Facility for Antiproton and Ion Research in Europe GmbH & GSI
Helmholtzzentrum für Schwerionenforschung GmbH

PETO'S PARADOX AND THE CANCER SUPPRESSION MECHANISMS OF AFRICAN ELEPHANTS

Anne Sophie Blokland

Utrecht University

1990675

Internship and Training Project Report
Darmstadt, Germany
GET_INvolved 2023-2024 (GI24-CC-XXX)

26 January 2024

Author(s)

Anne Sophie Blokland
Utrecht University
Heidelberglaan 8, 3584 CS Utrecht – The Netherlands
Email: sophie.blokland.sb@gmail.com

Project Mentor/Supervisor

Charlot Vandevoorde, Biophysics Department
GSI Helmholtzzentrum für Schwerionenforschung GmbH,
Planckstrasse 1, 64291 Darmstadt - Germany
Tel: +49 6159 71, Email: C.Vandevoorde@gsi.de

Programme Coordinator

Dr. Pradeep Ghosh
GSI Helmholtzzentrum für Schwerionenforschung GmbH &
Facility for Antiproton and Ion Research in Europe GmbH
Tel: +49 6159 71 3257, Fax: +49 6159 71 3916

Email: Pradeep.Ghosh@fair-center.eu

GET_INvolved 2023: (1990675)

Publisher: GSI Helmholtzzentrum für Schwerionenforschung GmbH,
Planckstraße. 1, 64291 Darmstadt - Germany
Published: January 2019

Declaration

I hereby declare that the work presented in this thesis has not been submitted for any other degree or professional qualification, and that it is the result of my own independent work.

Layman summary

Peto's paradox bekritiseerd de theoretische veronderstelling dat het risico op kanker zou moeten toenemen bij zoogdieren met een grotere lichaamsgrootte en een langere levensduur. De Afrikaanse olifant, het grootste landzoogdier, lijkt een unieke oplossing voor dit paradox te hebben door het hebben van 20 kopieën van het *TP53* tumor onderdrukkende gen, dat codeert voor het p53 eiwit. Ook zou het kunnen dat het recent ontdekte pseudo leukemie onderdrukkingsfactor gen (*LIF6*) hierbij een rol speelt. Deze bevindingen suggereren een verbeterd kanker beschermingsmechanisme met een gevoeligere of efficiëntere geprogrammeerde celdood (apoptose) als reactie op DNA-schade. In verband hiermee rijst er een vraag op, aangezien gemuteerde muizen met verhoogde p53 activiteit beschermd zijn tegen kankervorming maar vroegtijdige veroudering vertonen. Dit zou betekenen dat olifanten een verfijnde balans hebben ontwikkeld tussen de agressieve eliminatie van beschadigde cellen voordat ze kwaadaardig kunnen worden, en de uitputting van stamcellen die vroegtijdige veroudering tot gevolg zou kunnen hebben.

Deze studie vergelijkt de transcriptoom-analyse tussen menselijke (n=3) en Afrikaanse olifanten (n=3) perifere bloed mononucleaire cellen die zijn blootgesteld aan 2 Gy Co60 γ-straal bestraling om DNA-schade te induceren, evenals sham-bestraalde (0 Gy) controle samples. De analyse omvatte het uitvoeren van differentiële expressie analyse met behulp van de NOISeq-tool, gevolgd door een analyse om biologische processen op te sporen, gebaseerd op de Gene Ontology (GO) en Kyoto Encyclopedia of genen and genomen (KEGG) databases. Er werden duidelijke verschillen waargenomen in significante differentieel tot expressie gebrachte genen (DEGs) tussen menselijke (DEG: 7399) en olifantencellen (DEG: 2869), maar er waren ook enkele overlappende genen, zoals *PHLDA3*, *BBC3* en *CXCL11*, die voornamelijk geassocieerd zijn met pro-apoptotische en immuunreacties.

