TP53 copy number expansion is associated 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.

* 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, the evolution of extreme sensitivity to genotoxic stress, and a hyperactive TP53 signaling pathway in the elephant (Proboscidean) lineage. Furthermore we show that several of the *TP53* retrogenes (*TP53RTGs*) are transcribed and likely translated. While *TP53RTGs* do not appear to directly function as transcription factors, they do contribute to the enhanced sensitivity of elephant cells to DNA damage and the induction of apoptosis by regulating activity of the 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 (**Figure 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; 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; 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

103 paradox are a decrease in the copy number of oncogenes, an increase in the copy number of
104 tumor suppressor genes (Caulin and Maley, 2011; Leroi et al., 2003; Nunney, 1999), reduced
105 metabolic rates leading to decreased free radical production, reduced retroviral activity and load
106 (Katzourakis et al., 2014), increased immune surveillance, and selection for "cheater" tumors
107 that parasitize the growth of other tumors (Nagy et al., 2007), among many others. Naked mole
108 rats (*Heterocephalus glaber*), for example, which have very long lifespans for a small-bodied
109 organism evolved cells with extremely sensitive contact inhibition likely acting as a constraint on
110 tumor growth and metastasis (Seluanov et al., 2009; Tian et al., 2013). Similarly long-lived blind
111 mole rats (*Splanx sp.*) evolved an enhanced TP53-signaling and necrotic cell death
112 mechanisms that also likely constrains tumor growth (Ashur-Fabian et al., 2004; Avivi et al.,
113 2007; Avivi and Ashur-Fabian, 2005; Gorbunova et al., 2012; Manov et al., 2013). Thus while
114 some of the mechanisms that underlie cancer resistance in small long-lived mammals have
115 been identified, the mechanisms by which large bodied animals evolved enhanced cancer
116 resistance are unknown.

117 Here we use evolutionary genomics and comparative cell biology to explore the
118 mechanisms by which elephants, the largest extant land mammal (**Figure 1B**), have evolved
119 enhanced resistance to cancer. We found that the elephant genome encodes a single *TP53*
120 gene and 19 *TP53* retrogenes, several of which are transcribed and translated in elephant
121 tissues. Comparison of the African and Asian elephant *TP53* gene copy number with the copy
122 number in the genome of the extinct American mastodon, woolly mammoth, and Columbian
123 mammoth indicates that copy number increased relatively rapidly coincident with the evolution
124 of large body-sizes in the Proboscidean lineage. Finally we show that elephant cells have an
125 enhanced response to DNA-damage that is mediated by a hyperactive TP53 signaling pathway
126 and that this augmented TP53 signaling is dependent upon TP53 retrogenes and can be
127 transferred to the cells of other species through exogenous expression of elephant TP53
128 retrogenes. These results suggest that the origin of large body sizes, long lifespans, and
129 enhanced cancer resistance in the elephant lineage evolved at least in part through reinforcing
130 the anti-cancer mechanisms of the major 'guardian of the genome' TP53.
131
132
133
134

135 **RESULTS**

136 **Expansion of the *TP53* repertoire in Proboscideans**

137 We characterized *TP53* copy number in 61 Sarcopterygians (Lobe-finned fishes) with
138 draft or completed genomes, including large, long-lived mammals such as the African elephant
139 (*Loxodonta africana*), Bowhead (*Balaena mysticetus*) and Minke (*Balaenoptera acutorostrata*
140 *scammoni*) whales. We found that all Sarcopterygian genomes encoded a single *TP53* gene
141 and that some lineages also contained a few *TP53* retrogenes (*TP53RTG*), including
142 marsupials, Yangochiropteran bats, and Glires, in which 'processed' *TP53* pseudogenes have
143 previously been reported (Ciotta et al., 1995; Czosnek et al., 1984; Hulla, 1992; Tanooka et al.,
144 1995; Weghorst et al., 1995; Zakut-Houri et al., 1983). We also identified a single *TP53RTG*
145 gene in the lesser hedgehog tenrec (*Echinops telfairi*), which had been previously reported
146 (Belyi et al., 2010), rock hyrax (*Procavia capensis*), and West Indian manatee (*Trichechus*
147 *manatus*). The African elephant genome, however, encoded 19 *TP53RTG* genes (**Figure 2A**),
148 14 of which retain potential to encode truncated proteins (**Table 1**).

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

164

165 **The *TP53RTG* repertoire expanded through repeated segmental duplications**

166 Several mechanisms may have increased the *TP53RTG* copy number in the
167 Proboscidean lineage including serial retrotransposition from the *TP53* gene, serial
168 retrotransposition from the *TP53* and one or more daughter transcribed retrogenes, repeated
169 segmental duplications of chromosomal loci containing *TP53RTG* genes, or some combination
170 of these mechanisms. Consistent with copy number expansion through a single
171 retrotransposition event followed by repeated rounds of segmental duplication, we found that
172 each *TP53RTG* retrogene was flanked by nearly identical clusters of transposable elements
173 (**Figure 3A**) and embedded within a large genomic region with greater than 80% sequence
174 similarity (**Figure 3B**). Next we used progressiveMAUVE to align the 18 elephant contigs
175 containing *TP53RTG* retrogenes and found that they were all embedded within large locally
176 collinear blocks that span nearly the entire length of some contigs (**Figure 3C**), as expected for
177 segmental duplications.
178

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

180 Our observation that *TP53RTG* genes expanded through segmental duplications
181 suggests they may have a tree-like phylogenetic history that preserves information about when
182 in the evolution of Proboscideans the duplicates occurred. Therefore we assembled a dataset of
183 *TP53*/*TP53RTG* orthologs from 65 diverse mammals and jointly inferred the *TP53*/*TP53RTG*
184 gene tree and duplication dates in a Bayesian framework to determine if *TP53*/*TP53RTG* copy
185 number was correlated with body size evolution in Proboscideans. For comparison we also
186 inferred the phylogenetic history of *TP53*/*TP53RTG* genes using maximum likelihood and an
187 additional Bayesian method. We found all phylogenetic inference methods inferred that the
188 *TP53RTG* genes from elephant, hyrax, and manatee formed a well-supported sister clade to the
189 canonical genes from these species, whereas the tenrec *TP53* and *TP53RTG* genes formed a
190 separate well-supported clade (**Figure 4A** and **Figure 4-figure supplement 1**). These data
191 indicate that retrotransposition of *TP53* occurred independently in tenrecs and in the elephant,
192 hyrax, and manatee stem-lineage (Paenungulata), followed by expansion of *TP53RTG* genes in
193 the Proboscidean lineage.
194

195 Based on our time-calibrated phylogeny, we inferred that the initial retrotransposition of
196 the *TP53* gene in the Paenungulata stem-lineage occurred approximately 64 MYA (95% HPD =
197 62.3-66.2 MYA; **Figure 4A**). This was followed by a period of ~25 million years during which no
198 further retrotranspositions or segmental duplications were fixed in the genome, however, the
199 *TP53RTG* gene family rapidly expanded after ~40 MYA (95% HPD = 30.8-48.6 MYA; **Figure**
200 **4A**). To correlate *TP53*/*TP53RTG* copy number and the origin of large body sizes in
201 Proboscideans we estimated *TP53*/*TP53RTG* copy number through time and gathered data on
202 ancient Proboscidean body sizes from the literature (Evans et al., 2012; Smith et al., 2010). We
203 found that the increase in *TP53*/*TP53RTG* copy number in the Proboscidean lineage and
204 Proboscidean body size evolution closely mirrored each other (**Figure 4B**).
205
206 ***TP53RTG12* is transcribed from a transposable element derived promoter**

207 If expansion of the *TP53RTG* gene repertoire played a role in the resolution of Peto's
208 paradox during the evolution of large bodied Proboscideans, then one or more of the *TP53RTG*
209 genes should be transcribed. Therefore we generated RNA-Seq data from Asian elephant
210 dermal fibroblasts, African elephant term placental villus and adipose tissue, and used
211 previously published RNA-Seq data from Asian elephant PBMCs (Reddy et al., 2015) and
212 African elephant fibroblasts (Cortez et al., 2014) to determine if *TP53RTG* genes were
213 transcribed. We found that the *TP53* and *TP53RTG12* genes were robustly transcribed in all
214 samples, whereas *TP53RTG3* and *TP53RTG18* transcripts were much less abundant (**Figure**
215 **5A**). To confirm that the African and Asian elephant *TP53RTG* genes were transcribed, we
216 designed a set of PCR primers specific to the *TP53* and *TP53RTG* genes that flank a diagnostic
217 15-30bp deletion in *TP53RTG* genes (**Figure 5-figure supplement 1**) and used RT-PCR to
218 assay for expression in Elephant fibroblast cDNA generated from DNase treated RNA.
219 Consistent with transcription of the *TP53RTG* genes, we amplified PCR products at the
220 expected size for the *TP53* and *TP53RTG* transcripts but did not amplify PCR products from
221 negative control (no reverse transcriptase) samples (**Figure 5B**). Sanger sequencing of the
222 cloned PCR products confirmed transcription of *TP53RTG12* and *TP53RTG18/19* (**Figure 5-**
223 **figure supplement 1**) in African elephant and *TP53RTG12* and *TP53RTG13* in Asian elephant
224 fibroblast. We note that we used a Poly-T primer for cDNA synthesis, thus the amplification of
225 *TP53RTG* transcripts indicates that these transcripts are poly-adenylated.
226

227 Most retrogenes lack native regulatory elements such as promoters and enhancers to
228 initiate transcription, thus transcribed *TP53RTG* genes likely co-opted existing regulatory
229 elements or evolved regulatory elements *de novo*. To identify putative transcriptional start sites
230 and promoters of the highly expressed *TP53RTG12* gene we used geneid and GENESCAN to
231 computationally predict exons in the African elephant gene and mapped the African and Asian
232 elephant fibroblast RNA-Seq data onto scaffold_825, which encodes the *TP53RTG12* gene. We
233 found that both computational methods predicted an exon ~2kb upstream of the ENSEMBL
234 annotated *TP53RTG12* gene, within an RTE-type non-LTR retrotransposon (RTE1_LA) that we
235 annotated as Afrotherian-specific (**Figure 5C**). Consistent with this region encoding a
236 transcribed 5'-UTR, a peak of reads mapped within the predicted 5'-exon and within the
237 RTE1_LA retrotransposon (**Figure 5C**).
238
239 We attempted to identify the transcription start site of the *TP53RTG12* gene using
240 several 5'-RACE methods, however, we were unsuccessful in generating PCR products from
241 either African Elephant fibroblast or placenta cDNA, or Asian elephant fibroblast cDNA.
242 Therefore, we designed a set of 34 PCR primers tiled across the region of scaffold_825 that
243 encodes the *TP53RTG12* gene and used these primers to amplify PCR products from African
244 and Asian Elephant fibroblast cDNA generated from DNase treated RNA. We then
245 reconstructed the likely *TP53RTG12* promoter, transcription start site, and exon-intron structure
246 from the pattern of positive PCR products. These data suggest that the major transcription
247 initiation site of *TP53RTG12* is located within a RTE1_LA class transposable element (**Figure**
248 **5C**).
249
250 Next we tested the ability of the African and Asian elephant RTE1_LA elements and the
251 RTE1_LA consensus sequence (as a proxy for the ancestral RTE1_LA sequence) to function as
252 a promoter in transiently transfected African and Asian elephant fibroblasts when cloned into the
253 promoterless pGL4.10[*luc2*] luciferase reporter vector. We found that the African and Asian
254 elephant RTE1_LA elements increased luciferase expression 3.03-fold (t-test, P=2.41×10^{-8}) and
255 1.37-fold (t-test, P=2.60×10^{-4}), respectively, compared to empty vector controls (**Figure 5D**).
256 However, luciferase expression from the pGL4.10[*luc2*] vector containing the RTE1_LA
257 sequence was not significantly different than the empty vector control in either Asian (0.96-fold;
258 t-test, P=0.61) or African elephant fibroblasts (0.95-fold; t-test, P=0.37; **Figure 5D**). These data
259 indicate that transcription of *TP53RTG12* likely initiates within a RTE1_LA-derived promoter, but

260 that the ability of this RTE1_LA element to function as a promoter is not an ancestral feature of
261 RTE1_LA elements.
262
263 **Elephant cells likely express TP53RTG proteins and numerous isoforms of TP53**
264 To determine if *TP53RTG* transcripts are translated, we treated African elephant, Asian
265 elephant, and hyrax cells with 50 J/m^2 UV-C (to stabilize TP53) and the proteasome inhibitor
266 MG-132 (to block protein degradation), and assayed for TP53/TP53RTG proteins by Western
267 blotting total cell protein with a polyclonal TP53 antibody (FL-393) that we demonstrated
268 recognizes Myc-tagged TP53RTG12. We identified bands in both African and Asian elephant
269 and hyrax total cell protein at the expected size for the full length p53, Δ133p53β/γ, and p53ψ-
270 like isoforms of the TP53 protein (Khoury and Bourdon, 2010) as well as high molecular weight
271 bands corresponding to previously reported SDS denaturation resistant TP53 oligomers (Cohen
272 et al., 2008; Ottaggio et al., 2000) and (poly)ubiquitinated TP53 congugates (Sparks et al.,
273 2013) (**Figure 5E**). We also identified an elephant-specific band at the expected size for the
274 TP53RTG12 (19.6kDa) and TP53RTG19 (22.3kDa) proteins, suggesting that the *TP53RTG12*
275 and *TP53RTG19* transcripts are translated in elephant fibroblasts (**Figure 5E** and **Figure 5-**
276 **figure supplement 2**).
277
278 **Elephant cells have an enhanced TP53-dependent DNA-damage response**
279 Our observation that *TP53RTG* genes are expressed suggests that elephant cells may
280 have an altered TP53 signaling system compared to species without an expanded
281 *TP53/TP53RTG* gene repertoire. To directly test this hypothesis we transiently transfected
282 primary African elephant, Asian elephant, South African Rock hyrax (*Procavia capensis*
283 *capensis*), East African aardvark (*Orycteropus afer lademanni*), and Southern Three-banded
284 armadillo (*Tolypeutes matacus*) dermal fibroblasts with a luciferase reporter vector containing
285 two TP53 response elements (pGL4.38[*luc2p*/p53 RE/Hygro]) and Renilla control vector
286 (pGL4.74[*hRluc*/TK]). Next we used a dual luciferase reporter assay to measure activation of the
287 TP53 pathway in response to treatment with three DNA damage inducing agents (mitomycin C,
288 doxorubicin, or UV-C) or nutlin-3a, which inhibits the interaction between MDM2 and TP53 and
289 thus promotes TP53 signaling. We found that elephant cells generally up-regulated TP53
290 signaling in response to lower doses of each drug and UV-C than closely related species
291 without an expanded TP53 gene repertoire (**Figure 6A**), indicating elephant cells have evolved
292 an enhanced TP53 response.

293 To determine the consequences of an enhanced TP53 response we treated primary
294 African and Asian elephant, hyrax, aardvark, and armadillo dermal fibroblasts with mitomycin C,
295 doxorubicin, nutlin-3a, or UV-C and measured cell viability (live-cell protease activity),
296 cytotoxicity (dead-cell protease activity), and the induction apoptosis (caspase-3/7 activation)
297 using an ApoTox-Glo Triplex assay. Consistent with the results from the luciferase assay, we
298 found that lower doses of mitomycin C or doxorubicin induced apoptosis in elephant cells than
299 the other species (**Figure 6B**) and that the magnitude of the response was greater in elephant
300 than other species (**Figure 6A**). Similarly UV-C exposure generally induced more elephant cells
301 to undergo apoptosis than other species (**Figure 6A**). A striking exception to this trend was the
302 response of elephant cells to the MDM2 antagonist nutlin-3a, which elicited a strong TP53
303 transcriptional response (**Figure 6A**) but did not induce apoptosis (**Figure 6B**). Thus we
304 conclude that elephant cells generally upregulate TP53 signaling and apoptosis at lower levels
305 of DNA-damage than other species, but are resistant to nutlin 3-a induced apoptosis.
306
307 **TP53RTG genes are required for enhanced TP53 signaling and DNA-damage responses**
308 To test if *TP53RTG* genes are necessary for the enhanced TP53-dependent DNA-
309 damage response, we cotransfected African elephant fibroblasts with the pGL4.38[*luc2p*/p53
310 RE/Hygro] luciferase reporter vector, the pGL4.74[*hRluc*/TK] Renilla control vector, and either a
311 *TP53RTG*-specific siRNA or a scrambled siRNA control (**Figure 7A**). Next we used a dual
312 luciferase reporter assay to measure activation of the TP53 pathway in response to treatment
313 mitomycin C, doxorubicin, UV-C, or nutlin-3a. As expected given our previous results, we found
314 that African elephant fibroblasts transfected with control siRNA induced TP53 signaling in
315 response to each treatment (**Figure 7A**). In contrast, African elephant fibroblasts transfected
316 with *TP53RTG*-specific siRNA had significantly lower luciferase expression, and thus reduced
317 TP53 signaling, in response to either DNA-damaging agents (mitomycin C, doxorubicin, UV-C)
318 or MDM2 antagonism (nutlin-3a). *TP53RTG* knockdown also elevated baseline TP53 signaling
319 (**Figure 7B**). These data suggest that TP53RTG proteins have at least two distinct functions,
320 inhibiting TP53 signaling in the absence of inductive signals and potentiation of TP53 signaling
321 after the induction of DNA damage.
322
323
324
325

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

328 To test if TP53RTG12 is sufficient to mediate enhanced TP53 signaling and DNA-
329 damage responses we synthesized the African elephant *TP53RTG12* gene (with mouse codon
330 usage) and cloned it into the mammalian expression vector pcDNA3.1(+)/myc-His. We then
331 transiently transfected mouse 3T3-L1 cells with the *TP53RTG12* pcDNA3.1(+)/myc-His
332 expression vector and used the pGL4.38[*luc2P*/p53 RE/Hygro] reporter system and ApoToxGlo
333 assays to monitor activation of the TP53 signaling pathway and the induction of apoptosis in
334 response to treatment with mitomycin C, doxorubicin, nutlin-3a, or UV-C. We found that
335 heterologous expression of TP53RTG12 in mouse 3T3-L1 cells dramatically augmented
336 luciferase expression from the pGL4.38[*luc2P*/p53 RE/Hygro] reporter vector in response to
337 each treatment (**Figure 8A**) compared to empty vector controls, consistent with a enhancement
338 of the endogenous TP53 signaling pathway. Similarly, expression of TP53RTG12 significantly
339 augmented the induction of apoptosis in response to each treatment although the effects sizes
340 were modest (**Figure 8B**). These data indicate that TP53RTG12 acts via a trans-dominant
341 mechanism to enhance the induction of apoptosis by endogenous TP53 and that TP53RTG12 is
342 sufficient to recapitulate at least some of the enhanced sensitivity of elephant cells to DNA
343 damage. Furthermore, our observation that transfection with the *TP53RTG12*
344 pcDNA3.1(+)/myc-His expression vector augments TP53 signaling and apoptosis suggests that
345 the TP53RTG12 protein rather than transcript is responsible for these effects because the
346 *TP53RTG12* transgene was synthetized with mouse codon usage and is only 73% (394/535
347 nts) identical to the elephant *TP53RTG12* gene.
348

349 TP53RTG proteins are unlikely to directly regulate TP53 target genes because they lack
350 critical residues required for nuclear localization, tetramerization, and DNA-binding (**Figure 9A**).
351 Previous studies, for example, have shown the TP53 mutants lacking the tetramerization
352 domain and C-terminal tail are unable to bind DNA or transactivate luciferase expression from a
353 reporter vector containing TP53 response elements (Kim et al., 2012). Similarly the p53Ψ
354 isoform, which is truncated in the middle of the DNA binding domain and lacks the nuclear
355 localization signal and oligomerization domain, is unable to bind DNA and is transcriptionally
356 inactive (Senturk et al., 2014). Unexpectedly, we found that the GFP-tagged TP53RTG12
357 protein was both cytoplasmic and nuclear localized in transfected African elephant fibroblasts
358 (**Figure 9B**), suggesting it interacts with another nuclear localized protein to enter the nucleus.

