletion exon 1	Phyllodes breast tumor (12 yo)	19 yo, female
21 year old	Male	p.P177R	Osteosarcoma (18 yo) and colorectal cancer (21 yo)	34 yo, male
23 year old	Female	p.A347D	Pleomorphic high-grade sarcoma (19 yo)	23 yo, female
25 year old	Male	p.R248W	None	27 yo, male
27 year old	Female	p.I251L	None	24 yo, female
40 year old	Female	p.R196*	Adrenocortical carcinoma (2 yo) and invasive ductal carcinoma of the breast (37 yo) ductal carcinoma in situ (38 yo)	37 yo, female
45 year old	Female	p.A347D	None	44 yo, female
51 year old	Male	p.L194F	Lung Cancer	54 yo, male

Overall LFS population in the Cancer Genetics Study (N=51, living)				
LFS Patient Age Range (Mean / Median)	**LFS Patient Gender** (% Male)	**TP53 Mutation Types**	**LFS Cancer** (Yes vs. No)	**LFS Cancer Age Onset Range** (Mean / Median)
1-58 years old (25.3 / 25)	17 (34%) Male	15 mutations 3 unknown	22 (44%) yes 29 (56%) no	0- 52 year old (19.1 / 18)

Overall control population in the Cancer Genetics Study (N=910)	
Control Age Range (Mean / Median)	**Control Gender** (% Male)
1-92 years old (43.1/ 42)	255 (28%) Male

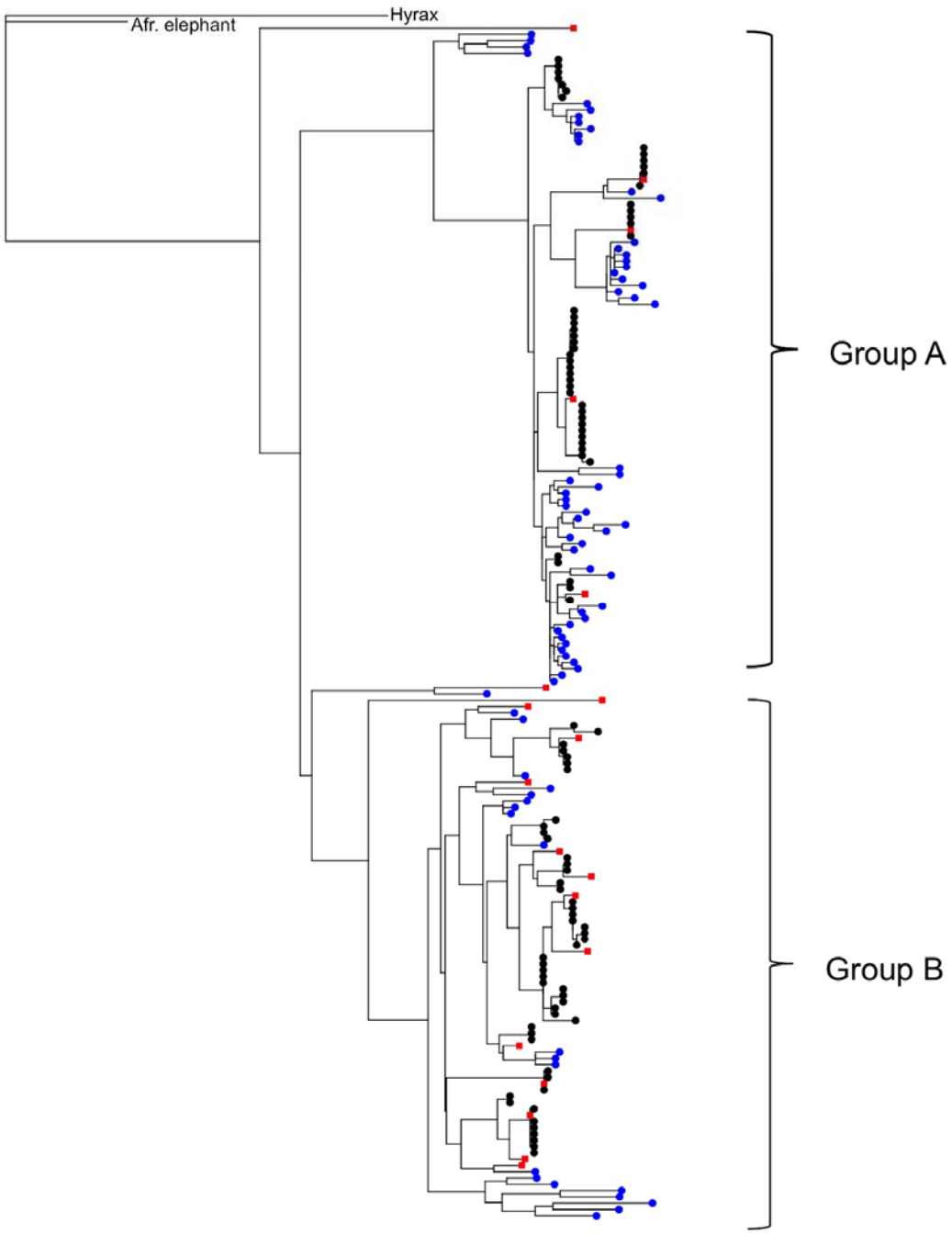
Living patients with Li-Fraumeni Syndrome (LFS) were identified through the Cancer Genetics Study as detailed above. Participants were selected to represent both affected and unaffected mutation carriers, and based on their availability to provide a blood sample. Age and gender matched participants from CGS without a personal history of cancer or a molecular diagnosis of a hereditary cancer predisposition were included as controls as detailed above.

eFigure 1. Logistic Regression Models for Tumor Incidence in Zoo Animals by Body Size, Lifespan, Mass Specific Basal Metabolic Rate (msBMR), and Combinations of These Variables



