TP53 copy number expansion correlates with the evolution of increased body size and an enhanced DNA damage response in elephants

Michael Sulak[‡1], Lindsey Fong[‡1], Katelyn Mika[1], Sravanthi Chigurupati[1], Lisa Yon[2], Nigel P. Mongan[2,3], Richard D. Emes[2,4], Vincent J. Lynch[*1‡]

1) Department of Human Genetics, The University of Chicago, Chicago, IL, 60637
2) School of Veterinary Medicine and Science, Faculty of Medicine and Health Sciences. University of Nottingham, Sutton Bonington Campus, Leicestershire, LE12 5RD. UK.
3) Department of Pharmacology, Weill Cornell Medical College, New York, NY, 10065
4) Advanced Data Analysis Centre, University of Nottingham UK.

[‡] Contributed equally

* To whom correspondence should be addressed: vjlynch@uchicago.edu

SUMMARY

A major constraint on the evolution of large body sizes in animals is an increased risk of developing cancer. There is no correlation, however, between body size and cancer risk. This lack of correlation is often referred to as 'Peto's Paradox'. Here we show that the elephant genome encodes 20 copies of the tumor suppressor gene *TP53* and that the increase in *TP53* copy number occurred coincident with the evolution of large body sizes in the elephant (Proboscidean) lineage. Furthermore we show that several of the *TP53* retrogenes are transcribed and translated and contribute to an enhanced sensitivity of elephant cells to DNA damage and the induction of apoptosis via a hyperactive TP53 signaling pathway. These results suggest that an increase in the copy number of *TP53* may have played a direct role in the evolution of very large body sizes and the resolution of Peto's paradox in Proboscideans.

INTRODUCTION

Lifespan and maximum adult body size are fundamental life history traits that vary considerably between species (Healy et al., 2014). The maximum lifespan among vertebrates, for example, ranges from over 211 years in the bowhead whale (*Balaena mysticetus*) to only 59 days in the pygmy goby (*Eviota sigillata*) whereas body sizes ranges from 136,000 kg in the blue whale (*Balaenoptera musculus*) to 0.5 g in the Eastern red-backed salamander (*Plethodon cinereus*) (Healy et al., 2014). Similar to other life history traits, such as body size and metabolic rate or body size and age at maturation, body size and lifespan are strongly correlated such that larger species tend to live longer than smaller species (**Fig. 1A**). While abiotic and biological factors have been proposed as major drivers of maximum body size evolution in animals, maximum body size within tetrapods appears to be largely determined by biology (Smith et al., 2010; Sookias et al., 2012). Mammals, for example, likely share biological constraints on the evolution of very large body sizes with rare breaks in those constraints underlying the evolution of gigantism in some lineages (Sookias et al., 2012), such as Proboscideans (elephants and their and extinct relatives), Cetaceans (whales), and the extinct hornless rhinoceros *Paraceratherium* ('Walter').

A major constraint on the evolution of large body sizes in animals is an increased risk of developing cancer. If all cells have a similar risk of malignant transformation and equivalent cancer suppression mechanisms, organism with many cells should have a higher risk of developing cancer than organisms with fewer cells; Similarly organisms with long lifespans have more time to accumulate cancer-causing mutations than organisms with shorter lifespans and therefore should be at an increased risk of developing cancer, a risk that is compounded in large bodied, long-lived organisms (Cairns, 1975; Caulin and Maley, 2011; Doll, 1971; Peto, 2015; R Peto, 1975). There are no correlations, however, between body size and cancer risk or lifespan and cancer risk across species (Leroi et al., 2003), this lack of correlation is often referred to as 'Peto's Paradox' (Caulin and Maley, 2011; R Peto, 1975). Epidemiological studies in wild populations of Swedish roe deer (*Capreolus capreolus*) and beluga whales (*Delphinapterus leucas*) in the highly polluted St. Lawrence estuary, for example, found cancer accounted for only 2% (Aguirre et al., 1999) and 27% (Martineau et al., 2002) of mortality, respectively, much lower than expected given body size of these species (Caulin and Maley, 2011).

Among the mechanisms large, long lived animals may have evolved that resolve Peto's

paradox are a decrease in the copy number of oncogenes, an increase in the copy number of tumor suppressor genes (Caulin and Maley, 2011; Leroi et al., 2003; Nunney, 1999), reduced metabolic rates leading to decreased free radical production, reduced retroviral activity and load (Katzourakis et al., 2014), increased immune surveillance, and selection for "cheater" tumors that parasitize the growth of other tumors (Nagy et al., 2007), among many others. Naked mole rats (*Heterocephalus glaber*), for example, which have very long lifespans for a small-bodied organism evolved cells with extremely sensitive contact inhibition likely acting as a constraint on tumor growth and metastasis (Seluanov et al., 2009; Tian et al., 2013). Similarly long-lived blind mole rats (*Splanx sp.*) evolved an enhanced TP53-depenent necrotic cell death mechanism that also likely constrains tumor growth (Gorbunova et al., 2012). Thus while some of the mechanisms that underlie cancer resistance in small long-lived mammals have been identified, the mechanisms by which large bodied animals evolved enhanced cancer resistance are unknown.

Here we use evolutionary genomics and comparative cell biology to explore the mechanisms by which elephants, the largest extant land mammal (**Fig. 1B**), have evolved enhanced resistance to cancer. We found that the elephant genome encodes a single *TP53* gene and 19 *TP53* retrogenes, several of which are transcribed and translated in elephant tissues. Comparison of the African and Asian elephant *TP53* gene copy number with the copy number in the genome of the extinct American mastodon, woolly mammoth, and Columbian mammoth indicates that copy number increased relatively rapidly coincident with the evolution of large body-sizes in the Proboscidean lineage. Finally we show that elephant cells have an enhanced response to DNA-damage that is mediated by a hyperactive TP53 signaling pathway and that this augmented TP53 signaling can be transferred to the cells of other species through exogenous expression of elephant TP53 retrogenes. These results suggest that the origin of large body sizes, long lifespans, and enhanced cancer resistance in the elephant lineage evolved at least in part through reinforcing the anti-cancer mechanisms of the major 'guardian of the genome' TP53.

RESULTS

Expansion of the *TP53* repertoire in Proboscideans

We characterized *TP53* copy number in 61 Sarcopterygian (Lobe-finned fishes)

genomes, including large, long-lived mammals such as the African elephant (*Loxodonta africana*), Bowhead (*Balaena mysticetus*) and Minke (*Balaenoptera acutorostrata scammoni*) whales. We found that all Sarcopterygian genomes encoded a single *TP53* gene and that some lineages also contained a few *TP53* retrogenes (*TP53RTG*), including marsupials, Yangochiropteran bats, and Glires, in which 'processed' *TP53* pseudogenes have previously been reported (Ciotta et al., 1995; Czosnek et al., 1984; Hulla, 1992; Tanooka et al., 1995; Weghorst et al., 1995; Zakut-Houri et al., 1983). We also identified a single *TP53RTG* gene in the lesser hedgehog tenrec (*Echinops telfairi*), rock hyrax (*Procavia capensis*), and West Indian manatee (*Trichechus manatus*), the African elephant genome, however, encoded 19 *TP53RTG* genes (**Fig. 2A**), 14 of which retain potential to encode truncated proteins (**Supplementary Fig. 1**).

To trace the expansion of *TP53RTG* genes in the Proboscidean lineage with greater phylogenetic resolution, we used three methods to estimate the minimum (1:1 orthology), average (normalized read depth), and maximum (gene tree reconciliation) *TP53*/*TP53RTG* copy number in the Asian elephant (*Elephas maximus*), extinct woolly (*Mammuthus primigenius*) and Columbian (*Mammuthus columbi*) mammoths, and the extinct American mastodon (*Mammut americanum*) using existing whole genome sequencing data (Enk et al., 2014; 2013; Rohland et al., 2010; Wilkie et al., 2013). As expected, we identified a single canonical *TP53* gene in these species and estimated the *TP53RTG* copy number in the Asian elephant genome to be 12-17, approximately 14 in both the Columbian and woolly mammoth genomes, and 3-8 in the 50,000-130,000 year old American mastodon genome (**Fig. 2B** and **Supplementary Fig. 2**). These data indicate that large-scale expansion of the *TP53RTG* gene family occurred in the Proboscidean lineage and suggest that *TP53RTG* copy number was lower in ancient Proboscideans such as the mastodon, which diverged from the elephant lineage ~25MYA (Rohland et al., 2010), than in recent species such as elephants and mammoths.

The *TP53RTG* repertoire expanded through repeated segmental duplications

Several mechanisms may have increased the *TP53RTG* copy number in the Proboscidean lineage, including serial retrotransposition from a single parent gene, serial retrotransposition from the original parent gene and one or more daughter transcribed retrogenes, repeated segmental duplications of chromosomal loci containing *TP53RTG* genes,

or some combination of these mechanisms. Consistent with copy number expansion through a single retrotransposition event followed by repeated rounds of segmental duplication, we found that each *TP53RTG* retrogene was flanked by nearly identical clusters of transposable elements (**Fig. 3A**) and embedded within a large genomic region with greater than 80% sequence similarity (**Fig. 3B**). Next we used progressiveMAUVE to align the 18 elephant contigs containing *TP53RTG* retrogenes and found that they were all embedded within large locally collinear blocks that span nearly the entire length of some contigs (**Fig. 3C**), as expected for segmental duplications.

***TP53RTG* copy number expansion is correlated with Proboscidean body size**

Our observation that *TP53RTG* genes expanded through segmental duplications suggests they may have a tree-like phylogenetic history that preserves information about when in the evolution of Proboscideans the duplicates occurred. Therefore we assembled a dataset of *TP53*/*TP53RTG* orthologs from 65 diverse mammals and jointly inferred the *TP53*/*TP53RTG* gene tree and duplication dates in a Bayesian framework to determine if *TP53*/*TP53RTG* copy number was correlated with body size evolution in Proboscideans. For comparison we also inferred the phylogenetic history of *TP53*/*TP53RTG* genes using maximum likelihood and an additional Bayesian method. We found all phylogenetic inference methods inferred that the *TP53RTG* genes from elephant, hyrax, and manatee formed a well-supported sister clade to the canonical genes from these species, whereas the tenrec *TP53* and *TP53RTG* genes formed a separate well-supported clade (**Fig. 4A** and **Supplementary Fig. 3**). These data indicate that retrotransposition of *TP53* occurred independently in tenrecs and in the elephant, hyrax, and manatee stem-lineage (Paenungulata), followed by expansion of *TP53RTG* genes in the Proboscidean lineage.

Based on our time-calibrated phylogeny, we inferred that the initial retrotransposition of the *TP53* gene in the Paenungulata stem-lineage occurred approximately 64 MYA (95% HPD = 62.3-66.2 MYA; **Fig. 4A**). This was followed by a period of ~25 million years during which no further retrotranspositions or segmental duplications were fixed in the genome, however, the *TP53RTG* gene family rapidly expanded after ~40 MYA (95% HPD = 30.8-48.6 MYA; **Fig. 4A**). To correlate *TP53*/*TP53RTG* copy number and the origin of large body sizes in Proboscideans we estimated *TP53*/*TP53RTG* copy number through time and gathered data on ancient

Proboscidean body sizes from the literature (Evans et al., 2012; Smith et al., 2010). We found that the increase in *TP53*/*TP53RTG* copy number in the Proboscidean lineage and Proboscidean body size evolution closely mirrored each other (**Fig. 4B**).

***TP53RTG12* is transcribed from a transposable element derived promoter**

If expansion of the *TP53RTG* gene repertoire played a role in the resolution of Peto's paradox during the evolution of large bodied Proboscideans, then one or more of the *TP53RTG* genes should be transcribed. Therefore we generated RNA-Seq data from Asian elephant dermal fibroblasts, African elephant term placental villus and adipose tissue, and used previously published RNA-Seq data from African elephant fibroblasts (Cortez et al., 2014) to determine if *TP53RTG* genes were transcribed. We found that the *TP53* and *TP53RTG12* genes were robustly transcribed in all samples, whereas *TP53RTG3* and *TP53RTG18* transcripts were much less abundant suggesting they are not actively transcribed (**Fig. 5A**). To confirm that the African and Asian elephant *TP53RTG12* genes were transcribed, we designed a set of PCR primers specific to the *TP53* and *TP53RTG12* genes that flank a 15-30bp deletion in *TP53RTG* genes (**Supplementary Fig. 4A**) and used PCR to assay for expression. Consistent with transcription of the *TP53RTG12* gene, we amplified PCR products at the expected size for the *TP53* and *TP53RTG12* transcripts from Elephant fibroblast cDNA generated from DNase treated RNA. In contrast, we did not amplify PCR products from negative control (no reverse transcriptase) samples (**Fig. 5B; Supplementary Fig. 4B/C**). Sanger sequencing of the cloned PCR products confirmed transcription of the *TP53RTG12* gene (**Supplementary Fig. 4D/E**). We note that we used a Poly-T primer for cDNA synthesis, thus the amplication of *TP53RTG12* PCR products indicates that the *TP53RTG12* is poly-adenylated.

Most retrogenes lack native regulatory elements such as promoters and enhancers to initiate transcription, thus transcribed *TP53RTG* genes likely co-opted existing regulatory elements or evolved regulatory elements *de novo*. To identify putative transcriptional start sites and promoters of the highly expressed *TP53RTG12* gene we used geneid and GENESCAN to computationally predict exons in the African elephant gene and mapped the African and Asian elephant fibroblast RNA-Seq data onto scaffold_825, which encodes the *TP53RTG12* gene. We found that both computational methods predicted an exon ~2kb upstream of the ENSEMBL annotated *TP53RTG12* gene within an RTE-type non-LTR retrotransposon (RTE1_LA) that we

annotated as Afrotherian-specific (**Fig. 5C**). Consistent with this region encoding a transcribed 5'-UTR, a peak of reads mapped within the predicted 5'-exon and within the RTE1_LA retrotransposon (**Fig. 5C**).

We attempted to identify the transcription start site of the *TP53RTG12* gene using several 5'-RACE methods, however, we were unsuccessful in generating PCR products from either African Elephant fibroblast or placenta cDNA, or Asian elephant fibroblast cDNA. Therefore, we designed a set of 34 PCR primers tiled across the region of scaffold_825 that encodes the *TP53RTG12* gene and used these primers to amplify PCR products from African and Asian Elephant fibroblast cDNA generated from DNase treated RNA. We then reconstructed the likely *TP53RTG12* promoter, transcription start site, and exon-intron structure from the pattern of positive PCR products. These data suggest that the major transcription initiation site of *TP53RTG12* is located within a RTE1_LA class transposable element (**Fig. 5C**).

Next we tested the ability of the African and Asian elephant RTE1_LA elements and the RTE1_LA consensus sequence (as a proxy for the ancestral RTE1_LA sequence) to function as a promoter when cloned into the promoterless pGL4.10[*luc2*] luciferase reporter vector in transiently transfected African and Asian elephant fibroblasts. We found that the African and Asian elephant RTE1_LA elements increased luciferase expression 3.03-fold (t-test, $P=2.41\times10^{-8}$) and 1.37-fold (t-test, $P=2.60\times10^{-4}$), respectively, compared to empty vector controls (**Fig. 5D**). However, luciferase expression from the pGL4.10[*luc2*] vector containing the RTE1_LA sequence was not significantly different than the empty vector control in either Asian (0.96-fold; t-test, $P=0.61$) or African elephant fibroblasts (0.95-fold; t-test, $P=0.37$; **Fig. 5D**). These data indicate that transcription of *TP53RTG12* likely initiates within a RTE1_LA-derived promoter, but that the ability of this RTE1_LA element to function as a promoter is not an ancestral feature of RTE1_LA elements.

Elephant cells express TP53RTG proteins and numerous isoforms of TP53

To determine if *TP53RTG* transcripts are translated, we treated African elephant, Asian elephant, and hyrax cells with 50 J/m^2 UV-C (to stabilize TP53) and the proteasome inhibitor MG-132 (to block protein degradation), and assayed for TP53/TP53RTG proteins by Western blotting total cell protein with a polyclonal TP53 antibody (FL-393) that we demonstrated recognizes Myc-tagged TP53RTG12 (**Fig. 8C**). We identified bands in both African and Asian

elephant and hyrax total cell protein at the expected size for the full length p53, Δ133p53β/γ, and p53ψ isoforms of canonical the TP53 protein (Khoury and Bourdon, 2010) as well as high molecular weight bands corresponding to previously reported SDS denaturation resistant TP53 oligomers (Cohen et al., 2008; Ottaggio et al., 2000) and (poly)ubiquitinated TP53 congugates (Sparks et al., 2013) (**Fig. 5E**). We also identified an elephant-specific band at the expected size for the TP53RTG12 (19.6kDa) protein suggesting that the *TP53RTG12* transcript is translated in elephant fibroblasts (**Fig. 5E**).

Elephant cells have an enhanced TP53-dependent DNA-damage response

Our observation that *TP53RTG* genes are expressed suggests that elephant cells may have an altered p53 signaling system compared to species without an expanded *TP53/TP53RTG* gene repertoire. To directly test this hypothesis we transiently transfected primary African elephant, Asian elephant, South African Rock hyrax (*Procavia capensis capensis*), East African aardvark (*Orycteropus afer lademanni*), and Southern Three-banded armadillo (*Tolypeutes matacus*) dermal fibroblasts with a luciferase reporter vector containing two TP53 response elements (pGL4.38[*luc2p*/p53 RE/Hygro]) and Renilla control vector (pGL4.74[*hRluc*/TK]). Next we used a dual luciferase reporter assay to measure activation of the TP53 pathway in response to treatment with three DNA damage inducing agents (mitomycin C, doxorubicin, or UV-C) or nutlin-3a, which inhibits the interaction between MDM2 and TP53. We found that elephant cells generally up-regulated TP53 signaling in response to lower doses of each drug and UV-C than closely related species without an expanded TP53 gene repertoire (**Fig. 6A**), indicating elephant cells have evolved an enhanced TP53 response.

To determine the consequences of an enhanced TP53 response we treated primary African and Asian elephant, hyrax, aardvark, and armadillo dermal fibroblasts with mitomycin C, doxorubicin, nutlin-3a, or UV-C and measured cell viability (live-cell protease activity), cytotoxicity (dead-cell protease activity), and the induction apoptosis (caspase-3/7 activation) using an ApoTox-Glo Triplex assay. Consistent with the results from the luciferase assay, we found that lower doses of mitomycin C or doxorubicin induced apoptosis in elephant cells than the other species (**Fig. 6B**) and that the magnitude of the response was greater in elephant than other species (**Fig. 6A**). Similarly UV-C exposure generally induced more elephant cells to undergo apoptosis than other species (**Fig. 6A**). A striking exception to this trend was the response of elephant cells to the MDM2 antagonist nutlin 3-a, which elicited a strong p53

transcriptional response (**Fig. 6A**) but did not induce apoptosis (**Fig. 6B**). Thus we conclude that elephant cells upregulate TP53 signaling and apoptosis at lower levels of DNA-damage than other species.

TP53RTG12 enhances TP53 signaling and DNA-damage responses via a transdominant mechanism

To test if TP53RTG12 is sufficient to mediate enhanced TP53 signaling and DNA-damage responses we synthesized the African elephant *TP53RTG12* gene (using mouse codon usage patterns) and cloned it into the mammalian expression vector pcDNA3.1(+)/myc-His. We then transiently transfected mouse 3T3-L1 cells with the *TP53RTG12* pcDNA3.1(+)/myc-His expression vector and used the pGL4.38[*luc2P*/p53 RE/Hygro] reporter system and ApoToxGlo assays to monitor activation of the TP53 signaling pathway and the induction of apoptosis in response to treatment with mitomycin C, doxorubicin, nutlin-3a, or UV-C. We found that heterologous expression of TP53RTG12 in mouse 3T3-L1 cells dramatically augmented luciferase expression from the pGL4.38[*luc2P*/p53 RE/Hygro] reporter vector in response to each treatment (**Fig. 7A**) compared to empty vector controls, consistent with a enhancement of the endogenous TP53 signaling pathway. Similarly, expression of TP53RTG12 significantly augmented the induction of apoptosis in response to each treatment (**Fig. 7B**). These data indicate that TP53RTG12 acts via a trans-dominant mechanism to enahnce the induction of apoptosis by endogenous TP53 and that TP53RTG12 is sufficient to recapitulate some of the enhanced sensitivity of elephant cells to DNA damage.