359 Despite relatively strong nuclear localization, however, cells co-transfected with the *TP53RTG12*
360 pcDNA3.1(+)/myc-His expression vector and the pGL4.38[*luc2P*/p53 RE/Hygro] luciferase
361 reporter vector did not have elevated luciferase expression compared to controls, suggesting
362 that TP53RTG12 is transcriptionally inactive or requires a cofactor to regulate target genes
363 (**Figure 9C**).
364
365 **A model for TP53RTG function – Decoy or guardian?**
366 While TP53RTG12 does not appear to directly regulate gene expression, many of the
367 TP53RTG proteins (including TP53RTG12) retain the MDM2 interaction motif in the
368 transactivation domain and dimerization sites in the DNA binding domain (**Figure 9A**). These
369 data suggest at least two non-exclusive models of TP53RTG action: 1) TP53RTG proteins may
370 act as 'decoys' for the MDM2 complex allowing the canonical TP53 protein to escape negative
371 regulation (**Figure 10A**); and 2) TP53RTG proteins may protect canonical TP53 from MDM2
372 mediated ubiqutination, which requires tetramerization (Kubbutat et al., 1998; Maki, 1999), by
373 dimerizing with canonical TP53 and thereby preventing the formation of tetramers (**Figure 10F**).
374
375 The decoy model depends on the ability of TP53RTG proteins to physically interact with
376 MDM2. Previous crystallographic studies of the TP53/MDM2 interaction have shown that a trio
377 of residues in TP53 (F19, W23, and L26) insert deeply into a hydrophobic cleft in MDM2, which
378 stabilizes the interaction (Kussie et al., 1996). We identified a W23G substitution in all TP53RTG
379 proteins at a site that is invariant for tryptophan in TP53 proteins including African and Asian
380 elephant TP53 (**Figure 10B**), suggesting that TP53RTG proteins may be unable to physically
381 interact with MDM2. To infer the structural and functional consequences of the TP53RTG W23G
382 substitution we generated a homology model of the elephant TP53RTG12/MDM2 complex using
383 I-TASSER/ModRefiner (Roy et al., 2010; Xu and Zhang, 2011; Zhang, 2008) and the crystal
384 structure of the MDM2/TP53 dimer as a template (Kussie et al., 1996). We found that the
385 TP53RTG12 transactivation domain was inferred to be a short α-helix (**Figure 10C**) and was
386 very similar to the template structure (RMSD: 1.756), however, the W23G substitution is
387 predicted to abolish crucial hydrophobic interactions between the amphipathic α-helix of TP53
388 and the hydrophobic cleft MDM2. Indeed, three methods (Pires et al., 2014) inferred that the
389 W23G substitution is destabilizing on the TP53RTG12/MDM2 interaction (mCSM ΔΔG=-2.42,
390 SDM ΔΔG=-5.46, DUET ΔΔG=-2.38) (**Figure 10D**). To experimentally test for an interaction
391 between TP53RTG12 and MDM2 we transiently transfected HEK-293 cells with the

392 TP53RTG12 pcDNA3.1(+)/myc-His expression vector, immunoprecipitated endogenous human
393 MDM2, and assayed for co-immunoprecipitation of TP53RTG12 by Western blotting. While we
394 efficiently co-immunoprecipitated endogenous human TP53 we did not co-immunoprecipitate
395 Myc-tagged TP53RTG12 (**Figure 10E** and **Figure 10-figure supplement 1**), consistent with a
396 lack of interaction between TP53RTG12 and MDM2.
397
398 Unlike the decoy model, the guardian model of TP53RTG function depends upon a
399 physical interaction between TP53RTG and TP53. The TP53 dimer is stabilized by hydrophobic
400 and polar interactions including a shell of nonpolar interactions formed by P177, H178, M243,
401 and G244 and a stabilization network next to the nonpolar layer formed by charged residues
402 from the two monomers (R181, E180, and R174). In addition, several polar and charged
403 residues nearby but not within the dimerization interface contribute to the stability of the
404 interacting monomers including D184 with R175, most of which are conserved in TP53RTG
405 proteins (**Figure 10G**). To infer if derived residues in the TP53RTG12 dimerization interface
406 might disrupt a physical interaction between TP53RTG12 and TP53 we generated a homology
407 model of the TP53RTG12/TP53 dimer using I-TASSER/ModRefiner (Roy et al., 2010; Xu and
408 Zhang, 2011; Zhang, 2008) and the crystal structure of the TP53 tetramer as a template
409 (Kitayner et al., 2006). We found that the TP53RTG12 dimerization interface was inferred to be
410 a 2-4 residue α-helix (**Figure 10H**) that was nearly identical to the template structure (RMSD:
411 0.605), suggesting TP53RTG12-specific residues in dimerization interface are unlikely to disrupt
412 the structure of the interface. Unlike the MDM2 interaction site, TP53RTG12-specific
413 substitutions were predicted to maintain intermolecular hydrophobic interactions with TP53
414 (**Figure 10H**). Consistent with maintenance of dimerization potential, the net ΔΔG of the derived
415 amino acid substitutions in the TP53RTG12 dimerization interface were under 2 (**Figure 10I**). To
416 experimentally test for an interaction between TP53RTG12 and TP53 we transiently transfected
417 HEK-293 cells with the TP53RTG12 pcDNA3.1(+)/myc-His expression vector,
418 immunoprecipitated TP53RTG12 with a Myc antibody, and assayed for co-immunoprecipitation
419 of endogenous human TP53 by Western blotting. We found that Myc-tagged TP53RTG12
420 efficiently co-immunoprecipitated endogenous human TP53 (**Figure 10J**). These data are
421 consistent with a physical interaction between TP53RTG12 and TP53 but not between
422 TP53RTG12 and MDM2, supporting the guardian model.
423
424

425 **DISCUSSION**

426 A major developmental and life history constraint on the evolution of large body sizes
427 and long lifespans in animals is an increased risk of developing cancer. There is no correlation,
428 however, between body size or lifespan and cancer risk because large and long-lived organisms
429 have evolved enhanced cancer suppression mechanisms that delay the development of cancer
430 until post-reproduction (when selection generally cannot act). This simple evolutionary rational
431 demands mechanistic explanations (Peto, 2015), which have thus far been elusive. Here we
432 show that the master tumor suppressor TP53, which is essential for preventing cancer because
433 it triggers proliferative arrest and apoptosis in response to a variety of stresses such as DNA
434 damage, was retroduplicated in the Paenungulate stem-lineage and rapidly increased in copy
435 number through repeated segmental duplications during with the evolution of Proboscideans.
436 The expansion of the *TP53RTG* gene family occurred coincident with the evolution of large body
437 sizes and enhanced sensitivity of elephant cells to genotoxic stress, suggesting that
438 Proboscideans resolved Peto's paradox at least in part through the evolution of augmented
439 TP53 signaling.
440

441 **Comparison to previous studies of elephant TP53**

442 Previous studies have suggested that the *TP53* gene family expanded in the elephant
443 lineage (Abegglen et al., 2015; Caulin and Maley, 2011; Caulin et al., 2015), however, these
444 studies did not establish the mechanism by which the *TP52RTG* gene family expanded. Several
445 potential mechanisms could have increased *TP52RTG* copy number, including serial
446 (independent) retrotranspostion from the parent *TP53* gene, a single retrotranspostion event
447 followed by repeated rounds of segmental duplication of *TP52RTG* containing loci,
448 retrotransposition of expressed *TP52RTG* genes, or some combination of these models. Each
449 model is associated with a distinct set of genomic 'fingerprints'. If copy number expanded
450 through independent retrotransposition events, for example, *TP53RTG* encoding regions of the
451 genome will not be homologous whereas the model of a single retrotranspostion event followed
452 by repeated rounds of segmental duplication predicts that the *TP53RTG* encoding loci will be
453 homolgous. Consistent with copy number expansion through a single retrotransposition event
454 followed by repeated rounds of segmental duplication, we found that flaking regions of each
455 *TP52RTG* locus were homologous and contained the same unique combination of transposable

456 elements. Indeed, the 3'-end of each duplicate terminates at a ~5kb long L1MB5 LINE element
457 suggesting that transposable element mediated recombination may have played a role in
458 promoting segmental duplication.

459 If *TP53RTG* copy number expansion played a causal role in evolution of enhanced
460 cancer resistance in elephants then the gene family must have expanded prior to or coincident
461 with the evolution of increased body sizes in Proboscideans rather than after the evolution of
462 large bodies but before the African and Asian elephant lineages diverged ~8 MYA (Rohland et
463 al., 2010; 2007). Thus dating the expansion of the *TP53RTG* gene family is essential for
464 determining if *TP53RTG* genes played a role in the evolution of enhanced cancer resistance in
465 elephants. Previous studies, however, did not establish when *TP53RTG* copy number expanded
466 in the evolution of Proboscideans. Fortunately our observation that copy number expansion
467 occurred through segmental duplications allowed us to use molecular phylogenetic methods to
468 date each duplication event. These data indicate that initial *TP53RTG* retrotransposition event
469 occurred in the Paenungulate stem-lineage ~64 MYA, followed rapid expansion after ~40 MYA.
470 We also found that the increase in *TP53RTG* copy number occurred coincident with the
471 evolution of large bodies in the Proboscidean lineage, implicating copy number expansion in the
472 resolution of Peto's paradox.

473 While Abegglen et al. (2015) found that elephant and human cells had different
474 sensitivities to ionizing radiation, their taxon sampling did not allow for polarizing which species
475 was different. Indeed, their taxon sampling does not allow for a Eutherian out-group. To
476 determine if elephants have a derived sensitivity to genotoxic stress, we compared the response
477 of fibroblasts from an African savannah elephant, an Asian elephant, and closely related out-
478 group species to DNA damage inducing agents. Our out-group species included the South
479 African Rock hyrax, the closest living relative of elephants, the East African aardvark, an
480 Afrotherian from the sister lineage of the Paenungulates, and the Southern Three-banded
481 armadillo, an Atlantogenatan from the sister lineage of the Afrotherians. Our results indicate that
482 elephant cells are particularly sensitive to genotoxic tress, suggesting that this sensitivity
483 evolved coincident with the evolution of large body sizes and an expanded *TP53* gene repertoire
484 in Proboscideans.
485
486 **An embarrassment of riches?**

487 Our observation that the elephant genome contains 19 *TP53RTG* genes raises
488 numerous questions: Do these loci, for example, encode functional genes or pseudogenes and
489 what processes underlie copy number expansion? Answers to these questions are essential for
490 understanding whether *TP53RTG* genes are casually related to the resolution of Peto's paradox
491 in Proboscideans or if they are irrelevant relicts of ancient transposition events, like so many
492 other pseudogenes that riddle mammalian genomes. Unfortunately it is difficult to answer these
493 questions.

494 Which *TP53RTG* loci encode functional genes and which encode pseudogenes is not
495 easy to infer. Many duplicate genes are preserved because they evolve tissue-specific or
496 developmental-stage specific expression patterns (subfunctionalization), new functions
497 (neofunctionalization) that resolve redundancy between duplicates, or reduced expression levels
498 that preserves correct expression dosage. Exhaustively characterizing gene expression in
499 elephant tissues is difficult because appropriate tissue samples are unavailable, therefore we
500 are unable to definitively determine which *TP53RTG* loci are transcribed. Our RNA-Seq and RT-
501 PCR/Sanger sequencing data indicate that at least *TP53RTG12*, *TP53RTG18/19*, and
502 *TP53RTG13* are transcribed in dermal fibroblasts. Abegglen et al. (2015) used RT-PCR and
503 Sanger sequencing to show two distinct transcripts were expressed in elephant PBMCs, but
504 they did not assign the loci to which these transcripts correspond. We analyzed the
505 chromatograms shown in Abegglen et al. eFigure 4 and found that the 185bp product is likely a
506 transcript from the *TP553RTG14* gene and the 201bp product is likely a transcript from the
507 *TP553RTG5* gene. Thus, our combined data suggest that at least five *TP53RTG* genes are
508 transcribed. Furthermore we did not observe *TP553RTG14* or *TP553RTG5* expression in
509 adipose, placenta, or fibroblasts suggesting that the expression of some *TP53RTG* genes is
510 tissue-specific.

511 The large number of *TP53RTG* loci in elephants combined with the observation that only
512 five are transcribed suggests that this gene family may evolve by a birth and death process, in
513 which new genes are created by duplication and some duplicates are maintained in the genome
514 whereas others become nonfunctional or deleted similar to other large genes families such as
515 histones (González-Romero et al., 2010; Rooney et al., 2002) and venom genes (Lynch, 2007).
516 Under this model selection acts to maintain a minimal number of functional copies (functional
517 copy number) rather than total copy number, the increase in total copy number is driven by the

518 total number of loci and the rates of duplication, loss, and fixation. Thus the overall increase in
519 total *TP53RTG* copy number may be a selectively neutral processes, driven simply by higher
520 rates of duplication and/or fixation than loss.
521
522 **Is TP53RTG12 a decoy, a guardian, or something else?**
523 Our results demonstrate that elephant cells induce TP53 signaling and trigger apoptosis
524 at lower thresholds of genotoxic stress than closely related species without an expanded *TP53*
525 repertoire and that this reduced sensitivity is dependent upon *TP53RTG* genes. Furthermore
526 heterologous expression of TP53RTG12 in mouse cells was sufficient to augment endogenous
527 TP53 signaling and recapitulate an elephant-like sensitivity to genotoxic stress, indicating that
528 TP53RTGs acts through a transdominant mechanism. While the mechanism of action is
529 unclear, *TP53RTG* genes may augment TP53 signaling through several non-exclusive
530 mechanisms including functioning as non-coding RNAs (Poliseno et al., 2010), protein 'decoys'
531 for the MDM2 complex that allowing canonical TP53 to escape negative regulation (Abegglen et
532 al., 2015), and protein 'guardians' that protect canonical TP53 from MDM2 mediated
533 ubiqutination.
534
535 Consistent with the guardian model, we found that Myc-tagged TP53RTG12 efficiently
536 co-immunoprecipitated with endogenous TP53 but not with endogenous MDM2. The lack of an
537 interaction between TP53RTG12 and MDM2 likely results from a W23G substitution that is
538 predicted to abolish a crucial hydrophobic interaction between the amphipathic α-helix of TP53
539 and the hydrophobic cleft MDM2. The W23G substitution is found in all TP53RTG proteins but
540 not African and Asian elephant TP53, indicating that it occurred before the expansion of the
541 *TP53RTG* gene family and likely prevents interaction of any TP53RTG protein and MDM2.
542 Previous studies have shown that tetramerization of TP53 is required for its efficient MDM2-
543 medaited ubiquitination (Kubbutat et al., 1998; Maki, 1999), suggesting that TP53RTG proteins
544 may dimerize with and protect TP53 from ubiquitination thereby contributing to a standing pool
545 of TP53 that is able to rapidly respond to DNA damage. Consistent with this mechanism,
546 transgenic mice with an increase in TP53 copy number (Garcia-Cao et al., 2002) or a
547 hypomorphic *Mdm2* allele have elevated basal TP53 activity and are resistant to tumor
548 formation (Mendrysa, 2006), indicating that shifting the TP53-MDM2 equilibrium away from
549 TP53 degradation can directly promote cancer resistance.
550

551 The 'decoy' model (Abegglen et al., 2015) has also been challenged because it would
552 allow for activation of the TP53 signaling pathway in the absence of DNA damage (Perez and
553 Komiya, 2016), which is generally lethal in animal models (Hoever et al., 1994; Lozano, 2010).
554 Thus if TP53RTG proteins allow TP53 to escape negative regulation by MDM2, how do
555 elephants tolerate elevated basal TP53 levels? Clearly further studies are required to
556 specifically test whether TP53RTG proteins interfere with the interaction between the MDM2
557 complex and TP53, protect TP53 from ubiquitination, or have other functions.
558
559 **What do extra *TP53* genes cost?**

560 Our observation that functional *TP53* duplicates only occur in the elephant lineage (and
561 perhaps some bats) suggests that increased *TP53* dosage has a cost. Previous studies found
562 that transgenic mice that overexpress *Trp53* were cancer resistant but had major life history
563 tradeoffs including slower pre- and post-natal growth rates and reduced size (Maier et al., 2004),
564 a shortened lifespan (Maier et al., 2004), accelerated aging (Tyner et al., 2002), and reduced
565 fertility (Allemand et al., 1999; Maier et al., 2004), as well as developmental tradeoffs including
566 reduced proliferation, cellularity, and atrophy across multiple organ and tissue systems (Dumble
567 et al., 2007; Maier et al., 2004), defective ureteric bud differentiation, and small kidneys (Godley
568 et al., 1996). Thus, increases in *TP53* copy number protects against cancer but appears to
569 come with the cost of developmental delays, accelerated aging, and reduced fertility (Campisi,
570 2003; Donehower, 2002; Ferbeyre and Lowe, 2002; Rodier et al., 2007).
571
572 Reduced male fertility appears to be a particularly expensive cost of increased *TP53*
573 dosage. Allemand et al. (1999), for example, generated transgenic lines of mice with one
574 (MTp53-176), two (MTp53-112), or 15 (MTp53-94) extra copies of the *Trp53* gene fused to the
575 inducible promoter of the metallothionein I (MT) gene. They found that transgenic males with the
576 highest *Trp53* dosage were nearly infertile because the majority of developing spermatids
577 underwent apoptosis before developing into mature sperm, males with intermediate dosage
578 were subfertile and produced sperm with abnormal morphologies (teratozoospermia) indicative
579 of defective terminal differentiation of postmeiotic cells, whereas males with the lowest dosage
580 were fertile. Similarly, Maier et al. (2004) generated a transgenic line that overexpresses the p44
581 isoform of *Trp53* (Δ40p53) and found that both males and females exhibited shortened
582 reproductive lifespans. Males, however, were more severely affected than females with a

583 catastrophic loss of sperm-producing cells and massive degeneration of the seminiferous
584 epithelium leading to a “Sertoli-cell only” phenotype.
585
586 In contrast to transgenic mice that either overexpress full length TP53 or the Δ40p53
587 isoform, transgenic 'super p53' mice with one (p53-tg) or two (p53-tg^{b}) additional copies of the
588 endogenous *TP53* locus have an enhanced DNA-damage response and are tumor resistant, yet
589 age normally and are fertile (Garcia-Cao et al., 2002; 2006; Matheu et al., 2007). Enhanced
590 tumor suppression and normal aging is also observed in *Mdm2*$^{puro/\Delta 7-12}$ transgenic mice, which
591 have one hypomorphic and one null allele of *Mdm2*, express ~30% of the wild-type level of
592 Mdm2, and have constitutively high TP53 activity (Mendrysa, 2006; Mendrysa et al., 2003).
593 These results suggest that the costs of increased *TP53* copy number are incurred above a
594 threshold of about 2-3 extra copies and can be reduced by maintaining normal regulation of
595 *TP53* transcription and negative post-transcriptional regulation by MDM2. Consistent with this
596 model, deletion of either *Mdm2* or *Mdm4* rescues *Trp53*-dependent embryonic lethality in mice
597 (Finch et al., 2002; Migliorini et al., 2002; Parant et al., 2001) and zebrafish (Chua et al., 2015).
598
599 Collectively these data indicate that increased *TP53* dosage comes with a cost, and
600 suggest that Proboscideans evolved a mechanism that reduced this cost, broke a major
601 developmental and evolutionary constraint on *TP53* copy number, or perhaps evaded paying
602 the cost of increased *TP53* copy number altogether. *TP53RTG* genes, for example, are
603 transcribed from a transposable element derived promoter that is evolutionarily younger than the
604 retrogenes. Thus, the initial *TP53RTG* genes were unlikely to be transcribed and incur a cost.
605 Similarly our phylogenetic analyses indicates that copy number expanded after the initial
606 *TP53RTG* gene acquired several mutations, including a premature stop codon that terminates
607 the protein before the DNA-binding domain, which likely reduces or completely eliminates the
608 costs associated with *TP53* target gene regulation. Although more detailed evolutionary and
609 comparative analyses are required to determine if *TP53RTG* genes incurred a cost, it is possible
610 that the costs were minimized because functional *TP53RTG* genes evolved through non-
611 functional intermediates, which accumulated loss of function mutations that minimized
612 redundancy with *TP53*.
613
614
615

649 **EXPERIMENTAL PROCEDURES**

650 **Identification of *TP53/PT53RTG* genes in Sarcopterygian genomes**

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

659

660 We thus used BLAT to characterize the *TP53* copy number in Human (*Homo sapiens*;
661 GRCh37/hg19), Chimp (*Pan troglodytes*; CSAC 2.1.4/panTro4), Gorilla (*Gorilla gorilla gorilla*;
662 gorGor3.1/gorGor3), Orangutan (*Pongo pygmaeus abelii*; WUGSC 2.0.2/ponAbe2), Gibbon
663 (*Nomascus leucogenys*; GGSC Nleu3.0/nomLeu3), Rhesus (*Macaca mulatta*; BGI
664 CR_1.0/rheMac3), Baboon (*Papio hamadryas*; Baylor Pham_1.0/papHam1), Marmoset
665 (*Callithrix jacchus*; WUGSC 3.2/calJac3), Squirrel monkey (*Saimiri boliviensis*; Broad/saiBol1),
666 Tarsier (*Tarsius syrichta*; Tarsius_syrichta2.0.1/tarSyr2), Bushbaby (*Otolemur garnettii*;
667 Broad/otoGar3), Mouse lemur (*Microcebus murinus*; Broad/micMur1), Chinese tree shrew
668 (*Tupaia chinensis*; TupChi_1.0/tupChi1), Squirrel (*Spermophilus tridecemlineatus*;
669 Broad/speTri2), Mouse (*Mus musculus*; GRCm38/mm10), Rat (*Rattus norvegicus*; RGSC
670 5.0/rn5), Naked mole-rat (*Heterocephalus glaber*; Broad HetGla_female_1.0/hetGla2), Guinea
671 pig (*Cavia porcellus*; Broad/cavPor3), Rabbit (*Oryctolagus cuniculus*; Broad/oryCun2), Pika
672 (*Ochotona princeps*; OchPri3.0/ochPri3), Kangaroo rat (*Dipodomys ordii*; Broad/dipOrd1),
673 Chinese hamster (*Cricetulus griseus*; C_griseus_v1.0/criGri1), Pig (*Sus scrofa*; SGSC
674 Sscrofa10.2/susScr3), Alpaca (*Vicugna pacos*; Vicugna_pacos-2.0.1/vicPac2), Dolphin
675 (*Tursiops truncatus*; Baylor Ttru_1.4/turTru2), Cow (*Bos taurus*; Baylor Btau_4.6.1/bosTau7),
676 Sheep (*Ovis aries*; ISGC Oar_v3.1/oviAri3), Horse (*Equus caballus*; Broad/equCab2), White
677 rhinoceros (*Ceratotherium simum*; CerSimSim1.0/cerSim1), Cat (*Felis catus*; ICGSC
678 Felis_catus 6.2/felCat5), Dog (*Canis lupus familiaris*; Broad CanFam3.1/canFam3), Ferret
679 *(Mustela putorius furo*; MusPutFur1.0/musFur1), Panda (*Ailuropoda melanoleuca*; BGI-
680 Shenzhen 1.0/ailMel1), Megabat (*Pteropus vampyrus*; Broad/pteVam1), Microbat (*Myotis*
681 *lucifugus*; Broad Institute Myoluc2.0/myoLuc2), Hedgehog (*Erinaceus europaeus*;