De oververtegenwoordiging analyse (ORA) op basis van de GO-database benadrukte meer patronen van immuunreacties, geassocieerd met neerwaarts gereguleerde genen in beide soorten. Bij mensen onthulde de genen set verrijkingsanalyse (GSEA) een verhoogde biogenesis van ribosomale sub eenheden en een geactiveerde p53 signalering. Het butanoaat- (of boterzuur) metabolisme werd significant onderdrukt bij

olifanten, wat geassocieerd is met glycolyse en glutamaat metabolisme en een centrale rol speelt in energieproductie. Het is interessant dat boterzuur een metaboliet is met een bekende anti-kankerrol tegen ontsteking en ongecontroleerd kwaadaardig gedrag van kankercellen, door de versnelling van apoptose en differentiatie.

Over het algemeen werden soort specifieke variaties in de DEGs en biologische processen geïdentificeerd die verband houden met de p53 signalering en apoptose gerelateerde biologische processen. Ondanks de beperkingen zoals hoge variabiliteit tussen replica's en limitatie in steekproefgrootte, biedt deze studie mogelijke suggesties en potentiële toekomstige onderzoeksrichtingen naar aanvullende biologische mechanismen die helpen bij het onderdrukken van kanker bij olifanten.

Abstract

Peto's paradox challenges the theoretical assumption that the risk of cancer should increase with larger body size and longer life span in mammalian species. The African elephant, the largest terrestrial mammal, seems to provide a unique solution to this paradox through its possession of 20 copies of the *TP53* tumour suppressor gene which encodes for the p53 protein, and the recently discovered leukemia inhibitory factor pseudogene (*LIF6)*. These findings suggest an enhanced cancer protection mechanism through more sensitive or efficient apoptotic processes in response to DNA damage. However, another question arises, since mutant mice with increased p53 activity are protected from tumorigenesis but show premature aging. This would mean that elephants have developed a sophisticated balance between the aggressive elimination of damaged cells before they can become cancerous, and stem cell exhaustion which could result in premature aging.

This study compares the transcriptomic response between human (n=3) and African elephant (n=3) peripheral blood mononuclear cells which have been exposed to 2 Gy Co60 γ-ray irradiation to induce DNA damage, as well as sham-irradiated (0 Gy) control samples. The analysis involved conducting differential expression analysis using the NOISeq tool, followed by pathway enrichment analysis, based on the Gene Ontology (GO) and Kyoto Encyclopedia of genes and genomes (KEGG) databases.

Distinct differences were observed in significant differential expressed genes (DEGs) between human (DEG: 7399) and elephant (DEG: 2869) cells, but there were also some overlapping genes, such as *PHLDA3*, *BBC3*, and *CXCL11*, mainly associated with pro-apoptotic and immune responses. Over-representation analysis based on the GO database highlighted more immune response patterns, associated with downregulated genes in both species. In humans, gene set enrichment analysis revealed increased ribosomal subunit biogenesis and activated p53 signalling. The butanoate (or butyrate) metabolism pathway was significantly supressed in elephants, which is associated with glycolysis and glutamate metabolism, having a central role in energy production. Interestingly, butyrate is a metabolite with a known anti-cancer role against inflammation and uncontrolled malignant behaviour of cancer cells, through the acceleration of apoptosis and differentiation.

Overall, species specific variations in pathways and DEGs related to the p53 signaling and apoptosis pathways were identified. Despite limitations such as high variability between replicates and sample size limitations, this study offers possible suggestions and potential future research directions into additional biological mechanisms which helps to suppress cancer in elephants.

Table of contents

Chapter 1: Introduction

1.1 Background

Cancer is one of the most harmful diseases on this planet and unique to all multicellular organisms. Cancer originates by the accumulation of genetic mutations in the DNA of the cell, frequently in the proto-oncogenic or tumour suppressor genes. In the case of proto-oncogenes, these mutations can lead to the overexpression of certain growth factors, while a lack of tumour suppressor proteins can contribute to the failure to regulate and inhibit cell division effectively. In both instances, cells can escape cell cycle control mechanisms and start to grow uncontrollably, spreading quickly for their own benefit, resulting in cancer (Nunney, 2013; Vincze et al., 2022).