TP53RTG proteins are unlikely to directly regulate TP53 target genes because they lack critical residues required for nuclear localization, tetramerization, and DNA-binding (**Fig. 8A**). Previous studies, for example, have shown the TP53 mutants lacking the tetramerization domain C-terminal tail are unable to bind DNA, chromatin, and transactivate luciferase expression from a reporter vector containing TP53 response elements (Kim et al., 2012). Consistent with these observations, we found that HEK-293 cells co-transfected with the *TP53RTG12* pcDNA3.1(+)/myc-His expression vector and the pGL4.38[*luc2P*/p53 RE/Hygro] luciferase reporter vector did not have elevated luciferase expression compared to controls, indicating that TP53RTG12 itself is transcriptionally inactive (**Fig. 8B**). Many of the TP53RTG proteins, however, retain the MDM2 interaction sites in the transactivation domain and the dimerization sites in the DNA binding domain, including TP53RTG12 (**Fig. 8A**).

These data suggest two non-exclusive models of TP53RTG action: 1) TP53RTG proteins may act as 'decoys' for the MDM2 complex allowing the canonical TP53 protein to escape negative regulation; and 2) TP53RTG proteins may protect canonical TP53 from MDM2 mediated ubiqutination, which requires tetramerization (Kubbutat et al., 1998; Maki, 1999), by dimerizing with canonical TP53 and thereby preventing the formation of tetramers. To test these hypotheses we transiently transfected HEK-293 cells with the TP53RTG12 pcDNA3.1(+)/myc-His expression vector, immunoprecipitated TP53RTG12 with an α-Myc antibody, and assayed for co-immunoprecipitation of MDM2 or TP53 by Western blotting. We found that Myc-tagged TP53RTG12 efficiently co-immunoprecipitated endogenous human TP53 but not MDM2 (**Fig. 8C**). Next we repeated this experiment but immunoprecipitated endogenous human MDM2, as expected we co-immunoprecipitated endogenous human TP53, but only weakly co-immunoprecipitated Myc-tagged TP53RTG12 if at all (**Fig. 8C**). These data suggest that TP53RTG12 does not directly regulate TP53 target genes and mediates its effects by binding canonical TP53, however, this interaction alone is not sufficient to initiate the TP53 signaling cascade and the induction of apoptosis (**Fig. 8D**).

DISCUSSION

A major developmental and life history constraint on the evolution of large body sizes and long lifespans in animals is an increased risk of developing cancer. There is no correlation, however, between body-size or lifespan and cancer risk across species because large and long-lived organisms have evolved enhanced cancer suppression mechanisms that generally delay the development of cancer until post-reproduction (when selection generally cannot act). This simple evolutionary rational demands mechanistic explanations (Peto, 2015), which have thus far been elusive. Sequencing and analysis of the bowhead whale genome, for example, identified numerous whale specific variants but did not identify obvious genomic changes likely to reduce cancer susceptibility (Keane et al., 2015). Similarly no strong candidate loci likely to underlie cancer resistance were identified in the genomes of Brandt's (*Myotis brandtii*) and David's (*Myotis davidii*) bats, which have small body sizes yet extremely long lifespans (Seim et al., 2013; Zhang et al., 2013). In contrast several mechanisms contributing to enhanced cancer resistance have been identified in naked (*Heterocephalus glaber*) and blind mole rats (*Spalax sp.*), which like bats have very long lifespans for a small-bodied organism, including extremely sensitive contact inhibition (Seluanov et al., 2009; Tian et al., 2013) and enhanced TP53-

depenent necrotic cell death (Gorbunova et al., 2012), respectively.

Here we show that retrogene of the master tumor suppressor TP53, which is essential for preventing cancer because it triggers proliferative arrest and apoptosis in response to a variety of stresses such as DNA damage, rapidly increased in copy number through repeated segmental duplications during with the evolution of large-bodied Proboscideans. Our results also demonstrate that elephant cells induce p53 signaling and trigger apoptosis at lower thresholds of genotoxic stress than closely related species without an expanded p53 repertoire, suggesting that Proboscideans resolved Peto's paradox at least in part through the evolution of augmented p53 signaling. Consistent with this hypothesis heterologous expression of TP53RTG12 in mouse cells was sufficient to augment endogenous p53 signaling and recapitulate an elephant-like sensitivity to genotoxic stress, indicating that TP53RTGs acts through a transdominant mechanism.

A transdominant mechanism suggests that TP53RTG12, which is transcriptionally inactive and therefore unlikely to directly regulate endogenous TP53 target genes, interacts with and modifies the function of proteins in the TP53 signaling pathway. Indeed we found that Myc-tagged TP53RTG12 efficiently co-immunoprecipitated with endogenous human TP53 but at best only weakly interacted with endogenous human MDM2 (despite conservation of the N-terminal MDM2 interaction residues in TP53RTG12). Previous studies have shown that the C-terminus of TP53 contributes to the physical interaction with MDM2 and that tetramerization of TP53 is required for its efficient MDM2-medaited ubiquitination (Kubbutat et al., 1998; Maki, 1999). These data suggest that TP53RTG proteins are unable to efficiently bind MDM2 because they lack the C-terminal tetramerization domain, which may also prevent ubiquitination of TP53RTG12 and TP53/TP53RTG12 dimers and contribute to a standing pool of TP53 that is able to rapidly respond to DNA damage. Consistent with this mechanism, transgenic mice with an increase in TP53 copy number (Garcia-Cao et al., 2002) or a hypomorphic *Mdm2* allele have elevated basal TP53 activity and are resistant to tumor formation (Mendrysa, 2006), indicating that shifting the TP53-MDM2 equilibrium away from TP53 degradation can directly promote cancer resistance.

While our data implicates the expansion of the *TP53RTG* gene repertoire in the evolution of enhanced cancer resistance in elephants, numerous details about the mechanisms that

underlie enhanced cancer resistance remain to be determined. Our data, for example, suggests that elephant cells initiate TP53 signaling and induce apoptosis at the earliest signal of DNA-damage but the precise molecular route to apoptosis is not clear. Elephant cells also appear to simultaneously express numerous isoforms of TP53, which also likely contributes to their sensitivity to DNA damage. These changes apparently evolved without major life history tradeoffs such as shortened lifespan, accelerated aging, or reduced fertility (Kang et al., 2009; Smith et al., 2012), however, elephants thus far are the only species identified with an increased copy number of functional TP53 genes suggesting indicating increased dosage of TP53 is generally deleterious (Tyner et al., 2002). Thus Proboscideans likely broke a major developmental and evolutionary constraint on TP53 copy number during the evolution of large body size.

EXPERIMENTAL PROCEDURES

Identification of *TP53/PT53RTG* genes in Sarcopterygian genomes

We used BLAT to search for *TP53* genes in 61 Sarcopterygian genomes using the human TP53 protein sequences as an initial query. After identifying the canonical *TP53* gene from each species, we used the nucleotide sequences corresponding to this *TP53* CDS as the query sequence for additional BLAT searches within that species genome. To further confirm the orthology of each *TP53* gene we used a reciprocal best BLAT approach, sequentially using the putative CDS of each *TP53* gene as a query against the human genome; in each case the query gene was identified as TP53. Finally we used the putative amino acid sequence of the TP53 protein as a query sequence in a BLAT search.

We thus used BLAT to characterize the *TP53* copy number in Human (*Homo sapiens*; GRCh37/hg19), Chimp (*Pan troglodytes*; CSAC 2.1.4/panTro4), Gorilla (*Gorilla gorilla gorilla*; gorGor3.1/gorGor3), Orangutan (*Pongo pygmaeus abelii*; WUGSC 2.0.2/ponAbe2), Gibbon (*Nomascus leucogenys*; GGSC Nleu3.0/nomLeu3), Rhesus (*Macaca mulatta*; BGI CR_1.0/rheMac3), Baboon (*Papio hamadryas*; Baylor Pham_1.0/papHam1), Marmoset (*Callithrix jacchus*; WUGSC 3.2/calJac3), Squirrel monkey (*Saimiri boliviensis*; Broad/saiBol1), Tarsier (*Tarsius syrichta*; Tarsius_syrichta2.0.1/tarSyr2), Bushbaby (*Otolemur garnettii*; Broad/otoGar3), Mouse lemur (*Microcebus murinus*; Broad/micMur1), Chinese tree shrew (*Tupaia chinensis*; TupChi_1.0/tupChi1), Squirrel (*Spermophilus tridecemlineatus*; Broad/speTri2), Mouse (*Mus musculus*; GRCm38/mm10), Rat (*Rattus norvegicus*; RGSC 5.0/rn5), Naked mole-rat (*Heterocephalus glaber*; Broad HetGla_female_1.0/hetGla2), Guinea pig (*Cavia porcellus*; Broad/cavPor3), Rabbit (*Oryctolagus cuniculus*; Broad/oryCun2), Pika (*Ochotona princeps*; OchPri3.0/ochPri3), Kangaroo rat (*Dipodomys ordii*; Broad/dipOrd1), Chinese hamster (*Cricetulus griseus*; C_griseus_v1.0/criGri1), Pig (*Sus scrofa*; SGSC Sscrofa10.2/susScr3), Alpaca (*Vicugna pacos*; Vicugna_pacos-2.0.1/vicPac2), Dolphin (*Tursiops truncatus*; Baylor Ttru_1.4/turTru2), Cow (*Bos taurus*; Baylor Btau_4.6.1/bosTau7), Sheep (*Ovis aries*; ISGC Oar_v3.1/oviAri3), Horse (*Equus caballus*; Broad/equCab2), White rhinoceros (*Ceratotherium simum*; CerSimSim1.0/cerSim1), Cat (*Felis catus*; ICGSC Felis_catus 6.2/felCat5), Dog (*Canis lupus familiaris*; Broad CanFam3.1/canFam3), Ferret *(Mustela putorius furo*; MusPutFur1.0/musFur1), Panda (*Ailuropoda melanoleuca*; BGI-Shenzhen 1.0/ailMel1), Megabat (*Pteropus vampyrus*; Broad/pteVam1), Microbat (*Myotis lucifugus*; Broad Institute Myoluc2.0/myoLuc2), Hedgehog (*Erinaceus europaeus*;

EriEur2.0/eriEur2), Shrew (*Sorex araneus*; Broad/sorAra2), Minke whale (*Balaenoptera acutorostrata scammoni*; balAcu1), Bowhead Whale (*Balaena mysticetus*; v1.0), Rock hyrax (*Procavia capensis*; Broad/proCap1), Sloth (*Choloepus hoffmanni*; Broad/choHof1), Elephant (*Loxodonta africana*; Broad/loxAfr3), Cape elephant shrew (*Elephantulus edwardii*; EleEdw1.0/eleEdw1), Manatee (*Trichechus manatus latirostris*; Broad v1.0/triMan1), Tenrec (*Echinops telfairi*; Broad/echTel2), Aardvark (*Orycteropus afer afer*; OryAfe1.0/oryAfe1), Armadillo (*Dasypus novemcinctus*; Baylor/dasNov3), Opossum (*Monodelphis domestica*; Broad/monDom5), Tasmanian devil (*Sarcophilus harrisii*; WTSI Devil_ref v7.0/sarHar1), Wallaby (*Macropus eugenii*; TWGS Meug_1.1/macEug2), Platypus (*Ornithorhynchus anatinus*; WUGSC 5.0.1/ornAna1), Medium ground finch (*Geospiza fortis*; GeoFor_1.0/geoFor1), Zebra finch (*Taeniopygia guttata*; WashU taeGut324/taeGut2), Budgerigar (*Melopsittacus undulatus*; WUSTL v6.3/melUnd1), Chicken (*Gallus gallus*; ICGSC Gallus_gallus-4.0/galGal4), Turkey (*Meleagris gallopavo*; TGC Turkey_2.01/melGal1), American alligator (*Alligator mississippiensis*; allMis0.2/allMis1), Painted turtle (*Chrysemys picta bellii*; v3.0.1/chrPic1), Lizard (*Anolis carolinensis*; Broad AnoCar2.0/anoCar2), X. tropicalis (*Xenopus tropicalis*; JGI 7.0/xenTro7), Coelacanth (*Latimeria chalumnae*; Broad/latCha1).

TP53/TP53RTG **copy number estimation in Proboscidean genomes**

Identification of TP53/TP53RTG genes in the Asian elephant genome

We used previously published whole genome shotgun sequencing data from an Asian elephant (*Elephas maximus*) generated on an Illumina HiSeq 2000{'Hou'} to estimate the *TP53/TP53RTG* copy number in the genome. For these analyses we combined data from two individual elephants (ERX334765 and ERX334764) into a single dataset of 151,482,390 76-nt paired end reads. We then mapped Asian elephant reads onto the African elephant *TP53/TP53RTG* genes using Bowtie2 in paired-end mode, with the local alignment and 'very sensitive' options and Cufflinks (version 0.0.7) to assemble mapped reads into genes which we treated as putative Asian elephant orthologs. We identified 14 putative 1:1 orthologous *TP53RTG* genes in the Asian elephant genome, including *TP53RTG1*, *TP53RTG2*, *TP53RTG3*, *TP53RTG4*, *TP53RTG5*, *TP53RTG6*, *TP53RTG9*, *TP53RTG10*, *TP53RTG11*, *TP53RTG13*, *TP53RTG14*, *TP53RTG17*, *TP53RTG18*, and *TP53RTG19*.

Identification of TP53/TP53RTG genes in the Columbian and Woolly mammoth genomes

We used previously published whole genome shotgun sequencing data from an ~11,000 year old Columbian mammoth (*Mammuthus columbi*) generated on an Illumina HiSeq 1000{'Sparks'} to estimate the *TP53/TP53RTG* copy number in the genome. For these analyses we combined data from three individual mammoths (SRX329134, SRX329135, SRX327587, SRX327586, SRX327583, SRX327582) into a single dataset of 158,704,819 102-nt paired end reads. We then mapped Asian elephant reads onto the African elephant *TP53/TP53RTG* genes using Bowtie2 in paired-end mode, with the local alignment and 'very sensitive' options, and Cufflinks (version 0.0.7) to assemble mapped reads into genes which we treated as putative Asian elephant orthologs. We identified 14 putative 1:1 orthologous *TP53RTG* genes in the Columbian mammoth genome, including *TP53RTG10*, *TP53RTG11*, *TP53RTG13*, *TP53RTG14*, *TP53RTG16*, *TP53RTG17*, *TP53RTG19*, *TP53RTG2*, *TP53RTG3*, *TP53RTG4*, *TP53RTG5*, *TP53RTG6*, *TP53RTG8*, *TP53RTG9*.

Identification of TP53/TP53RTG genes in the American mastodon genome

We used previously published whole genome shotgun sequencing data generated on a Roche 454 Genome Sequencer (GS FLX) to estimate the *TP53* copy number in a 50,000-130,000 year old American mastodon (*Mammut americanum*){'Langmead'}. To estimate *TP53* copy number we combined 454 sequencing data from 'Library A' (SRX015822) and 'Library B' (SRX015823) into a single dataset of 518,925 reads. Reads were converted from FASTQ to FASTA and aligned to African elephant *TP53* gene and retrogene sequences using Lastz. We used the 'Roche-454 90% identity' mapping mode, not reporting matches lower than 90% identity. Three American mastodon reads were identified that mapped to at least one *TP53* gene/retrogene. The likely identity of these reads were determined by using BLAT to map their location in the African elephant (Broad/loxAfr3) genome; mastodon reads mapping to a single elephant gene with >98% identity were considered likely mastodon orthologs. This strategy mapped one read to the *TP53TRG8* retrogene, one read to the either the *TP53TRG3* or *TP53RTG10* retrogenes, and one read to the *TP53TRG6* retrogene.

Gene tree reconciliation

For gene tree reconciliation putative Asian elephant, Columbian mammoth, and woolly mammoth *TP53*/*TP53RTG* orthologs were manually placed into the African elephant *TP53*/*TP53RTG* gene tree. We then used Notung v2.6 (Trapnell et al., 2012) to reconcile the gene and species trees, and interpreted Asian elephant-, Columbian mammoth-, and wooly

mammoth-specific gene losses as unsampled *TP53RTG* genes if the putative divergence date of the ortholog predated the African-(Asian elephant/Mammoth) divergence. We thus inferred two unsampled genes (*TP53RTG7* and *TP53RTG8*) in the Asian elephant genome, and three unsampled genes (*TP53RTG18*, *TP53RTG7* and *TP53RTG1*) in the Columbian mammoth genome. Note that this method will estimate the *minimum* number of *TP53*/*TP53RTG* genes in the genome because our mapping strategy cannot identify lineage specific duplications after each species diverged from the African elephant lineage. Finally we used Notung v2.6 (Wilkie et al., 2013) to reconcile the gene and species trees, and interpreted lineage-specific gene losses as unsampled *TP53RTG* genes (i.e., genes present in the genome but missing from the aDNA sequencing data). This process was modified slightly for American mastodon, such that 'lost' *TP53RTG* genes younger than the mastodon-elephant split were considered elephant-specific paralogs rather than unsampled genes present in the mastodon genome.

Normalized read depth

We also used normalized mapped read depth to estimate *TP53*/*TP53RTG* copy number. Unlike gene tree reconciliation, copy number estimates based on read depth cannot resolve the orthology of specific *TP53*/*TP53RTG* copies but can estimate the total number of *TP53*/*TP53RTG* copies in the genome and thus provide evidence of lineage specific copy number changes. For this analyses we mapped Asian elephant-, Columbian mammoth-, and woolly mammoth-specific reads onto the African elephant *TP53/TP53RTG* genes using Bowtie 2 (Langmead and Salzberg, 2012) with the local alignment and 'very sensitive' options and Cufflinks (version 0.0.7) (Trapnell et al., 2012) to assemble mapped reads into genes. We then summed the read depth ('coverage') of *TP53RTG* genes and normalized this summed *TP53RTG* read depth to the read depth of five single copy genomic regions equal in length to average *TP53RTG* gene length. This ratio represents an estimate of the *TP53/TP53RTG* copy number in the genome.

Correlation between *TP53/TP53RTG* copy number and body size evolution

Relative divergence (duplication) times of the TP53 retrogenes were estimated using the TP53 alignment described above and BEAST (v1.7.4) (Rohland et al., 2010). We used the general time reversible model (GTR), empirical nucleotide frequencies (+F), a proportion of invariable sites estimated from the data (+I), four gamma distributed rate categories (+G), an uncorrelated lognormal relaxed molecular clock to model substitution rate variation across

lineages, a Yule speciation tree prior, uniform priors for the GTR substitution parameters, gamma shape parameter, proportion of invariant sites parameter, and nucleotide frequency parameter. We used an Unweighted Pair Group Arithmetic Mean (UPGMA) starting tree.

To obtain posterior distributions of estimated divergence times, we use five node calibrations modeled as normal priors (standard deviation = 1) to constrain the age of the root nodes for the Eutheria (104.7 MYA), Laurasiatheria (87.2 MYA), Boreoeutherian (92.4 MYA), Atlantogenatan (103 MYA), and Paenungulata (64.2 MYA); divergence dates were obtained from www.timetree.org using the 'Expert Result' divergence dates. The analysis was run for 5 million generations and sampled every 1000 generations with a burn-in of 1 million generations; convergence was assessed using Tracer, which indicated convergence was reached rapidly (within 100,000 generations). Proboscidean body size data were obtained from previously published studies on mammalian body size evolution (Chen et al., 2000).