682 EriEur2.0/eriEur2), Shrew (*Sorex araneus*; Broad/sorAra2), Minke whale (*Balaenoptera*
683 *acutorostrata scammoni*; balAcu1), Bowhead Whale (*Balaena mysticetus*; v1.0), Rock hyrax
684 (*Procavia capensis*; Broad/proCap1), Sloth (*Choloepus hoffmanni*; Broad/choHof1), Elephant
685 (*Loxodonta africana*; Broad/loxAfr3), Cape elephant shrew (*Elephantulus edwardii*;
686 EleEdw1.0/eleEdw1), Manatee (*Trichechus manatus latirostris*; Broad v1.0/triMan1), Tenrec
687 (*Echinops telfairi*; Broad/echTel2), Aardvark (*Orycteropus afer afer*; OryAfe1.0/oryAfe1),
688 Armadillo (*Dasypus novemcinctus*; Baylor/dasNov3), Opossum (*Monodelphis domestica*;
689 Broad/monDom5), Tasmanian devil (*Sarcophilus harrisii*; WTSI Devil_ref v7.0/sarHar1), Wallaby
690 (*Macropus eugenii*; TWGS Meug_1.1/macEug2), Platypus (*Ornithorhynchus anatinus*; WUGSC
691 5.0.1/ornAna1), Medium ground finch (*Geospiza fortis*; GeoFor_1.0/geoFor1), Zebra finch
692 (*Taeniopygia guttata*; WashU taeGut324/taeGut2), Budgerigar (*Melopsittacus undulatus*;
693 WUSTL v6.3/melUnd1), Chicken (*Gallus gallus*; ICGSC Gallus_gallus-4.0/galGal4), Turkey
694 (*Meleagris gallopavo*; TGC Turkey_2.01/melGal1), American alligator (*Alligator mississippiensis*;
695 allMis0.2/allMis1), Painted turtle (*Chrysemys picta bellii*; v3.0.1/chrPic1), Lizard (*Anolis*
696 *carolinensis*; Broad AnoCar2.0/anoCar2), X. tropicalis (*Xenopus tropicalis*; JGI 7.0/xenTro7),
697 Coelacanth (*Latimeria chalumnae*; Broad/latCha1).
698
699 ***TP53/TP53RTG* copy number estimation in Proboscidean genomes**
700 *Identification of TP53/TP53RTG genes in the Asian elephant genome*
701 We used previously published whole genome shotgun sequencing data from an Asian
702 elephant (*Elephas maximus*) generated on an Illumina HiSeq 2000{'Hou'} to estimate the
703 *TP53/TP53RTG* copy number in the genome. For these analyses we combined data from two
704 individual elephants (ERX334765 and ERX334764) into a single dataset of 151,482,390 76-nt
705 paired end reads. We then mapped Asian elephant reads onto the African elephant
706 *TP53/TP53RTG* genes using Bowtie2 in paired-end mode, with the local alignment and 'very
707 sensitive' options and Cufflinks (version 0.0.7) to assemble mapped reads into genes which we
708 treated as putative Asian elephant orthologs. We identified 14 putative 1:1 orthologous
709 *TP53RTG* genes in the Asian elephant genome, including *TP53RTG1*, *TP53RTG2*, *TP53RTG3*,
710 *TP53RTG4*, *TP53RTG5*, *TP53RTG6*, *TP53RTG9*, *TP53RTG10*, *TP53RTG11*, *TP53RTG13*,
711 *TP53RTG14*, *TP53RTG17*, *TP53RTG18*, and *TP53RTG19*.
712
713 *Identification of TP53/TP53RTG genes in the Columbian and Woolly mammoth genomes*

714 We used previously published whole genome shotgun sequencing data from an ~11,000
715 year old Columbian mammoth (*Mammuthus columbi*) generated on an Illumina HiSeq
716 1000{'Sparks'} to estimate the *TP53/TP53RTG* copy number in the genome. For these analyses
717 we combined data from three individual mammoths (SRX329134, SRX329135, SRX327587,
718 SRX327586, SRX327583, SRX327582) into a single dataset of 158,704,819 102-nt paired end
719 reads. We then mapped Asian elephant reads onto the African elephant *TP53/TP53RTG* genes
720 using Bowtie2 in paired-end mode, with the local alignment and 'very sensitive' options, and
721 Cufflinks (version 0.0.7) to assemble mapped reads into genes which we treated as putative
722 Asian elephant orthologs. We identified 14 putative 1:1 orthologous *TP53RTG* genes in the
723 Columbian mammoth genome, including *TP53RTG10*, *TP53RTG11*, *TP53RTG13*, *TP53RTG14*,
724 *TP53RTG16*, *TP53RTG17*, *TP53RTG19*, *TP53RTG2*, *TP53RTG3*, *TP53RTG4*, *TP53RTG5*,
725 *TP53RTG6*, *TP53RTG8*, *TP53RTG9*.
726
727 *Identification of TP53/TP53RTG genes in the American mastodon genome*
728 We used previously published whole genome shotgun sequencing data generated on a
729 Roche 454 Genome Sequencer (GS FLX) to estimate the *TP53* copy number in a 50,000-
730 130,000 year old American mastodon (*Mammut americanum*){'Langmead'}. To estimate *TP53*
731 copy number we combined 454 sequencing data from 'Library A' (SRX015822) and 'Library B'
732 (SRX015823) into a single dataset of 518,925 reads. Reads were converted from FASTQ to
733 FASTA and aligned to African elephant *TP53* gene and retrogene sequences using Lastz. We
734 used the 'Roche-454 90% identity' mapping mode, not reporting matches lower than 90%
735 identity. Three American mastodon reads were identified that mapped to at least one *TP53*
736 gene/retrogene. The likely identity of these reads were determined by using BLAT to map their
737 location in the African elephant (Broad/loxAfr3) genome; mastodon reads mapping to a single
738 elephant gene with >98% identity were considered likely mastodon orthologs. This strategy
739 mapped one read to the *TP53TRG8* retrogene, one read to the either the *TP53TRG3* or
740 *TP53RTG10* retrogenes, and one read to the *TP53TRG6* retrogene.
741
742 *Gene tree reconciliation*
743 For gene tree reconciliation putative Asian elephant, Columbian mammoth, and woolly
744 mammoth *TP53*/*TP53RTG* orthologs were manually placed into the African elephant
745 *TP53*/*TP53RTG* gene tree. We then used Notung v2.6 (Trapnell et al., 2012) to reconcile the
746 gene and species trees, and interpreted Asian elephant-, Columbian mammoth-, and wooly

747 mammoth-specific gene losses as unsampled *TP53RTG* genes if the putative divergence date
748 of the ortholog predated the African-(Asian elephant/Mammoth) divergence. We thus inferred
749 two unsampled genes (*TP53RTG7* and *TP53RTG8*) in the Asian elephant genome, and three
750 unsampled genes (*TP53RTG18*, *TP53RTG7* and *TP53RTG1*) in the Columbian mammoth
751 genome. Note that this method will estimate the *minimum* number of *TP53*/*TP53RTG* genes in
752 the genome because our mapping strategy cannot identify lineage specific duplications after
753 each species diverged from the African elephant lineage. Finally we used Notung v2.6 (Wilkie et
754 al., 2013) to reconcile the gene and species trees, and interpreted lineage-specific gene losses
755 as unsampled *TP53RTG* genes (i.e., genes present in the genome but missing from the aDNA
756 sequencing data). This process was modified slightly for American mastodon, such that 'lost'
757 *TP53RTG* genes younger than the mastodon-elephant split were considered elephant-specific
758 paralogs rather than unsampled genes present in the mastodon genome.
759
760 *Normalized read depth*
761 We also used normalized mapped read depth to estimate *TP53*/*TP53RTG* copy number.
762 Unlike gene tree reconciliation, copy number estimates based on read depth cannot resolve the
763 orthology of specific *TP53*/*TP53RTG* copies but can estimate the total number of
764 *TP53*/*TP53RTG* copies in the genome and thus provide evidence of lineage specific copy
765 number changes. For this analyses we mapped Asian elephant-, Columbian mammoth-, and
766 woolly mammoth-specific reads onto the African elephant *TP53/TP53RTG* genes using Bowtie
767 2 (Langmead and Salzberg, 2012) with the local alignment and 'very sensitive' options and
768 Cufflinks (version 0.0.7) (Trapnell et al., 2012) to assemble mapped reads into genes. We then
769 summed the read depth ('coverage') of *TP53RTG* genes and normalized this summed
770 *TP53RTG* read depth to the read depth of five single copy genomic regions equal in length to
771 average *TP53RTG* gene length. This ratio represents an estimate of the *TP53/TP53RTG* copy
772 number in the genome.
773
774 **Correlation between *TP53/TP53RTG* copy number and body size evolution**
775 Relative divergence (duplication) times of the TP53 retrogenes were estimated using the
776 TP53 alignment described above and BEAST (v1.7.4) (Rohland et al., 2010). We used the
777 general time reversible model (GTR), empirical nucleotide frequencies (+F), a proportion of
778 invariable sites estimated from the data (+I), four gamma distributed rate categories (+G), an
779 uncorrelated lognormal relaxed molecular clock to model substitution rate variation across

780 lineages, a Yule speciation tree prior, uniform priors for the GTR substitution parameters,
781 gamma shape parameter, proportion of invariant sites parameter, and nucleotide frequency
782 parameter. We used an Unweighted Pair Group Arithmetic Mean (UPGMA) starting tree.
783
784 To obtain posterior distributions of estimated divergence times, we use five node
785 calibrations modeled as normal priors (standard deviation = 1) to constrain the age of the root
786 nodes for the Eutheria (104.7 MYA), Laurasiatheria (87.2 MYA), Boreoeutherian (92.4 MYA),
787 Atlantogenatan (103 MYA), and Paenungulata (64.2 MYA); divergence dates were obtained
788 from www.timetree.org using the 'Expert Result' divergence dates. The analysis was run for 5
789 million generations and sampled every 1000 generations with a burn-in of 1 million generations;
790 convergence was assessed using Tracer, which indicated convergence was reached rapidly
791 (within 100,000 generations). Proboscidean body size data were obtained from previously
792 published studies on mammalian body size evolution (Chen et al., 2000).
793
794 **Inference of *TP53/TP53RTG* copy number expansion through segmental duplications**
795 We also observed that the genomic region surrounding each *TP53RTG* gene contained
796 blocks of homolgous transposable element insertions, suggesting that these regions are
797 segmental duplications. To confirm this observation we used MUSCLE (Edgar, 2004) to align an
798 approximately 20kb region surrounding each *TP53RTG* gene and found that conservation within
799 this region was very high, again suggesting these regions are relatively recent segmental
800 duplications. To identify if the contigs on which *TP53RTG* genes are located contained locally
801 collinear blocks (LCBs), as expected for segmental duplications, we aligned contigs using
802 progressiveMAUVE (Darling et al., 2004) as implemented in Genious (v6.1.2).
803
804 **Phylogenetic analyses of *TP53/TP53RTG* genes**
805 We generated a dataset of *TP53* orthologs from 65 diverse mammals identified from
806 GenBank, and included the *TP53* genes and retrogenes we identified from the African elephant,
807 hyrax, manatee, tenrec, cape elephant shrew, and armadillo genomes. Nucleotide sequences
808 were aligned using the MAFFT algorithm (Katoh and Standley, 2014) and the FFT refinement
809 strategy implemented in the GUIDANCE webserver (Penn et al., 2010). Alignment confidence
810 was assessed with 100 bootstrap replicates and ambiguously aligned sites (under the default
811 GUIDANCE exclusion rule) removed prior to phylogenetic analyses; lineage specific insertions
812 and deletions were also removed prior to phylogenetic analyses.

813

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

825

826 **Gene expression data (RNA-Seq)**

827 To determine if *TP53RTG* genes were expressed, we generated RNA-Seq data from
828 term placental villus and adipose tissue from African elephants (*Loxodonta africana*) and
829 primary dermal fibroblasts from Asian elephants (*Elepas maximus*). Briefly, sequencing libraries
830 were prepared using standard Illumina protocols with poly(A) selection, and sequenced as
831 100bp single-end reads on a HiSeq2000. We also used previously published RNA-Seq data
832 generated from primary fibroblasts (isolated from ear clips) from a male (GSM1227965) and
833 female (GSM1227964) African elephant generated on an Illumina Genome Analyzer IIx (101
834 cycles, single end) and a male (GSM1278046) African elephant generated on an Illumina HiSeq
835 2000 (101 cycles, single end) (Cortez et al., 2014) and combined these reads into a single
836 dataset of 138,954,285 reads. Finally we used previously published 100bp single-end reads
837 generated on a HiSeq2000 to identify *TP53RTG* transcripts from Asian elephant PBMCs
838 (SRX1423033) (Reddy et al., 2015).

839

840 Reads were aligned to a custom built elephant reference gene set generated by
841 combining the sequences of the canonical *TP53* gene and *TP53RTG* genes with the ENSEMBL
842 African elephant (*Loxodonta africana*) CDS gene build
843 (Loxodonta_africana.loxAfr3.75.cds.all.fa) with Bowtie2 (Langmead and Salzberg, 2012). Bowtie
844 2 settings were: 1) both local alignment and end-to-end mapping; 2) preset option: sensitive; 3)
845 Trim n-bases from 5' of each read: 0; and 4) Trim n-bases from 3' of each read: 0. Transcript

846 assembly and FPKM estimates were generated with Cufflinks (version 0.0.7) (Trapnell et al.,
847 2012) using aligned reads from Bowtie2, non-default parameters included quartile normalization
848 and multi-read correction. Finally we transformed FPKM estimates into transcripts per million
849 (TPM), TPM=(FPKM per gene/sum FPKM all genes)x10^6, and defined genes with TPM≥2 as
850 expressed (Li et al., 2010; Wagner et al., 2012; 2013).
851
852 The TP53RTG genes are 80.0-82.7% identical to TP53 at the nucleotide level, with 204-
853 231 total nucleotide differences compared to TP53. This level of divergence allows for many
854 reads to be uniquely mapped to each gene, there will also be significant read mapping
855 uncertainty in regions of these genes with few nucleotide differences. However, if read mapping
856 uncertainty was leading to false positive mappings of *TP53* derived reads to *TP53RTG* genes
857 we would expect to observe expression of many *TP53RTG* genes, rather the robust expression
858 of a single (*TP53RTG12*) gene. We also counted the number of uniquely mapped reads to each
859 *TP53RTG* gene and *TP53*. We found that 0-8 reads were uniquely mapped to most TP53RTG
860 genes, except *TP53RTG12* which had ~115 uniquely mapped reads across samples and TP53
861 which had ~3000 uniquely mapped reads across tissue samples. Thus we conclude that read
862 mapping uncertainty has not adversely affected our RNA-Seq analyses.
863
864 **Gene expression data (PCR and Sanger sequencing)**
865 We further confirmed expression of *TP53RTG12* transcripts in elephant cells through
866 RT-PCR, taking advantage of differences between *TP53RTG12* and *TP53* sequences to design
867 two aligned primer sets, one TP53-specific, the other *TP53RTG12* specific. *TP53RTG12*
868 primers were: 5´ *ggg gaa act cct tcc tga ga* 3´ (forward) and 5´ *cca gac aga aac gat agg tg* 3´
869 (reverse). TP53 primers were: 5´ *atg gga act cct tcc tga ga* 3´ (forward) and 5´ *cca gac gga aac*
870 *cat agg tg* 3´ (reverse). The *TP53* amplicon is expected to be 251 bps in length, while deletions
871 present in the *TP53RTG12* sequence lead to a smaller projected amplicon size of 220 bps.
872
873 Total RNA was extracted from cultured *Loxodonta* and *Elephas* fibroblasts (RNAeasy Plus
874 Mini kit, Qiagen), then DNase treated (Turbo DNA-*free* kit, Ambion) and reverse-transcribed
875 using an olgio-dT primer for cDNA synthesis (Maxima H Minus First Strand cDNA Synthesis kit,
876 Thermo Scientific). Control RT reactions were otherwise processed identically, except for the
877 omission of reverse transcriptase from the reaction mixture. RT products were PCR-amplified
878 for 45 cycles of 94°/20 seconds, 56°/30 seconds, 72°/30 seconds using a BioRad CFX96 Real

879 Time qPCR detection system and SYBR Green master mix (QuantiTect, Qiagen). PCR products
880 were electrophoresed on 3% agarose gels for 1 hour at 100 volts, stained with SYBR safe, and
881 imaged in a digital gel box (ChemiDoc MP, BioRad) to visualize relative amplicon sizes. PCR
882 products were also directly sequenced at the University of Chicago Genomics core facility,
883 confirming projected product sizes and sequence identities.
884
885 **Gene prediction, promoter annotation, and TP53RTG transcript identification**
886 We used geneid v1.2 (http://genome.crg.es/software/geneid/geneid.html) to infer if the
887 *TP53RTG12* gene contained a non-coding exon 5' to the predicted ATG start codon. For gene
888 structure prediction we used the full-length scaffold_825 from African elephant (Broad/loxAfr3)
889 sequence and forced an internal exon where the *TP53RTG12* gene is encoded in scaffold_825.
890 Geneid identified a putative exon 5' to the *TP53RTG12* gene from nucleotides 1761-1935 and
891 an exon 3' from nucleotides 6401-6776. We also used GENESCAN (http://genes.mit.edu/cgi-
892 bin/genscanw_py.cgi) to predict the location of exons in the full-length scaffold_825 sequence
893 and identified a putative exon from nucleotides 1750-1986. We next used Bowtie2 (Langmead
894 and Salzberg, 2012) to map African and Asian elephant fibroblast RNA-Seq data onto African
895 elephant scaffold_825 with the default settings.
896
897 **Taxonomic distribution of the RTE1_LA retrotransposon**
898 The RTE1_LA non-LTR retrotransposon has previously been described from the African
899 elephant genome, these elements are generally more than 90% identical to the consensus
900 RTE1_LA sequence but less than 70% identical to other mammalian RTEs
901 (http://www.girinst.org/2006/vol6/issue3/RTE1_LA.html). These data suggest that the RTE_LA
902 element has relatively recently expanded in the elephant genome. To determine the taxonomic
903 distribution of the RTE1_LA element we used BLAT to search the lesser hedgehog tenrec
904 (*Echinops telfairi*), rock hyrax (*Procavia capensis*), West Indian manatee (*Trichechus manatus*),
905 armadillo (*Dasypus novemcinctus*), and sloth (*Choloepus hoffmanni*) genomes. We identified
906 numerous copies of the RTE_LA element in the genome of the Afrotherians, but not the
907 Xenarthran genomes. These data indicate that the RTE_LA element is Afrotherian-specific,
908 rather than Elephant-specific.
909
910
911

912 **TP53/TP53RTG Western blotting**

913 Elephant and Hyrax primary fibroblasts (San Diego Zoo, "Frozen Zoo") were grown to
914 confluency in 10cm dishes at 37°C/5% CO_2 in a culture medium consisting of FGM/EMEM (1:1)
915 supplemented with insulin, FGF, 6% FBS and Gentamicin/Amphotericin B (FGM-2, singlequots,
916 Clonetics/Lonza). Culture medium was removed from dishes just prior to UV treatment and
917 returned to cells shortly afterwards. Experimental cells were exposed to 50J/m^2 UV-C radiation
918 in a crosslinker (Stratalinker 2400, Stratagene), while control cells passed through media
919 changes but were not exposed to UV. A small volume (~3mL) of PBS covered fibroblasts at the
920 time of UV exposure. To inhibit TP53 proteolysis, MG-132 (10μM) was added to experimental
921 cell medium 1 hour prior to UV exposure, and maintained until the time of cell lysis. 5h post-UV
922 treatment, cells were briefly rinsed in PBS, then lysed and boiled in 2x SDS-PAGE sample
923 buffer. Lysates were separated via SDS-PAGE on 10% gels for 1h at 140 volts, then
924 electrophoretically transferred to PVDF membranes (1h at 85 volts). Membranes were blocked
925 for 1h in 5% milk in TBST and incubated overnight at 4°C with rabbit polyclonal TP53 antibodies
926 (FL-393, Santa Cruz Biotechnology, and ab131442, Abcam). Blots were washed 3x in TBST,
927 incubated with HRP-conjugated, anti-rabbit IgG 2° antibodies for 1 h at RT, and washed 4 more
928 times in TBST. Protein bands were visualized via enhanced chemiluminescence (BioRad
929 Clarity), and imaged in a digital gel box (Chemidoc MP, BioRad). Western blots were replicated
930 three independent times.

931

932 **Co-immunoprecipitation**

933 Human HEK-293 cells were grown to 80% confluency in 20cm dishes at 37°C/5% CO_2 in
934 a culture medium consisting of DMEM supplemented with 10% FBS. At 80% confluency, cells
935 were transiently transfected with the *TP53RTG12* pcDNA3.1(+)/myc-His expression vector.
936 After 16 hours the transfection media was removed and replaced with fresh DMEM, and the
937 cells were incubated an additional 24 hours before harvesting. After removing DMEM and
938 washing cells twice with PBS, 1 mL ice-cold lysis buffer (20 mM Tris, pH 8.0, 40 mM KCl, 10
939 mM MgCl2, 10% glycerol, 1% Triton X-100, 1x Complete EDTA-free protease inhibitor cocktail
940 (Roche), 1x PhosSTOP (Roche)) was added to each plate and cells were harvested by scraping
941 with a rubber spatula. Cells were then incubated on ice for 30 minutes in 420 mM NaCl. Whole
942 cell lysate was cleared by centrifugation at 10,000 rpm for 30 minutes at 4°C, and supernatant
943 was transferred to a clean microfuge tube. After equilibrating protein concentrations, 1 mL of

944 sample was mixed with 40 mL of α-MDM2 or α-Myc antibody conjugated agarose beads
945 (Sigma) pre-washed with TNT buffer (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 0.05% Triton X-
946 100), and rotated overnight at 4°C. The following day, samples were treated with 50 U DNase
947 (Roche) and 2.5 µg RNase (Roche) for 60 minutes at room temperature, as indicated. Samples
948 were washed 3x with 1 mL wash buffer (150 mM NaCl, 0.5% Triton X-100). After the final wash,
949 agarose beads were resuspended in elution buffer (500 mM Tris pH 7.5, 1 M NaCl), and boiled
950 to elute immunoprecipitated complexes. Eluted protein was run on Bis-tris gels, probed with
951 antibodies and visualized by Chemi-luminescence. Serial Westerns were performed for each
952 antibody following chemical stripping and re-blocking. Antibodies were from Santa Cruz: MDM2
953 (SMP14) sc-965, lot #J2314; p53 (Fl-393) sc-6243, lot # D0215; c-Myc (9E10) sc-40, lot #
954 G2413.
955
956 **ApoTox-Glo Viability/Cytotoxicity/Apoptosis Analyses**
957 Primary fibroblasts were grown to 80% confluency in T-75 culture flasks at 37°C/5% CO_2
958 in a culture medium consisting of FGM/EMEM (1:1) supplemented with insulin, FGF, 6% FBS
959 and Gentamicin/Amphotericin B (FGM-2, singlequots, Clonetics/Lonza). 10^4 cells were seeded
960 into each well of an opaque bottomed 96-well plate, leaving a column with no cells (background
961 control); each 96-well plate contained paired elephant and hyrax samples. Serial dilutions of
962 Doxorubicin (0uM, 0.5uM, 1.0uM, 5uM, 10uM and 50uM), Mitomycin c (0uM, 0.5uM, 1.0uM,
963 5uM, 10uM and 50uM), and Nutlin-3a (0uM, 0.5uM, 1.0uM, 5uM, 10uM and 50uM) and 90%
964 culture media were added to each well such that there were four biological replicates for each
965 condition. After 18 hrs of incubation with each drug, cell viability, cytotoxicity, and caspase-3/7
966 activity were measured using the ApoTox-Glo Triplex Assay (Promega) in a GloMax-Multi+
967 Reader (Promega). Data were standardized to no-drug control cells. For UV-C treatment, culture
968 medium was removed from wells prior to UV treatment and returned to cells shortly afterwards.
969 Experimental cells were exposed to 50J/m^2 UV-C radiation in a crosslinker (Stratalinker 2400,
970 Stratagene), while control cells passed through media changes but were not exposed to UV. A
971 small volume (~30uL) of PBS covered fibroblasts at the time of UV exposure. Cell viability,
972 cytotoxicity, and caspase-3/7 activity were measured using the ApoTox-Glo Triplex Assay
973 (Promega) in a GloMax-Multi+ Reader (Promega) 6, 12, 28.5 and 54.5 hours after UV-C
974 treatment. Data were standardized to no UV-C exposure control cells. ApoTox-Glo Triplex
975 Assays were replicated three independent times.
976

977 **Luciferase assays**

978 Primary fibroblasts were grown to 80% confluency in T-75 culture flasks at 37°C/5% CO_2
979 in a culture medium consisting of FGM/EMEM (1:1) supplemented with insulin, FGF, 6% FBS
980 and Gentamicin/Amphotericin B (FGM-2, singlequots, Clonetics/Lonza). At confluency, cells
981 were trypsinized, centrifuged at 90g for 10min and resuspended in nucleofection/supplement
982 solution and incubated for 15min. After incubation 1ug of the pGL4.38[*luc2p*/p53 RE/Hygro]
983 luciferase reporter vector and 100ng of the pGL4.74[*hRluc/*TK] Renilla reporter vector were
984 transiently transfected into elephant and hyrax cells using the Amaxa Basic Nucleofector Kit
985 (Lonza) using protocol T-016. Immediately following nuleofection, 10^4 cells were seeded into
986 each well of an opaque bottomed 96-well plate, leaving a column with no cells (background
987 control); each 96-well plate contained paired elephant and hyrax samples. 24 hours after
988 nucleofection cells were treated with either vehicle control, Doxorubicin, Mitomycin c, Nutlin-3a,
989 or 50J/m^2 UV-C. Luciferase expression was assayed 18 hours after drug/UV-C treatment cells,
990 using the Dual-Luciferase Reporter Assay System (Promega) in a GloMax-Multi+ Reader
991 (Promega). For all experiments luciferase expression was standardized to Renilla expression to
992 control for differences nucleofection efficiency across samples; Luc./Renilla data is standardized
993 to (Luc./Renilla) expression in untreated control cells. Each luciferase experiment was replicated
994 three independent times.
995

996 **siRNA experiments**

997 siRNAs designed to specifically-target TP53RTG were validated via qRT-PCR using the two
998 primer sets described earlier, which amplify either TP53RTG or canonical TP53 cDNAs, to
999 confirm specificity and efficacy of knockdown. Sequences of the three TP53RTG-specific
1000 siRNAs used are as follows: 1) 5'-CAGCGGAGGCAGUAGAUGAUU-3', 2) 5'-
1001 GGCUCAAGGAAUAUCAGAAUU-3', 3) 5'-CAGCAGCGGAGGCAGUAGAUU-3' (Dharmacon).
1002 *Loxodonta* fibroblasts were transfected with siRNAs using Lipofectamine LTX, and tested 48-72
1003 hours later for either TP53 response via luciferase assay or for cell viability/toxicity/apoptosis via
1004 ApoTox-Glo assay.
1005
1006
1007
1008
1009

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).