If one assumes an equal somatic mutation rate across mammalian cells, it is expected that organisms with a large body size (more cells) and/or a longer life span (more cell divisions) will have a higher risk of getting cancer (Fleming, Creevy and Promislow, 2011; Wirén et al., 2014). However, this is not observed across different mammalian species, where the occurrence of cancer does not seem to increase with an animal's body size and lifespan. Although elephants have 100 times more cells than humans, and whales an even greater 1000 times more cells than humans, no increase in cancer incidence has been reported (Caulin et al., 2015). Intriguingly, mice which have 1000 times fewer cells than humans and demonstrate a considerably higher susceptibility to cancer. This lack of correlation between body size and cancer risk is known as the Peto's paradox (Peto et al., 1975; Vincze et al., 2022). Various explanations have been proposed to explain this paradox, such as the role of pathogen-driven oncogenesis and the higher cancer risk in carnivores, particularly those with low microbiome diversity (Vincze et al., 2022). Evolution seems to have changed parameters of somatic evolution in order to equalise cancer risks across the mammalian kingdom, ignoring the order of magnitude differences in body size or lifespan (Noble, Kaltz and Hochberg, 2015).

Researchers found a distinct expression of the *TP53* gene in proboscideans, comprising the elephants and extinct related species. In particular, 20 copies of the *TP53* were found in the living African elephant (*Loxodonta africana*), which could potentially explain their unique cancer resistance characteristics (Abegglen et al., 2015). Humans possess a single *TP53* gene with two alleles that are frequently mutated throughout an individual's

lifespan. The gene plays a central role in cancer suppression across all species and encodes for the protein p53 (Bartas et al., 2021). This protein is essential in the apoptosis pathway, helping to eliminate damaged cells, and in maintaining genomic stability by its role in the activation of DNA repair proteins and its regulation of proteins such as p21 and Mdm2 (Abegglen et al., 2015). On the one hand, the expression of p53 leads to increased levels of p21, which can induce cell cycle arrest (Abbas and Dutta, 2009). On the other hand, Mdm2 is an inhibitor of p53, and the overexpression of the *MDM2* gene potentially leads to uncontrolled cell proliferation. A mutated *TP53* allele is associated with family cancer syndromes or inherited mutations, such as Li Fraumeni syndrome, which often plays a major role in cancer predisposition in humans (Bartas et al., 2021). While one copy of the *TP53* gene in African elephants is orthologous to *TP53* found in other mammals, the additional 19 *TP53* copies present in elephants are retrogenes and could potentially enhance the p53 response mechanisms. *TP53* retrogene 9 was identified as the most highly expressed retrogene, capable of inducing apoptosis at the mitochondria, independent of transcriptional activity (Preston et al., 2023). In addition, the *TP53* retrogene 12 seems to be transcribed in African elephant fibroblasts and could enhance the p53 signaling pathway and DNA-damage responses (Sulak et al., 2016). Unfortunately, it remains uncertain whether the *TP53* retrogenes in African elephants produce functional proteins or merely serve as non-functional pseudogenes, which makes it hard to understand if they are casually related to the solution of the Peto's paradox (Abegglen et al., 2015; Sulak et al., 2016; Preston et al., 2023). Adding extra copies of the *TP53* gene to mice resulted in them being more protected to cancer but aging faster (Tyner et al., 2002). This accelerated aging process could lead to depletion or reduction of the function of stem cells, which are essential for tissue regeneration and repair. Nonetheless, another mouse model was developed with one or two extra copies of normal *TP53* genes, which are expressed from the same promotor and regulated in the same fashion as the endogenous *TP53*. This resulted in cancer-resistant mice that age at a normal rate (García-Cao et al., 2002).

In addition to the identification of the *TP53* retrogenes, 7-11 additional copies of leukemia inhibitory factors (*LIF*) were found in African elephants, of which most lack regulatory elements and are not expressed (Vazquez et al., 2018). There is one exception: the pro-apoptotic pseudogene *LIF6*. *LIF6* is hardly active in elephant cells

under normal conditions, but when DNA damage occurs, the *TP53* gene steps in to increase the activity of *LIF6*, which in turn leads to enhanced cell death. Moreover, Abbeglen et al., (2015) found an increased p21 expression in elephants after inducing DNA damage with radiation exposure, which is presumably activated by p53.