Inference of *TP53/TP53RTG* copy number expansion through segmental duplications

We also observed that the genomic region surrounding each *TP53RTG* gene contained blocks of homolgous transposable element insertions, suggesting that these regions are segmental duplications. To confirm this observation we used MUSCLE (Edgar, 2004) to align an approximately 20kb region surrounding each *TP53RTG* gene and found that conservation within this region was very high, again suggesting these regions are relatively recent segmental duplications. To identify if the contigs on which *TP53RTG* genes are located contained locally collinear blocks (LCBs), as expected for segmental duplications, we aligned contigs using progressiveMAUVE (Darling et al., 2004) as implemented in Genious (v6.1.2).

Phylogenetic analyses of *TP53/TP53RTG* genes

We generated a dataset of *TP53* orthologs from 65 diverse mammals identified from GenBank, and included the *TP53* genes and retrogenes we identified from the African elephant, hyrax, manatee, tenrec, cape elephant shrew, and armadillo genomes. Nucleotide sequences were aligned using the MAFFT algorithm (Katoh and Standley, 2014) and the FFT refinement strategy implemented in the GUIDANCE webserver (Penn et al., 2010). Alignment confidence was assessed with 100 bootstrap replicates and ambiguously aligned sites (under the default GUIDANCE exclusion rule) removed prior to phylogenetic analyses; lineage specific insertions and deletions were also removed prior to phylogenetic analyses.

TP53 phylogenies were inferred using maximum likelihood implemented in PhyML (v3.1) (Guindon et al., 2010) using a general time reversible model (GTR), empirical nucleotide frequencies (+F), a proportion of invariable sites estimated from the data (+I), four gamma distributed rate categories (+G), and using the best of NNI and SPR branch moves during the topology search. Branch supports were assessed using the aBayes, aLRT, SH-like, and Chi2-based 'fast' methods as well as 1000 bootstrap replicates (Guindon et al., 2010). We also used MrBayes (v3.2.2 x64) (Hulsenbeck and Ronquist, 2001)for Bayesian tree inference using the GTR+I+G model; the analysis was run with 2 simultaneous runs of 4 chains for 4 million generations with the chains sampled every 1000 generations. We used Tracer (v1.5) to assess when the chains reached stationarity, which occurred around generation 100,000. At completion of the run the PRSF value was 1.00 indicating the analyses had reached convergence.

Gene expression data (RNA-Seq)

To determine if *TP53RTG* genes were expressed, we generated RNA-Seq data from term placental villus and adipose tissue from African elephants (*Loxodonta africana*) and primary dermal fibroblasts from Asian elephants (*Elepas maximus*). Briefly, sequencing libraries were prepared using standard Illumina protocols with poly(A) selection, and sequenced as 100bp single-end reads on a HiSeq2000. We also used previously published RNA-Seq data generated from primary fibroblasts (isolated from ear clips) from a male (GSM1227965) and female (GSM1227964) African elephant generated on an Illumina Genome Analyzer IIx (101 cycles, single end) and a male (GSM1278046) African elephant generated on an Illumina HiSeq 2000 (101 cycles, single end) (Cortez et al., 2014) and combined these reads into a single dataset of 138,954,285 reads.

Reads were aligned to a custom built elephant reference gene set generated by combining the sequences of the canonical *TP53* gene and *TP53RTG* genes with the ENSEMBL African elephant (*Loxodonta africana*) CDS gene build (Loxodonta_africana.loxAfr3.75.cds.all.fa) with Bowtie2 (Langmead and Salzberg, 2012). Bowtie 2 settings were: 1) both local alignment and end-to-end mapping; 2) preset option: sensitive; 3) Trim n-bases from 5' of each read: 0; and 4) Trim n-bases from 3' of each read: 0. Transcript assembly and FPKM estimates were generated with Cufflinks (version 0.0.7) (Trapnell et al., 2012) using aligned reads from Bowtie2, non-default parameters included quartile normalization

and multi-read correction. Finally we transformed FPKM estimates into transcripts per million (TPM), TPM=(FPKM per gene/sum FPKM all genes)x10^6, and defined genes with TPM≥2 as expressed (Li et al., 2010; Wagner et al., 2012; 2013).

The TP53RTG genes are 80.0-82.7% identical to TP53 at the nucleotide level, with 204-231 total nucleotide differences compared to TP53. This level of divergence allows for many reads to be uniquely mapped to each gene, there will also be significant read mapping uncertainty in regions of these genes with few nucleotide differences. However, if read mapping uncertainty was leading to false positive mappings of TP53 derived reads to TP53RTG genes we would expect to observe expression of many TP53RTG genes, rather the robust expression of a single (TP53RTG12) gene. We also counted the number of uniquely mapped reads to each TP53RTG gene and TP53. We found that 0-8 reads were uniquely mapped to most TP53RTG genes, except TP53RTG12 which had ~115 uniquely mapped reads and TP53 which had ~3000 uniquely mapped reads across our tissue samples. Thus we conclude that read mapping uncertainty has not adversely affected our RNA-Seq analyses.

Gene expression data (PCR and Sanger sequencing)

We further confirmed expression of TP53RTG12 transcripts in elephant cells through RT-PCR, taking advantage of differences between TP53RTG12 and TP53 sequences to design two aligned primer sets, one TP53-specific, the other TP53RTG12 specific. TP53RTG12 primers were: 5´ *ggg gaa act cct tcc tga ga* 3´ (forward) and 5´ *cca gac aga aac gat agg tg* 3´ (reverse). TP53 primers were: 5´ *atg gga act cct tcc tga ga* 3´ (forward) and 5´ *cca gac gga aac cat agg tg* 3´ (reverse). The TP53 amplicon is expected to be 251 bps in length, while deletions present in the TP53RTG12 sequence lead to a smaller projected amplicon size of 220 bps.

Total RNA was extracted from cultured *Loxodonta* and *Elephas* fibroblasts (RNAeasy Plus Mini kit, Qiagen), then DNase treated (Turbo DNA-*free* kit, Ambion) and reverse-transcribed using an olgio-dT primer for cDNA synthesis (Maxima H Minus First Strand cDNA Synthesis kit, Thermo Scientific). Control RT reactions were otherwise processed identically, except for the omission of reverse transcriptase from the reaction mixture. RT products were PCR-amplified for 45 cycles of 94°/20 seconds, 56°/30 seconds, 72°/30 seconds using a BioRad CFX96 Real Time qPCR detection system and SYBR Green master mix (QuantiTect, Qiagen). PCR products were electrophoresed on 3% agarose gels for 1 hour at 100 volts, stained with SYBR safe, and

imaged in a digital gel box (ChemiDoc MP, BioRad) to visualize relative amplicon sizes. PCR products were also directly sequenced at the University of Chicago Genomics core facility, confirming projected product sizes and sequence identities.

Gene prediction, promoter annotation, and TP53RTG transcript identification

We used geneid v1.2 (http://genome.crg.es/software/geneid/geneid.html) to infer if the *TP53RTG12* gene contained a non-coding exon 5' to the predicted ATG start codon. For gene structure prediction we used the full-length scaffold_825 from African elephant (Broad/loxAfr3) sequence and forced an internal exon where the *TP53RTG12* gene is encoded in scaffold_825. Geneid identified a putative exon 5' to the *TP53RTG12* gene from nucleotides 1761-1935 and an exon 3' from nucleotides 6401-6776. We also used GENESCAN (http://genes.mit.edu/cgi-bin/genscanw_py.cgi) to predict the location of exons in the full-length scaffold_825 sequence and identified a putative exon from nucleotides 1750-1986. We next used Bowtie2 (Langmead and Salzberg, 2012) to map African and Asian elephant fibroblast RNA-Seq data onto African elephant scaffold_825 with the default settings.

Taxonomic distribution of the RTE1_LA retrotransposon

The RTE1_LA non-LTR retrotransposon has previously been described from the African elephant genome, these elements are generally more than 90% identical to the consensus RTE1_LA sequence but less than 70% identical to other mammalian RTEs (http://www.girinst.org/2006/vol6/issue3/RTE1_LA.html). These data suggest that the RTE_LA element has relatively recently expanded in the elephant genome. To determine the taxonomic distribution of the RTE1_LA element we used BLAT to search the lesser hedgehog tenrec (*Echinops telfairi*), rock hyrax (*Procavia capensis*), West Indian manatee (*Trichechus manatus*), armadillo (*Dasypus novemcinctus*), and sloth (*Choloepus hoffmanni*) genomes. We identified numerous copies of the RTE_LA element in the genome of the Afrotherians, but not the Xenarthran genomes. These data indicate that the RTE_LA element is Afrotherian-specific, rather than Elephant-specific.

TP53/TP53RTG Western blotting

Elephant and Hyrax primary fibroblasts (San Diego Zoo, "Frozen Zoo") were grown to confluency in 10cm dishes at 37°C/5% CO_2 in a culture medium consisting of FGM/EMEM (1:1) supplemented with insulin, FGF, 6% FBS and Gentamicin/Amphotericin B (FGM-2, singlequots, Clonetics/Lonza). Culture medium was removed from dishes just prior to UV treatment and

returned to cells shortly afterwards. Experimental cells were exposed to 50J/m^2 UV-C radiation in a crosslinker (Stratalinker 2400, Stratagene), while control cells passed through media changes but were not exposed to UV. A small volume (~3mL) of PBS covered fibroblasts at the time of UV exposure. To inhibit TP53 proteolysis, MG-132 (10μM) was added to experimental cell medium 1 hour prior to UV exposure, and maintained until the time of cell lysis. 5h post-UV treatment, cells were briefly rinsed in PBS, then lysed and boiled in 2x SDS-PAGE sample buffer. Lysates were separated via SDS-PAGE on 10% gels for 1h at 140 volts, then electrophoretically transferred to PVDF membranes (1h at 85 volts). Membranes were blocked for 1h in 5% milk in TBST and incubated overnight at 4°C with rabbit polyclonal TP53 antibodies (FL-393, Santa Cruz Biotechnology, and ab131442, Abcam). Blots were washed 3x in TBST, incubated with HRP-conjugated, anti-rabbit IgG 2° antibodies for 1 h at RT, and washed 4 more times in TBST. Protein bands were visualized via enhanced chemiluminescence (BioRad Clarity), and imaged in a digital gel box (Chemidoc MP, BioRad). Western blots were replicated three independent times.

Co-immunoprecipitation

Human HEK-293 cells were grown to 80% confluency in 20cm dishes at 37°C/5% CO_2 in a culture medium consisting of DMEM supplemented with 10% FBS. At 80% confluency, cells were transiently transfected with the *TP53RTG12* pcDNA3.1(+)/myc-His expression vector. After 16 hours the transfection media was removed and replaced with fresh DMEM, and the cells were incubated an additional 24 hours before harvesting. After removing DMEM and washing cells twice with PBS, 1 mL ice-cold lysis buffer (20 mM Tris, pH 8.0, 40 mM KCl, 10 mM MgCl2, 10% glycerol, 1% Triton X-100, 1x Complete EDTA-free protease inhibitor cocktail (Roche), 1x PhosSTOP (Roche)) was added to each plate and cells were harvested by scraping with a rubber spatula. Cells were then incubated on ice for 30 minutes in 420 mM NaCl. Whole cell lysate was cleared by centrifugation at 10,000 rpm for 30 minutes at 4°C, and supernatant was transferred to a clean microfuge tube. After equilibrating protein concentrations, 1 mL of sample was mixed with 40 mL of α-MDM2 or α-Myc antibody conjugated agarose beads (Sigma) pre-washed with TNT buffer (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 0.05% Triton X-100), and rotated overnight at 4°C. The following day, samples were treated with 50 U DNase (Roche) and 2.5 µg RNase (Roche) for 60 minutes at room temperature, as indicated. Samples were washed 3x with 1 mL wash buffer (150 mM NaCl, 0.5% Triton X-100). After the final wash,

agarose beads were resuspended in elution buffer (500 mM Tris pH 7.5, 1 M NaCl), and boiled to elute immunoprecipitated complexes. Eluted protein was run on Bis-tris gels, probed with antibodies and visualized by Chemi-luminescence. Serial Westerns were performed for each antibody following chemical stripping and re-blocking. Antibodies were from Santa Cruz: MDM2 (SMP14) sc-965, lot #J2314; p53 (Fl-393) sc-6243, lot # D0215; c-Myc (9E10) sc-40, lot # G2413.

ApoTox-Glo Viability/Cytotoxicity/Apoptosis Analyses

Primary fibroblasts were grown to 80% confluency in T-75 culture flasks at 37°C/5% CO_2 in a culture medium consisting of FGM/EMEM (1:1) supplemented with insulin, FGF, 6% FBS and Gentamicin/Amphotericin B (FGM-2, singlequots, Clonetics/Lonza). 10^4 cells were seeded into each well of an opaque bottomed 96-well plate, leaving a column with no cells (background control); each 96-well plate contained paired elephant and hyrax samples. Serial dilutions of Doxorubicin (0uM, 0.5uM, 1.0uM, 5uM, 10uM and 50uM), Mitomycin c (0uM, 0.5uM, 1.0uM, 5uM, 10uM and 50uM), and Nutlin-3a (0uM, 0.5uM, 1.0uM, 5uM, 10uM and 50uM) and 90% culture media were added to each well such that there were four biological replicates for each condition. After 18 hrs of incubation with each drug, cell viability, cytotoxicity, and caspase-3/7 activity were measured using the ApoTox-Glo Triplex Assay (Promega) in a GloMax-Multi+ Reader (Promega). Data were standardized to no-drug control cells. For UV-C treatment, culture medium was removed from wells prior to UV treatment and returned to cells shortly afterwards. Experimental cells were exposed to 50J/m^2 UV-C radiation in a crosslinker (Stratalinker 2400, Stratagene), while control cells passed through media changes but were not exposed to UV. A small volume (~30uL) of PBS covered fibroblasts at the time of UV exposure. Cell viability, cytotoxicity, and caspase-3/7 activity were measured using the ApoTox-Glo Triplex Assay (Promega) in a GloMax-Multi+ Reader (Promega) 6, 12, 28.5 and 54.5 hours after UV-C treatment. Data were standardized to no UV-C exposure control cells. ApoTox-Glo Triplex Assays were replicated three independent times.

Luciferase assays

Primary fibroblasts were grown to 80% confluency in T-75 culture flasks at 37°C/5% CO_2 in a culture medium consisting of FGM/EMEM (1:1) supplemented with insulin, FGF, 6% FBS and Gentamicin/Amphotericin B (FGM-2, singlequots, Clonetics/Lonza). At confluency, cells

were trypsinized, centrifuged at 90g for 10min and resuspended in nucleofection/supplement solution and incubated for 15min. After incubation 1ug of the pGL4.38[*luc2p*/p53 RE/Hygro] luciferase reporter vector and 100ng of the pGL4.74[*hRluc/*TK] Renilla reporter vector were transiently transfected into elephant and hyrax cells using the Amaxa Basic Nucleofector Kit (Lonza) using protocol T-016. Immediately following nuleofection, 10^4 cells were seeded into each well of an opaque bottomed 96-well plate, leaving a column with no cells (background control); each 96-well plate contained paired elephant and hyrax samples. 24 hours after nucleofection cells were treated with either vehicle control, Doxorubicin, Mitomycin c, Nutlin-3a, or 50J/m^2 UV-C. Luciferase expression was assayed 18 hours after drug/UV-C treatment cells, using the Dual-Luciferase Reporter Assay System (Promega) in a GloMax-Multi+ Reader (Promega). For all experiments luciferase expression was standardized to Renilla expression to control for differences nucleofection efficiency across samples; Luc./Renilla data is standardized to (Luc./Renilla) expression in untreated control cells. Each luciferase experiment was replicated three independent times.

AUTHOR CONTRIBUTIONS

MS, LF, KM, LY, NPM, RDE, and SC performed experiments. VJL and MS designed the study, performed genomic analyses, and wrote the manuscript. All authors have read and approve the manuscript.

ACKNOWLEDGEMENTS

We acknowledge the financial support of The University of Chicago (VJL) and the University of Nottingham (NPM, LY, RDE). RDE was additionally supported by funding through the Advanced Data Analysis Centre. African Elephant Adipose tissue samples used for mRNA-seq were kindly provided through collaboration with T. Allen (The Paul Mellon Laboratory, Newmarket, Suffolk, United Kingdom) and F. Stansfield (The Elephant Research and Conservation Unit, Savé Valley Conservancy, Harare, Zimbabwe).

REFERENCES

Aguirre, A.A., Bröjer, C., and Mörner, T. (1999). Descriptive epidemiology of roe deer mortality in Sweden. J. Wildl. Dis. *35*, 753–762.

Cairns, J. (1975). Mutation selection and the natural history of cancer. Nature *255*, 197–200.

Caulin, A.F., and Maley, C.C. (2011). Peto“s Paradox: evolution”s prescription for cancer prevention. Trends Ecol. Evol. (Amst.) *26*, 175–182.

Chen, K., Durand, D., and Farach-Colton, M. (2000). NOTUNG: a program for dating gene duplications and optimizing gene family trees. J. Comput. Biol. *7*, 429–447.

Ciotta, C., Dogliotti, E., and Bignami, M. (1995). Mutation analysis in two newly identified rat p53 pseudogenes. Mutagenesis *10*, 123–128.

Cohen, M., Wuillemin, C., and Bischof, P. (2008). Trophoblastic p53 is stabilised by a cis-trans isomerisation necessary for the formation of high molecular weight complexes involving the N-terminus of p53. Biochimie *90*, 855–862.

Cortez, D., Marin, R., Toledo-Flores, D., Froidevaux, L., Liechti, A., Waters, P.D., Grutzner, F., and Kaessmann, H. (2014). Origins and functional evolution of Y chromosomes across mammals. Nature *508*, 488–493.

Czosnek, H.H., Bienz, B., Givol, D., Zakut-Houri, R., Pravtcheva, D.D., Ruddle, F.H., and Oren, M. (1984). The gene and the pseudogene for mouse p53 cellular tumor antigen are located on different chromosomes. Mol. Cell. Biol. *4*, 1638–1640.

Darling, A.C.E., Mau, B., Blattner, F.R., and Perna, N.T. (2004). Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Res. *14*, 1394–1403.

Doll, R. (1971). The age distribution of cancer: implications for models of carcinogenesis. Journal of the Royal Statistical Society Series a (….

Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res.

Enk, J.M., Devault, A.M., Kuch, M., and Murgha, Y.E. (2014). Ancient whole genome enrichment using baits built from modern DNA. Molecular Biology and ….

Enk, J., Rouillard, J.-M., and Poinar, H. (2013). Quantitative PCR as a predictor of aligned ancient DNA read counts following targeted enrichment. BioTechniques *55*, 300–309.

Evans, A.R., Jones, D., Boyer, A.G., Brown, J.H., Costa, D.P., Ernest, S.K.M., Fitzgerald, E.M.G., Fortelius, M., Gittleman, J.L., Hamilton, M.J., et al. (2012). The maximum rate of mammal evolution. *109*, 4187–4190.

Garcia-Cao, I., García-Cao, M., Martín-Caballero, J., Criado, L.M., Klatt, P., Flores, J.M., Weill, J.-C., Blasco, M.A., and Serrano, M. (2002). “Super p53” mice exhibit enhanced DNA damage

response, are tumor resistant and age normally. Embo J. *21*, 6225–6235.

Gorbunova, V., Hine, C., Tian, X., Ablaeva, J., Gudkov, A.V., Nevo, E., and Seluanov, A. (2012). Cancer resistance in the blind mole rat is mediated by concerted necrotic cell death mechanism. Proc. Natl. Acad. Sci. U.S.a. *109*, 19392–19396.

Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst. Biol. *59*, 307–321.

Healy, K., Guillerme, T., Finlay, S., Kane, A., Kelly, S.B.A., McClean, D., Kelly, D.J., Donohue, I., Jackson, A.L., and Cooper, N. (2014). Ecology and mode-of-life explain lifespan variation in birds and mammals. Proc. Biol. Sci. *281*, 20140298.