1043 **REFERENCES**

1044

1045 Abegglen, L.M., Caulin, A.F., Chan, A., Lee, K., Robinson, R., Campbell, M.S., Kiso, W.K.,
1046 Schmitt, D.L., Waddell, P.J., Bhaskara, S., et al. (2015). Potential Mechanisms for Cancer
1047 Resistance in Elephants and Comparative Cellular Response to DNA Damage in Humans.
1048 Jama *314*, 1850–1860.

1049 Aguirre, A.A., Bröjer, C., and Mörner, T. (1999). Descriptive epidemiology of roe deer mortality
1050 in Sweden. J. Wildl. Dis. *35*, 753–762.

1051 Allemand, I., Anglo, A., Jeantet, A.Y., Cerutti, I., and May, E. (1999). Testicular wild-type p53
1052 expression in transgenic mice induces spermiogenesis alterations ranging from differentiation
1053 defects to apoptosis. Oncogene *18*, 6521–6530.

1054 Ashur-Fabian, O., Avivi, A., Trakhtenbrot, L., Adamsky, K., Cohen, M., Kajakaro, G., Joel, A.,
1055 Amariglio, N., Nevo, E., and Rechavi, G. (2004). Evolution of p53 in hypoxia-stressed Spalax
1056 mimics human tumor mutation. Proc. Natl. Acad. Sci. U.S.a. *101*, 12236–12241.

1057 Avivi, A., Ashur-Fabian, O., Joel, A., Trakhtenbrot, L., Adamsky, K., Goldstein, I., Amariglio, N.,
1058 Rechavi, G., and Nevo, E. (2007). P53 in blind subterranean mole rats--loss-of-function versus
1059 gain-of-function activities on newly cloned Spalax target genes. Oncogene *26*, 2507–2512.

1060 Avivi, A., and Ashur-Fabian, O. (2005). p53: A Key Player in Tumoral and Evolutionary
1061 Adaptation: A Lesson from the Israeli Blind Subterranean Mole Rat. Cell Cycle *4*, 368–372.

1062 Belyi, V.A., Ak, P., Markert, E., and Wang, H. (2010). The origins and evolution of the p53 family
1063 of genes. Cold Spring ….

1064 Cairns, J. (1975). Mutation selection and the natural history of cancer. Nature *255*, 197–200.

1065 Campisi, J. (2003). Cancer and ageing: rival demons? Nat. Rev. Cancer.

1066 Caulin, A.F., and Maley, C.C. (2011). Peto“s Paradox: evolution”s prescription for cancer
1067 prevention. Trends Ecol. Evol. (Amst.) *26*, 175–182.

1068 Caulin, A.F., Graham, T.A., Wang, L.-S., and Maley, C.C. (2015). Solutions to Peto's paradox
1069 revealed by mathematical modelling and cross-species cancer gene analysis. Philos. Trans. R.
1070 Soc. Lond., B, Biol. Sci. *370*.

1071 Chen, K., Durand, D., and Farach-Colton, M. (2000). NOTUNG: a program for dating gene
1072 duplications and optimizing gene family trees. J. Comput. Biol. *7*, 429–447.

1073 Chua, J.S., Liew, H.P., Guo, L., and Lane, D.P. (2015). Tumor-specific signaling to p53 is
1074 mimicked by Mdm2 inactivation in zebrafish: insights from mdm2 and mdm4 mutant zebrafish.
1075 Oncogene *34*, 5933–5941.

1076 Ciotta, C., Dogliotti, E., and Bignami, M. (1995). Mutation analysis in two newly identified rat p53
1077 pseudogenes. Mutagenesis *10*, 123–128.

1078 Cohen, M., Wuillemin, C., and Bischof, P. (2008). Trophoblastic p53 is stabilised by a cis-trans
1079 isomerisation necessary for the formation of high molecular weight complexes involving the N-
1080 terminus of p53. Biochimie *90*, 855–862.

1081 Cortez, D., Marin, R., Toledo-Flores, D., Froidevaux, L., Liechti, A., Waters, P.D., Grutzner, F.,
1082 and Kaessmann, H. (2014). Origins and functional evolution of Y chromosomes across
1083 mammals. Nature *508*, 488–493.

1084 Czosnek, H.H., Bienz, B., Givol, D., Zakut-Houri, R., Pravtcheva, D.D., Ruddle, F.H., and Oren,
1085 M. (1984). The gene and the pseudogene for mouse p53 cellular tumor antigen are located on
1086 different chromosomes. Mol. Cell. Biol. *4*, 1638–1640.

1087 Darling, A.C.E., Mau, B., Blattner, F.R., and Perna, N.T. (2004). Mauve: multiple alignment of
1088 conserved genomic sequence with rearrangements. Genome Res. *14*, 1394–1403.

1089 Doll, R. (1971). The age distribution of cancer: implications for models of carcinogenesis.
1090 Journal of the Royal Statistical Society Series a (….

1091 Donehower, L.A. (2002). Does p53 affect organismal aging? J. Cell. Physiol. *192*, 23–33.

1092 Dumble, M., Moore, L., Chambers, S.M., Geiger, H., Van Zant, G., Goodell, M.A., and
1093 Donehower, L.A. (2007). The impact of altered p53 dosage on hematopoietic stem cell
1094 dynamics during aging. Blood *109*, 1736–1742.

1095 Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high
1096 throughput. Nucleic Acids Res.

1097 Enk, J.M., Devault, A.M., Kuch, M., and Murgha, Y.E. (2014). Ancient whole genome
1098 enrichment using baits built from modern DNA. Molecular Biology and ….

1099 Enk, J., Rouillard, J.-M., and Poinar, H. (2013). Quantitative PCR as a predictor of aligned
1100 ancient DNA read counts following targeted enrichment. BioTechniques *55*, 300–309.

1101 Evans, A.R., Jones, D., Boyer, A.G., Brown, J.H., Costa, D.P., Ernest, S.K.M., Fitzgerald,
1102 E.M.G., Fortelius, M., Gittleman, J.L., Hamilton, M.J., et al. (2012). The maximum rate of
1103 mammal evolution. *109*, 4187–4190.

1104 Ferbeyre, G., and Lowe, S.W. (2002). The price of tumour suppression? Nature *415*, 26–27.

1105 Finch, R.A., Donoviel, D.B., Potter, D., Shi, M., Fan, A., Freed, D.D., Wang, C.-Y., Zambrowicz,
1106 B.P., Ramirez-Solis, R., Sands, A.T., et al. (2002). mdmx is a negative regulator of p53 activity
1107 in vivo. Cancer Res. *62*, 3221–3225.

1108 Garcia-Cao, I., García-Cao, M., Martín-Caballero, J., Criado, L.M., Klatt, P., Flores, J.M., Weill,
1109 J.-C., Blasco, M.A., and Serrano, M. (2002). “Super p53” mice exhibit enhanced DNA damage
1110 response, are tumor resistant and age normally. Embo J. *21*, 6225–6235.

1111 Garcia-Cao, I., García-Cao, M., Tomás-Loba, A., Martín-Caballero, J., Flores, J.M., Klatt, P.,
1112 Blasco, M.A., and Serrano, M. (2006). Increased p53 activity does not accelerate telomere-
1113 driven ageing. EMBO Rep.

1114 Godley, L.A., Kopp, J.B., and Eckhaus, M. (1996). Wild-type p53 transgenic mice exhibit altered
1115 differentiation of the ureteric bud and possess small kidneys. Genes & ….

1116 González-Romero, R., Rivera-Casas, C., Ausió, J., Méndez, J., and Eirín-López, J.M. (2010).
1117 Birth-and-death long-term evolution promotes histone H2B variant diversification in the male
1118 germinal cell line. Mol. Biol. Evol. *27*, 1802–1812.

1119 Gorbunova, V., Hine, C., Tian, X., Ablaeva, J., Gudkov, A.V., Nevo, E., and Seluanov, A. (2012).
1120 Cancer resistance in the blind mole rat is mediated by concerted necrotic cell death mechanism.
1121 Proc. Natl. Acad. Sci. U.S.a. *109*, 19392–19396.

1122 Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010).
1123 New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the
1124 performance of PhyML 3.0. Syst. Biol. *59*, 307–321.

1125 Healy, K., Guillerme, T., Finlay, S., Kane, A., Kelly, S.B.A., McClean, D., Kelly, D.J., Donohue,
1126 I., Jackson, A.L., and Cooper, N. (2014). Ecology and mode-of-life explain lifespan variation in
1127 birds and mammals. Proc. Biol. Sci. *281*, 20140298.

1128 Hoever, M., Clement, J.H., Wedlich, D., Montenarh, M., and Knöchel, W. (1994).
1129 Overexpression of wild-type p53 interferes with normal development in Xenopus laevis
1130 embryos. Oncogene *9*, 109–120.

1131 Hulla, J.E. (1992). The rat genome contains a p53 pseudogene: detection of a processed
1132 pseudogene using PCR. PCR Methods Appl. *1*, 251–254.

1133 Hulsenbeck, J.P., and Ronquist, F. (2001). MRBAYES: Bayesian inference of phylogenetic
1134 trees. Bioinformatics *17*, 754–755.

1135 Katoh, K., and Standley, D.M. (2014). MAFFT: iterative refinement and additional methods.
1136 Methods Mol. Biol. *1079*, 131–146.

1137 Katzourakis, A., Magiorkinis, G., Lim, A.G., Gupta, S., Belshaw, R., and Gifford, R. (2014).
1138 Larger Mammalian body size leads to lower retroviral activity. PLoS Pathog. *10*, e1004214.

1139 Khoury, M.P., and Bourdon, J.C. (2010). The isoforms of the p53 protein. Cold Spring Harbor
1140 ….

1141 Kim, H., Kim, K., Choi, J., Heo, K., Baek, H.J., Roeder, R.G., and An, W. (2012). p53 requires
1142 an intact C-terminal domain for DNA binding and transactivation. J. Mol. Biol. *415*, 843–854.

1143 Kitayner, M., Rozenberg, H., Kessler, N., Rabinovich, D., Shaulov, L., Haran, T.E., and
1144 Shakked, Z. (2006). Structural basis of DNA recognition by p53 tetramers. Mol. Cell *22*, 741–
1145 753.

1146 Kubbutat, M., Ludwig, R.L., and Ashcroft, M. (1998). Regulation of Mdm2-directed degradation
1147 by the C terminus of p53. Molecular and Cellular ….

1148 Kussie, P.H., Gorina, S., Marechal, V., Elenbaas, B., Moreau, J., Levine, A.J., and Pavletich,
1149 N.P. (1996). Structure of the MDM2 oncoprotein bound to the p53 tumor suppressor

1150 transactivation domain. Science *274*, 948–953.

1151 Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat.
1152 Methods.

1153 Leroi, A.M., Koufopanou, V., and Burt, A. (2003). Cancer selection. Nat. Rev. Cancer *3*, 226–
1154 231.

1155 Li, B., Ruotti, V., Stewart, R.M., Thomson, J.A., and Dewey, C.N. (2010). RNA-Seq gene
1156 expression estimation with read mapping uncertainty. Bioinformatics *26*, 493–500.

1157 Lozano, G. (2010). Mouse models of p53 functions. Cold Spring Harbor Perspectives in Biology.

1158 Lynch, V.J. (2007). Inventing an arsenal: adaptive evolution and neofunctionalization of snake
1159 venom phospholipase A2 genes. BMC Evol. Biol. *7*, 2.

1160 Maier, B., Gluba, W., Bernier, B., Turner, T., Mohammad, K., Guise, T., Sutherland, A., Thorner,
1161 M., and Scrable, H. (2004). Modulation of mammalian life span by the short isoform of p53.
1162 Genes Dev. *18*, 306–319.

1163 Maki, C.G. (1999). Oligomerization is required for p53 to be efficiently ubiquitinated by MDM2. J.
1164 Biol. Chem.

1165 Manov, I., Hirsh, M., Iancu, T.C., Malik, A., Sotnichenko, N., Band, M., Avivi, A., and Shams, I.
1166 (2013). Pronounced cancer resistance in a subterranean rodent, the blind mole-rat, Spalax: in
1167 vivo and in vitro evidence. BMC Biol. *11*, 91.

1168 Martineau, D., Lemberger, K., Dallaire, A., Labelle, P., Lipscomb, T.P., Michel, P., and
1169 Mikaelian, I. (2002). Cancer in wildlife, a case study: beluga from the St. Lawrence estuary,
1170 Québec, Canada. Environ. Health Perspect. *110*, 285–292.

1171 Matheu, A., Maraver, A., Klatt, P., Flores, I., Garcia-Cao, I., Borras, C., Flores, J.M., Viña, J.,
1172 Blasco, M.A., and Serrano, M. (2007). Delayed ageing through damage protection by the
1173 Arf/p53 pathway. Nature *448*, 375–379.

1174 Mendrysa, S.M. (2006). Tumor suppression and normal aging in mice with constitutively high
1175 p53 activity. Genes Dev. *20*, 16–21.

1176 Mendrysa, S.M., McElwee, M.K., Michalowski, J., O'Leary, K.A., Young, K.M., and Perry, M.E.
1177 (2003). mdm2 Is critical for inhibition of p53 during lymphopoiesis and the response to ionizing
1178 irradiation. Mol. Cell. Biol. *23*, 462–472.

1179 Migliorini, D., Lazzerini Denchi, E., Danovi, D., Jochemsen, A., Capillo, M., Gobbi, A., Helin, K.,
1180 Pelicci, P.G., and Marine, J.-C. (2002). Mdm4 (Mdmx) regulates p53-induced growth arrest and
1181 neuronal cell death during early embryonic mouse development. Mol. Cell. Biol. *22*, 5527–5538.

1182 Nagy, J.D., Victor, E.M., and Cropper, J.H. (2007). Why don“t all whales have cancer? A novel
1183 hypothesis resolving Peto”s paradox. Integr. Comp. Biol. *47*, 317–328.

1184 Nunney, L. (1999). Lineage selection and the evolution of multistage carcinogenesis. Proc. Biol.

1185 Sci. *266*, 493–498.

1186 Ottaggio, L., Bozzo, S., Moro, F., Sparks, A., Campomenosi, P., Miele, M., Bonatti, S., Fronza,
1187 G., Lane, D.P., and Abbondandolo, A. (2000). Defective nuclear localization of p53 protein in a
1188 Chinese hamster cell line is associated with the formation of stable cytoplasmic protein
1189 multimers in cells with gene amplification. Carcinogenesis *21*, 1631–1638.

1190 Parant, J., Chavez-Reyes, A., Little, N.A., Yan, W., Reinke, V., Jochemsen, A.G., and Lozano,
1191 G. (2001). Rescue of embryonic lethality in Mdm4-null mice by loss of Trp53 suggests a
1192 nonoverlapping pathway with MDM2 to regulate p53. Nat. Genet. *29*, 92–95.

1193 Penn, O., Privman, E., Ashkenazy, H., and Landan, G. (2010). GUIDANCE: a web server for
1194 assessing alignment confidence scores. Nucleic Acids ….

1195 Perez, R.P., and Komiya, T. (2016). TP53 Gene and Cancer Resistance in Elephants. Jama
1196 *315*, 1789–1790.

1197 Peto, R. (2015). Quantitative implications of the approximate irrelevance of mammalian body
1198 size and lifespan to lifelong cancer risk. Philos. Trans. R. Soc. Lond., B, Biol. Sci. *370*.

1199 Pires, D., Ascher, D.B., and Blundell, T.L. (2014). DUET: a server for predicting effects of
1200 mutations on protein stability using an integrated computational approach. Nucleic Acids Res.

1201 Poliseno, L., Salmena, L., Zhang, J., Carver, B., Haveman, W.J., and Pandolfi, P.P. (2010). A
1202 coding-independent function of gene and pseudogene mRNAs regulates tumour biology. Nature
1203 *465*, 1033–1038.

1204 R Peto, F.J.R.P.N.L.L.L.J.C. (1975). Cancer and ageing in mice and men. British Journal of
1205 Cancer *32*, 411.

1206 Reddy, P.C., Sinha, I., Kelkar, A., Habib, F., Pradhan, S.J., Sukumar, R., and Galande, S.
1207 (2015). Comparative sequence analyses of genome and transcriptome reveal novel transcripts
1208 and variants in the Asian elephant Elephas maximus. J. Biosci. *40*, 891–907.

1209 Rodier, F., Campisi, J., and Bhaumik, D. (2007). Two faces of p53: aging and tumor
1210 suppression. Nucleic Acids Res.

1211 Rohland, N., Malaspinas, A.-S., Pollack, J.L., Slatkin, M., Matheus, P., and Hofreiter, M. (2007).
1212 Proboscidean mitogenomics: chronology and mode of elephant evolution using mastodon as
1213 outgroup. PLoS Biol. *5*, e207.

1214 Rohland, N., Reich, D., Mallick, S., Meyer, M., Green, R.E., Georgiadis, N.J., Roca, A.L., and
1215 Hofreiter, M. (2010). Genomic DNA sequences from mastodon and woolly mammoth reveal
1216 deep speciation of forest and savanna elephants. PLoS Biol. *8*, e1000564.

1217 Rooney, A.P., Piontkivska, H., and Nei, M. (2002). Molecular evolution of the nontandemly
1218 repeated genes of the histone 3 multigene family. Mol. Biol. Evol. *19*, 68–75.

1219 Roy, A., Kucukural, A., and Zhang, Y. (2010). I-TASSER: a unified platform for automated
1220 protein structure and function prediction. Nat Protoc *5*, 725–738.

1221 Seluanov, A., Hine, C., Azpurua, J., Feigenson, M., Bozzella, M., Mao, Z., Catania, K.C., and
1222 Gorbunova, V. (2009). Hypersensitivity to contact inhibition provides a clue to cancer resistance
1223 of naked mole-rat. Proc. Natl. Acad. Sci. U.S.a. *106*, 19352–19357.

1224 Senturk, S., Yao, Z., Camiolo, M., Stiles, B., Rathod, T., Walsh, A.M., Nemajerova, A., Lazzara,
1225 M.J., Altorki, N.K., Krainer, A., et al. (2014). p53Ψ is a transcriptionally inactive p53 isoform able
1226 to reprogram cells toward a metastatic-like state. Proc. Natl. Acad. Sci. U.S.a. *111*, E3287–
1227 E3296.

1228 Smith, F.A., Boyer, A.G., Brown, J.H., Costa, D.P., Dayan, T., Ernest, S.K.M., Evans, A.R.,
1229 Fortelius, M., Gittleman, J.L., Hamilton, M.J., et al. (2010). The evolution of maximum body size
1230 of terrestrial mammals. Science *330*, 1216–1219.

1231 Sookias, R.B., Benson, R., and Butler, R.J. (2012). Biology, not environment, drives major
1232 patterns in maximum tetrapod body size through time. Biology Letters.

1233 Sparks, A., Dayal, S., Das, J., Robertson, P., Menendez, S., and Saville, M.K. (2013). The
1234 degradation of p53 and its major E3 ligase Mdm2 is differentially dependent on the proteasomal
1235 ubiquitin receptor S5a. Oncogene.

1236 Tanooka, H., Ootsuyama, A., Shiroishi, T., and Moriwaki, K. (1995). Distribution of the p53
1237 pseudogene among mouse species and subspecies. Mammalian Genome.

1238 Tian, X., Azpurua, J., Hine, C., Vaidya, A., Myakishev-Rempel, M., Ablaeva, J., Mao, Z., Nevo,
1239 E., Gorbunova, V., and Seluanov, A. (2013). High-molecular-mass hyaluronan mediates the
1240 cancer resistance of the naked mole rat. Nature *499*, 346–349.