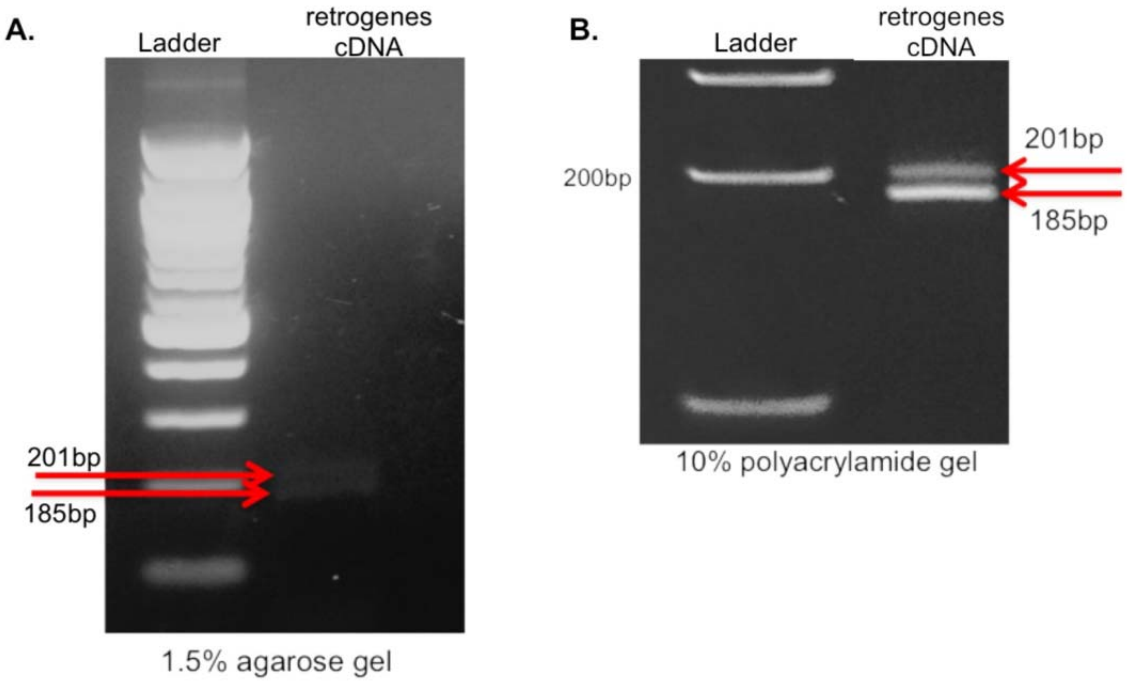
Top left box displays log(mass), top middle displays log(mass)+msBMR, top right displays lifespan + msBMR, bottom left displays lifespan(years), bottom middle displays msBMR, and bottom right displays log(mass*lifespan)+msBMR. Note that the msBMR is smaller in larger animals so the line appears reversed in the msBMR figure, with larger animals represented on the left side of this graph. Animals with non-null values for the x-axis variables are in each respective plot and all data is provided in eTable 1.

eFigure 2. Maximum Likelihood Phylogeny of Asian Elephant Retrogenes



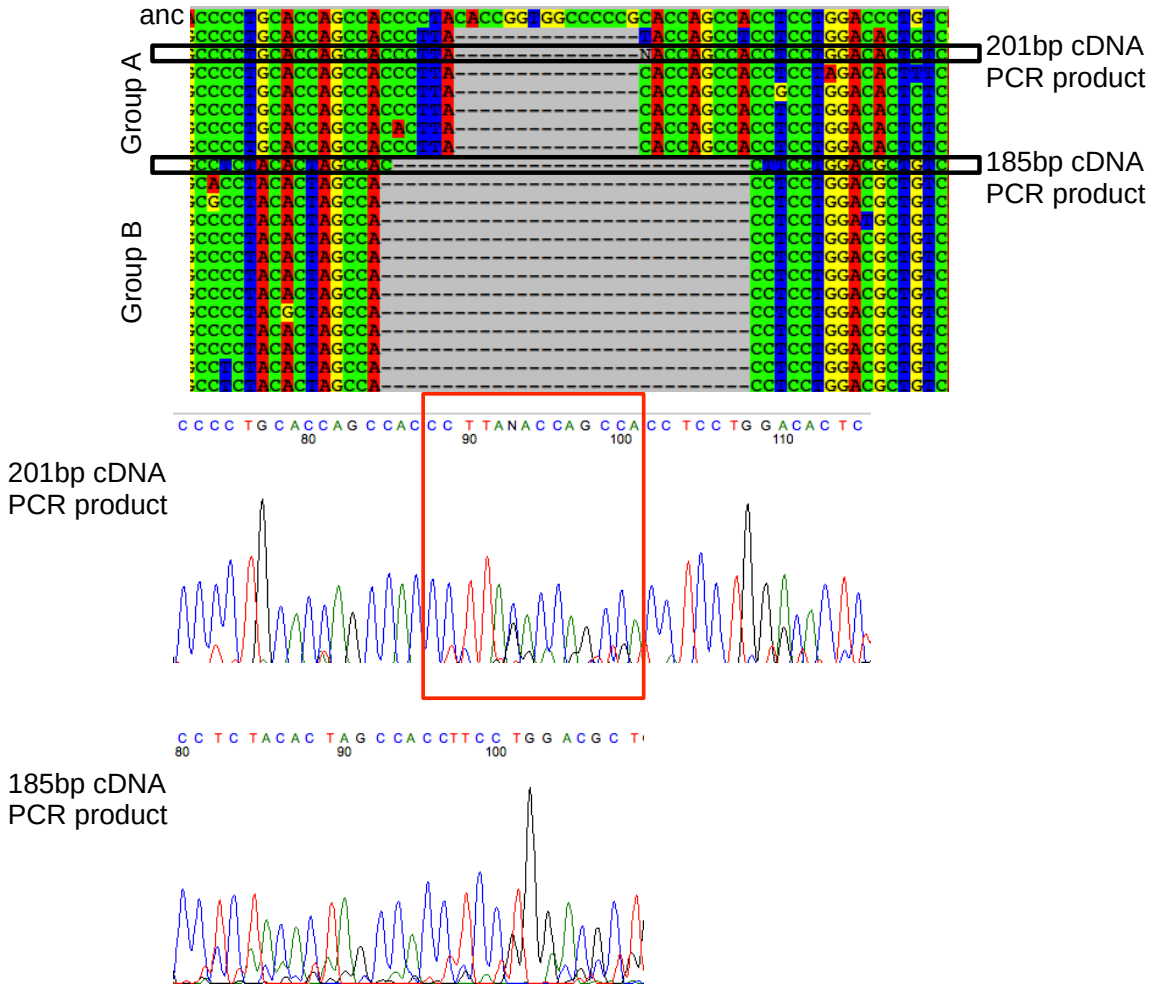
The capillary sequenced clones from the Asian elephant (blue circles), African elephant (black circles), and published sequences from GenBank (red squares) are graphed to show both Group A and Group B retrogenes similar to the African elephant genome. All sequenced clones are from this study. The branch labeled 'Afr .elephant' is the coding sequence of the African elephant ancestral *TP53* and 'hyrax' represents the coding sequences from the hyrax *TP53*. DNA from one elephant was used to produce the clones to avoid polymorphisms between individual elephants.