Hulla, J.E. (1992). The rat genome contains a p53 pseudogene: detection of a processed pseudogene using PCR. PCR Methods Appl. *1*, 251–254.

Hulsenbeck, J.P., and Ronquist, F. (2001). MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics *17*, 754–755.

Kang, H.-J., Feng, Z., Sun, Y., Atwal, G., Murphy, M.E., Rebbeck, T.R., Rosenwaks, Z., Levine, A.J., and Hu, W. (2009). Single-nucleotide polymorphisms in the p53 pathway regulate fertility in humans. Proc. Natl. Acad. Sci. U.S.a. *106*, 9761–9766.

Katoh, K., and Standley, D.M. (2014). MAFFT: iterative refinement and additional methods. Methods Mol. Biol. *1079*, 131–146.

Katzourakis, A., Magiorkinis, G., Lim, A.G., Gupta, S., Belshaw, R., and Gifford, R. (2014). Larger Mammalian body size leads to lower retroviral activity. PLoS Pathog. *10*, e1004214.

Keane, M., Semeiks, J., Webb, A.E., Li, Y.I., Quesada, V., Craig, T., Madsen, L.B., van Dam, S., Brawand, D., Marques, P.I., et al. (2015). Insights into the evolution of longevity from the bowhead whale genome. Cell Rep *10*, 112–122.

Khoury, M.P., and Bourdon, J.C. (2010). The isoforms of the p53 protein. Cold Spring Harbor ….

Kim, H., Kim, K., Choi, J., Heo, K., Baek, H.J., Roeder, R.G., and An, W. (2012). p53 requires an intact C-terminal domain for DNA binding and transactivation. J. Mol. Biol. *415*, 843–854.

Kubbutat, M., Ludwig, R.L., and Ashcroft, M. (1998). Regulation of Mdm2-directed degradation by the C terminus of p53. Molecular and Cellular ….

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods.

Leroi, A.M., Koufopanou, V., and Burt, A. (2003). Cancer selection. Nat. Rev. Cancer *3*, 226–231.

Li, B., Ruotti, V., Stewart, R.M., Thomson, J.A., and Dewey, C.N. (2010). RNA-Seq gene

expression estimation with read mapping uncertainty. Bioinformatics *26*, 493–500.

Maki, C.G. (1999). Oligomerization is required for p53 to be efficiently ubiquitinated by MDM2. J. Biol. Chem.

Martineau, D., Lemberger, K., Dallaire, A., Labelle, P., Lipscomb, T.P., Michel, P., and Mikaelian, I. (2002). Cancer in wildlife, a case study: beluga from the St. Lawrence estuary, Québec, Canada. Environ. Health Perspect. *110*, 285–292.

Mendrysa, S.M. (2006). Tumor suppression and normal aging in mice with constitutively high p53 activity. Genes Dev. *20*, 16–21.

Nagy, J.D., Victor, E.M., and Cropper, J.H. (2007). Why don“t all whales have cancer? A novel hypothesis resolving Peto”s paradox. Integr. Comp. Biol. *47*, 317–328.

Nunney, L. (1999). Lineage selection and the evolution of multistage carcinogenesis. Proc. Biol. Sci. *266*, 493–498.

Ottaggio, L., Bozzo, S., Moro, F., Sparks, A., Campomenosi, P., Miele, M., Bonatti, S., Fronza, G., Lane, D.P., and Abbondandolo, A. (2000). Defective nuclear localization of p53 protein in a Chinese hamster cell line is associated with the formation of stable cytoplasmic protein multimers in cells with gene amplification. Carcinogenesis *21*, 1631–1638.

Penn, O., Privman, E., Ashkenazy, H., and Landan, G. (2010). GUIDANCE: a web server for assessing alignment confidence scores. Nucleic Acids ….

Peto, R. (2015). Quantitative implications of the approximate irrelevance of mammalian body size and lifespan to lifelong cancer risk. Philos. Trans. R. Soc. Lond., B, Biol. Sci. *370*.

R Peto, F.J.R.P.N.L.L.L.J.C. (1975). Cancer and ageing in mice and men. British Journal of Cancer *32*, 411.

Rohland, N., Reich, D., Mallick, S., Meyer, M., Green, R.E., Georgiadis, N.J., Roca, A.L., and Hofreiter, M. (2010). Genomic DNA sequences from mastodon and woolly mammoth reveal deep speciation of forest and savanna elephants. PLoS Biol. *8*, e1000564.

Seim, I., Fang, X., Xiong, Z., Lobanov, A.V., and Huang, Z. (2013). Genome analysis reveals insights into physiology and longevity of the Brandt's bat Myotis brandtii. Nature.

Seluanov, A., Hine, C., Azpurua, J., Feigenson, M., Bozzella, M., Mao, Z., Catania, K.C., and Gorbunova, V. (2009). Hypersensitivity to contact inhibition provides a clue to cancer resistance of naked mole-rat. Proc. Natl. Acad. Sci. U.S.a. *106*, 19352–19357.

Smith, F.A., Boyer, A.G., Brown, J.H., Costa, D.P., Dayan, T., Ernest, S.K.M., Evans, A.R., Fortelius, M., Gittleman, J.L., Hamilton, M.J., et al. (2010). The evolution of maximum body size of terrestrial mammals. Science *330*, 1216–1219.

Smith, K.R., Hanson, H.A., Mineau, G.P., and Buys, S.S. (2012). Effects of BRCA1 and BRCA2 mutations on female fertility. Proc. Biol. Sci. *279*, 1389–1395.

Sookias, R.B., Benson, R., and Butler, R.J. (2012). Biology, not environment, drives major patterns in maximum tetrapod body size through time. Biology Letters.

Sparks, A., Dayal, S., Das, J., Robertson, P., Menendez, S., and Saville, M.K. (2013). The degradation of p53 and its major E3 ligase Mdm2 is differentially dependent on the proteasomal ubiquitin receptor S5a. Oncogene.

Tanooka, H., Ootsuyama, A., Shiroishi, T., and Moriwaki, K. (1995). Distribution of the p53 pseudogene among mouse species and subspecies. Mammalian Genome.

Tian, X., Azpurua, J., Hine, C., Vaidya, A., Myakishev-Rempel, M., Ablaeva, J., Mao, Z., Nevo, E., Gorbunova, V., and Seluanov, A. (2013). High-molecular-mass hyaluronan mediates the cancer resistance of the naked mole rat. Nature *499*, 346–349.

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., Pimentel, H., Salzberg, S.L., Rinn, J.L., and Pachter, L. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc *7*, 562–578.

Tyner, S.D., Venkatachalam, S., Choi, J., Jones, S., Ghebranious, N., Igelmann, H., Lu, X., Soron, G., Cooper, B., Brayton, C., et al. (2002). p53 mutant mice that display early ageing-associated phenotypes. Nature *415*, 45–53.

Wagner, G.P., Kin, K., and Lynch, V.J. (2012). Measurement of mRNA abundance using RNA-seq data: RPKM measure is inconsistent among samples. Theory Biosci. *131*, 281–285.

Wagner, G.P., Kin, K., and Lynch, V.J. (2013). A model based criterion for gene expression calls using RNA-seq data. Theory Biosci. *132*, 159–164.

Weghorst, C.M., Buzard, G.S., Calvert, R.J., and Hulla, J.E. (1995). Cloning and sequence of a processed p53 pseudogene from rat: a potential source of false “mutations” in PCR fragments of tumor DNA. Gene.

Wilkie, G.S., Davison, A.J., Watson, M., Kerr, K., Sanderson, S., Bouts, T., Steinbach, F., and Dastjerdi, A. (2013). Complete genome sequences of elephant endotheliotropic herpesviruses 1A and 1B determined directly from fatal cases. J. Virol. *87*, 6700–6712.

Zakut-Houri, R., Oren, M., Bienz, B., Lavie, V., Hazum, S., and Givol, D. (1983). A single gene and a pseudogene for the cellular tumour antigen p53. Nature *306*, 594–597.

Zhang, G., Cowled, C., Shi, Z., Huang, Z., Bishop-Lilly, K.A., Fang, X., Wynne, J.W., Xiong, Z., Baker, M.L., Zhao, W., et al. (2013). Comparative analysis of bat genomes provides insight into the evolution of flight and immunity. Science *339*, 456–460.

FIGURE LEGENDS

Figure 1. Body size evolution in vertebrates.

(A) Relationship between body mass (g) and lifespan (years) among 2,556 vertebrates. Blue line shows the linear regression between log (body mass) and log (lifespan), R^2=0.32.

(B) Body size comparison between living (African and Asian elephants) and extinct (Steppe mammoth) Proboscideans, Cetaceans (Minke whale), and the extinct hornless rhinoceros *Paraceratherium* ('Walter'), and humans.

Figure 2. Expansion of the *TP53RTG* gene repertoire in Proboscideans.

(A) *TP53* copy number in 61 Sarcopterygian (Lobe-finned fish) genomes. Clade names are shown for lineages in which the genome encodes more than one *TP53* gene or pseudogene.

(B) Estimated *TP53*/*TP53RTG* copy number inferred from complete genome sequencing data (WGS, purple), 1:1 orthology (green), gene tree reconciliation (blue), and normalized read depth from genome sequencing data (red). Whiskers on normalized read depth copy number estimates show the 95% confidence interval of the estimate.

Figure 3. *TP53RTG* copy number increased through segmental duplications.

(A) Organization of the *TP53* and *TP53RTG* loci in African elephant. The *TP53*/*TP53RTG* gene tree is shown at the left. The location of homologous transposable elements that flank the *TP53RTG* genes are shown and labeled. *TP53RTG* genes with intact start codons are labeled with arrows, stop codons are shown in red.

(B) Multiple sequence alignment (MUSCLE) of elephant *TP53RTG* containing contigs. The location of the *TP53RTG* genes are shown with a blue bar. Sites are color coded according to their conservation (see inset key).

(C) ProgressiveMAUVE alignment of elephant *TP53RTG* containing contigs. Colored boxes shown the location of collinear blocks, lines connect homologous collinear blocks on different contigs.

Figure 4. *TP53RTG* copy number is correlated with body size evolution in Proboscideans.

(A) Time calibrated Bayesian phylogeny of *TP53*/*TP53RTG* genes. *TP53RTG* genes are shown in blue, the 95% highest posterior density (HPD) of estimated divergence dates

are shown as red bars, nodes with a posterior probability (PP)>0.95 are labeled with closed circles whereas nodes with a PP≤0.95 are labeled with open circles. The period corresponding to the expansion of the *TP53RTG* gene repertoire is shown in a grey.

(B) *TP53*/*TP53RTG* copy number (blue) and Proboscidean body size (red) increases through time are correlated.

Figure 5. *TP53RTG12* is transcribed and translated.

(A) Transcription of elephant *TP53* and *TP53RTG* genes in dermal fibroblasts, white adipose, and placental villi. RNA-Seq data are shown as mean transcripts per million (TPM) with 95% confidence intervals of TPM value. Blue bars show TPM estimates from 'end-to-end' read mapping and gray bars shown 'local' read mapping.

(B) qRT-PCR products generated Asian (left, blue sqaure) and African (right, light blue square) elephant fibroblast cDNA using primers specific to TP53 and TP53RTG12. cDNA was generated from DNaseI-treated RNA. No reverse transcriptase (no RT) controls for each qPCR reaction are shown, end point PCR products are shown.

(C) Coverage of mapped reads from Asian (dark blue) and African (light blue) elephant fibroblast RNA-Seq data across the region of scaffold_885 encoding the *TP53RTG12* gene. The location of TP53RTG12 exons predicted from geneid and GENESCAN are shown in blue introns are shown as lines with arrows indicating the direction of transcription. Gray bars show the location of transposable elements around the *TP53RTG12* gene, darker gray indicates high sequence similarity to the consensus of each element. PCR tiles across this region are shown for African (Lox.) and Asian (Ele.) elephants, PCR primers generating amplicons are shown in red. The inferred *TP53RTG12* transcript is shown below. 1 kb scale shown from position 1 of African elephant (Broad/loxAfr3) scaffold_825.

(D) Relative luciferase (Luc.) expression in Asian and African fibroblasts transfected with either the promoterless pGL4.10[*luc2*] luciferase reporter vector (empty vector), pGL4.10 containing the RTE_LA consensus sequences (Consensus), pGL4.10 containing the RTE_LA from Asian elephant (Ele. RTE_LA), or pGL4.10 containing the RTE_LA from African elephant (Lox. RTE_LA). Results are shown as fold difference in Luc. expression standardized to empty vector and *Renilla* controls. n=16, Wilcoxon P-values.

(E) Western blot of total cell protein isolated from South African Rock hyrax, Asian elephant (*Elephas*), and African elephant (*Loxodona*) dermal fibroblasts. -, control cells. +, cells treated with 50 J/m^2 UV-C and the proteasome inhibitor MG-132. The name and predicted molecular weights of TP53 isoforms are shown.

Figure 6. Elephant cells have enhanced TP53 signaling and are hyper-responsive to DNA damage.

(A) Relative luciferase (Luc.) expression in African elephant, Asian elephant, hyrax, aardvark, and armadillo fibroblasts transfected with the pGL4.38[*luc2p*/p53 RE/Hygro] reporter vector and treated with either mitomycin c, doxorubicin, nutlin-3a, or UV-C. Data is shown as fold difference in Luc. expression 18hrs after treatment standardized to species paired empty vector and *Renilla* controls. n=12, mean±SEM.

(B) Relative capsase-3/7 (Cas3/7) activity in African elephant, Asian elephant, hyrax, aardvark, and armadillo treated with either mitomycin c, doxorubicin, nutlin-3a, or UV-C. Data is shown as fold difference in Cas3/7 activity 18hrs after treatment standardized to species paired untreated controls. n=12, mean±SEM.

Figure 7. TP53RTG12 mediates enhanced TP53 signaling and DNA-damage responses.

(A) Relative luciferase (Luc.) expression in mouse 3T3-L1 cells co-transfected with either the pGL4.38[*luc2P*/p53 RE/Hygro] Luc. reporter vector, *TP53RTG12* pcDNA3.1(+)/myc-His expression vector, or empty pcDNA3.1(+)/myc-His and treated with either mitomycin c, doxorubicin, nutlin-3a, or UV-C. Data is shown as fold difference in Luc. expression 18hrs after treatment standardized to cells transfected with only pGL4.38[*luc2P*/p53 RE/Hygro] and *Renilla* controls. n=12, mean±SEM. ***, Wilcoxon P>0.001.

(B) Relative capsase-3/7 (Cas3/7) activity in mouse 3T3-L1 cells transfected with either the *TP53RTG12* pcDNA3.1(+)/myc-His expression vector or empty pcDNA3.1(+)/myc-His and treated with either mitomycin c, doxorubicin, nutlin-3a, or UV-C. Data is shown as fold difference in Cas3/7 activity 18hrs after treatment standardized to mock transfected. n=12, mean±SEM. ***, Wilcoxon P>0.001.

Figure 8. TP53RTG12 interacts with canonical TP53 but not MDM2

(A) Domain structure of the elephant TP53 and TP53RTG proteins. Cartoons show the location of the transactivation domain (light blue) including residues that mediate the interaction with MDM2 (dark blue) and nuclear export sequence (NES), the DNA binding domain (light green) including residues that mediate dimerization (dark green) and DNA binding, the nuclear localization sequence (NLS), and the tetramerization domain (pink).

(B) Relative luciferase (Luc.) expression in HEK-293 cells co-transfected with either the TP53RTG12 pcDNA3.1(+)/myc-His expression vector (RTG12+) or empty pcDNA3.1(+)/myc-His expression vector (Control) and the pGL4.38[*luc2P*/p53 RE/Hygro] reporter vector. Data is shown as fold difference in Luc. expression 18hrs after transfection standardized to empty vector and *Renilla* controls. n=10, Wilcoxon P=0.97.

(C) HEK-293 cells were transiently transfected with the TP53RTG12 pcDNA3.1(+)/myc-His expression vector and total cell protein immunoprecipitated with either an α-Myc (IP: RTG12-Myc) or α-MDM2 (IP: MDM2) antibody. Co-immunoprecipitation of Myc-tagged TP53RTG12, TP53, or MDM2 were assayed by serial Western blotting with α-Myc, α-TP53, and α-MDM2 antibodies after chemically stripping the blot, respectively.

(C) Model of TP53RTG12 function. Under normal conditions TP53 is negatively regulate by the MDM complex which ubiquitinates TP53 tetramers leading to proteosomal degradation. Upon DNA damage TP53 is phosphorylated, preventing interaction with the MDM complex and activating downstream TP53 signaling. In elephant cells TP53RTG12 may dimerize with canonical TP53 and block formation of TP53 tetramers. These TP53RTG12/TP53 dimers cannot be ubiquitinated generating a pool of TP53 proteins to rapidly respond to lower levels or DNA damage and stress than other species.

Supplementary Figure 1. Domain organization of TP53 and TP53RTG proteins. Domain structure of the elephant TP53 and TP53RTG proteins. Cartoons show the location of the transactivation domain including the MDM2 interaction motif and nuclear export sequence (NES), the DNA binding domain including the dimerization motif and residues critical for DNA binding, the nuclear localization sequence (NLS), and the tetramerization domain.

Supplementary Figure 2. Reconciled *TP53/TP53RTG* gene trees. Reconciled *TP53RTG* gene trees for Columbian mammoth, woolly mammoth, Asian elephant, and American mastodon.

Supplementary Figure 3. TP53/TP53RTG gene trees. Maximum likelihood gene trees with aBayes, Chi-Square, aLRT, and SH-like branch supports. Bayesian tree inferred by MrBayes. Numbers along branches indicate node supports from each method. Note that the full dataset included 65 diverse species and only the Paeunngulate clade is shown. TP53 gene trees were inferred using a general time reversible model (GTR), empirical nucleotide frequencies (+F), a proportion of invariable sites estimated from the data (+I), four gamma distributed rate categories (+G), and using the best of NNI and SPR branch moves during the topology search (for the maximum likelihood tree).

Supplementary Figure 4. PCR and Sanger sequencing confirm *TP53RTG12* is transcribed in elephant fibroblasts.

(A) Multiple sequence alignment (MSA) of African elephant *TP53* and *TP53RTG1-TP53RTG19*. Numbering (shown at top) is relative to the canonical TP53 transcript in African elephant (ENSLAFT00000007484), sequence conservation in this region is shown in Bits, as a logo, and percent identity (green, 100% identity of paralogs; yellow, <100% identity of paralogs). Bases in each gene are color-coded (blue, C; red, A; yellow, G; and T, green). The *TP53RTG*-specific deletion is shown as a black line, and the location of the PCR primers than span the deletion are shown as arrows.

(B) Chromatogram from Sanger sequencing of the *TP53RTG12* qRT-PCR amplicon confirm the *TP53RTG12* is transcribed in African elephant fibroblasts. Bases that flank the *TP53RTG*-specific deletion are highlighted in red.

(C) Chromatogram from Sanger sequencing of the *TP53RTG12* qRT-PCR amplicon confirm the *TP53RTG12* is transcribed in Asian elephant fibroblasts. Bases that flank the *TP53RTG*-specific deletion are highlighted in red.

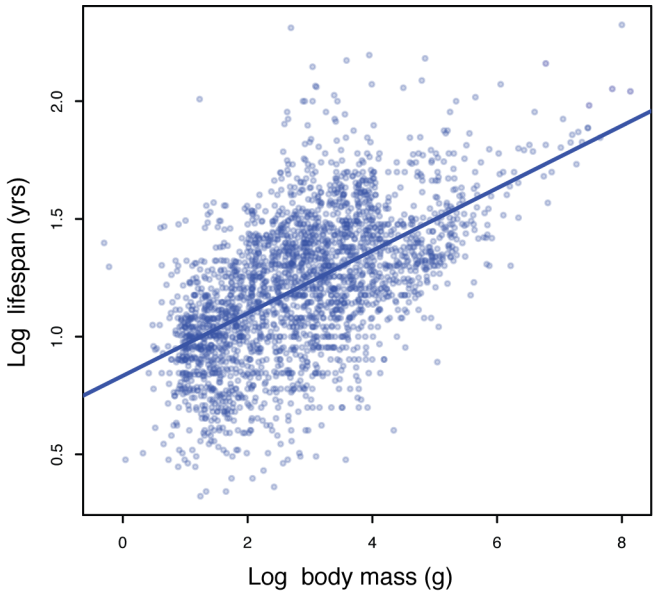
Supplementary Figure 5. Unedited Western blots shown in Fig. 5E.