1241 Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., Pimentel, H., Salzberg, S.L.,
1242 Rinn, J.L., and Pachter, L. (2012). Differential gene and transcript expression analysis of RNA-
1243 seq experiments with TopHat and Cufflinks. Nat Protoc *7*, 562–578.

1244 Tyner, S.D., Venkatachalam, S., Choi, J., Jones, S., Ghebranious, N., Igelmann, H., Lu, X.,
1245 Soron, G., Cooper, B., Brayton, C., et al. (2002). p53 mutant mice that display early ageing-
1246 associated phenotypes. Nature *415*, 45–53.

1247 Wagner, G.P., Kin, K., and Lynch, V.J. (2012). Measurement of mRNA abundance using RNA-
1248 seq data: RPKM measure is inconsistent among samples. Theory Biosci. *131*, 281–285.

1249 Wagner, G.P., Kin, K., and Lynch, V.J. (2013). A model based criterion for gene expression
1250 calls using RNA-seq data. Theory Biosci. *132*, 159–164.

1251 Weghorst, C.M., Buzard, G.S., Calvert, R.J., and Hulla, J.E. (1995). Cloning and sequence of a
1252 processed p53 pseudogene from rat: a potential source of false “mutations” in PCR fragments of
1253 tumor DNA. Gene.

1254 Wilkie, G.S., Davison, A.J., Watson, M., Kerr, K., Sanderson, S., Bouts, T., Steinbach, F., and
1255 Dastjerdi, A. (2013). Complete genome sequences of elephant endotheliotropic herpesviruses
1256 1A and 1B determined directly from fatal cases. J. Virol. *87*, 6700–6712.

1257 Xu, D., and Zhang, Y. (2011). Improving the physical realism and structural accuracy of protein

models by a two-step atomic-level energy minimization. Biophys. J. *101*, 2525–2534.

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, Y. (2008). I-TASSER server for protein 3D structure prediction. BMC Bioinformatics *9*, 40.

1290 **FIGURE LEGENDS**

1291 **Figure 1. Body size evolution in vertebrates.**

1292 (A) Relationship between body mass (g) and lifespan (years) among 2,556 vertebrates. Blue
1293 line shows the linear regression between log (body mass) and log (lifespan), R^2=0.32.

1294 (B) Body size comparison between living (African and Asian elephants) and extinct (Steppe
1295 mammoth) Proboscideans, Cetaceans (Minke whale), and the extinct hornless
1296 rhinoceros *Paraceratherium* ('Walter'), and humans.

1297

1298 **Figure 2. Expansion of the *TP53RTG* gene repertoire in Proboscideans.**

1299 (A) *TP53* copy number in 61 Sarcopterygian (Lobe-finned fish) genomes. Clade names are
1300 shown for lineages in which the genome encodes more than one *TP53* gene or
1301 pseudogene.

1302 (B) Estimated *TP53*/*TP53RTG* copy number inferred from complete genome sequencing
1303 data (WGS, purple), 1:1 orthology (green), gene tree reconciliation (blue), and
1304 normalized read depth from genome sequencing data (red). Whiskers on normalized
1305 read depth copy number estimates show the 95% confidence interval of the estimate.

1306

1307 **Figure 2-figure supplement 1. Reconciled *TP53*/*TP53RTG* gene trees.** Reconciled
1308 *TP53RTG* gene trees for Columbian mammoth, woolly mammoth, Asian elephant, and
1309 American mastodon.

1310

1311 **Figure 3. *TP53RTG* copy number increased through segmental duplications.**

1312 (A) Organization of the *TP53* and *TP53RTG* loci in African elephant. The *TP53*/*TP53RTG*
1313 gene tree is shown at the left. The location of homologous transposable elements that
1314 flank the *TP53RTG* genes are shown and I Abegglen ed. *TP53RTG* genes with intact
1315 start codons are labeled with arrows, stop codons are shown in red.

1316 (B) Multiple sequence alignment (MUSCLE) of elephant *TP53RTG* containing contigs. The
1317 location of the *TP53RTG* genes are shown with a blue bar. Sites are color coded
1318 according to their conservation (see inset key).

1319 (C) ProgressiveMAUVE alignment of elephant *TP53RTG* containing contigs. Colored boxes
1320 shown the location of collinear blocks, lines connect homologous collinear blocks on
1321 different contigs.

1322

1323 **Figure 4. *TP53RTG* copy number is correlated with body size evolution in**
1324 **Proboscideans.**

1325 (A) Time calibrated Bayesian phylogeny of *TP53*/*TP53RTG* genes. *TP53RTG* genes are
1326 shown in blue, the 95% highest posterior density (HPD) of estimated divergence dates
1327 are shown as red bars, nodes with a posterior probability (PP)>0.95 are labeled with
1328 closed circles whereas nodes with a PP≤0.95 are labeled with open circles. The period
1329 corresponding to the expansion of the *TP53RTG* gene repertoire is shown in a grey.

1330 (B) *TP53*/*TP53RTG* copy number (blue) and Proboscidean body size (red) increases
1331 through time are correlated.

1332

1333 **Figure 4-figure supplement 1. TP53/TP53RTG gene trees.** Maximum likelihood gene trees
1334 with aBayes, Chi-Square, aLRT, and SH-like branch supports. Bayesian tree inferred by
1335 MrBayes. Numbers along branches indicate node supports from each method. Note that the full
1336 dataset included 65 diverse species and only the Paeunngulate clade is shown. TP53 gene
1337 trees were inferred using a general time reversible model (GTR), empirical nucleotide
1338 frequencies (+F), a proportion of invariable sites estimated from the data (+I), four gamma
1339 distributed rate categories (+G), and using the best of NNI and SPR branch moves during the
1340 topology search (for the maximum likelihood tree).

1341

1342 **Figure 5. *TP53RTG12* is transcribed and translated.**

1343 (A) Transcription of elephant *TP53* and *TP53RTG* genes in dermal fibroblasts, white
1344 adipose, and placental villi. RNA-Seq data are shown as mean transcripts per million
1345 (TPM) with 95% confidence intervals of TPM value. Blue bars show TPM estimates from
1346 'end-to-end' read mapping and gray bars shown 'local' read mapping.

1347 (B) qRT-PCR products generated Asian (left, blue sqaure) and African (right, light blue
1348 square) elephant fibroblast cDNA using primers specific to TP53 and TP53RTG12.
1349 cDNA was generated from DNaseI-treated RNA. No reverse transcriptase (no RT)
1350 controls for each qPCR reaction are shown, end point PCR products are shown.

1351 (C) Coverage of mapped reads from Asian (dark blue) and African (light blue) elephant
1352 fibroblast RNA-Seq data across the region of scaffold_885 encoding the *TP53RTG12*
1353 gene. The location of TP53RTG12 exons predicted from geneid and GENESCAN are

1354 shown in blue introns are shown as lines with arrows indicating the direction of
1355 transcription. Gray bars show the location of transposable elements around the
1356 *TP53RTG12* gene, darker gray indicates high sequence similarity to the consensus of
1357 each element. PCR tiles across this region are shown for African (Lox.) and Asian (Ele.)
1358 elephants, PCR primers generating amplicons are shown in red. The inferred
1359 *TP53RTG12* transcript is shown below. 1 kb scale shown from position 1 of African
1360 elephant (Broad/loxAfr3) scaffold_825.

1361 (D) Relative luciferase (Luc.) expression in Asian and African fibroblasts transfected with
1362 either the promoterless pGL4.10[*luc2*] luciferase reporter vector (empty vector), pGL4.10
1363 containing the RTE_LA consensus sequences (Consensus), pGL4.10 containing the
1364 RTE_LA from Asian elephant (Ele. RTE_LA), or pGL4.10 containing the RTE_LA from
1365 African elephant (Lox. RTE_LA). Results are shown as fold difference in Luc. expression
1366 standardized to empty vector and *Renilla* controls. n=16, Wilcoxon P-values.

1367 (E) Western blot of total cell protein isolated from South African Rock hyrax, Asian elephant
1368 (*Elephas*), and African elephant (*Loxodona*) dermal fibroblasts. -, control cells. +, cells
1369 treated with 50 J/m^2 UV-C and the proteasome inhibitor MG-132. The name and
1370 predicted molecular weights of TP53 isoforms are shown.

1371

1372 **Figure 5-figure supplement 1. PCR and Sanger sequencing confirm *TP53RTG12* is**
1373 **transcribed in elephant fibroblasts.**

1374 (A) Multiple sequence alignment (MSA) of African elephant *TP53* and *TP53RTG1-*
1375 *TP53RTG19*. Numbering (shown at top) is relative to the canonical TP53 transcript in
1376 African elephant (ENSLAFT00000007484), sequence conservation in this region is
1377 shown in Bits, as a logo, and percent identity (green, 100% identity of paralogs; yellow,
1378 <100% identity of paralogs). Bases in each gene are color-coded (blue, C; red, A; yellow,
1379 G; and T, green). The *TP53RTG*-specific deletion is shown as a black line, and the
1380 location of the PCR primers than span the deletion are shown as arrows.

1381 (B) Chromatogram from Sanger sequencing of the *TP53RTG12* qRT-PCR amplicon confirm
1382 the *TP53RTG12* is transcribed in African elephant fibroblasts. Bases that flank the
1383 *TP53RTG*-specific deletion are highlighted in red.

1384 (C) Chromatogram from Sanger sequencing of the *TP53RTG12* qRT-PCR amplicon confirm
1385 the *TP53RTG12* is transcribed in Asian elephant fibroblasts. Bases that flank the
1386 *TP53RTG*-specific deletion are highlighted in red.

1387

1388 **Figure 5-figure supplement 2. Unedited Western blots shown in Fig. 5E.**

1389 (A) Western blot of total cell protein isolated from African elephant and hyrax fibroblasts. -,
1390 control cells. +, cells treated with 50 J/m^2 UV-C and MG-132. 1° & 2°, probed with
1391 primary antibody against TP53 and secondary HRP conjugated antibody. 2° only, probed
1392 only with secondary HRP conjugated antibody

1393 (B) Western blot of total cell protein isolated from Asian elephant and hyrax fibroblasts. -,
1394 control cells. +, cells treated with 50 J/m^2 UV-C and MG-132. 1° & 2°, probed with
1395 primary antibody against TP53 and secondary HRP conjugated antibody. 2° only, probed
1396 only with secondary HRP conjugated antibody

1397

1398

1399 **Figure 6. Elephant cells have enhanced TP53 signaling and are hyper-responsive to DNA**
1400 **damage.**

1401 (A) Relative luciferase (Luc.) expression in African elephant, Asian elephant, hyrax,
1402 aardvark, and armadillo fibroblasts transfected with the pGL4.38[*luc2p*/p53 RE/Hygro]
1403 reporter vector and treated with either mitomycin c, doxorubicin, nutlin-3a, or UV-C. Data
1404 is shown as fold difference in Luc. expression 18hrs after treatment standardized to
1405 species paired empty vector and *Renilla* controls. n=12, mean±SD.

1406 (B) Relative capsase-3/7 (Cas3/7) activity in African elephant, Asian elephant, hyrax,
1407 aardvark, and armadillo treated with either mitomycin c, doxorubicin, nutlin-3a, or UV-C.
1408 Data is shown as fold difference in Cas3/7 activity 18hrs after treatment standardized to
1409 species paired untreated controls. n=12, mean±SD.

1410

1411 **Figure 7. TP53RTG genes are required for enhanced TP53 signaling and DNA-damage**
1412 **responses.**

1413 (A) Expression of TP53RTG and TP53 transcripts in African elephant fibroblasts treated with
1414 an siRNA to knockdown the expression of TP53RTG genes (red) or a scrambled
1415 (Control) siRNA (blue). Results are shown as fold-change in TP53RTG and TP53
1416 transcript abundance relative to transcript abundance in scrambled siRNA control cells.
1417 The TP53RTG siRNA efficiently reduces the expression of TP53RTG transcripts, but
1418 does not reduce the expression of TP53 transcripts.

1419 (B) Relative luciferase (Luc.) expression in African elephant fibroblasts transfected with the
1420 pGL4.38[*luc2p*/p53 RE/Hygro] reporter vector and treated with an siRNA to knockdown
1421 the expression of TP53RTG genes (red) or a scrambled (negative control) siRNA (blue),
1422 and treated with either mitomycin c, doxorubicin, nutlin-3a, or UV-C. Data is shown as
1423 fold difference in Luc. expression 18hrs after treatment standardized to *Renilla* controls
1424 and no treatment. n>4, mean±SD. **, Wilcoxon P>0.01. ***, Wilcoxon P>0.001.
1425
1426 **Figure 7-figure supplement 1. The TP53RTG siRNA knocks downs TP53RTG transcripts**
1427 **but not TP53 transcripts.** African elephant fibroblasts were transfected with either an siRNA to
1428 knockdown the expression of TP53RTG genes (red) or a scrambled (negative control) siRNA
1429 (blue). TP53RTG and TP53 transcripts were assayed by qRT-PCR 48 hours later.
1430
1431 **Figure 8. TP53RTG12 enhances TP53 signaling and DNA-damage responses.**

1432 (A) Relative luciferase (Luc.) expression in mouse 3T3-L1 cells co-transfected with either the
1433 pGL4.38[*luc2P*/p53 RE/Hygro] Luc. reporter vector, *TP53RTG12* pcDNA3.1(+)/myc-His
1434 expression vector, or empty pcDNA3.1(+)/myc-His and treated with either mitomycin c,
1435 doxorubicin, nutlin-3a, or UV-C. Data is shown as fold difference in Luc. expression
1436 18hrs after treatment standardized to cells transfected with only pGL4.38[*luc2P*/p53
1437 RE/Hygro] and *Renilla* controls. n=12, mean±SD. ***, Wilcoxon P>0.001.
1438 (B) Relative capsase-3/7 (Cas3/7) activity in mouse 3T3-L1 cells transfected with either the
1439 *TP53RTG12* pcDNA3.1(+)/myc-His expression vector or empty pcDNA3.1(+)/myc-His
1440 and treated with either mitomycin c, doxorubicin, nutlin-3a, or UV-C. Data is shown as
1441 fold difference in Cas3/7 activity 18hrs after treatment standardized to mock transfected.
1442 n=12, mean±SD. ***, Wilcoxon P>0.001.
1443 **Figure 9. TP53RTG proteins are unlikely to directly regulate TP53 target genes.**

1444 (A) Domain structure of TP53 and TP53RTG proteins.
1445 (B) Localization of TP53RTG12-GFP in African elephant fibroblasts.
1446 (C) Relative luciferase (Luc.) expression in African elephant fibroblasts transfected with co-
1447 transfected with either the *TP53RTG12* pcDNA3.1(+)/myc-His expression vector and the
1448 pGL4.38[*luc2P*/p53 RE/Hygro] luciferase reporter vector, or empty pcDNA3.1(+)/myc-His
1449 and the pGL4.38[*luc2P*/p53 RE/Hygro] luciferase reporter vector. Data is shown as fold

difference in Luc. expression 48hrs after transfection standardized to *Renilla* and cells transfected with pcDNA3.1(+)/myc-His and pGL4.38[*luc2P*/p53 RE/Hygro]. n=10.

Figure 10. TP53RTG12 interacts with TP53 but not MDM2

(A) Decoy 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 the MDM2 complex, allowing TP53 to escape negative regulation.

(B) Logo of the MDM2 interaction motif from 99 mammalian TP53 proteins (upper) and TP53RTG proteins (lower).

(C) Structure model of the MDM2/TP53 (left) and MDM2/TP53RTG12 (right) interaction. MDM2 is shown in tan with hydrophobic residues that mediate the interaction with TP53 as spheres, TP53 in gray with W23 as a sphere, and TP53RTG12 shown in blue with G23 as a sphere.

(D) Predicted effects of the W23G substitution (ΔΔG) on the stability of the MDM2/TP53RTG12 interaction estimated with mCSM, SDM, and DUET.

(E) HEK-293 cells were transiently transfected with the TP53RTG12 pcDNA3.1(+)/myc-His expression vector and total cell protein immunoprecipitated with an α-MDM2 antibody. Co-immunoprecipitation of Myc-tagged TP53RTG12 and TP53 were assayed by Western blotting with α-Myc, α-TP53, and α-MDM2 antibodies after chemically stripping the blot, respectively. Although TP53 was co-immunoprecipitated, Myc-tagged TP53RTG12 was not.

(F) Rather than interfering with the interaction between the MDM complex and TP53 as in the 'Decoy" model, 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.

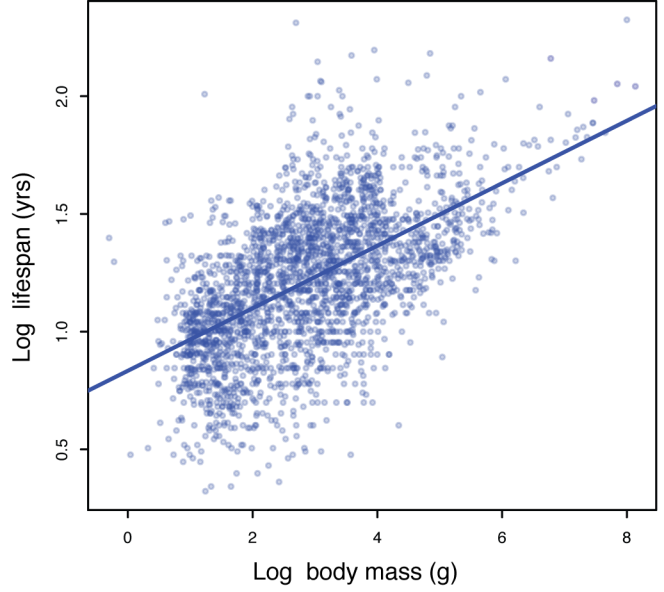
(G) Logo of the dimerization domain of TP53 from 99 mammals (upper) and TP53RTG proteins (lower).

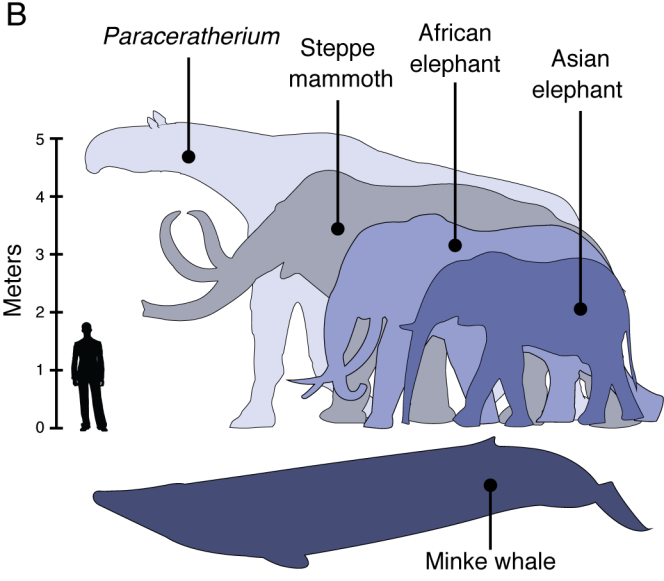
1481 (H) Model of the TP53/TP53RTG12 interaction. Residues critical for dimerization are shown
1482 as spheres, note sites involved in dimerization are conserved in TP53RTG proteins.
1483 (I) Predicted effects of the W23G substitution (ΔΔG) on the stability of the
1484 MDM2/TP53RTG12 interaction estimated with mCSM, SDM, and DUET.
1485 (J) HEK-293 cells were transiently transfected with the TP53RTG12 pcDNA3.1(+)/myc-His
1486 expression vector and total cell protein immunoprecipitated with an α-Myc antibody. Co-
1487 immunoprecipitation of TP53 and MDM2 were assayed by serial Western blotting with α-
1488 Myc, α-TP53, and α-MDM2 antibodies after chemically stripping the blot, respectively.
1489
1490 **Figure 10-figure supplement 1. Uncropped Western blots shown in Fig. 8B.** HEK-293 cells
1491 were transiently transfected with the TP53RTG12 pcDNA3.1(+)/myc-His expression vector and
1492 total cell protein immunoprecipitated with either an α-Myc (IP: RTG12-Myc) or α-MDM2 (IP:
1493 MDM2) antibody. Co-immunoprecipitation of Myc-tagged TP53RTG12, TP53, or MDM2 were
1494 assayed by Western blotting with α-Myc, α-TP53, and α-MDM2 antibodies respectively.
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513

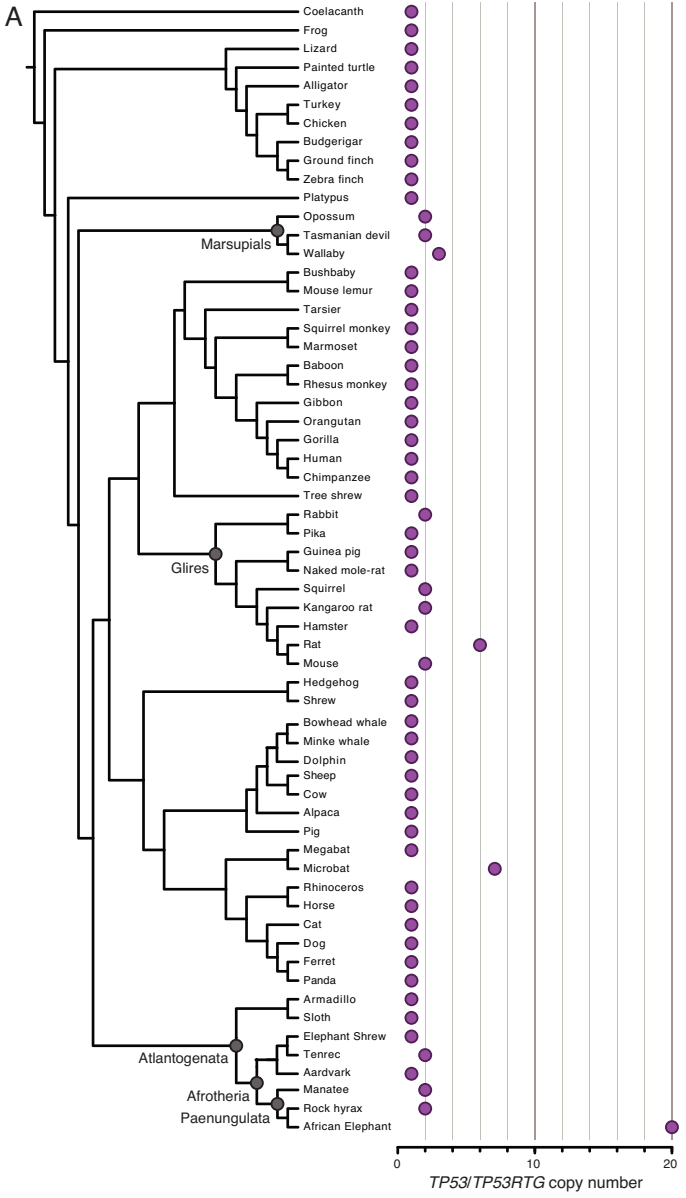
1514 Table 1. Summary information for African elephant *TP53*/*TP53RTG* genes.
1515