eFigure 3. TP53 Retrogene Transcription



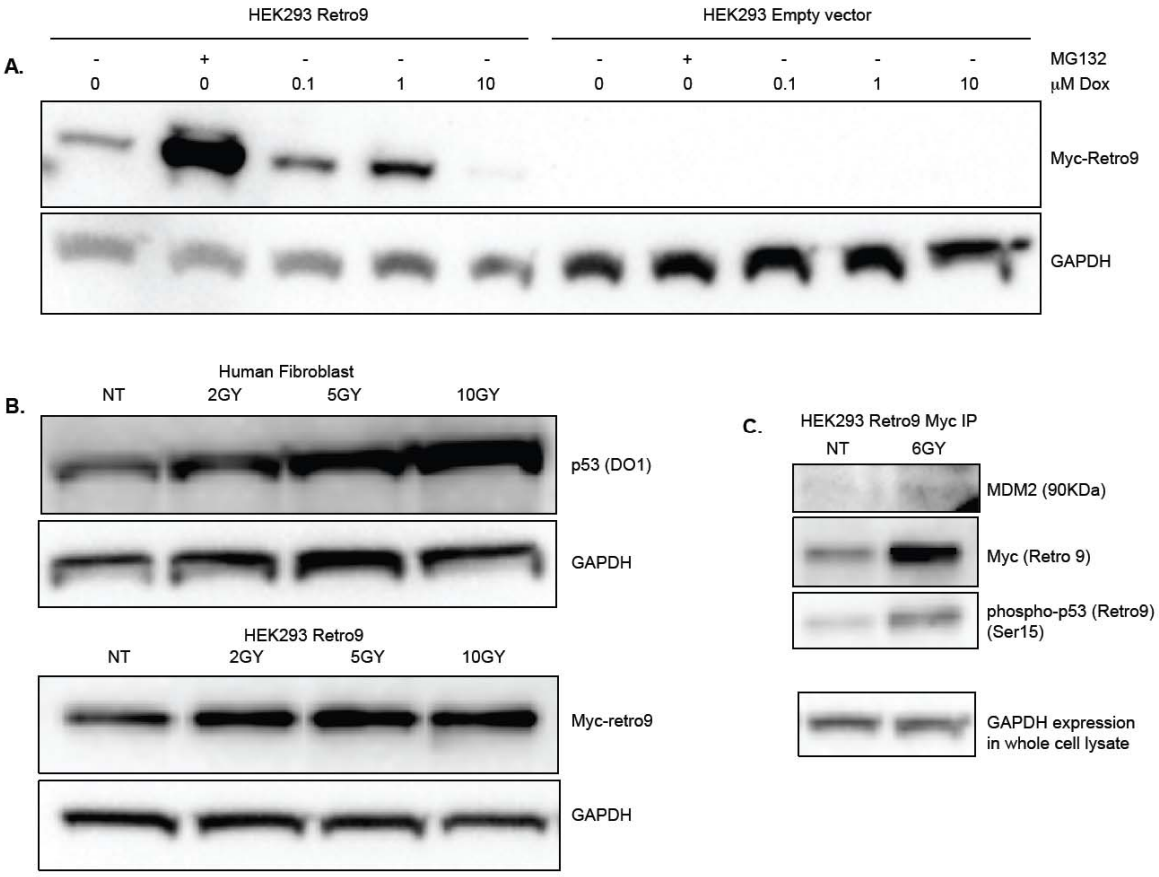
(A) Shown here is the PCR product from cDNA prepared from the African elephant RNA sample at 5hr after treatment with 2Gy IR run on 1.5% agarose and (B) 10% polyacrylamide for better separation and clarity of the bands. Red arrows in panel A and B point to bands or PCR products that are 201bp and 185bp long which are the expected sizes produced by the retrogenes for Group A and Group B *TP53* elephant retrogenes, respectively (sequence shown in eFigure 4). RNA shown is from one individual elephant. The left lanes are the ladder and the right lanes show the PCR product from the elephant retrogene cDNA. This gel shows the 5-hour time point, and bands were seen in all time points and in both treated and untreated samples.