Elephants may have evolved a more advanced regulation of apoptosis and DNA damage response mechanisms, presumably controlled by the p53 signaling pathway and enhanced by extra copies of *TP53* retrogenes. Compared to human cells, elephants showed an increased sensitivity to apoptosis induced by DNA damage, which indicated a stronger response to genotoxic stress and a lower threshold that triggers p53 mediated cell death (Abegglen et al., 2015). This can be seen as a possible evolutionary adaptation to prevent cancer through the efficient removal of mutated cells at a pre-cancerous stage. However, many hypotheses exist regarding the genes and mechanisms that may have contributed to the evolution of enhanced cancer resistance in African elephants. In addition, the sophisticated balance between preventing cancer through apoptosis and avoiding stem cell exhaustion or premature aging is an element among many others to unravel. Hence, this study aims to look for mechanisms and important genes that might be involved in this tight balancing act and effective cancer suppression of African elephants. Additionally, this study aims to substantiate the previously explored role of the *TP53* retrogenes, *LIF6* pseudogene, and other key genes identified in African elephants.

1.2 Report structure

The study aims to further explain the mechanisms underlying the effective cancer resistance observed in African elephants by performing bioinformatics analysis. Initially, bulk RNA sequencing (RNAseq) was conducted on elephant and human peripheral blood mononuclear cells that had been previously exposed to 2 Gray Co60 γ-ray irradiation (this part of the study was previously performed at iThemba LABS in 2021-2022). Differential expressed genes (DEGs) were identified based on the RNAseq results through comparative differential expression analysis. Subsequently, pathway enrichment analysis (PEA) was performed on both the elephant and human DEGs, revealing distinct differences in gene ontology (GO) terms and Kyoto Encyclopedia of genes and genomes (KEGG) pathways between both species, which will be further discussed.

Chapter 2: Methodology

2.1 Introduction

In this section, the sample selection criteria are clarified, followed by a description of the sequencing techniques and data processing methods. Next, the approaches to identify differential expressed genes (DEGs) and to conduct pathway enrichment analysis (PEA) are elaborated in detail.

2.2 Sample collection and irradiation

Peripheral blood samples were collected from three adult human males (*Homo sapiens*) and three adult elephant males (*Loxodonta africana*), between the age of 24-40 years old. Elephant samples were collected at Botlierskop Private Game Reserve in 2022 (before the start of this internship), in the Western Cape, South Africa by a research team affiliated with NRF-iThemba LABS (Cape Town, South Africa). Ethical approval was obtained from the Animal Research Ethics Committee (AREC) of the University of the Western Cape (UWC) (South Africa), with reference AR21/6/4. In addition, landowner permission letters were obtained prior to elephant sample collection. Samples were obtained from elephants who underwent a scheduled veterinary intervention (vaccination or replacement of tracking collar) requiring sedation via a single intramuscular injection of Thianil® (Wildlife Pharmaceuticals Pty LTD., Nelspruit, South Africa). The sample collection piggybacked on these interventions, which limited unnecessary sedation risks to the elephants for the sole purpose of this research project. Prior to the collection of the human peripheral blood samples informed consent was obtained. These experiments were approved by the South African Human Sciences Research Council Research Ethics Committee (reference number: HSRC REC 3/23/10/19).

Using a density gradient cell separation medium (Histopaque-1077, Merck, Modderfontein, Johannesburg, South Africa) with a 2 (diluted blood) over 1 (gradient medium) ratio, peripheral blood mononuclear cells (PMBCs) were separated from whole blood. After centrifuging the layered solution at 900 ×g for 20 min, the plasma-density medium interface was transferred to a fresh tube and rinsed twice with phosphate-buffered saline (PBS), followed by centrifugation at 900 ×g for 10 min. After the separation, cells were diluted in RPMI medium (Lonza, Walkersville, MD, USA), which