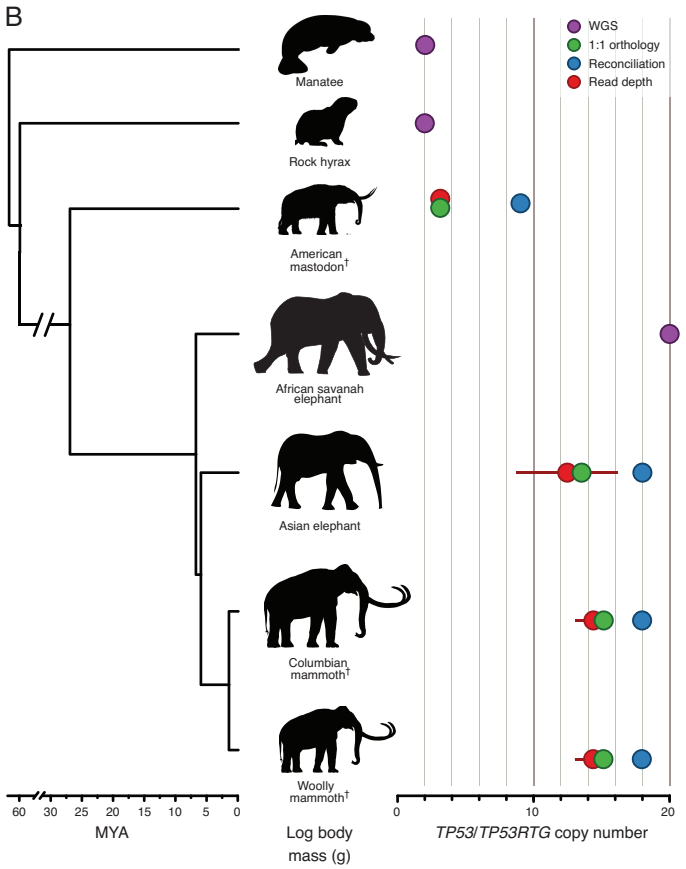
Gene ID	ENSEMBL ID	Scaffold (loxAfr3)	Chromosome (loxAfr4)	Coding potential	ORF size
TP53	ENSLAFG00000007483	47	Chr 11	Yes	392aa
TP53RTG1	ENSLAFG00000025553	175	Unmapped	No	N/A
TP53RTG2	N/A	217	Unmapped	Yes	134aa
TP53RTG3	ENSLAFG00000027474	406	Unmapped	Yes	79aa
TP53RTG4	N/A	627	Unmapped	Yes	134aa
TP53RTG5	ENSLAFG00000027348	221	Unmapped	Yes	162aa
TP53RTG6	N/A	76	Chr 27	Yes	123aa
TP53RTG7	N/A	208	Unmapped	No	N/A
TP53RTG8	ENSLAFG00000027820	294	Unmapped	Yes	210aa
TP53RTG9	ENSLAFG00000027669	786	Unmapped	No	N/A
TP53RTG10	ENSLAFG00000030555	221	Unmapped	Yes	210aa
TP53RTG11	N/A	281		Yes	203aa
TP53RTG12	ENSLAFG00000028299	825	Unmapped	Yes	180aa
TP53RTG13	ENSLAFG00000032042	458	Unmapped	No	N/A
TP53RTG14	ENSLAFG00000026238	928	Unmapped	Yes	210aa
TP53RTG15	ENSLAFG00000027365	656	Unmapped	Yes	210aa
TP53RTG16	ENSLAFG00000030880	378	Unmapped	No	N/A
TP53RTG17	ENSLAFG00000028692	552	Unmapped	Yes	111aa
TP53RTG18	N/A	498	Unmapped	Yes	111aa
TP53RTG19	ENSLAFG00000032258	342	Unmapped	Yes	210aa

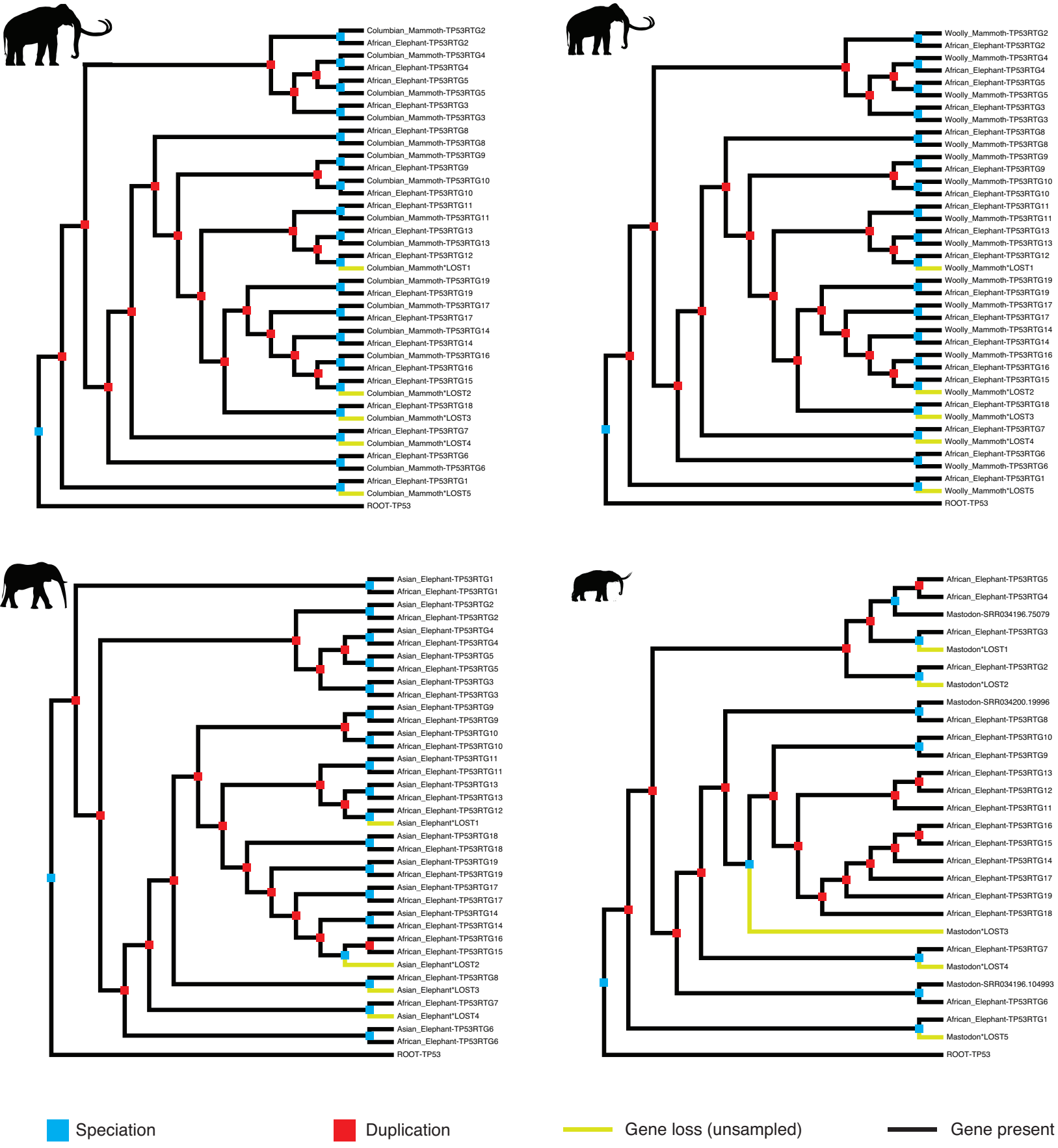
1516
1517

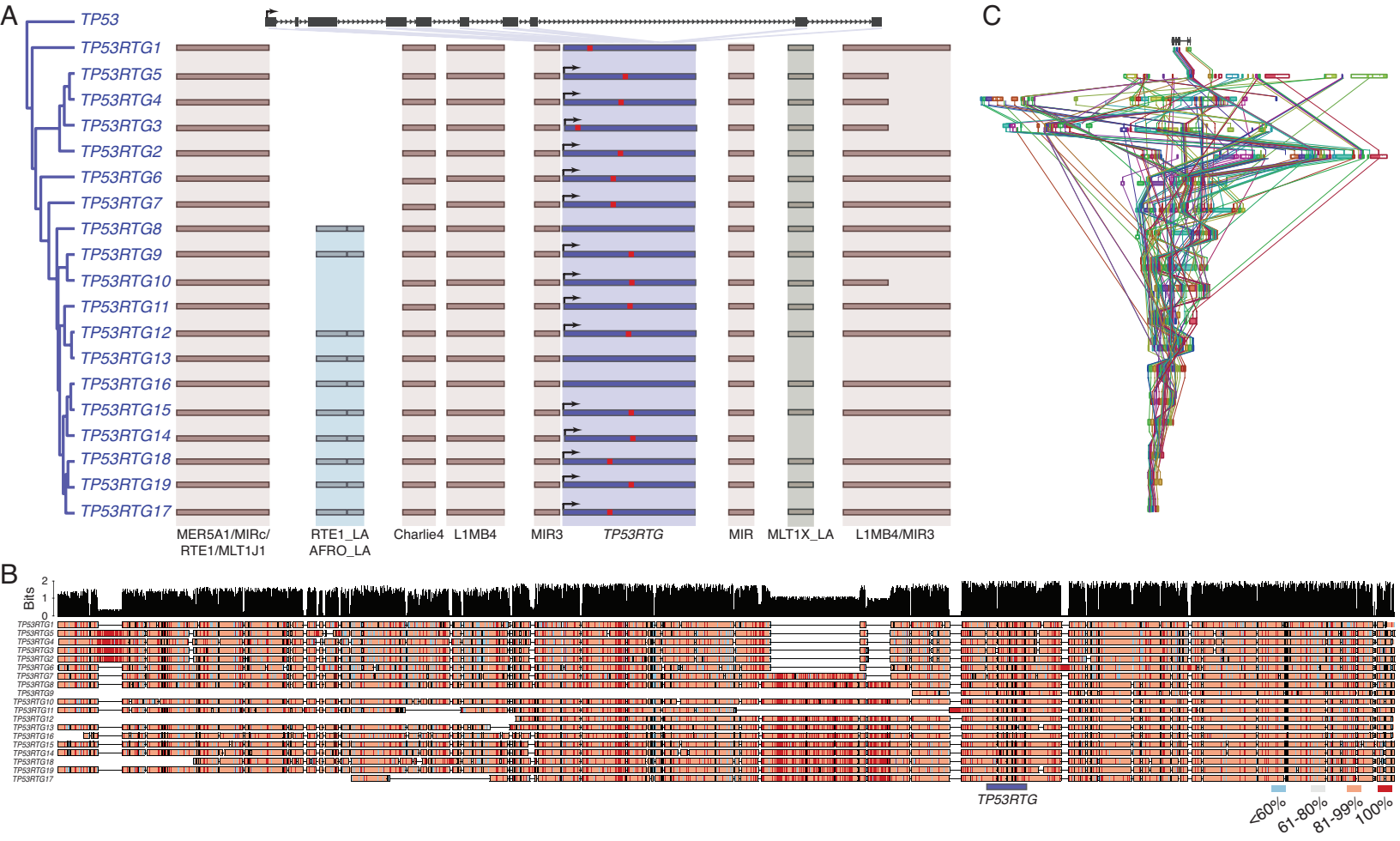

Log lifespan (yrs)
0.5
1.0
1.5
2.0
0
2
4
6
8
Log body mass (g)

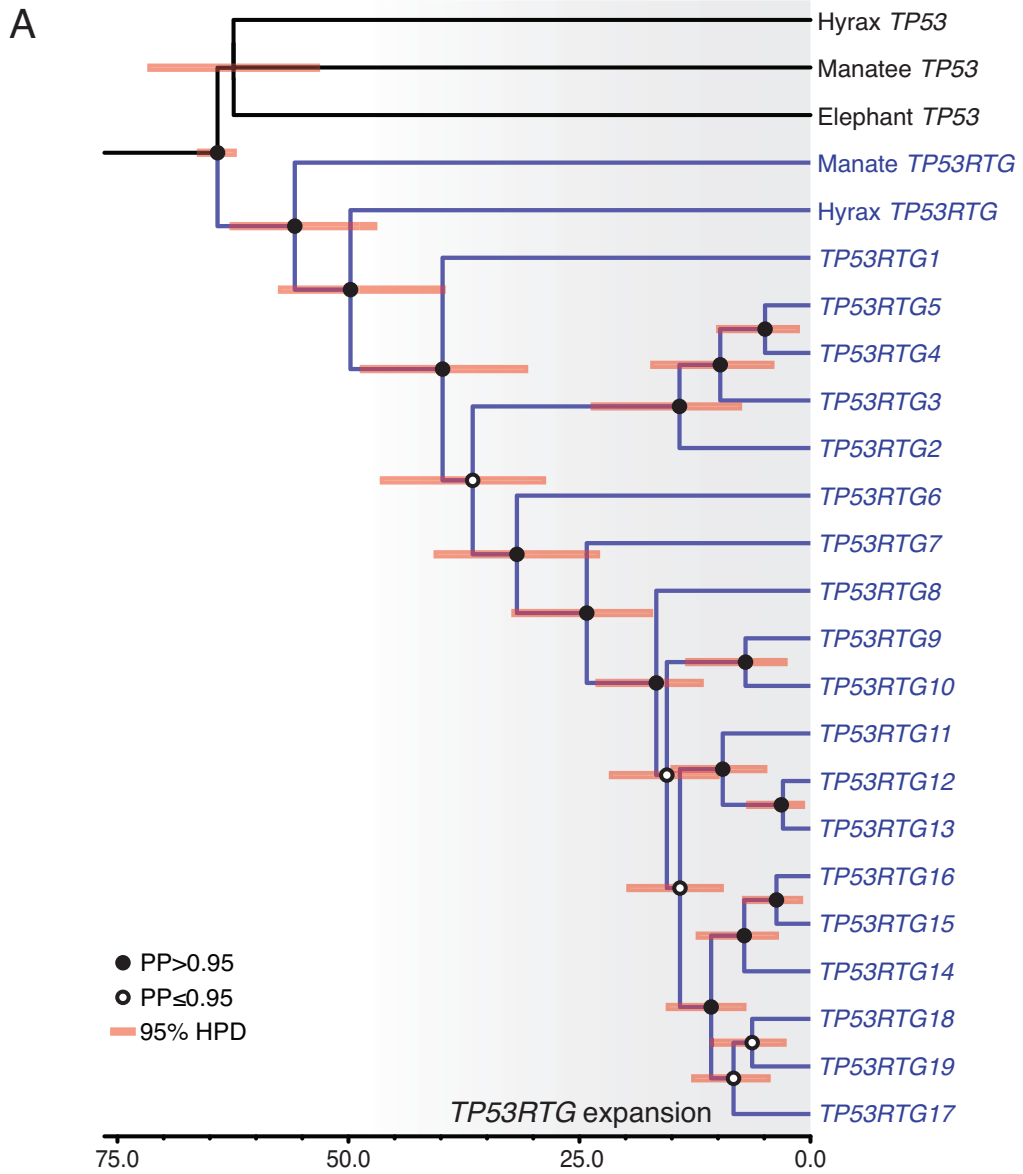

B
Paraceratherium
Steppe mammoth
African elephant
Asian elephant
Meters
0
1
2
3
4
5
Minke whale

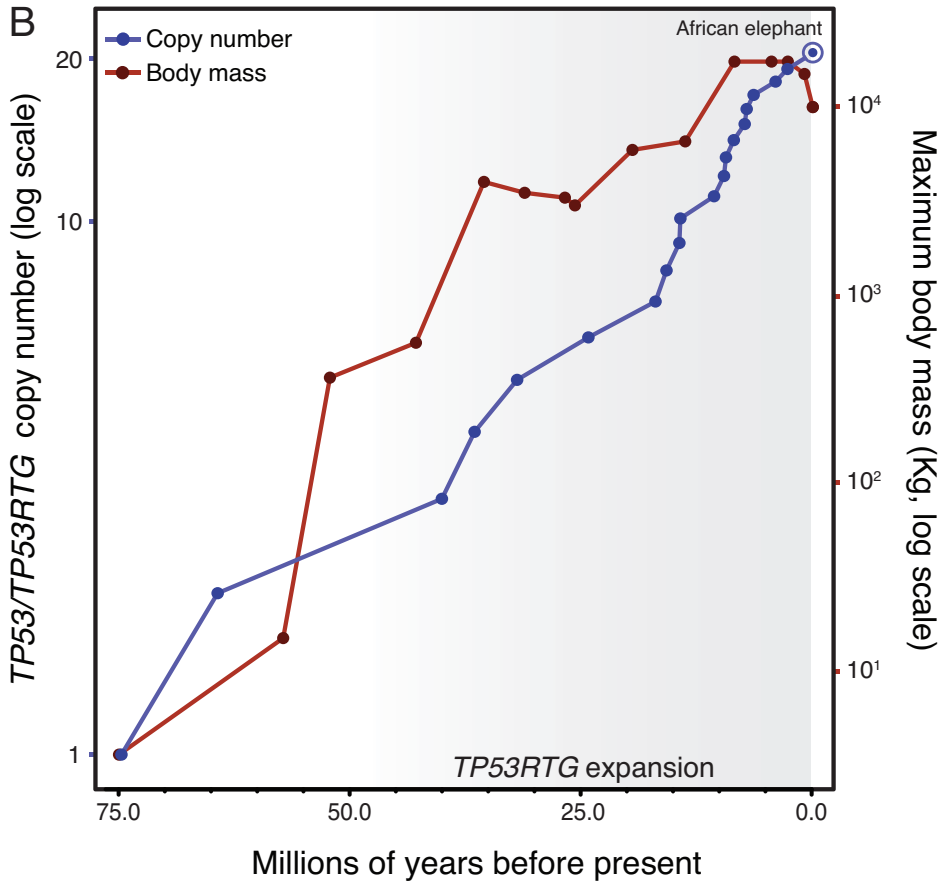

A
Coelacanth
Frog
Lizard
Painted turtle
Alligator
Turkey
Chicken
Budgerigar
Ground finch
Zebra finch
Platypus
Opossum
Tasmanian devil
Wallaby
Marsupials
Bushbaby
Mouse lemur
Tarsier
Squirrel monkey
Marmoset
Baboon
Rhesus monkey
Gibbon
Orangutan
Gorilla
Human
Chimpanzee
Tree shrew
Rabbit
Pika
Guinea pig
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
Atlantogenata
Manatee
Afrotheria
Rock hyrax
Paenungulata
African Elephant
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

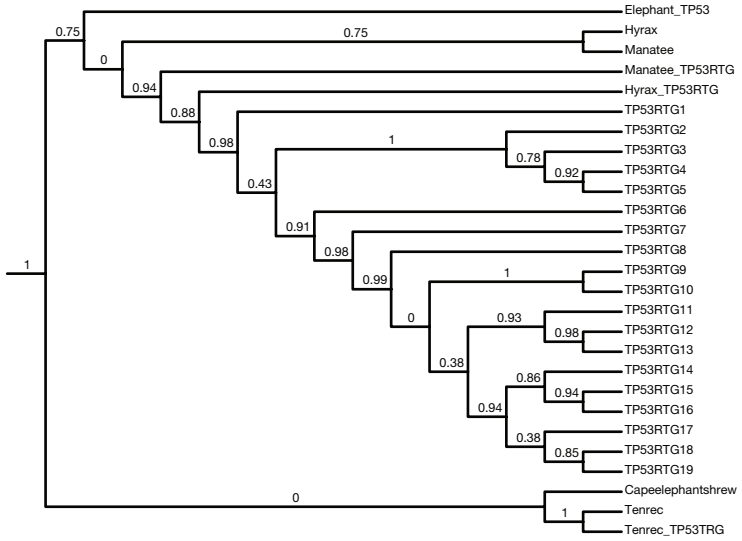

Columbian_Mammoth-TP53RTG2
African_Elephant-TP53RTG2
Columbian_Mammoth-TP53RTG4
African_Elephant-TP53RTG4
African_Elephant-TP53RTG5
Columbian_Mammoth-TP53RTG5
African_Elephant-TP53RTG3
Columbian_Mammoth-TP53RTG3
African_Elephant-TP53RTG8
Columbian_Mammoth-TP53RTG8
Columbian_Mammoth-TP53RTG9
African_Elephant-TP53RTG9
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


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%

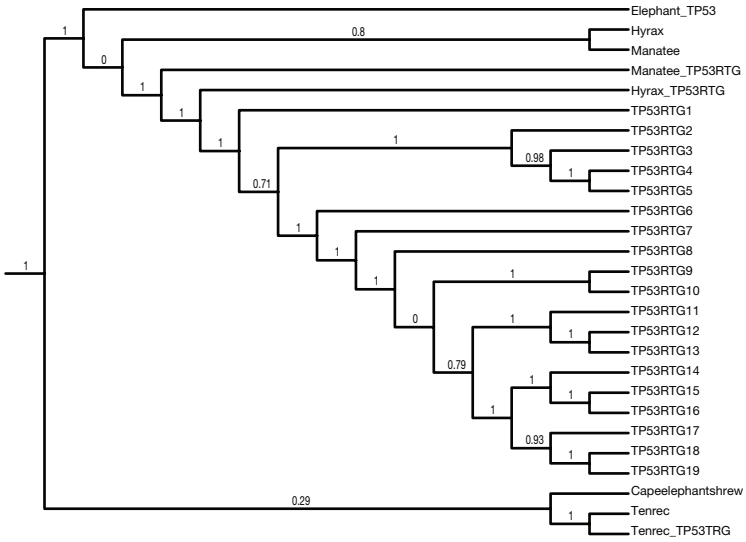

A
Hyrax TP53
Manatee TP53
Elephant TP53
Manate TP53RTG
Hyrax TP53RTG
TP53RTG1
TP53RTG5
TP53RTG4
TP53RTG3
TP53RTG2
TP53RTG6
TP53RTG7
TP53RTG8
TP53RTG9
TP53RTG10
TP53RTG11
TP53RTG12
TP53RTG13
TP53RTG16
TP53RTG15
TP53RTG14
TP53RTG18
TP53RTG19
TP53RTG17
PP>0.95
PP≤0.95
95% HPD
TP53RTG expansion
75.0
50.0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