eFigure 4. Capillary Sequencing of Excised Gel Bands From PCR of Elephant cDNA



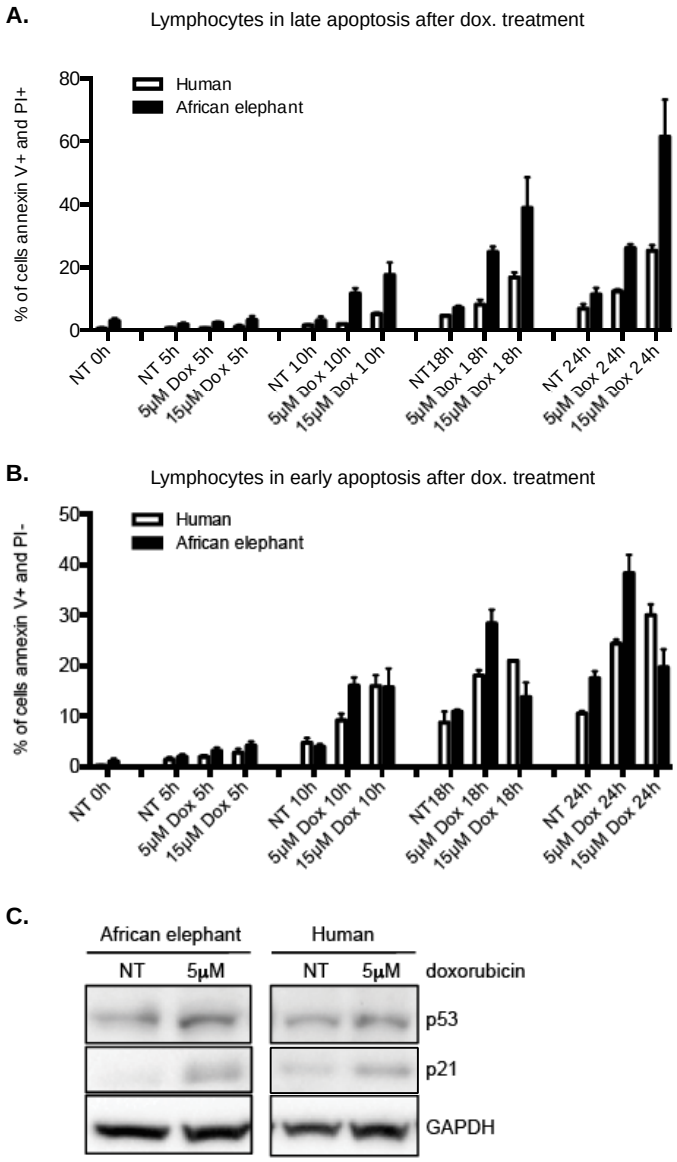
The top alignment is a screenshot from Seaview showing a segment of the ancestral *TP53* coding sequence, retrogene sequences from Groups A (201bp product) and B (185pb product), and the aligned PCR product sequences (outlined by black rectangles) from the excised gel bands from the 10% polyacrylamide gel (eFigure 3). The sequences traces are shown below with the red box indicating the deleted region from the retrogenes in Group B relative to those in Group A.

eFigure 5. Elephant p53 Retrogene 9 Protein Expression in Transfected HEK293 and DNA Damage Response in Human Fibroblasts



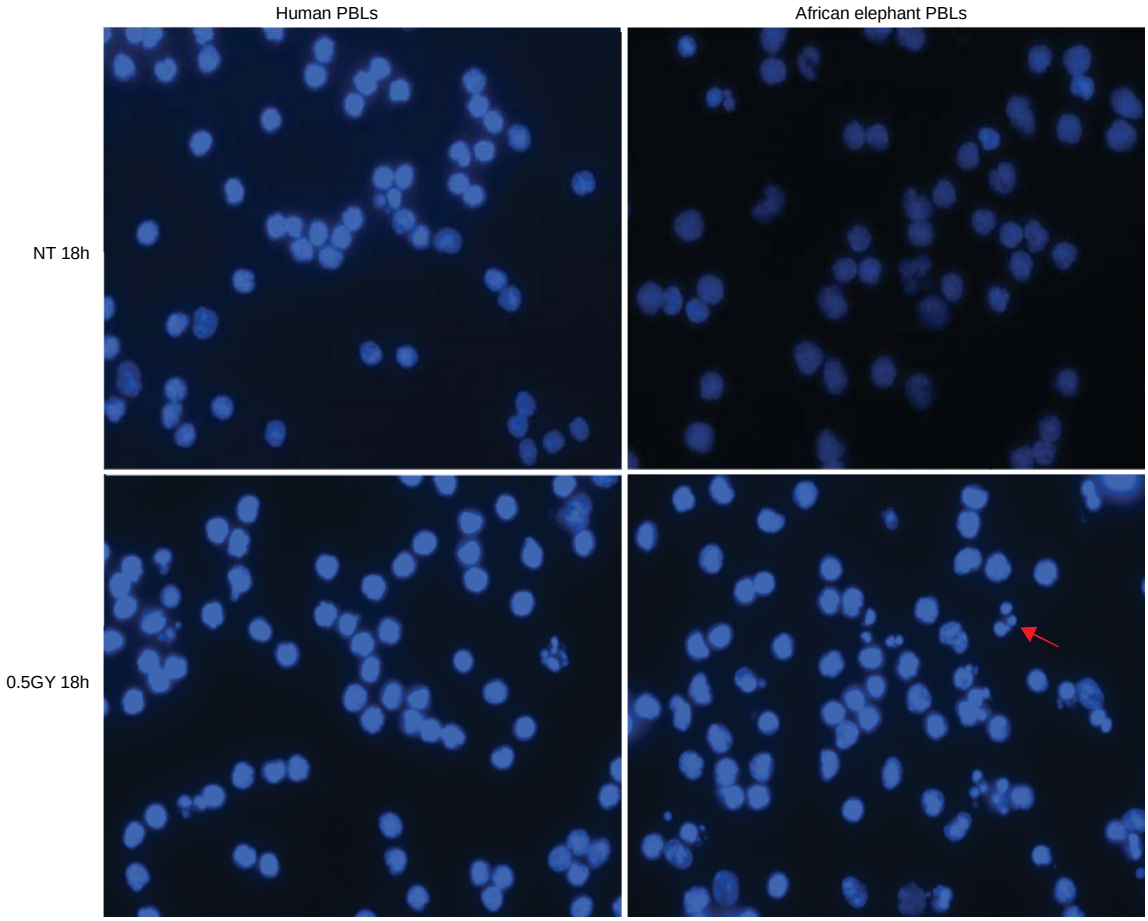
(A) This western blot shows the expression of Myc-tagged retrogene 9 (retro9; GenBank ID: KF715863) protein after transfection in HEK293 cells and the increased Retro9 expression when the cells were treated with either a proteasome inhibitor (MG132) or a DNA damaging agent (doxorubicin) for 5 hours. Cells transfected with empty vector were included as a negative control for the Myc antibody. (B) These western blots show increased expression of Myc-retro9 protein in HEK293 cells transfected with retro9 5 hours after DNA damage from IR (bottom panel), similar to the increase in activated p53 observed in human fibroblasts treated with IR (top panel). (C) This western blot shows immunoprecipitation (IP) of Myc-retrogene 9 using an antibody to Myc from lysates of HEK293 retro9 cells that were either not treated (NT) or cultured for 2 hours post 6GY IR. When probed with an antibody to phospho-p53 (Ser15), a band of the same size as Myc-retro9 was observed. In addition, a 90 KDa band was observed, indicating that Mdm2 co-immunoprecipitated with Myc-retro9.