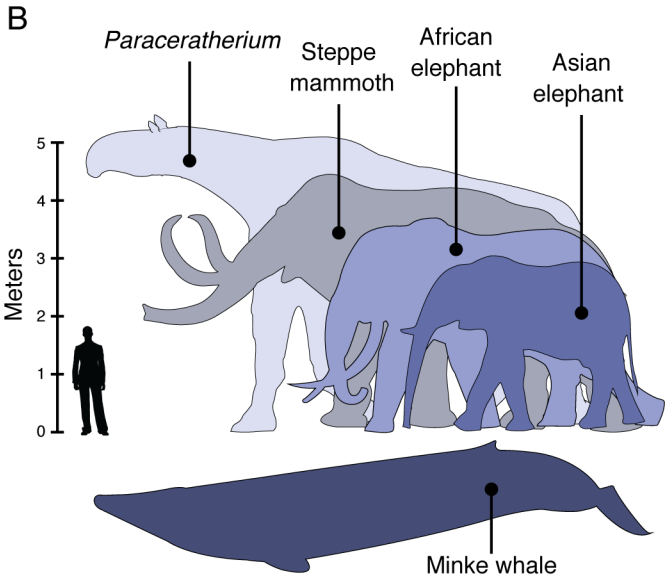
(A) Western blot of total cell protein isolated from African elephant and hyrax fibroblasts. -, control cells. +, cells treated with 50 J/m^2 UV-C and MG-132. 1° & 2°, probed with

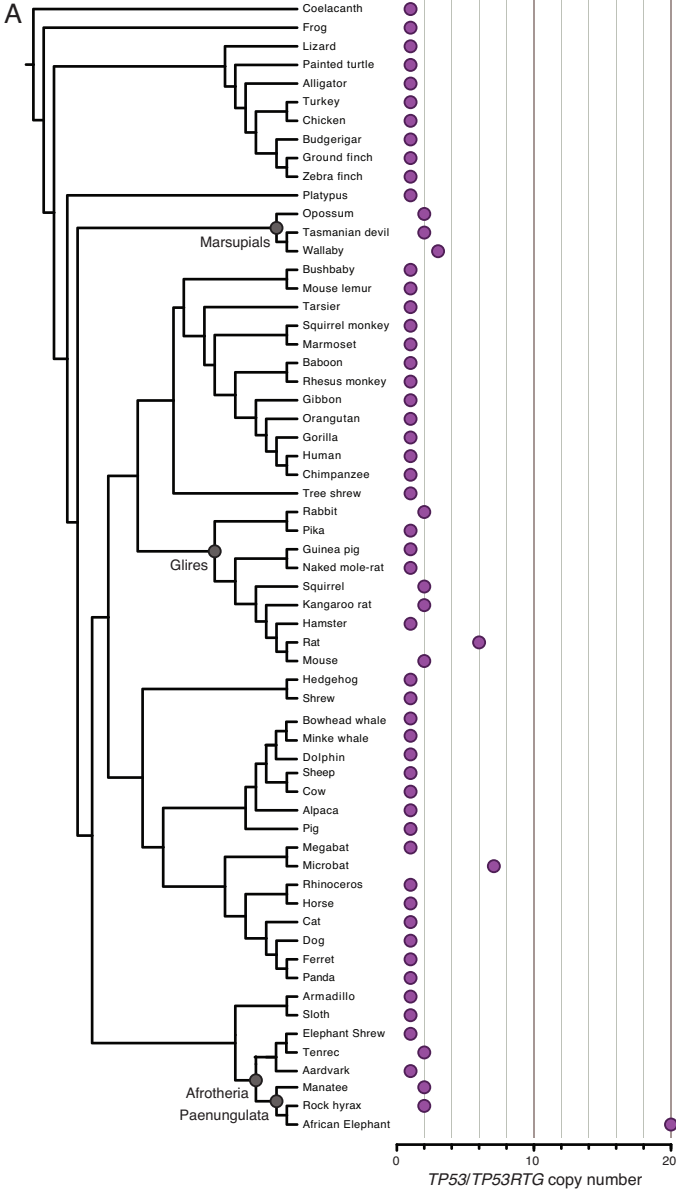
primary antibody against TP53 and secondary HRP conjugated antibody. 2° only, probed only with secondary HRP conjugated antibody

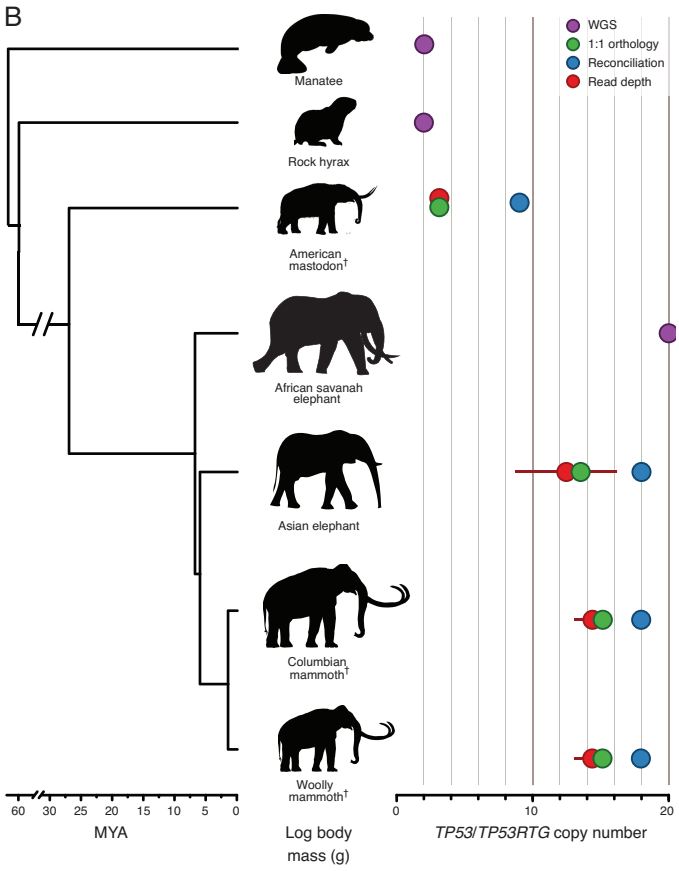
(B) Western blot of total cell protein isolated from Asian elephant and hyrax fibroblasts. -, control cells. +, cells treated with 50 J/m^2 UV-C and MG-132. 1° & 2°, probed with primary antibody against TP53 and secondary HRP conjugated antibody. 2° only, probed only with secondary HRP conjugated antibody

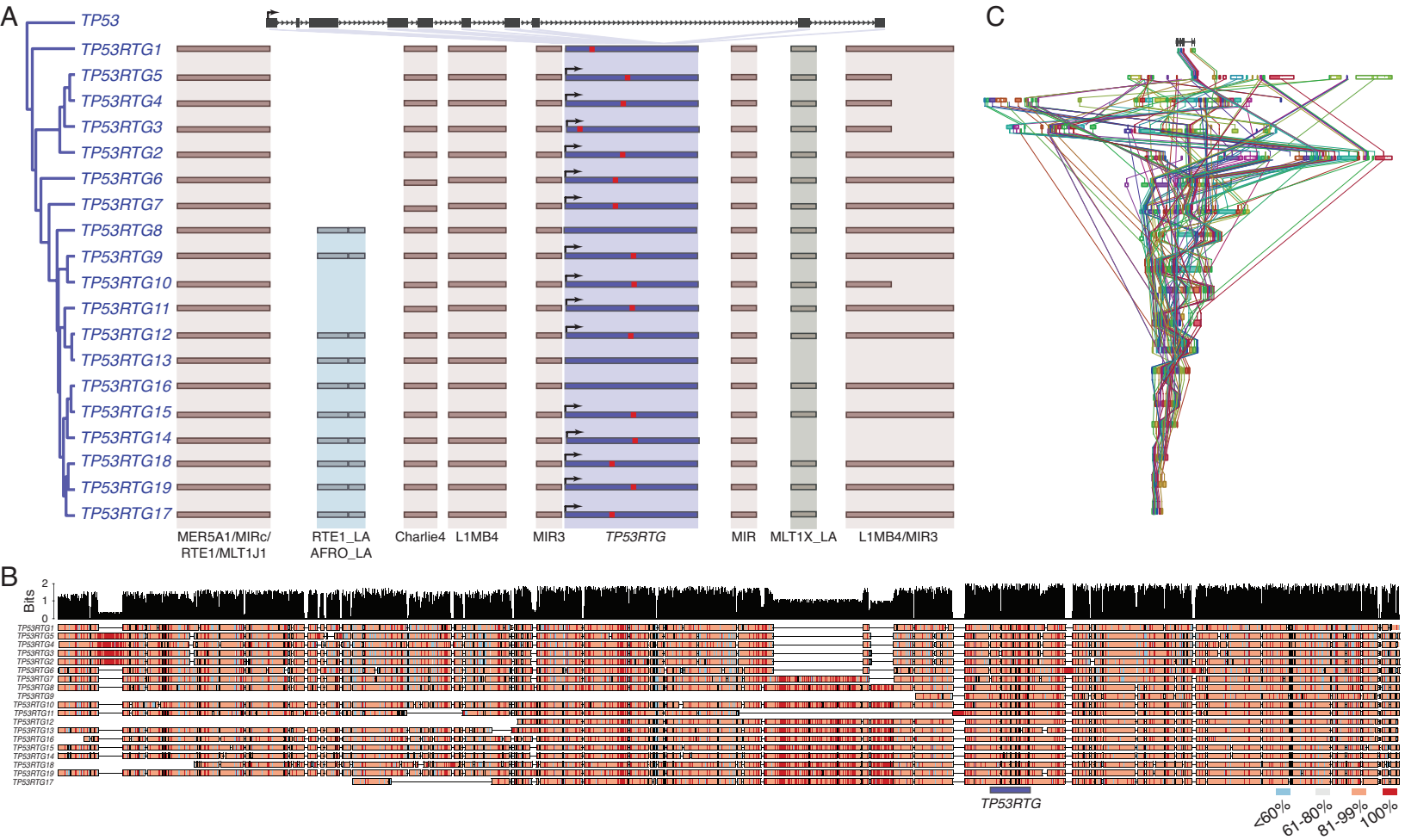
Supplementary Figure 6. Uncropped Western blots shown in Fig. 8B. HEK-293 cells were transiently transfected with the TP53RTG12 pcDNA3.1(+)/myc-His expression vector and total cell protein immunoprecipitated with either an α-Myc (IP: RTG12-Myc) or α-MDM2 (IP: MDM2) antibody. Co-immunoprecipitation of Myc-tagged TP53RTG12, TP53, or MDM2 were assayed by Western blotting with α-Myc, α-TP53, and α-MDM2 antibodies respectively.

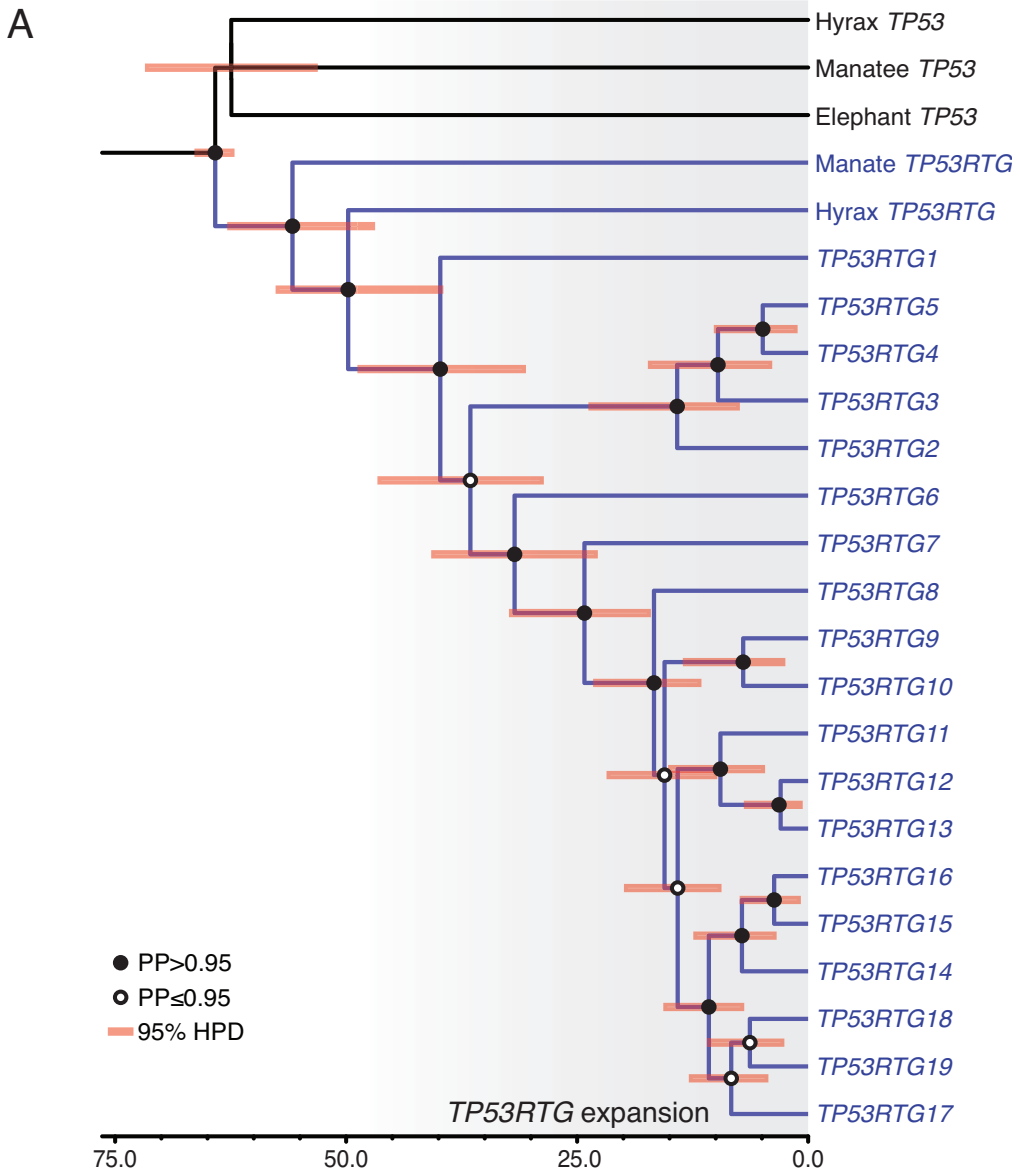

Log lifespan (yrs)
0.5
1.0
1.5
2.0
0
2
4
6
8
Log body mass (g)

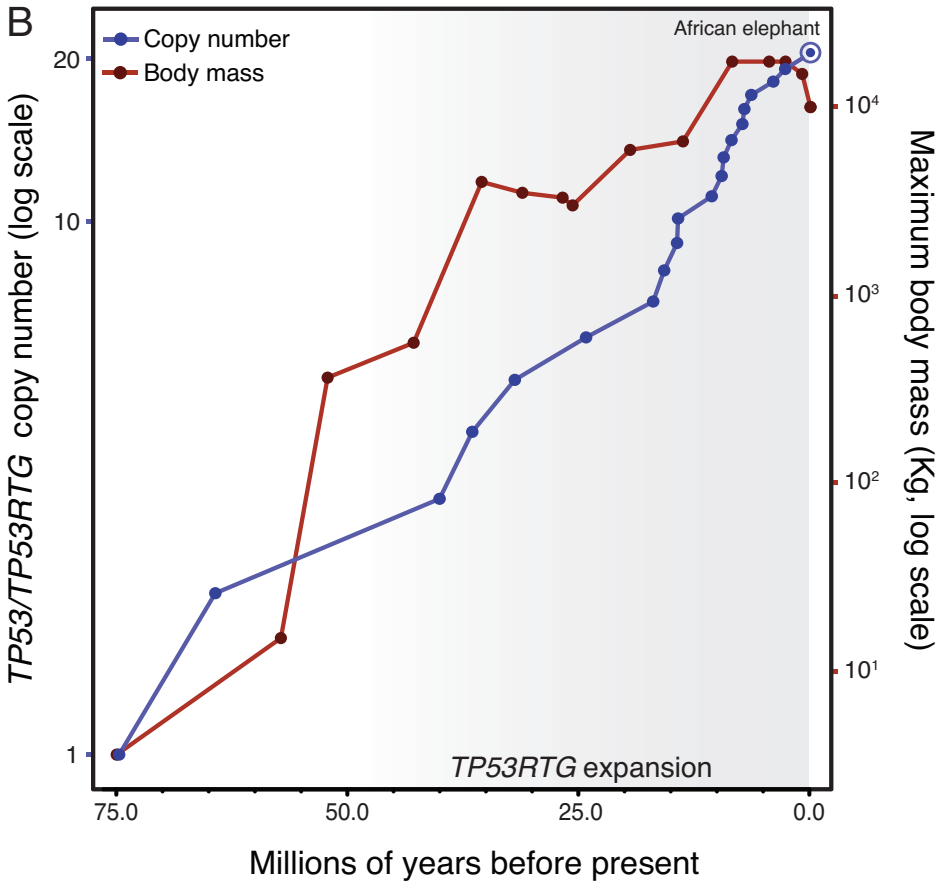

B
Paraceratherium
Steppe mammoth
African elephant
Asian elephant
Meters
0
1
2
3
4
5
Minke whale

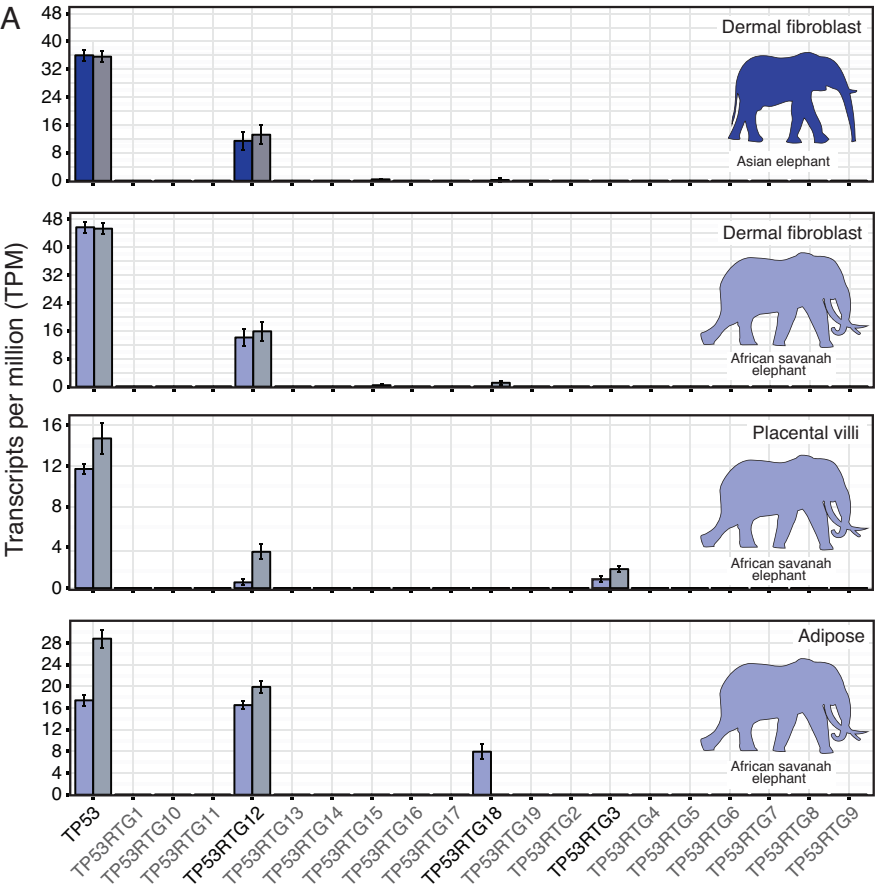

A
Coelacanth
Frog
Lizard
Painted turtle
Alligator
Turkey
Chicken
Budgerigar
Ground finch
Zebra finch
Platypus
Opossum
Tasmanian devil
Wallaby
Marsupials
Bushbaby
Mouse lemur
Tarsier
Squirrel monkey
Marmoset
Baboon
Rhesus monkey
Gibbon
Orangutan
Gorilla
Human
Chimpanzee
Tree shrew
Rabbit
Pika
Guinea pig
Naked mole-rat
Glires
Squirrel
Kangaroo rat
Hamster
Rat
Mouse
Hedgehog
Shrew
Bowhead whale
Minke whale
Dolphin
Sheep
Cow
Alpaca
Pig
Megabat
Microbat
Rhinoceros
Horse
Cat
Dog
Ferret
Panda
Armadillo
Sloth
Elephant Shrew
Tenrec
Aardvark
Manatee
Rock hyrax
African Elephant
Afrotheria
Paenungulata
0
10
20
TP53/TP53RTG copy number