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

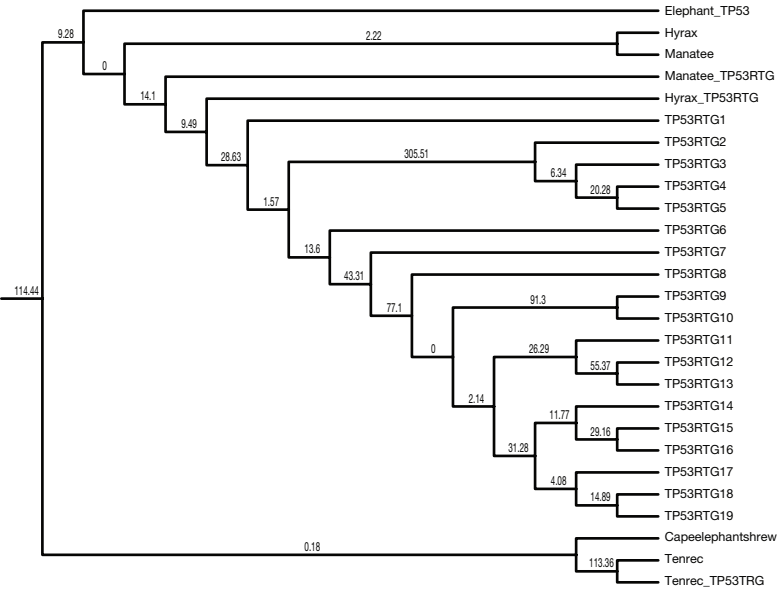
aBayes



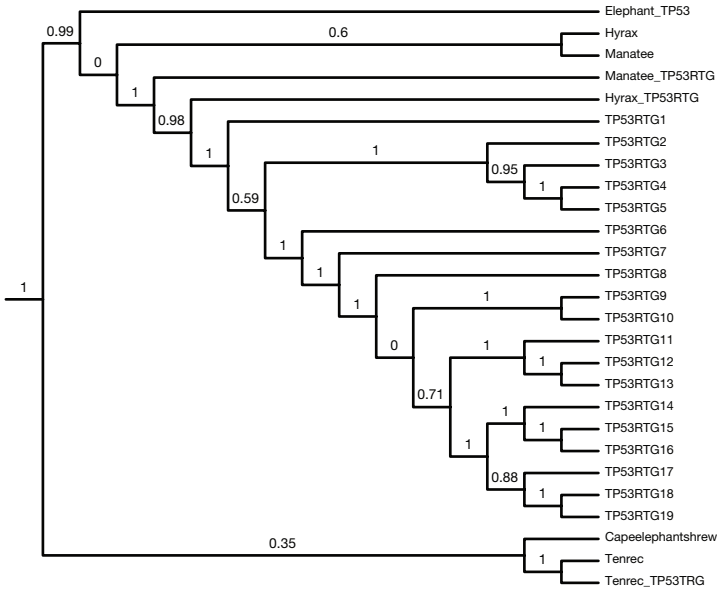
Chi-Square



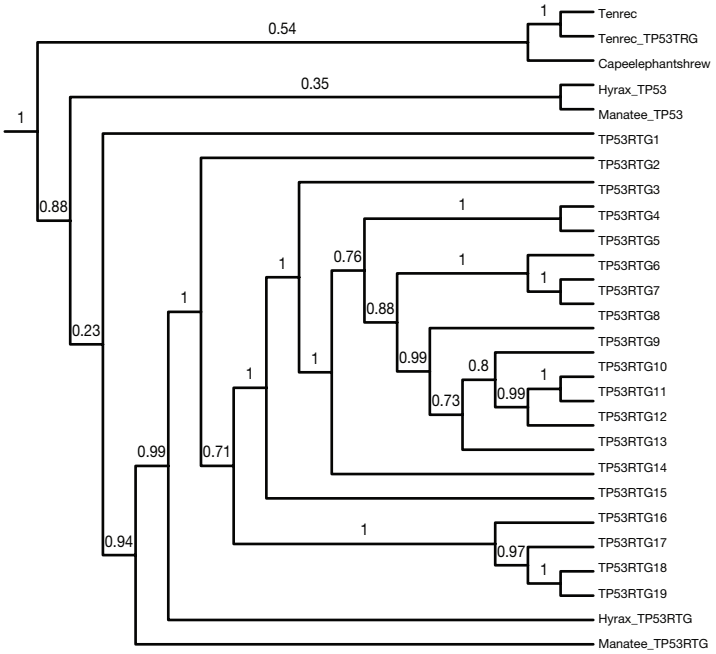
aLRT

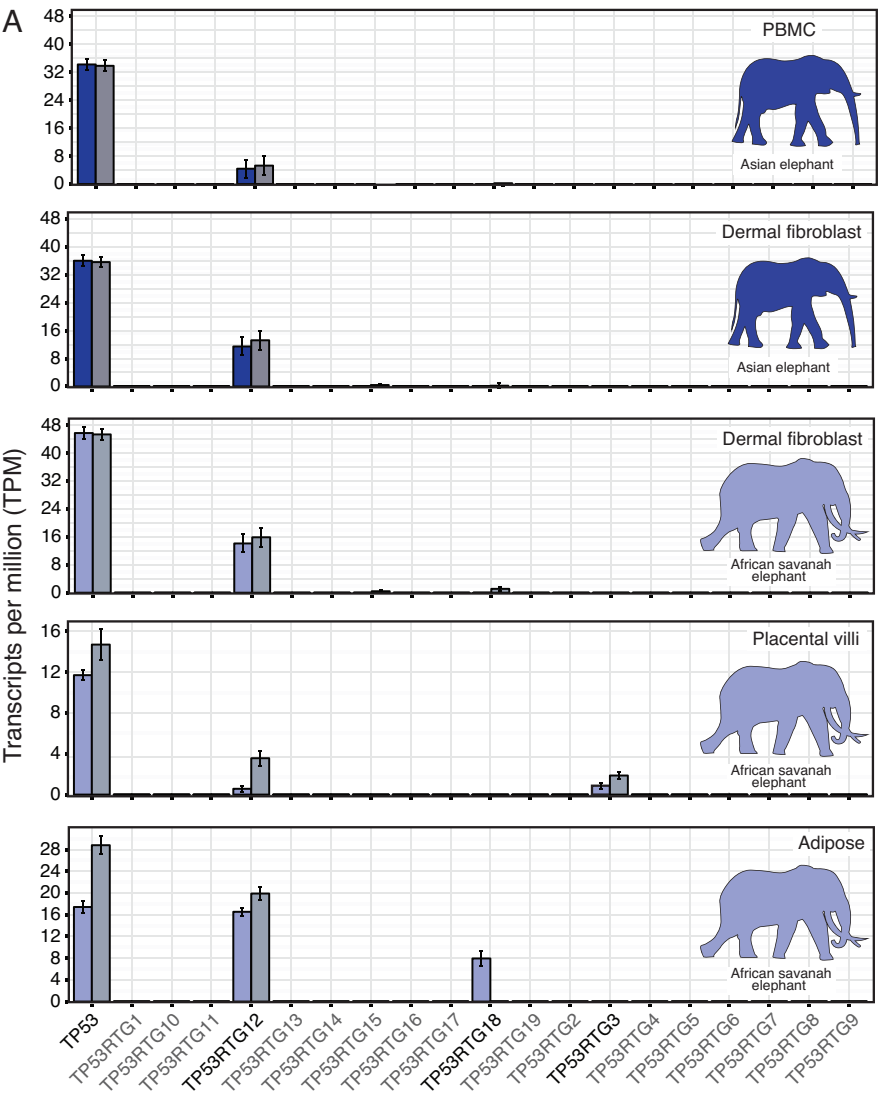


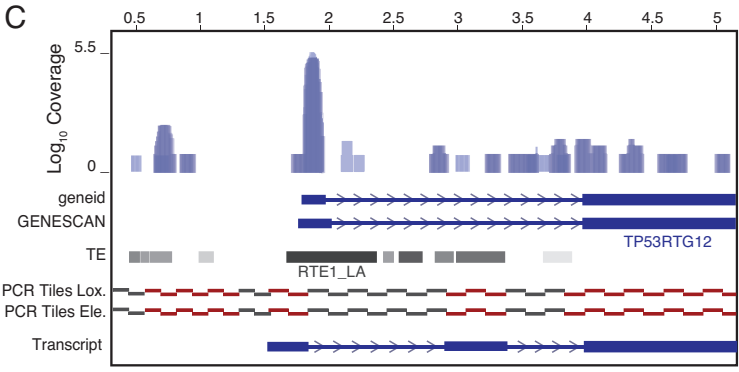
SH-Like

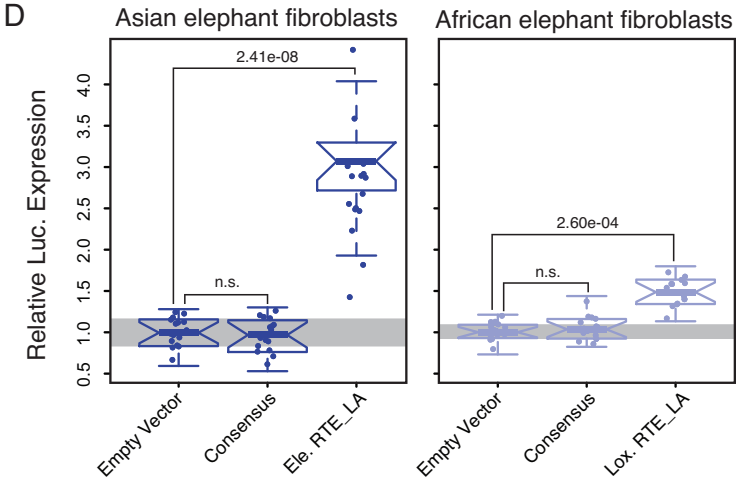


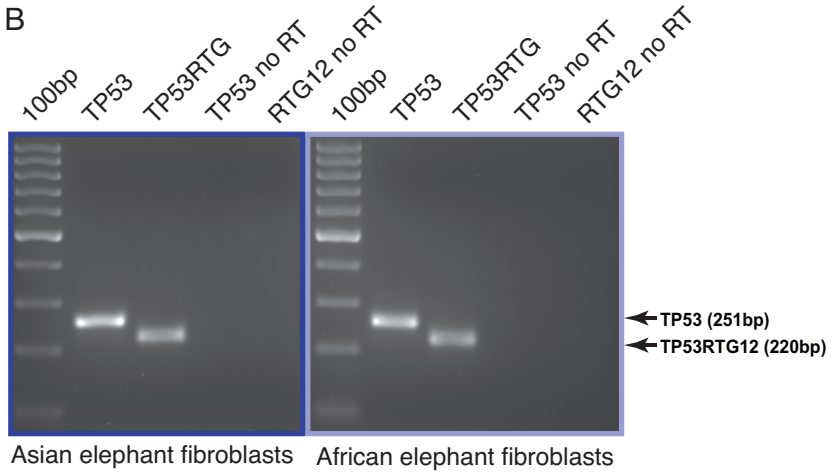
MrBayes

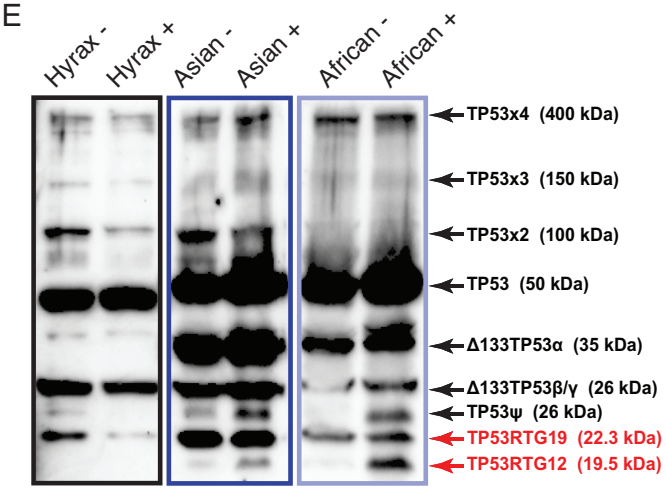


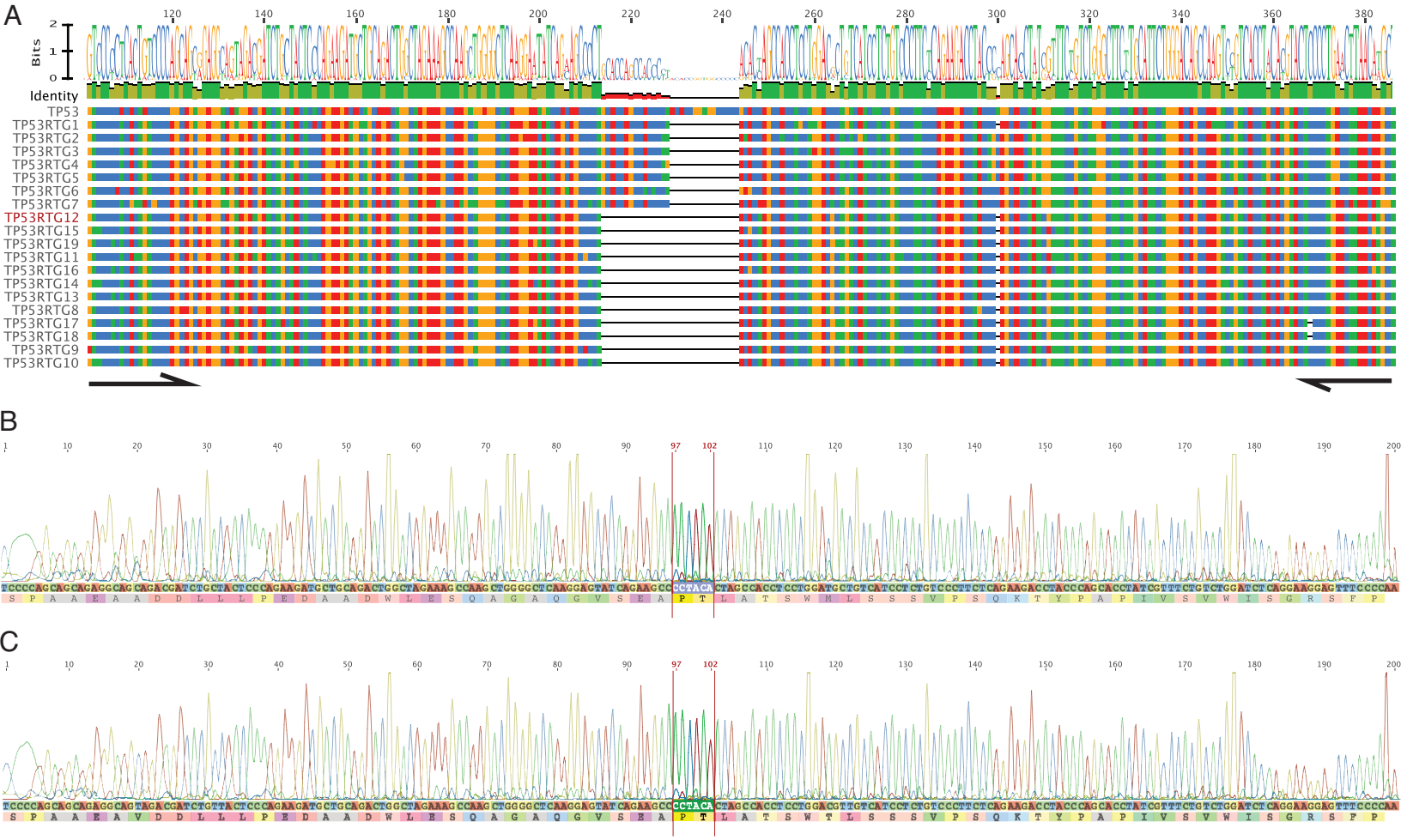

A
PBMC
Asian elephant
Dermal fibroblast
Asian elephant
Dermal fibroblast
African savanah elephant
Placental villi
African savanah elephant
Adipose
African savanah elephant
Transcripts per million (TPM)
TP53
TP53RTG1
TP53RTG10
TP53RTG11
TP53RTG12
TP53RTG13
TP53RTG14
TP53RTG15
TP53RTG16
TP53RTG17
TP53RTG18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

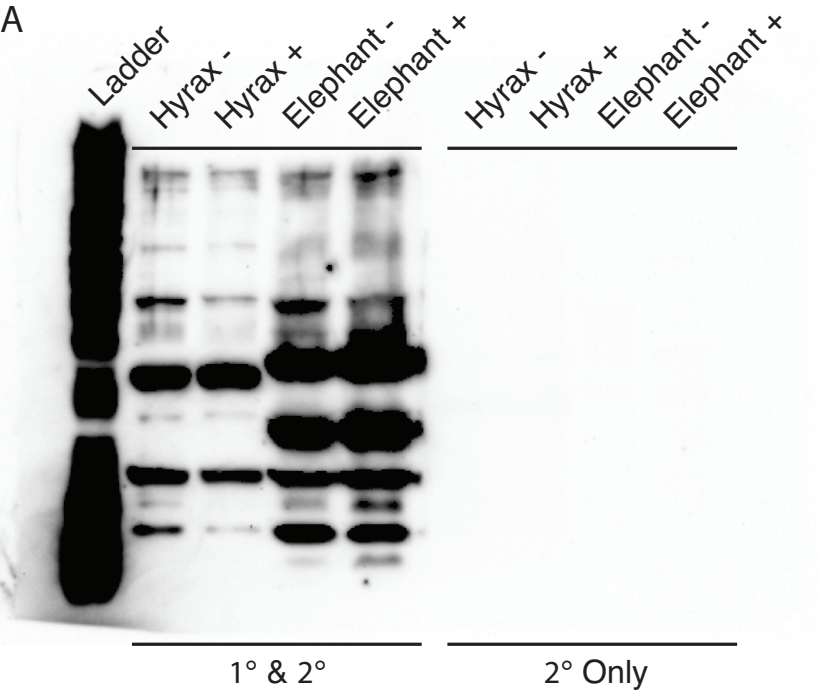

C
Log10 Coverage
geneid
GENESCAN
TE
RTE1_LA
TP53RTG12
PCR Tiles Lox.
PCR Tiles Ele.
Transcript

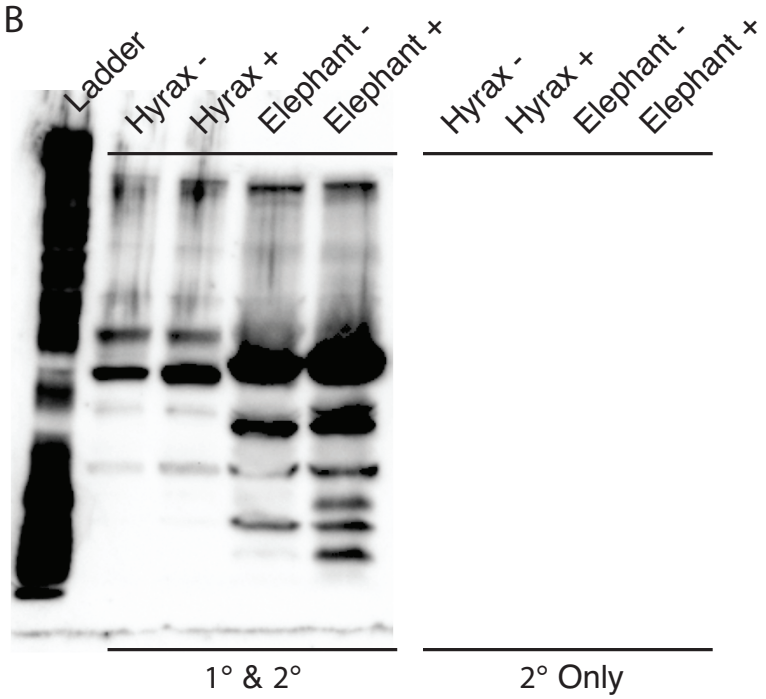

D
Asian elephant fibroblasts
African elephant fibroblasts
Relative Luc. Expression
2.41e-08
n.s.
2.60e-04
n.s.
Empty Vector
Consensus
Ele. RTE_LA
Empty Vector
Consensus
Lox. RTE_LA

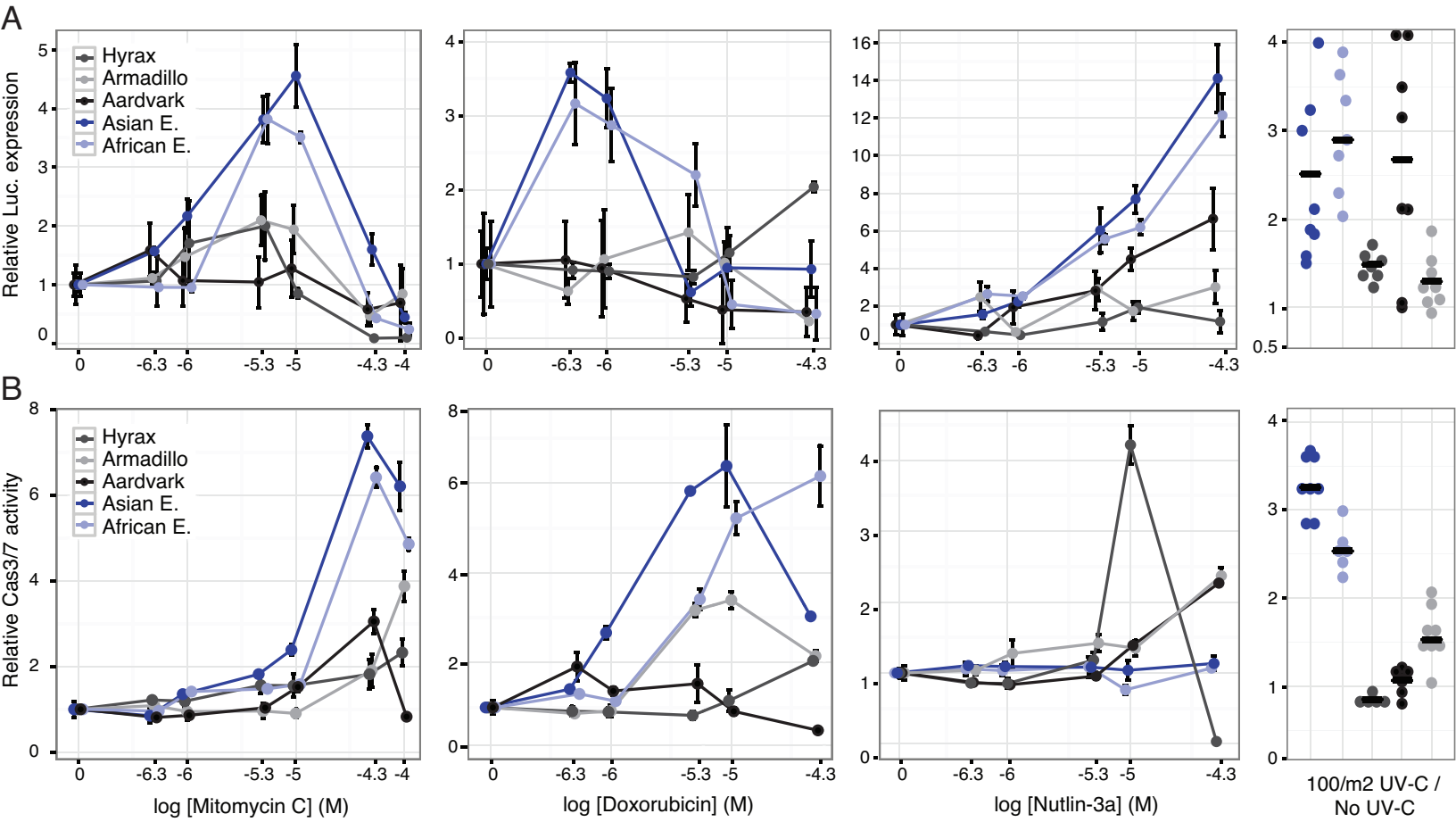

B
100bp
TP53
TP53RTG
TP53 no RT
RTG12 no RT
100bp
TP53
TP53RTG
TP53 no RT
RTG12 no RT
TP53 (251bp)
TP53RTG12 (220bp)
Asian elephant fibroblasts
African elephant fibroblasts