eFigure 6. African Elephant Lymphocytes and p53-Dependent Response to Doxorubicin Treatment



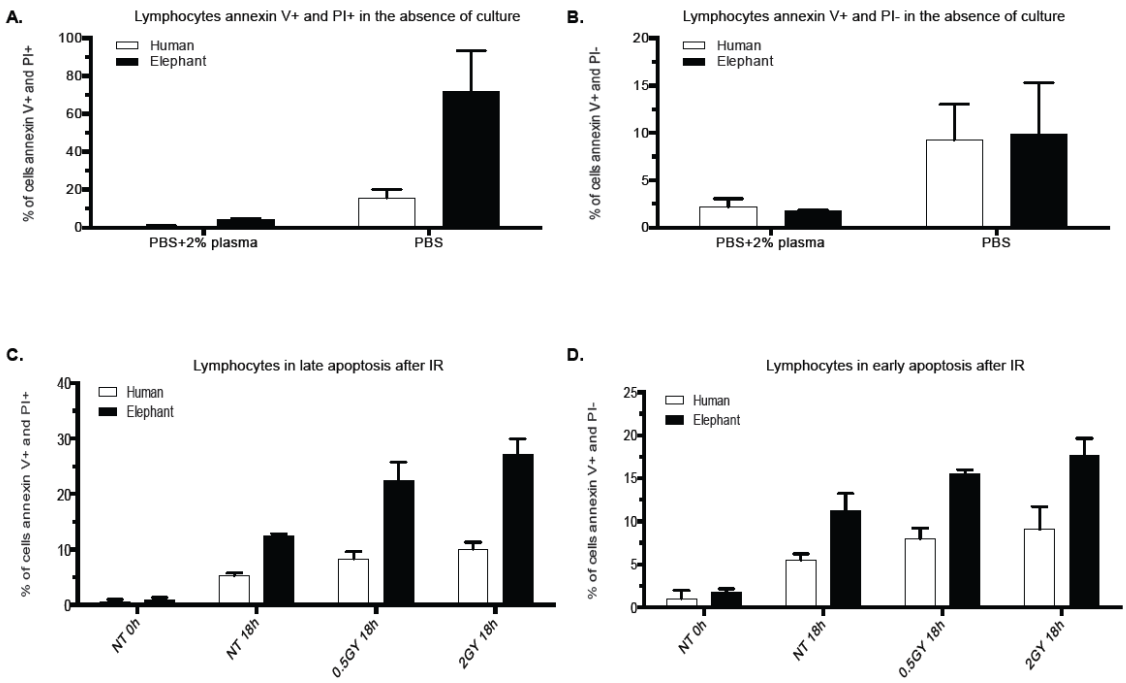
(A) Percentage of late stage apoptotic cells and (B) early apoptotic cells are shown. The error bars represent the 95% confidence interval. Significant differences were observed at different times points with a two-sided *t*-test without assuming consistent variance (Late stage apoptosis: *P*<0.001 to 0.0071, Early stage apoptosis: *P*<0.001 to 0.033). (C) The western blot shows induction of p21 at 5 hours in African elephant and human lymphocytes with doxorubicin treatment.

eFigure 7. DAPI Staining for Visual Representation of Apoptotic Cells in Human vs. African Elephant PBLs



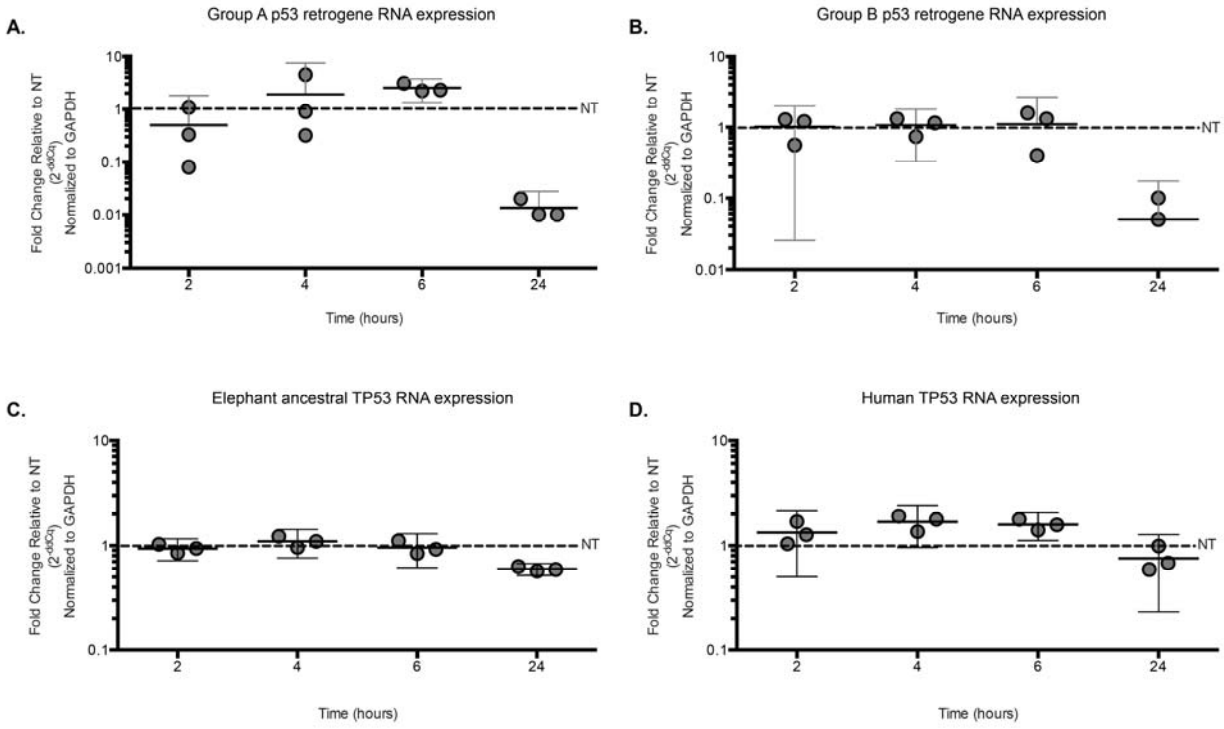
Few apoptotic cells are seen in the no treatment panels, while the 0.5GY panels show more apoptotic elephant cells compared to human cells as indicated by nuclear blebbing (red arrow). Magnification is 60x.