was supplemented with 10% fetal calf serum (Gibco, Dun Laoghaire, Dublin, Ireland), and 1% penicillin and streptomycin (Lonza, Walkersville, MD, USA) in a total volume of 2 mL. The tubes were exposed to γ-ray irradiation at doses of 0 Gy (which served as the control) and 2 Gy using a Cobalt-60 Theratron 780 teletherapy unit. RNA was extracted from these samples at four hours post-irradiation using the RNeasy Mini Kit (Qiagen). Afterwards, the RNA samples were preserved in a frozen state at -80°C for further analysis. Once all samples were collected, shipment on dry ice was arranged through Sequentia Biotech to Macrogen (South Korea) for RNA sequencing. Upon arrival, the RNA integrity number (RIN) was measured using an Agilent Technologies 2100 Bioanalyser as a quality control check for the RNA integrity. These values are represented in **Table 1.** Although not all RIN values were greater than or equal to 7, the decision was made to proceed with the library preparation step. Eventually, bulk RNA sequencing was performed on both human (n = 6) and elephant (n = 6) samples belonging to 2 experimental groups, 0 Gy and 2 Gy.

Table 1. *RIN values of the Homo sapiens (n=6) and Loxodonta africana (n=6).*

#	SAMPLE NAME	RIN VALUE
HOMO SAPIENS		
1	D1-RF_0Gy_4H	7.1
2	D1-RF_2Gy_4H	6.4
3	D2-PB_0Gy_4H	7.9
4	D2-PB_2Gy_4H	7.9
5	D3-R_0Gy_4H	7.3
6	D3-R_2Gy_4H	7.2
LOXODONTA AFRICANA		
1	E1-14_0Gy_4H	6.2
2	E1-14_2Gy_4H	6.6
3	E2-15_0Gy_4H	5.6
4	E2-15_2Gy_4H	6.0
5	E3-16_0Gy_4H	7.3
6	E3-16_2Gy_4H	7.3

2.3 Sequencing information

Bulk RNA sequencing was performed by Macrogen, using the Illumina NovaSeq 6000 system sequencing technique. Each sample was processed to generate paired-end reads

of 150 base pairs in length, with a depth of 30 million reads per sample. The TruSeq RNA library prep kit was used for RNA-seq library preparation.

2.4 Read pre-processing

The quality of the sequencing data was first evaluated using FASTQC software to ensure the reads met quality standards (Andrews, S., 2010). Next, the software TRIMMOMATIC was used to remove adapters and low-quality bases from the reads (Bolger, Lohse and Usadel, 2014), applying criteria of a minimum read length of 35 base pairs and a base quality threshold score of 25.

2.5 Read annotation

After read processing and expression quantification of the reads, read annotation was performed. This step assigns identified reads with specific genes, which is essential for subsequent pathway analysis and was done using R version 4.3.1 (RStudio Team, 2020).

2.5.1 Homo sapiens

The high-quality reads were aligned to the human reference genome (GRCh38 v108) using the STAR aligner software (version 2.7.9a) (Dobin et al., 2013). FeatureCounts (version 2.0.0) was used to quantify gene expression values as raw fragment counts (Liao, Smyth and Shi, 2014).

2.5.2 Loxodonta africana

First, an index from the transcriptome data was build. Subsequently, transcript quantification was conducted on both forward and reverse reads using Kallisto, a read aligner classified as a pseudo aligner (Bray et al., 2016).

2.6 Differential expression analysis

Statistical analyses of the RNA sequencing data were performed with the NOISeq package available through Bioconductor using R version 4.2.2 (Tarazona et al., 2011; 2015; RStudio Team, 2020). NOISeq is commonly applied in the differential expression analysis of replicates that have a relatively high variability. It has a slightly higher rate of the identification of false positives DEGs compared to other methods, but it can identify broader range of DEGs under such variable conditions. The differential expression analysis included the comparison between samples irradiated with 2 Gy and the control

group (0 Gy) from both species. Finally, DEGs were considered significantly differentially expressed when the false discovery rate (FDR) fell below 0.05.

2.6.1 Sample exclusion

The overall quality of the experiment was assessed based on the similarity between replicates. To further subtract noise from the data, the function ARSyNseq was used which applies an ANOVA analysis (Nueda, Ferrer and Conesa, 2012). Subsequently, the HTSFilter package was chosen to remove lowly expressed genes within the samples (Rau et al., 2013).

2.6.2 Volcano plots

To visualise the differential expressed genes the EnhancedVolcano package from Bioconductor was used (Blighe, Rana and Lewis, 2023). This package optimises label placement and avoids cluttering.