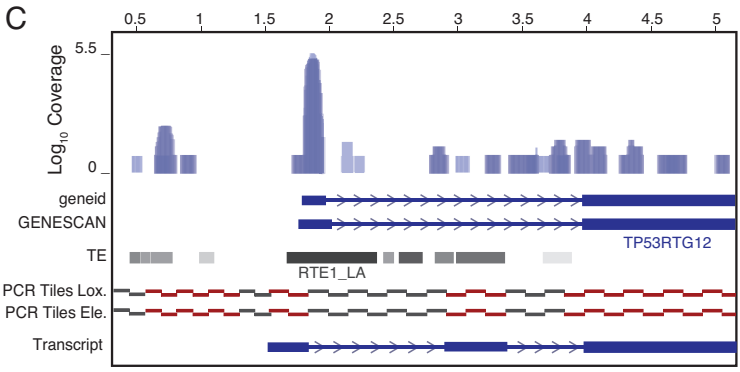

B
WGS
1:1 orthology
Reconciliation
Read depth
Manatee
Rock hyrax
American mastodon†
African savanah elephant
Asian elephant
Columbian mammoth†
Woolly mammoth†
60
30
25
20
15
10
5
0
MYA
Log body mass (g)
0
10
20
TP53/TP53RTG copy number

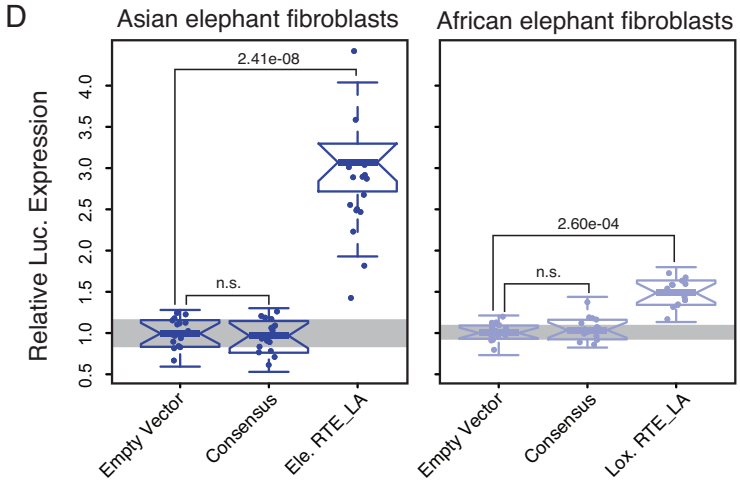

A
TP53
TP53RTG1
TP53RTG5
TP53RTG4
TP53RTG3
TP53RTG2
TP53RTG6
TP53RTG7
TP53RTG8
TP53RTG9
TP53RTG10
TP53RTG11
TP53RTG12
TP53RTG13
TP53RTG16
TP53RTG15
TP53RTG14
TP53RTG18
TP53RTG19
TP53RTG17
MER5A1/MIRc/
RTE1/MLT1J1
RTE1_LA
AFRO_LA
Charlie4
L1MB4
MIR3
TP53RTG
MIR
MLT1X_LA
L1MB4/MIR3
C
B
Bits
2
1
0
TP53RTG
<60%
61-80%
81-99%
100%

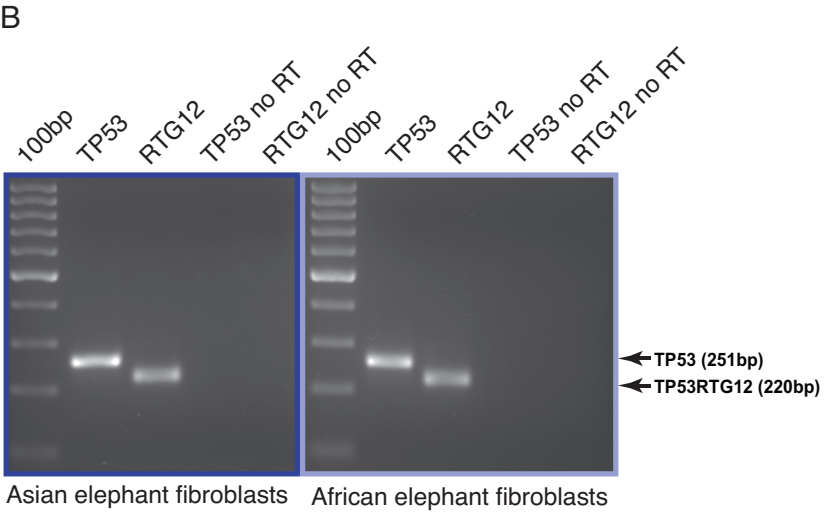

A
Hyrax TP53
Manatee TP53
Elephant TP53
Manate TP53RTG
Hyrax TP53RTG
TP53RTG1
TP53RTG5
TP53RTG4
TP53RTG3
TP53RTG2
TP53RTG6
TP53RTG7
TP53RTG8
TP53RTG9
TP53RTG10
TP53RTG11
TP53RTG12
TP53RTG13
TP53RTG16
TP53RTG15
TP53RTG14
TP53RTG18
TP53RTG19
TP53RTG17
PP>0.95
PP≤0.95
95% HPD
TP53RTG expansion
75.0
50.0
25.0
0.0

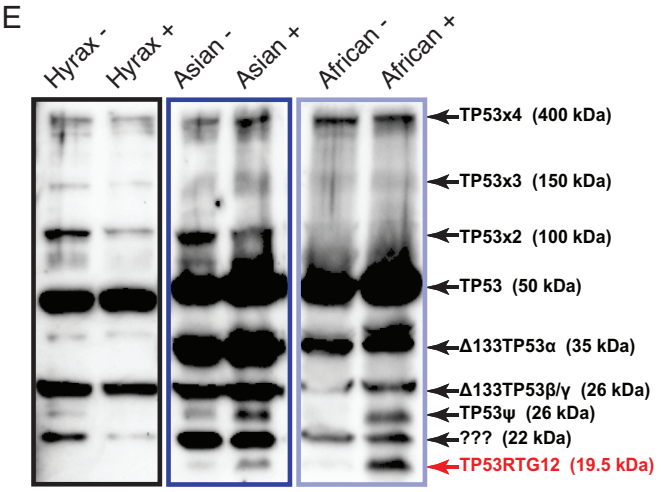

B
Copy number
Body mass
African elephant
TP53/TP53RTG copy number (log scale)
Maximum body mass (Kg, log scale)
20
10
1
10⁴
10³
10²
10¹
TP53RTG expansion
75.0
50.0
25.0
0.0
Millions of years before present

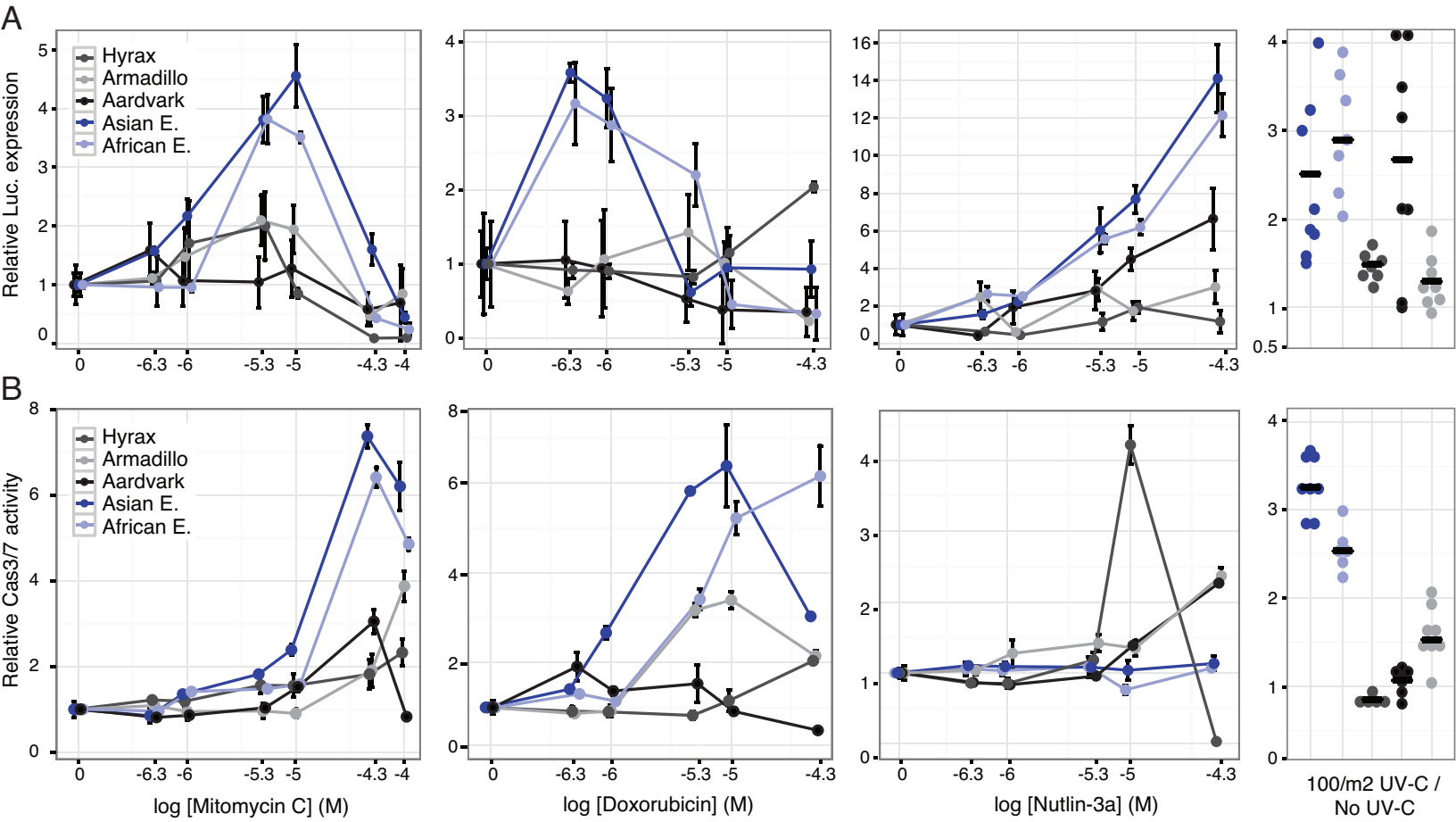

A
Transcripts per million (TPM)
Dermal fibroblast
Asian elephant
Dermal fibroblast
African savanah elephant
Placental villi
African savanah elephant
Adipose
African savanah elephant
TP53
TP53RTG1
TP53RTG10
TP53RTG11
TP53RTG12
TP53RTG13
TP53RTG14
TP53RTG15
TP53RTG16
TP53RTG17
TP53RTG18
TP53RTG19
TP53RTG2
TP53RTG3
TP53RTG4
TP53RTG5
TP53RTG6
TP53RTG7
TP53RTG8
TP53RTG9

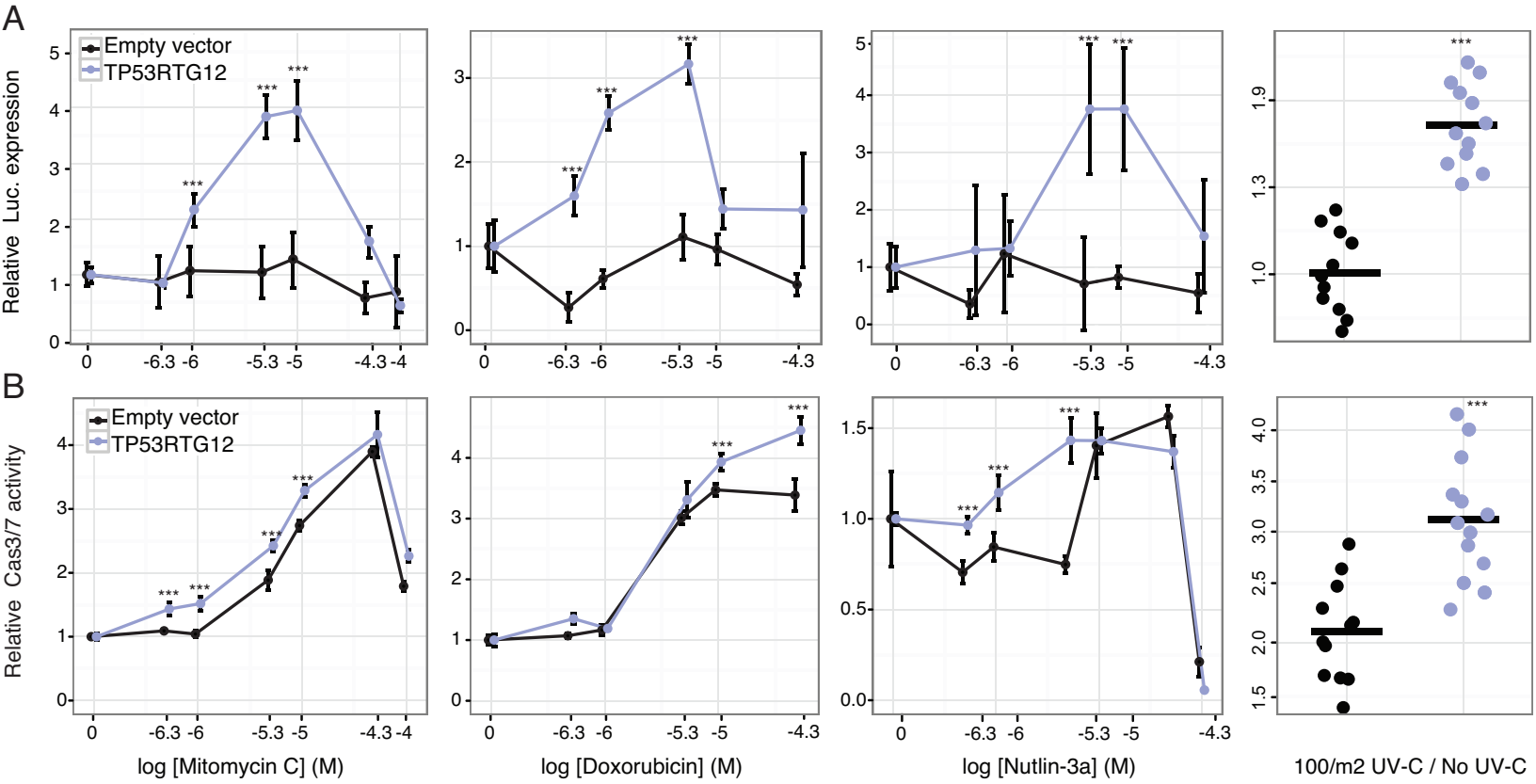

C
Log10 Coverage
geneid
GENESCAN
TE
RTE1_LA
TP53RTG12
PCR Tiles Lox.
PCR Tiles Ele.
Transcript

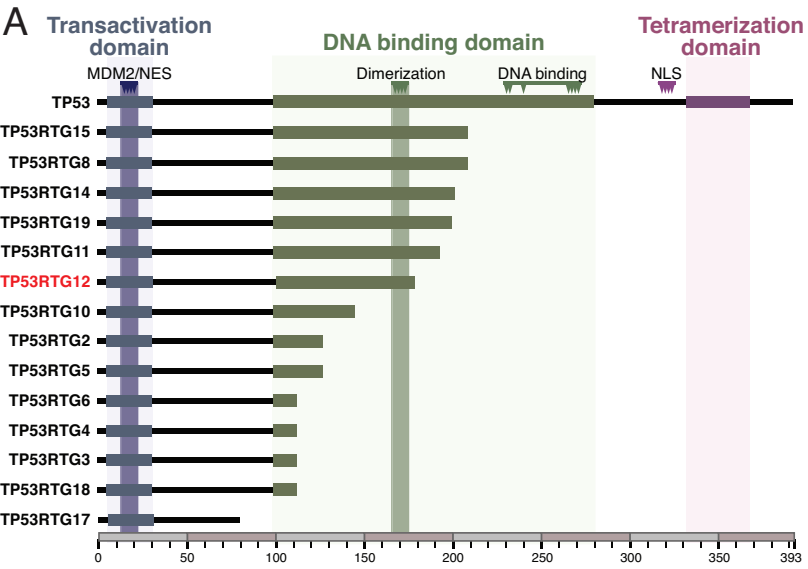

D
Asian elephant fibroblasts
African elephant fibroblasts
Relative Luc. Expression
2.41e-08
n.s.
2.60e-04
n.s.
Empty Vector
Consensus
Ele. RTE_LA
Empty Vector
Consensus
Lox. RTE_LA

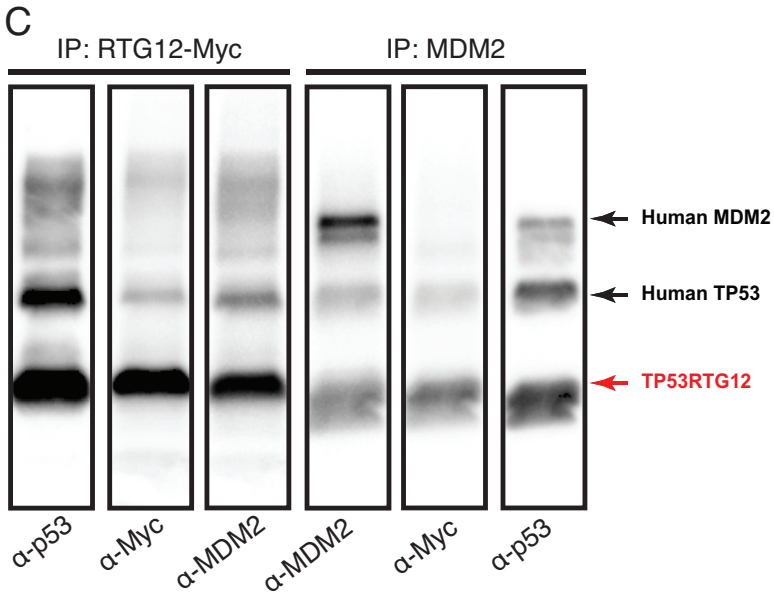

B
100bp
TP53
RTG12
TP53 no RT
RTG12 no RT
100bp
TP53
RTG12
TP53 no RT
RTG12 no RT
TP53 (251bp)
TP53RTG12 (220bp)
Asian elephant fibroblasts
African elephant fibroblasts