eFigure 8. Lymphocyte Wash Conditions



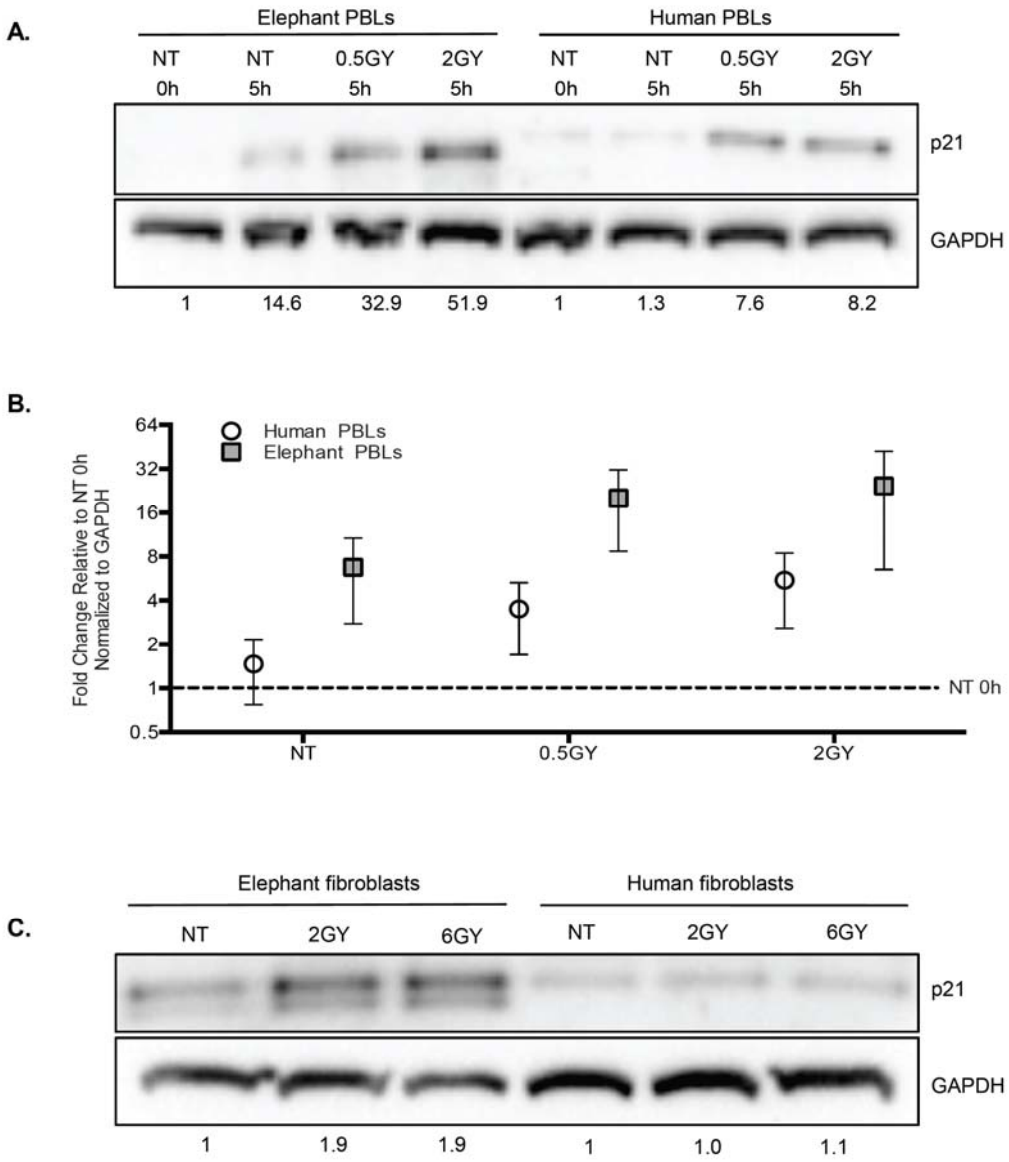
(A) As displayed, more lymphocytes stained for late apoptosis (annexin V [AV]+ and propidium iodide [PI]+) after being washed in PBS compared to PBS+2% autologous plasma, indicating that elephant lymphocytes were more sensitive to PBS (*P*<0.001). (B) No significant difference in early apoptotic cells (AV+/PI-) was observed between human and elephant cells washed in either PBS or PBS+2% plasma. (C) PBLs treated with the indicated doses of IR were cultured for 18 hours, and as shown, the elephant cells washed in the presence of protein still showed a significant increase in late apoptosis (AV+/PI+) compared to human cells (NT: *P*<0.001, 0.5GY: *P*<0.001, 18GY: *P*<0.001). (D) Also shown in this figure, more elephant lymphocytes stained positive for early apoptosis (AV+/PI-) when washed in the presence of protein (NT: *P*<0.001, 0.5GY: *P*<0.001, 18GY: *P*=0.001). Significance was measured by 2-sided, unpaired *t*-test without assuming equal variance, and error bars represent the 95% CI for each experiment performed in triplicate.