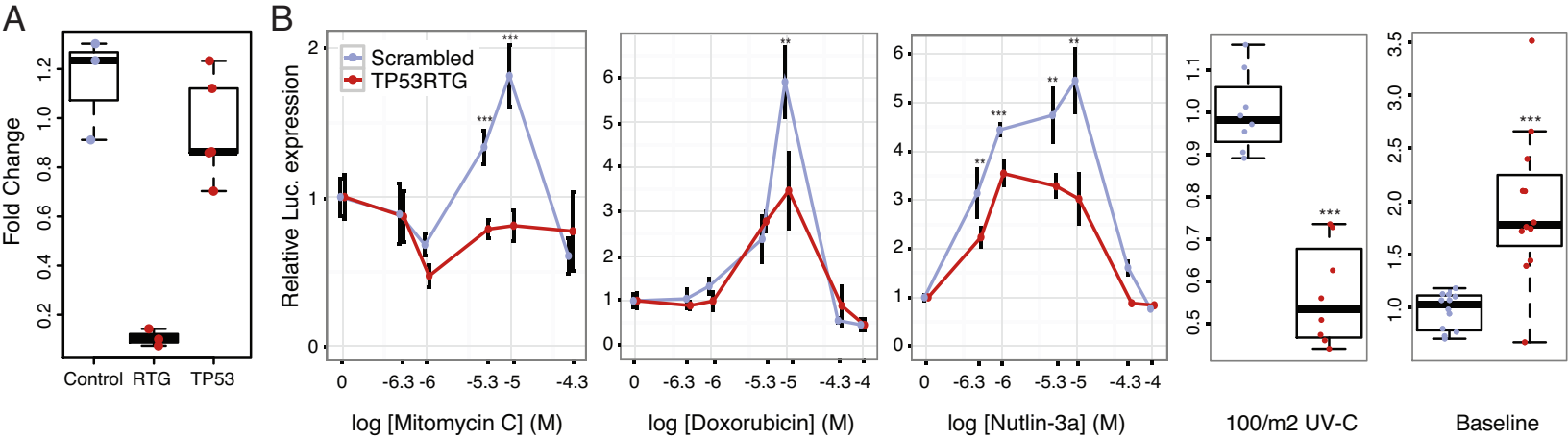

E
Hyrax -
Hyrax +
Asian -
Asian +
African -
African +
TP53x4 (400 kDa)
TP53x3 (150 kDa)
TP53x2 (100 kDa)
TP53 (50 kDa)
Δ133TP53α (35 kDa)
Δ133TP53β/γ (26 kDa)
TP53ψ (26 kDa)
TP53RTG19 (22.3 kDa)
TP53RTG12 (19.5 kDa)

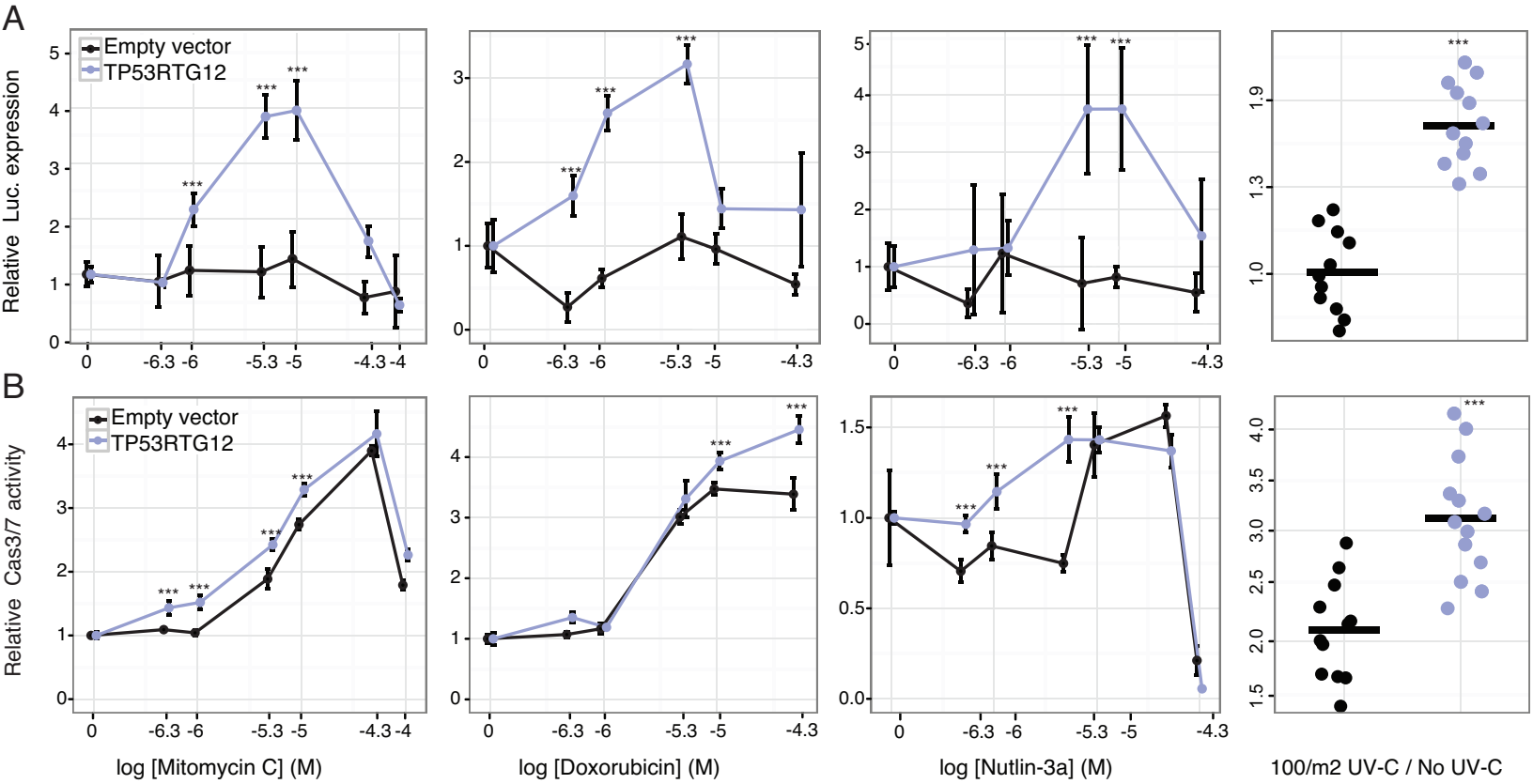

A
Bits
Identity
TP53
TP53RTG1
TP53RTG2
TP53RTG3
TP53RTG4
TP53RTG5
TP53RTG6
TP53RTG7
TP53RTG12
TP53RTG15
TP53RTG19
TP53RTG11
TP53RTG16
TP53RTG14
TP53RTG13
TP53RTG8
TP53RTG17
TP53RTG18
TP53RTG9
TP53RTG10
B
C

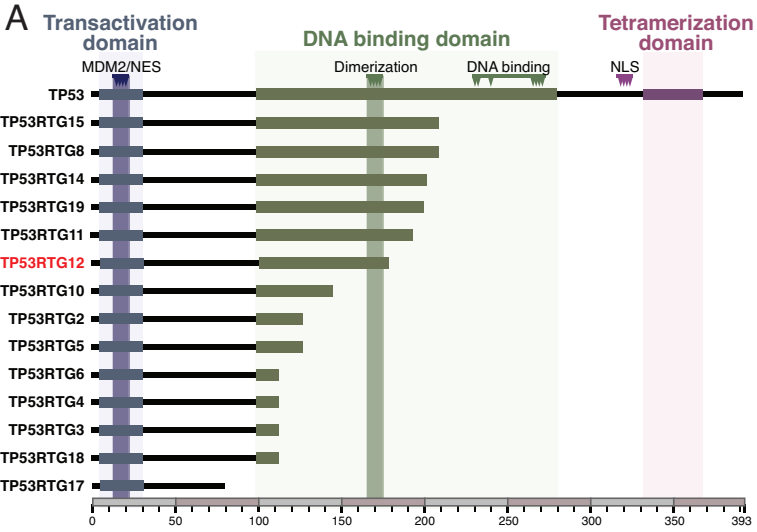

A
Ladder
Hyrax -
Hyrax +
Elephant -
Elephant +
Hyrax -
Hyrax +
Elephant -
Elephant +
1° & 2°
2° Only

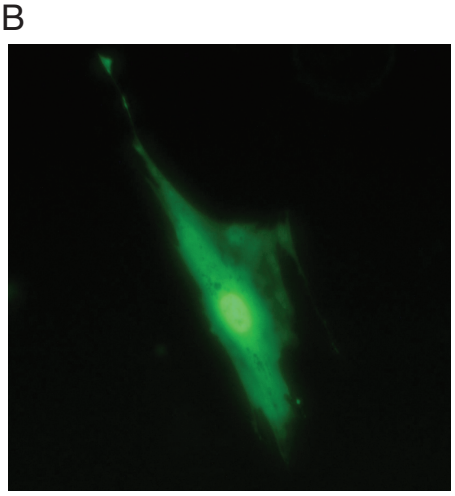

B
Ladder
Hyrax -
Hyrax +
Elephant -
Elephant +
Hyrax -
Hyrax +
Elephant -
Elephant +
1° & 2°
2° Only

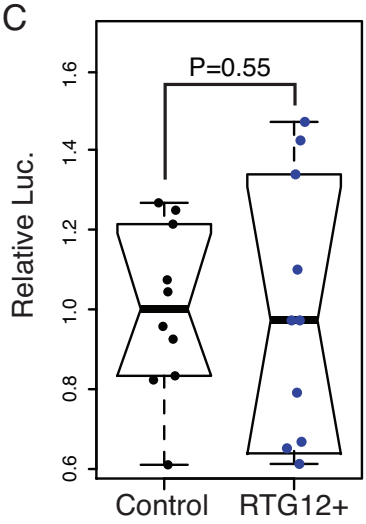

A
B
Hyrax
Armadillo
Aardvark
Asian E.
African E.
Relative Luc. expression
Relative Cas3/7 activity
0
-6.3
-6
-5.3
-5
-4.3
-4
log [Mitomycin C] (M)
log [Doxorubicin] (M)
log [Nutlin-3a] (M)
100/m2 UV-C /
No UV-C

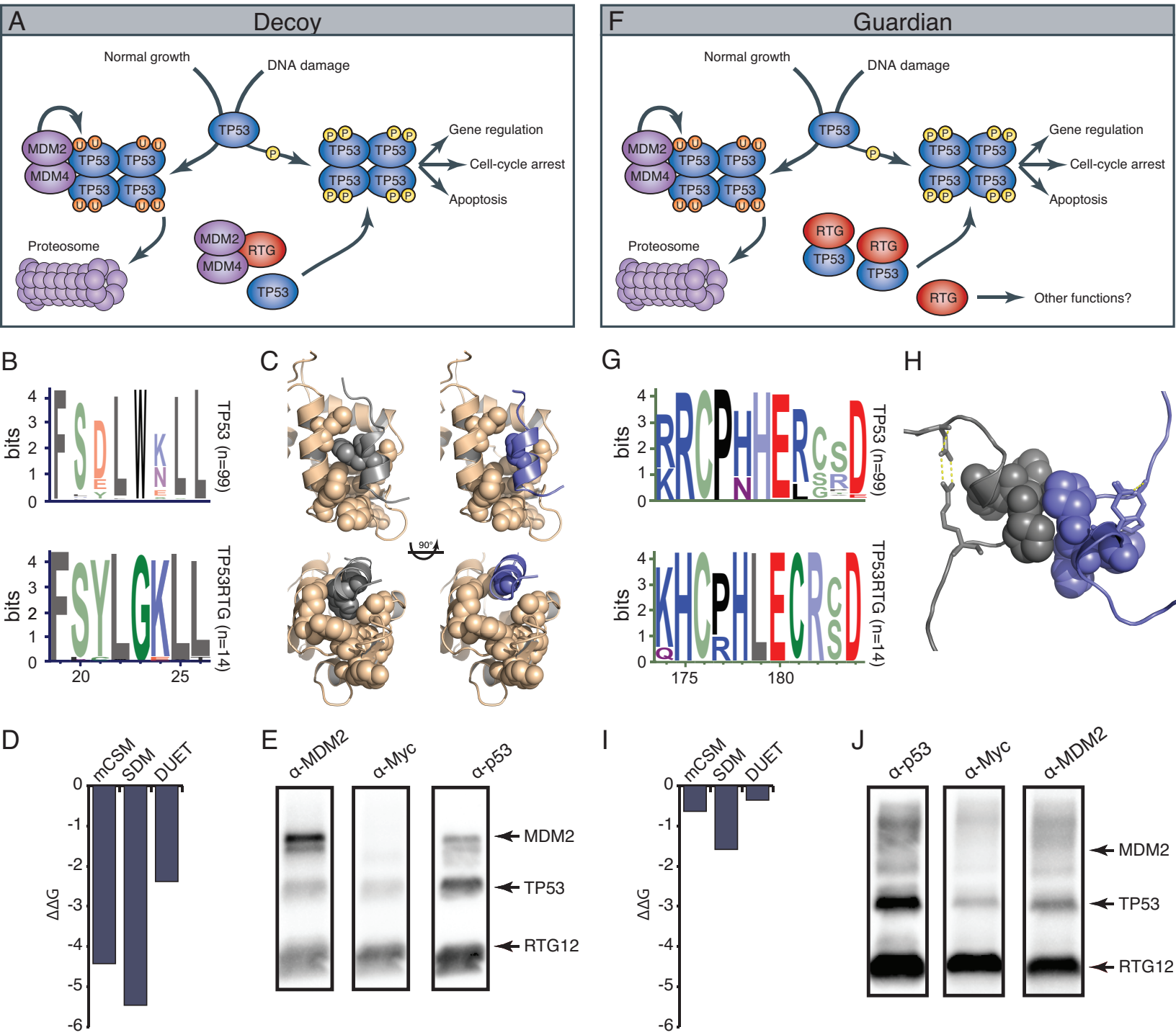

A
Fold Change
Control
RTG
TP53
B
Scrambled
TP53RTG
Relative Luc. expression
log [Mitomycin C] (M)
log [Doxorubicin] (M)
log [Nutlin-3a] (M)
100/m2 UV-C
Baseline

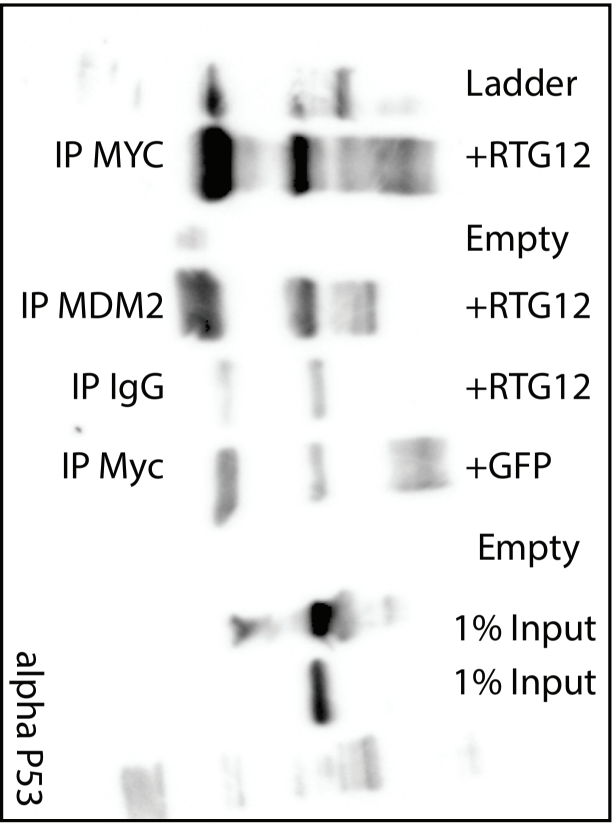

A
B
Empty vector
TP53RTG12
Relative Luc. expression
Relative Cas3/7 activity
log [Mitomycin C] (M)
log [Doxorubicin] (M)
log [Nutlin-3a] (M)
100/m2 UV-C / No UV-C
***

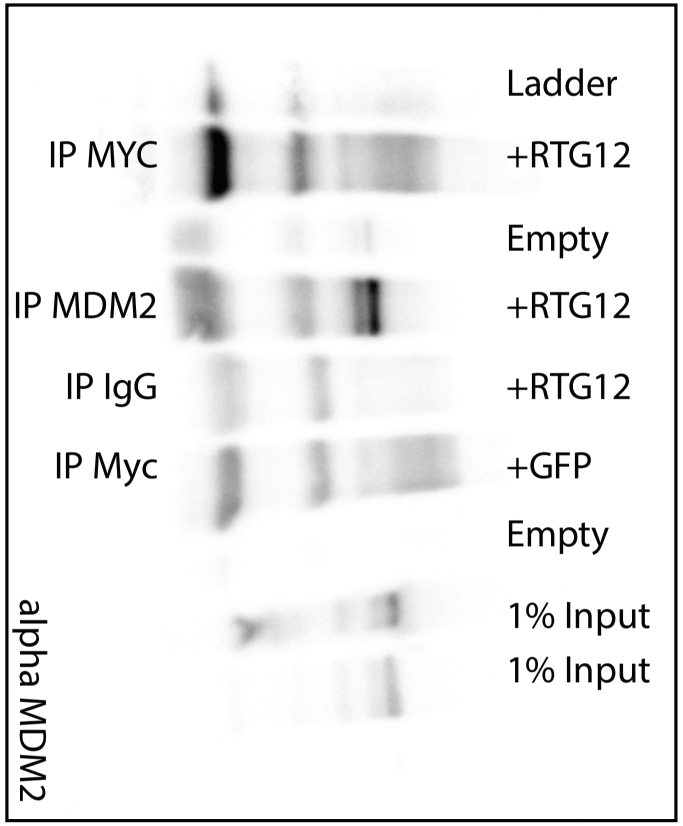

A
Transactivation domain
MDM2/NES
DNA binding domain
Dimerization
DNA binding
Tetramerization domain
NLS
TP53
TP53RTG15
TP53RTG8
TP53RTG14
TP53RTG19
TP53RTG11
TP53RTG12
TP53RTG10
TP53RTG2
TP53RTG5
TP53RTG6
TP53RTG4
TP53RTG3
TP53RTG18
TP53RTG17
0
50
100
150
200
250
300
350
393

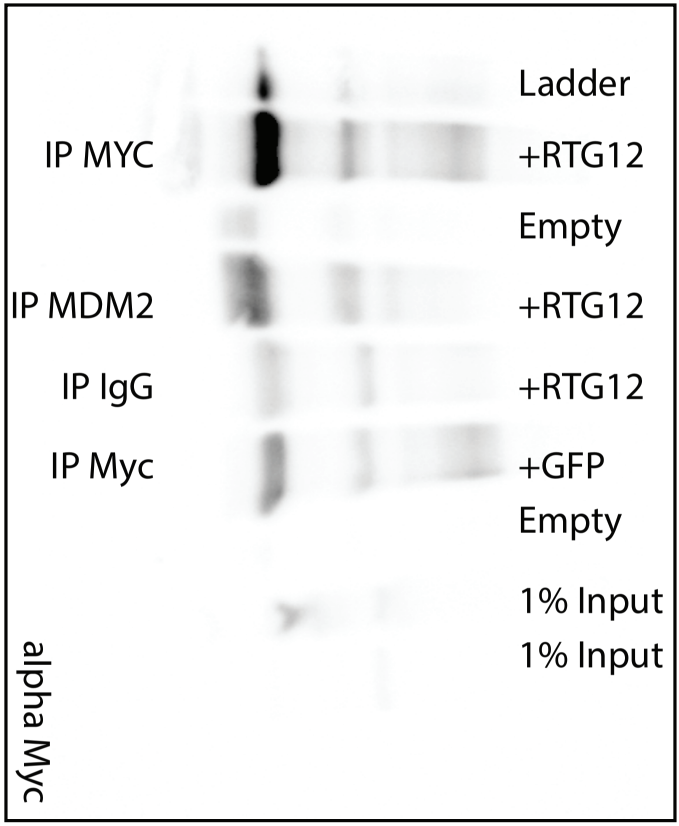

B


C
P=0.55
Relative Luc.
0.6
0.8
1.0
1.2
1.4
1.6
Control
RTG12+


A
Decoy
Normal growth
DNA damage
TP53
MDM2
MDM4
TP53
Proteosome
Gene regulation
Cell-cycle arrest
Apoptosis
RTG
F
Guardian
Other functions?
B
bits
TP53 (n=99)
TP53RTG (n=14)
C
90°
G
H
D
mCSM
SDM
DUET
ΔΔG
E
α-MDM2
α-Myc
α-p53
MDM2
TP53
RTG12
I
J


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