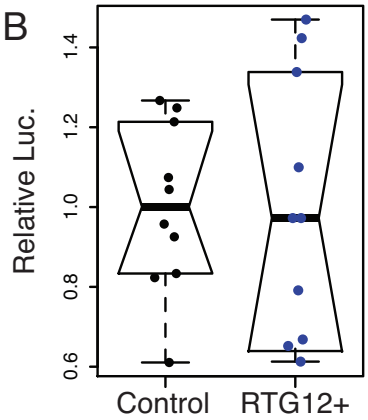

E
Hyrax -
Hyrax +
Asian -
Asian +
African -
African +
TP53x4 (400 kDa)
TP53x3 (150 kDa)
TP53x2 (100 kDa)
TP53 (50 kDa)
Δ133TP53α (35 kDa)
Δ133TP53β/γ (26 kDa)
TP53ψ (26 kDa)
??? (22 kDa)
TP53RTG12 (19.5 kDa)

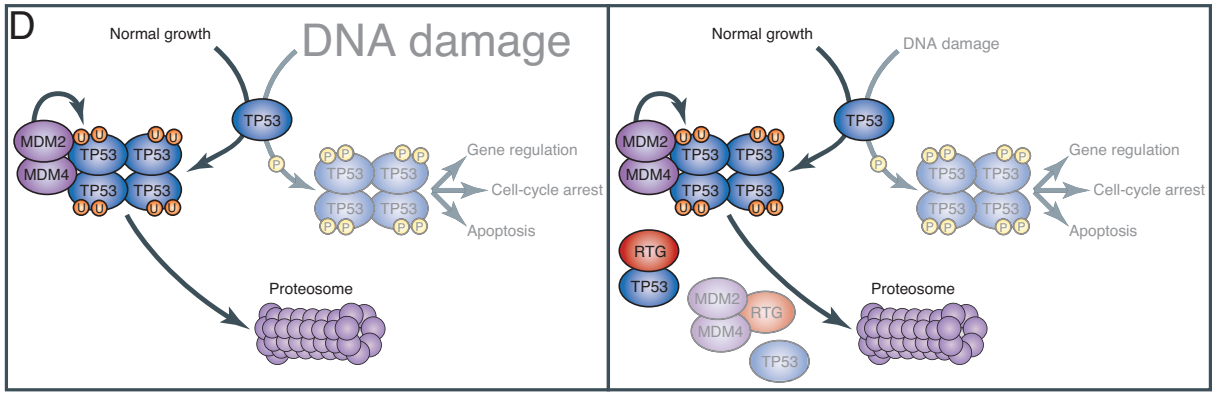

A
B
Hyrax
Armadillo
Aardvark
Asian E.
African E.
Relative Luc. expression
Relative Cas3/7 activity
0
-6.3
-6
-5.3
-5
-4.3
-4
log [Mitomycin C] (M)
log [Doxorubicin] (M)
log [Nutlin-3a] (M)
100/m2 UV-C /
No UV-C


A
B
Empty vector
TP53RTG12
Relative Luc. expression
Relative Cas3/7 activity
log [Mitomycin C] (M)
log [Doxorubicin] (M)
log [Nutlin-3a] (M)
100/m2 UV-C / No UV-C


A
Transactivation domain
MDM2/NES
DNA binding domain
Dimerization
DNA binding
Tetramerization domain
NLS
TP53
TP53RTG15
TP53RTG8
TP53RTG14
TP53RTG19
TP53RTG11
TP53RTG12
TP53RTG10
TP53RTG2
TP53RTG5
TP53RTG6
TP53RTG4
TP53RTG3
TP53RTG18
TP53RTG17
0
50
100
150
200
250
300
350
393

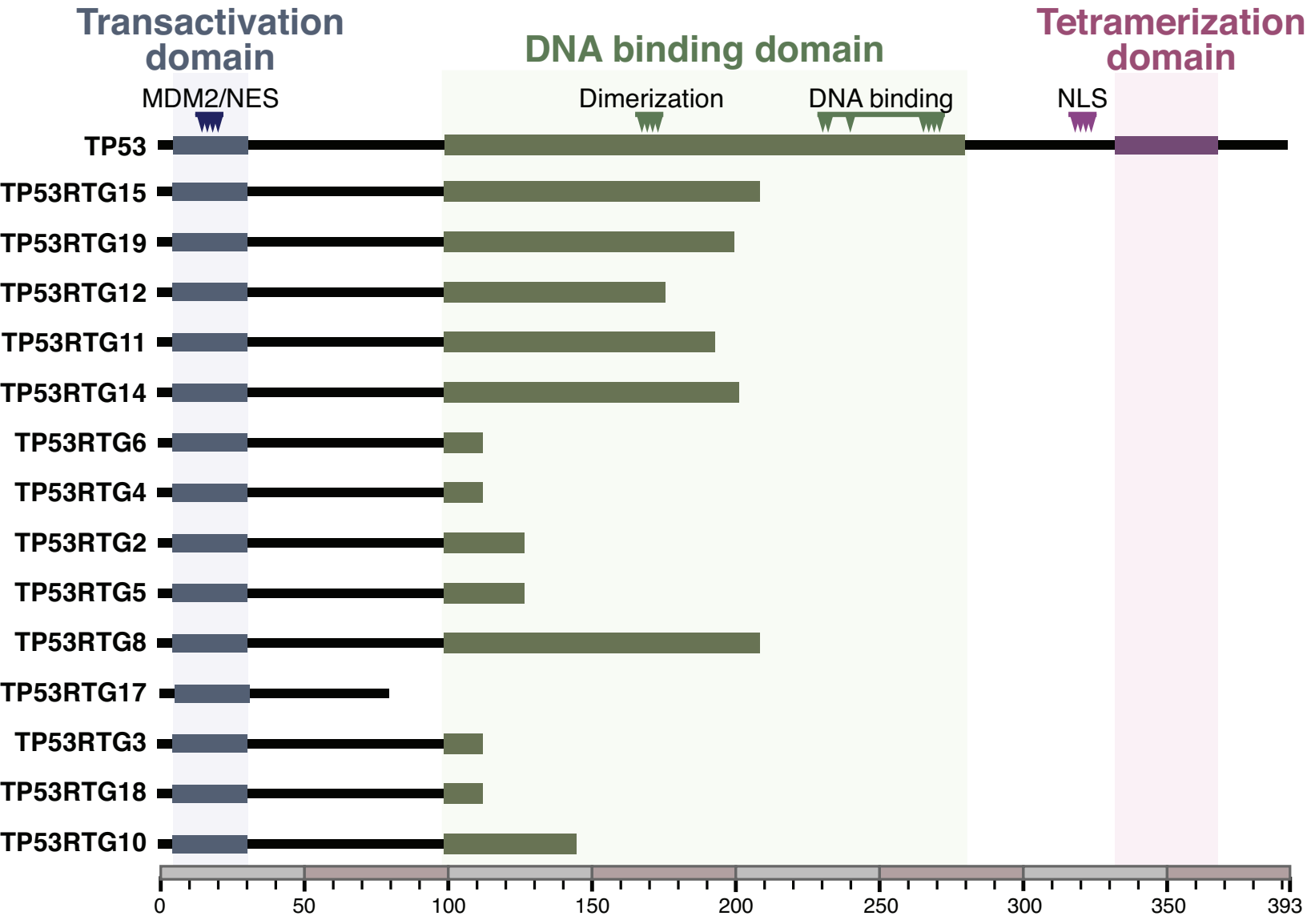

C
IP: RTG12-Myc
IP: MDM2
Human MDM2
Human TP53
TP53RTG12
α-p53
α-Myc
α-MDM2
α-MDM2
α-Myc
α-p53

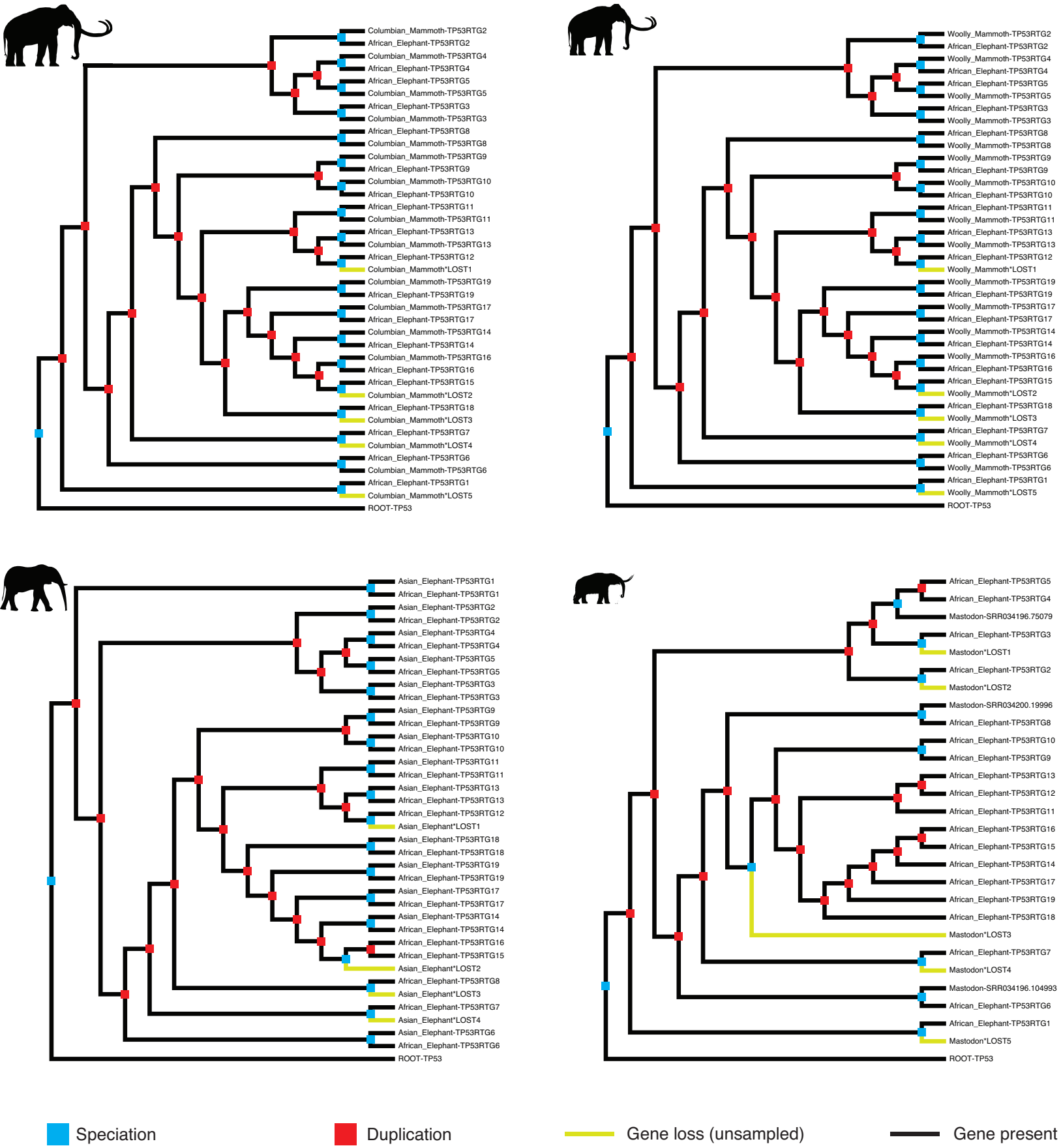

B
Relative Luc.
0.6
0.8
1.0
1.2
1.4
Control
RTG12+


D
Normal growth
DNA damage
TP53
MDM2
MDM4
Gene regulation
Cell-cycle arrest
Apoptosis
Proteosome
RTG

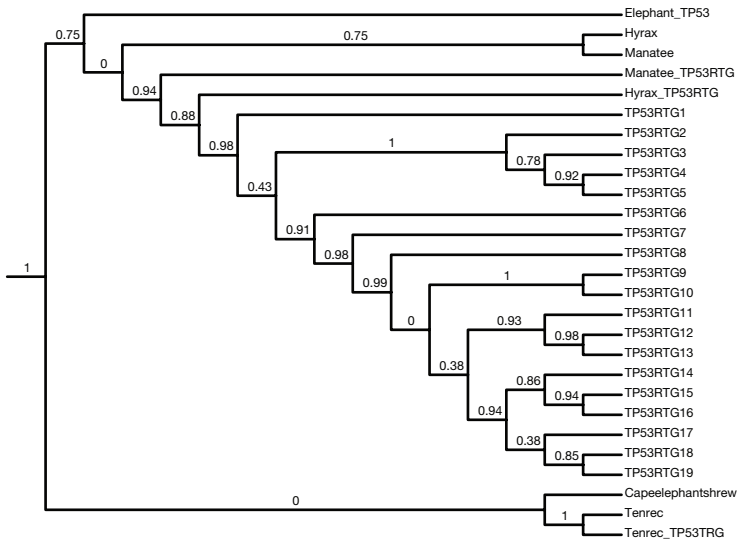
Supplementary table 1.

Gene ID	ENSEMBL ID	Scaffold	Coding potential	ORF size
TP53	ENSLAFG00000007483	47	Yes	392aa
TP53RTG1	ENSLAFG00000025553	175	No	N/A
TP53RTG2	N/A	217	Yes	134aa
TP53RTG3	ENSLAFG00000027474	406	Yes	79aa
TP53RTG4	N/A	627	Yes	134aa
TP53RTG5	ENSLAFG00000027348	221	Yes	162aa
TP53RTG6	N/A	76	Yes	123aa
TP53RTG7	N/A	208	No	N/A
TP53RTG8	ENSLAFG00000027820	294	Yes	210aa
TP53RTG9	ENSLAFG00000027669	786	No	N/A
TP53RTG10	ENSLAFG00000030555	221	Yes	210aa
TP53RTG11	N/A	281	Yes	203aa
TP53RTG12	ENSLAFG00000028299	825	Yes	180aa
TP53RTG13	ENSLAFG00000032042	458	No	N/A
TP53RTG14	ENSLAFG00000026238	928	Yes	210aa
TP53RTG15	ENSLAFG00000027365	656	Yes	210aa
TP53RTG16	ENSLAFG00000030880	378	No	N/A
TP53RTG17	ENSLAFG00000028692	552	Yes	111aa
TP53RTG18	N/A	498	Yes	111aa
TP53RTG19	ENSLAFG00000032258	342	Yes	210aa

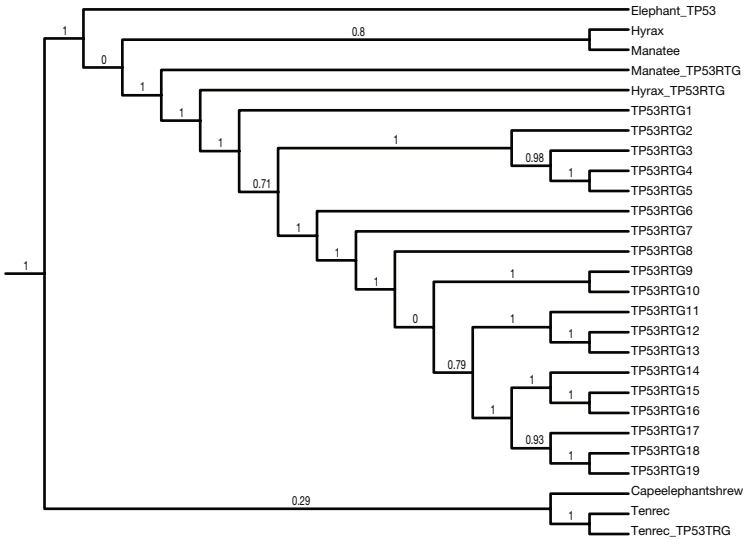

Transactivation domain
DNA binding domain
Tetramerization domain
MDM2/NES
Dimerization
DNA binding
NLS
TP53
TP53RTG15
TP53RTG19
TP53RTG12
TP53RTG11
TP53RTG14
TP53RTG6
TP53RTG4
TP53RTG2
TP53RTG5
TP53RTG8
TP53RTG17
TP53RTG3
TP53RTG18
TP53RTG10
0
50
100
150
200
250
300
350
393


Columbian_Mammoth-TP53RTG2
African_Elephant-TP53RTG2
Columbian_Mammoth-TP53RTG4
African_Elephant-TP53RTG4
African_Elephant-TP53RTG5
Columbian_Mammoth-TP53RTG5
African_Elephant-TP53RTG3
Columbian_Mammoth-TP53RTG3
African_Elephant-TP53RTG8
Columbian_Mammoth-TP53RTG8
Columbian_Mammoth-TP53RTG9
African_Elephant-TP53RTG9
Columbian_Mammoth-TP53RTG10
African_Elephant-TP53RTG10
African_Elephant-TP53RTG11
Columbian_Mammoth-TP53RTG11
African_Elephant-TP53RTG13
Columbian_Mammoth-TP53RTG13
African_Elephant-TP53RTG12
Columbian_Mammoth*LOST1
Columbian_Mammoth-TP53RTG19
African_Elephant-TP53RTG19
Columbian_Mammoth-TP53RTG17
African_Elephant-TP53RTG17
Columbian_Mammoth-TP53RTG14
African_Elephant-TP53RTG14
Columbian_Mammoth-TP53RTG16
African_Elephant-TP53RTG16
African_Elephant-TP53RTG15
Columbian_Mammoth*LOST2
African_Elephant-TP53RTG18
Columbian_Mammoth*LOST3
African_Elephant-TP53RTG7
Columbian_Mammoth*LOST4
African_Elephant-TP53RTG6
Columbian_Mammoth-TP53RTG6
African_Elephant-TP53RTG1
Columbian_Mammoth*LOST5
ROOT-TP53
Woolly_Mammoth-TP53RTG2
African_Elephant-TP53RTG2
Woolly_Mammoth-TP53RTG4
African_Elephant-TP53RTG4
African_Elephant-TP53RTG5
Woolly_Mammoth-TP53RTG5
African_Elephant-TP53RTG3
Woolly_Mammoth-TP53RTG3
African_Elephant-TP53RTG8
Woolly_Mammoth-TP53RTG8
Woolly_Mammoth-TP53RTG9
African_Elephant-TP53RTG9
Woolly_Mammoth-TP53RTG10
African_Elephant-TP53RTG10
African_Elephant-TP53RTG11
Woolly_Mammoth-TP53RTG11
African_Elephant-TP53RTG13
Woolly_Mammoth-TP53RTG13
African_Elephant-TP53RTG12
Woolly_Mammoth*LOST1
Woolly_Mammoth-TP53RTG19
African_Elephant-TP53RTG19
Woolly_Mammoth-TP53RTG17
African_Elephant-TP53RTG17
Woolly_Mammoth-TP53RTG14
African_Elephant-TP53RTG14
Woolly_Mammoth-TP53RTG16
African_Elephant-TP53RTG16
African_Elephant-TP53RTG15
Woolly_Mammoth*LOST2
African_Elephant-TP53RTG18
Woolly_Mammoth*LOST3
African_Elephant-TP53RTG7
Woolly_Mammoth*LOST4
African_Elephant-TP53RTG6
Woolly_Mammoth-TP53RTG6
African_Elephant-TP53RTG1
Woolly_Mammoth*LOST5
ROOT-TP53
Asian_Elephant-TP53RTG1
African_Elephant-TP53RTG1
Asian_Elephant-TP53RTG2
African_Elephant-TP53RTG2
Asian_Elephant-TP53RTG4
African_Elephant-TP53RTG4
Asian_Elephant-TP53RTG5
African_Elephant-TP53RTG5
Asian_Elephant-TP53RTG3
African_Elephant-TP53RTG3
Asian_Elephant-TP53RTG9
African_Elephant-TP53RTG9
Asian_Elephant-TP53RTG10
African_Elephant-TP53RTG10
Asian_Elephant-TP53RTG11
African_Elephant-TP53RTG11
Asian_Elephant-TP53RTG13
African_Elephant-TP53RTG13
African_Elephant-TP53RTG12
Asian_Elephant*LOST1
Asian_Elephant-TP53RTG18
African_Elephant-TP53RTG18
Asian_Elephant-TP53RTG19
African_Elephant-TP53RTG19
Asian_Elephant-TP53RTG17
African_Elephant-TP53RTG17
Asian_Elephant-TP53RTG14
African_Elephant-TP53RTG14
African_Elephant-TP53RTG16
African_Elephant-TP53RTG15
Asian_Elephant*LOST2
African_Elephant-TP53RTG8
Asian_Elephant*LOST3
African_Elephant-TP53RTG7
Asian_Elephant*LOST4
Asian_Elephant-TP53RTG6
African_Elephant-TP53RTG6
ROOT-TP53
African_Elephant-TP53RTG5
African_Elephant-TP53RTG4
Mastodon-SRR034196.75079
African_Elephant-TP53RTG3
Mastodon*LOST1
African_Elephant-TP53RTG2
Mastodon*LOST2
Mastodon-SRR034200.19996
African_Elephant-TP53RTG8
African_Elephant-TP53RTG10
African_Elephant-TP53RTG9
African_Elephant-TP53RTG13
African_Elephant-TP53RTG12
African_Elephant-TP53RTG11
African_Elephant-TP53RTG16
African_Elephant-TP53RTG15
African_Elephant-TP53RTG14
African_Elephant-TP53RTG17
African_Elephant-TP53RTG19
African_Elephant-TP53RTG18
Mastodon*LOST3
African_Elephant-TP53RTG7
Mastodon*LOST4
Mastodon-SRR034196.104993
African_Elephant-TP53RTG6
African_Elephant-TP53RTG1
Mastodon*LOST5
ROOT-TP53
Speciation
Duplication
Gene loss (unsampled)
Gene present