eFigure 9. mRNA Expression of Retrogenes and Ancestral *TP53*



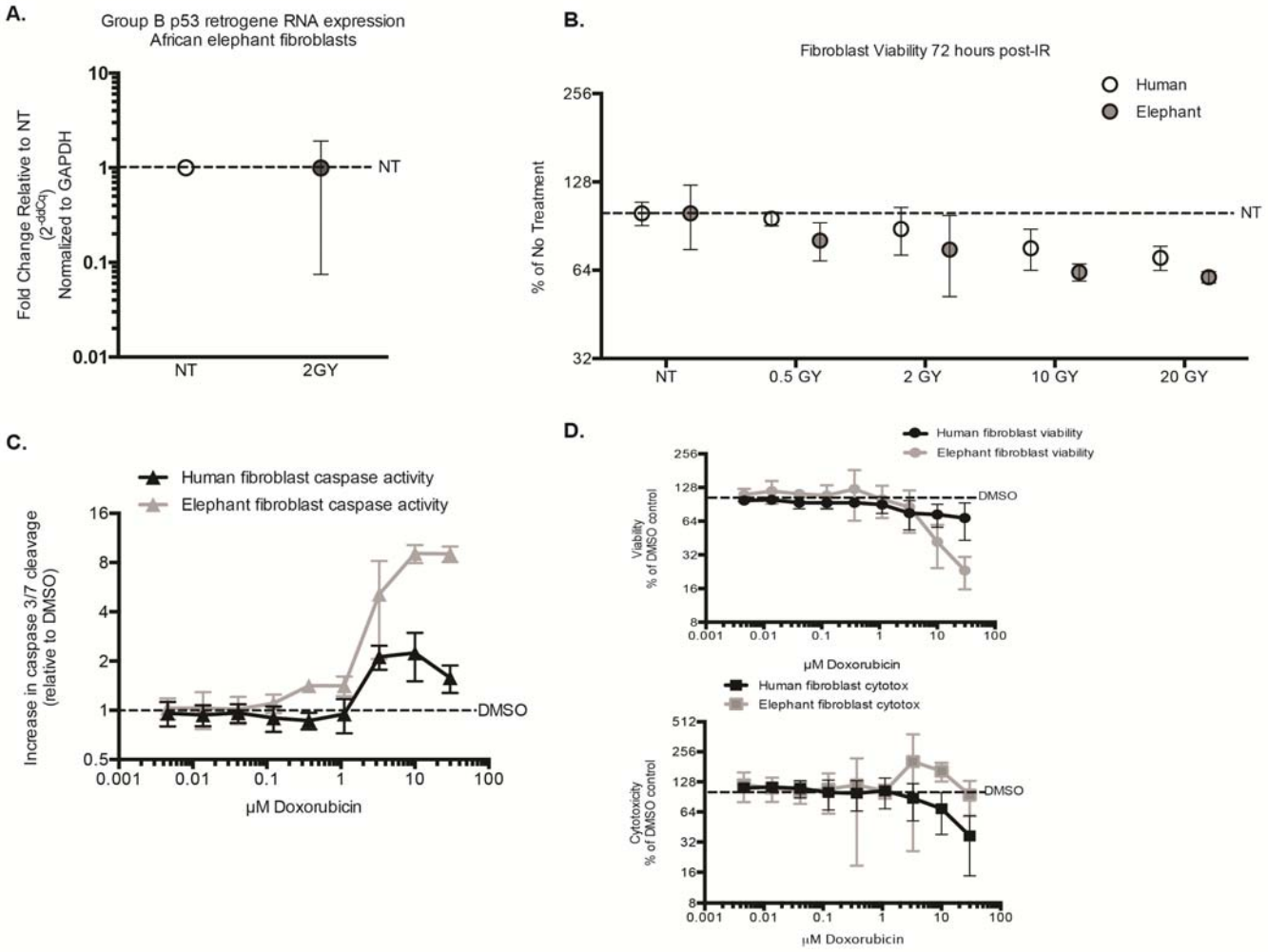
(A) Group A *TP53* retrogene gene expression and (B) Group B *TP53* retrogene gene expression are plotted, showing detection of cDNA for both groups in treated and untreated elephants PBLs. Both *TP53* retrogene groups significantly decreased expression 24 hours after IR treatment (Group A: $P<0.001$, Group B: $P<0.001$). (C) mRNA expression of the ancestral *TP53* gene remained relatively stable in response to 2GY IR in the African elephant as shown in this plot, but its expression did decrease at 24 hours ($P<0.001$). (D) Ancestral *TP53* gene expression was seen to increase slightly in human PBLs with IR (2 hours: $P=0.15$, 4 hours: $P=0.015$, 6 hours: $P=0.0057$) and then decrease slightly at 24 hours ($P=0.11$). Data is displayed as the fold change in expression relative to the untreated (NT) cells at each given time point. Significant changes in expression compared to the NT samples was computed with a two-sided *t*-test, error bars represent the 95% CI for each experiment performed in triplicate, and the dashed horizontal line shows the labeled data marker for the NT reference value.

eFigure 10. p21 Expression in Elephant Cells Compared to Human Cells



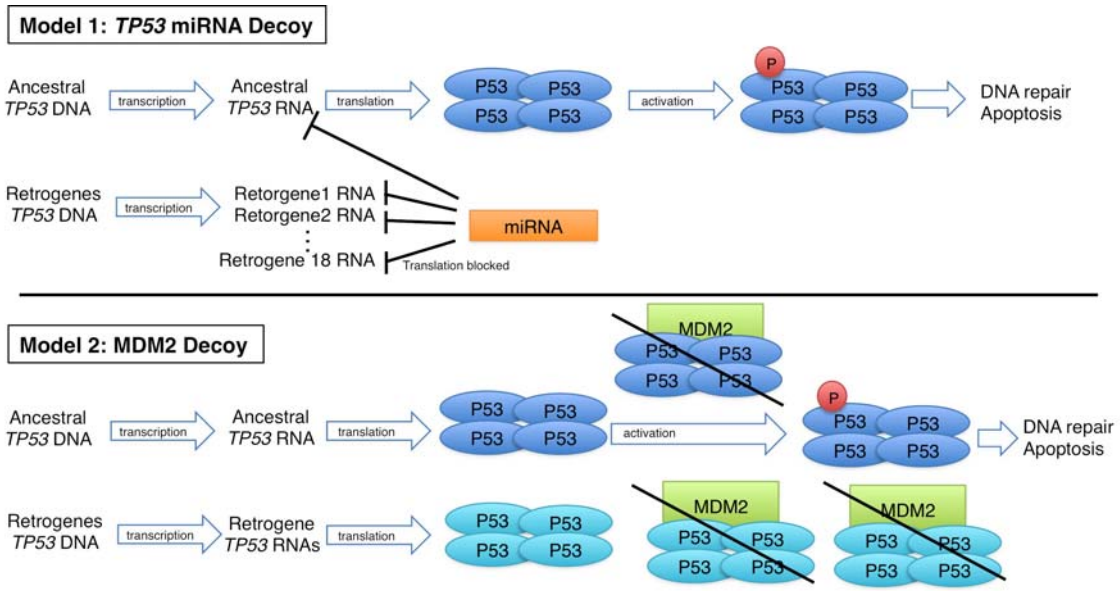
(A) Western blots are displayed for PBLs exposed to the indicated doses of IR after 5 hours, showing that elephant PBLs expressed more p21 protein compared to human PBLs. (B) Shown in this panel are the combined data points for p21 gene expression at the 5 hour time point from a western blot experiment performed in quadruplicate and a western blot experiment performed in triplicate. Two-sided, unpaired *t*-tests were run without assuming consistent variance (NT: P=0.0078, 0.5GY: P=0.0042, 2GY: P=0.025), error bars indicate 95% CI. (C) Elephant fibroblasts expressed more p21 protein as measured by western blot compared to human fibroblasts after 2 hours in response to irradiation. Numbers below western blots indicate the increase in fold change after treatment relative to no treatment at 0h for each individual.

eFigure 11. *TP53* Retrogenes Expression and DNA Damage Response in African Elephant Fibroblasts Compared to Human Fibroblasts



(A) Elephant fibroblasts express Group B retrogenes as measured by qPCR and expression did not increase with 2 Gray (GY) ionizing irradiation (IR) relative to untreated cells (NT). (B) A significant decrease in viability was observed for elephant fibroblasts compared to human fibroblasts 72 hours after treatment with the indicated doses of IR (0.5GY: *P*=0.0102, 10GY: *P*=0.018, 20GY=0.0046). (C) Treatment of elephant fibroblasts for 18 hours with doxorubicin led to a greater increase in caspase activity compared to human fibroblasts (10µM and 30µM: *P*<0.001). (D) Elephant fibroblast cytotoxicity increased (top panel (10µM: *P*=0.0051, 30µM: *P*=0.0018) and viability decreased (bottom panel (10µM: *P*<0.001, 30µM: *P*=0.0044), consistent with apoptosis. For all panels, *P* values were calculated using 2-sided, unpaired *t*-tests without assuming consistent variance. Error bars indicate 95% CI.

eFigure 12. Two Decoy Models Proposed for *TP53* Retrogenes in Elephants



As shown in this cartoon, each decoy model could act independently or in combination to increase activated p53 protein available to respond to DNA damage. Model 1 shows the retrogene *TP53* RNAs acting as decoys for circulating miRNA, which normally targets ancestral *TP53*. The miRNAs bind the abundant *TP53* retrogenes and signal them for degradation, allowing the ancestral *TP53* mRNA to escape degradation and translate into protein to respond to DNA damage. Model 2 depicts another decoy scenario whereby the *TP53* retrogenes form p53 protein tetramers. The translated retrogene proteins act as decoys for Mdm2 that targets them for degradation by the proteasome, enabling the accumulation of ancestral p53 tetramers. These escaped tetramers can then be activated to respond to DNA damage and contribute to increased apoptosis. This second model is supported by the findings that Mdm2 co-precipitated with *TP53* retrogene 9 when transfected into HEK293 cells (eFigure 6).

eReferences

1. Griner LA. *Pathology of Zoo Animals.* San Diego, CA: Zoological Society of San Diego; 1983.
2. de Magalhães JP, Costa J. A database of vertebrate longevity records and their relation to other life-history traits. *Journal of evolutionary biology.* 2009;22(8):1770-1774.
3. Elephant Encyclopedia. www.elephant.se 1995-2014.
4. ACS. *American Cancer Society. Cancer Facts & Figures 2015.* 2015.
5. Ferlay J, Soerjomataram I, Dikshit R, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer.* 2015;136(5):E359-386.
6. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 2004;32(5):1792-1797.
7. Lindblad-Toh K, Garber M, Zuk O, et al. A high-resolution map of human evolutionary constraint using 29 mammals. *Nature.* 2011;478(7370):476-482.
8. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics.* 2009;25(14):1754-1760.
9. McKenna A, Hanna M, Banks E, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 2010;20(9):1297-1303.
10. DePristo MA, Banks E, Poplin R, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet.* 2011;43(5):491-498.
11. Wilson AJ, Holson E, Wagner F, et al. The DNA damage mark pH2AX differentiates the cytotoxic effects of small molecule HDAC inhibitors in ovarian cancer cells. *Cancer Biol Ther.* 2011;12(6):484-493.
12. Louis N, Evelegh C, Graham FL. Cloning and sequencing of the cellular-viral junctions from the human adenovirus type 5 transformed 293 cell line. *Virology*. 1997;233:423-429.
13. Berk AJ. Recent lessons in gene expression, cell cycle control, and cell biology from adenovirus. *Oncogene*. 2005;24:7673-7685.