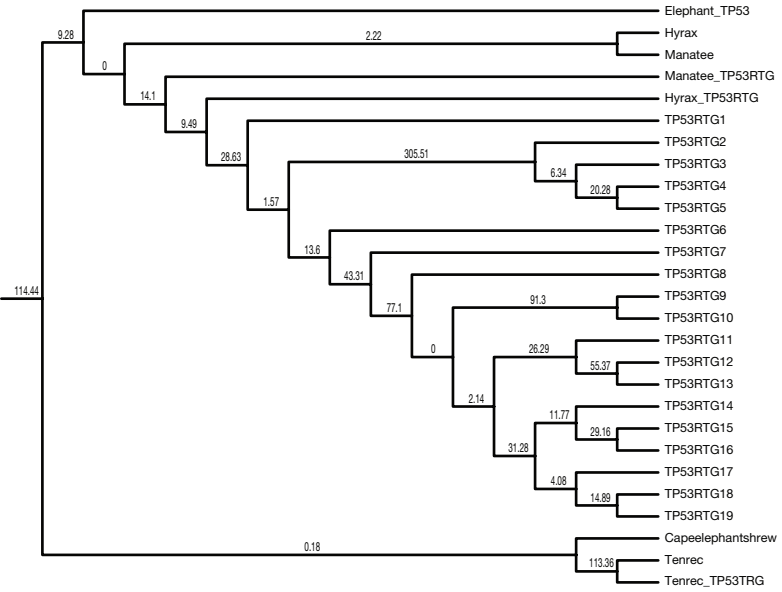
aBayes



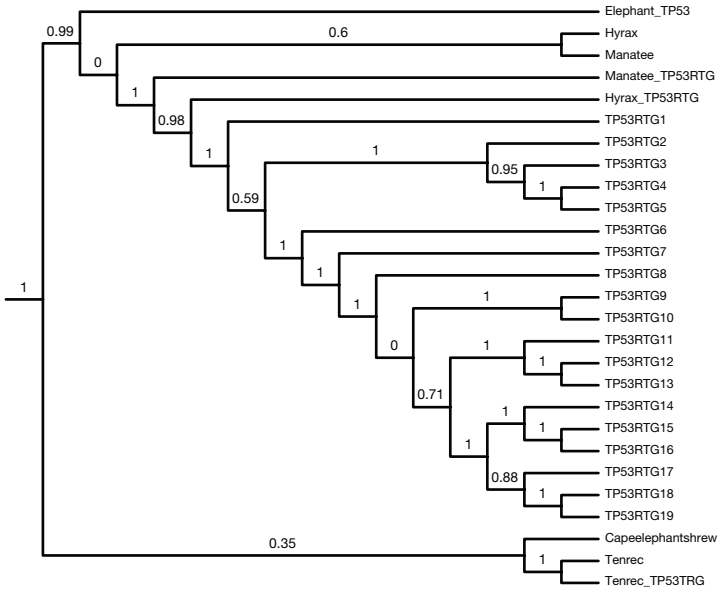
Chi-Square



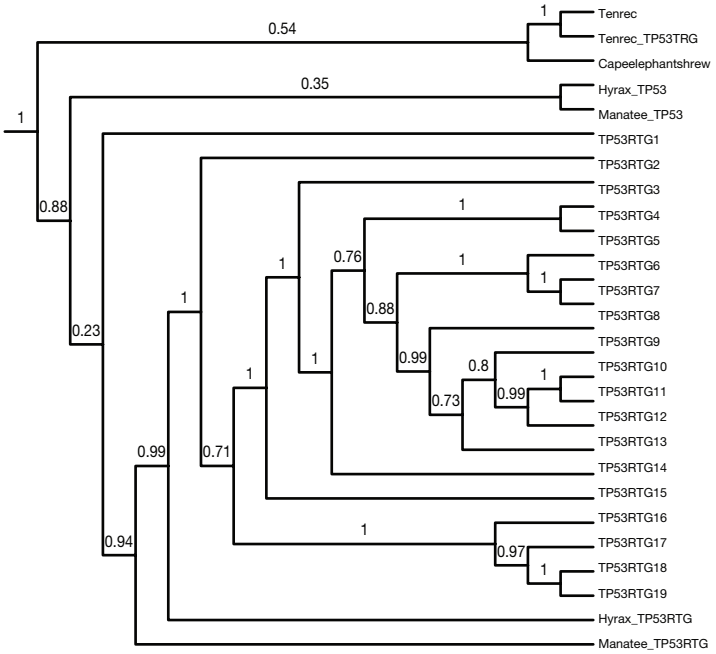
aLRT

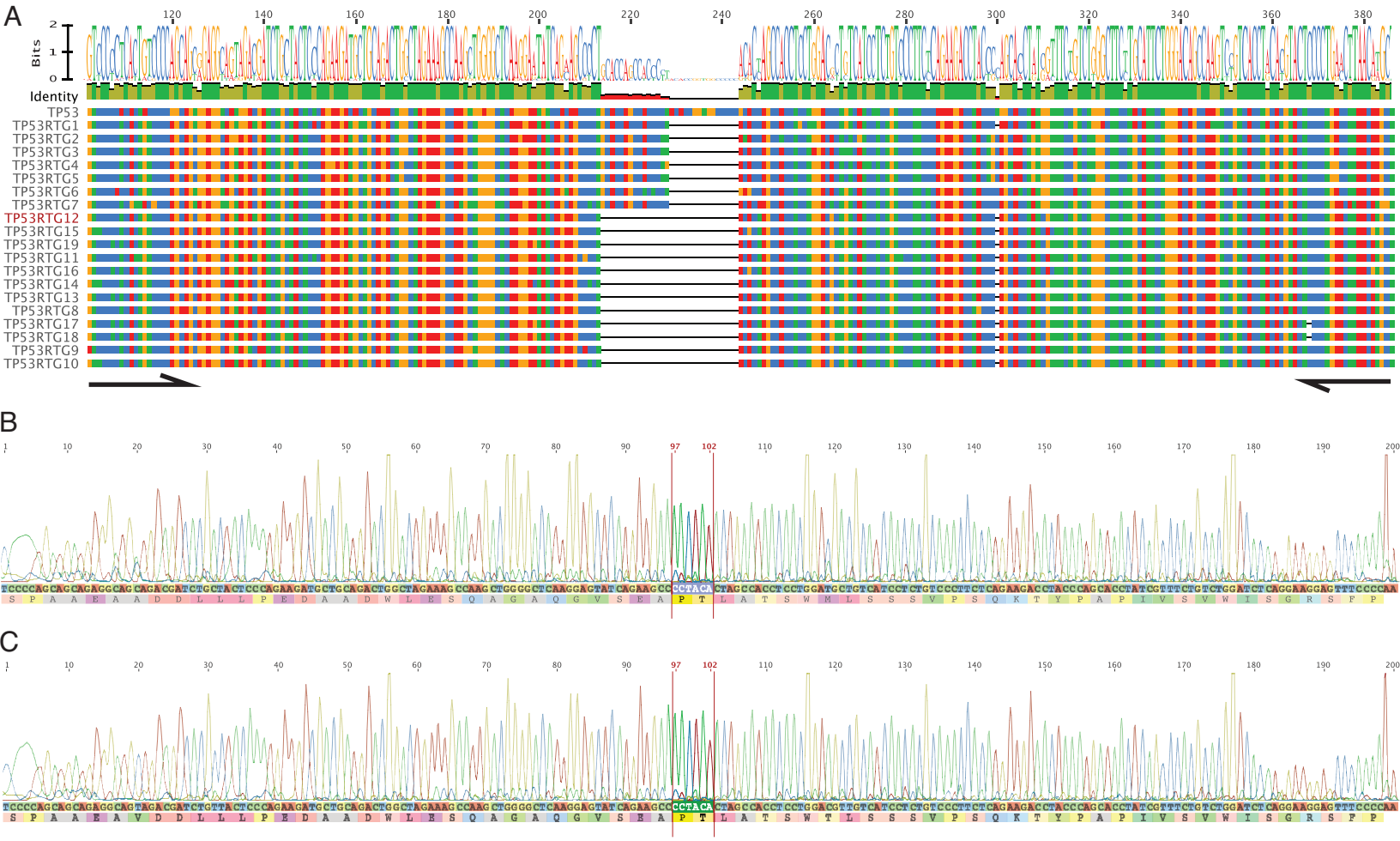


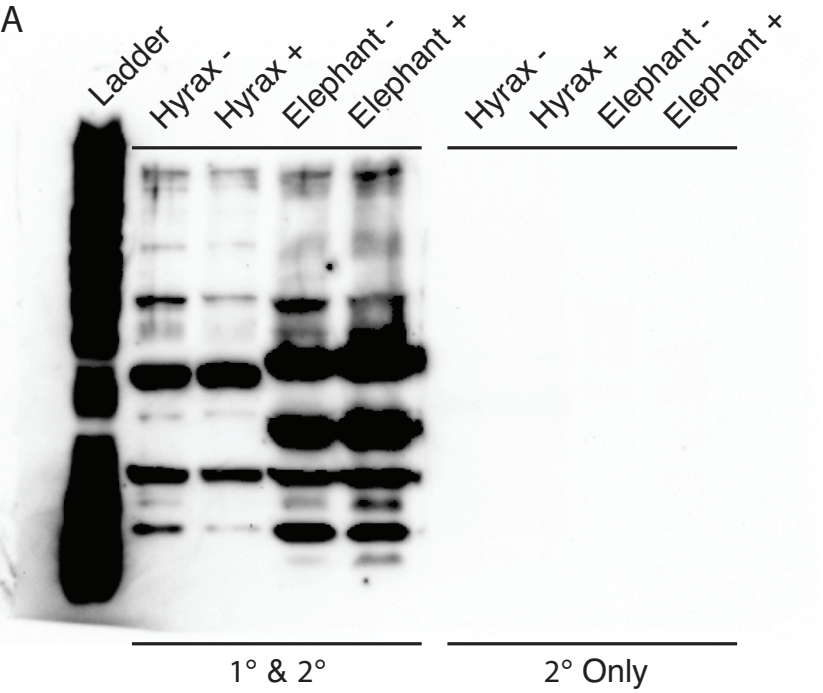
SH-Like

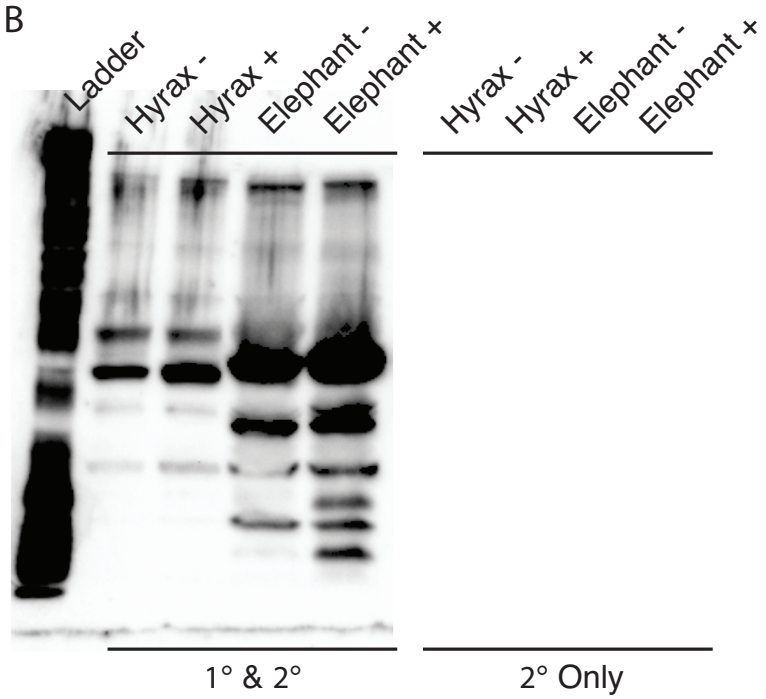


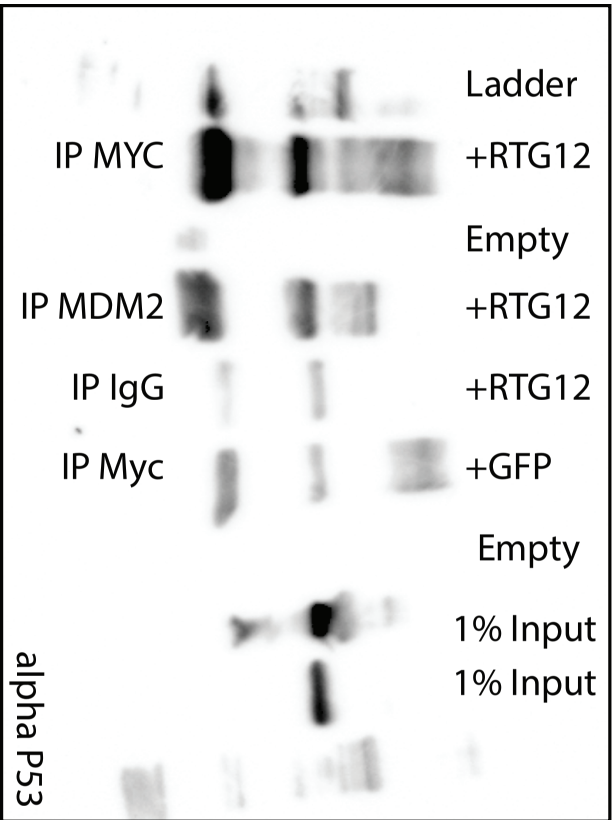
MrBayes

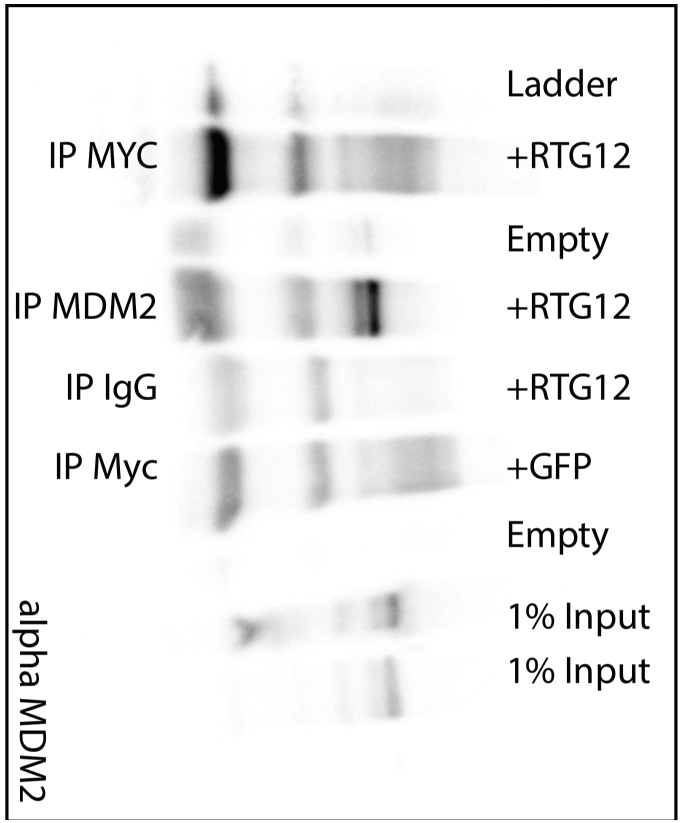


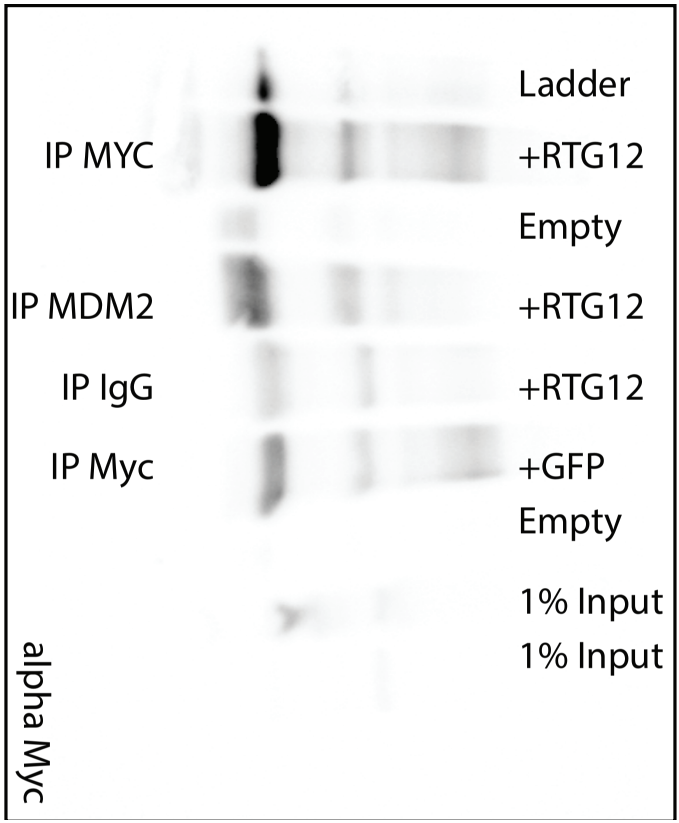

A
Bits
Identity
TP53
TP53RTG1
TP53RTG2
TP53RTG3
TP53RTG4
TP53RTG5
TP53RTG6
TP53RTG7
TP53RTG12
TP53RTG15
TP53RTG19
TP53RTG11
TP53RTG16
TP53RTG14
TP53RTG13
TP53RTG8
TP53RTG17
TP53RTG18
TP53RTG9
TP53RTG10
B
C


A
Ladder
Hyrax -
Hyrax +
Elephant -
Elephant +
Hyrax -
Hyrax +
Elephant -
Elephant +
1° & 2°
2° Only


B
Ladder
Hyrax -
Hyrax +
Elephant -
Elephant +
Hyrax -
Hyrax +
Elephant -
Elephant +
1° & 2°
2° Only


Ladder
IP MYC
+RTG12
Empty
IP MDM2
+RTG12
IP IgG
+RTG12
IP Myc
+GFP
Empty
1% Input
1% Input
alpha P53


Ladder
IP MYC
+RTG12
Empty
IP MDM2
+RTG12
IP IgG
+RTG12
IP Myc
+GFP
Empty
1% Input
1% Input
alpha MDM2


Ladder
IP MYC
+RTG12
Empty
IP MDM2
+RTG12
IP IgG
+RTG12
IP Myc
+GFP
Empty
1% Input
1% Input
alpha Myc