2.6.3 Venn diagrams

The package VennDiagram version 1.7.3 was used to create Venn diagrams used in the study (Chen and Boutros, 2011).

2.7 Gene annotation

Following differential expression analysis, gene annotation was performed to assign gene names and descriptions to the identified DEGs. OrgDb annotation packages were required for subsequent pathway analysis with clusterProfiler.

2.7.1 Homo sapiens

For human gene annotation, the org.Hs.eg.db package from Bioconductor was used (Carlson, 2019). This is a database for human genes that provides a mapping between the gene identifiers used in the sequencing data (Ensembl IDs) and the actual (widely known) gene names, along with functional information about each gene.

2.7.2 Loxodonta africana

Elephant reads were annotated using unique elephant Ensembl IDs. Since there was no complete OrgDb annotation package available for elephants, the elephant Ensembl gene IDs were converted to their corresponding human ortholog. This conversion was done using the EnsDb.Hsapiens.v79 package from Bioconductor (Rainer, 2017). Subsequently, the converted human Ensembl IDs were functionally annotated with the same

org.Hs.eg.db package from Bioconductor and further used for functional annotation and PEA.

2.8 Functional annotation

After the read annotation process, gene functional annotations were assigned to the genes using two accurate and frequently used databases: Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) (Ashburner et al., 2000; Kanehisa et al., 2023). This approach ensured a broad understanding of the functional roles of the identified genes within biological mechanisms. Gene annotation (see Section **2.7**), functional annotation and pathway enrichment analysis (see Section **2.9**) are seamlessly integrated within the same function of Clusterprofiler. The steps to perform pathway enrichment analysis are explained in the next section of this methodology.

2.8.1 Gene ontology

The GO database serves as a computational representation of categorising genes based on their functions across diverse organisms (Thomas, 2017). GO is categorised into three aspects for describing gene functions: molecular function (MF), cellular component (CC), and biological process (BP) (Ashburner et al., 2000; Wu et al., 2021; The Gene Ontology Consortium et al., 2023). Each GO annotation establishes a connection between a particular gene and a corresponding GO term.

2.8.2 Kyoto Encyclopedia of genes and genomes

KEGG covers a wide range of resources. It merely serves as a reference knowledge base, which is designed to give deep sight information in the biological interpretation of identified genes (Kanehisa and Goto, 2000; Kanehisa, 2019; Kanehisa et al., 2023). Moreover, it provides insights into the pathways and complex interactions of genes within biological systems.

2.9 Pathway enrichment analysis

To identify significant associations based on the differential expressed genes in this study, pathway enrichment analysis (PEA) (also called functional enrichment analysis) was performed. This tool can determine signaling or biological pathways, based on statistical methods including accurate and current gene biological annotations (Wijesooriya et al., 2022). As functional annotation and pathway enrichment analysis happen within the same function, the R package clusterProfiler 4.0 was used for the

implementation functional enrichment tools (Yu et al., 2012; Wu et al., 2021). The two functional enrichment tools used in this study are over-representation analysis (ORA) and functional class scoring (FCS). For both tools, pathways or terms were considered significant when the outcomes of the adjusted p-value (p-value underwent multiple comparison testing) and the q-value (adjusted p-values using an optimised FDR approach) were lower than 0.05, considering the correction for multiple comparisons.

2.9.1 Over-representation analysis

In ORA, DEGs that are significantly differentially expressed and/or meet specified fold change criteria are examined against a predefined pathway or background gene list (Wu et al., 2021). This tool uses statistical tests to determine whether the count of DEGs in the specific gene set is higher than would be expected by chance, compared to the background gene list. In ORA, upregulated and downregulated are examined separately.

2.9.2 Gene set enrichment analysis

GSEA in the most common FCS tool. GSEA uses gene-level statistics or log2 fold changes obtained from the results of differential expression analysis (Wu et al., 2021). This tool assesses whether background gene sets associated with specific biological pathways are enriched among the positive or negative log2 fold changes of the generated gene sets associated (ranked gene list). GSEA includes all identified DEGs without applying filters, analysing both upregulated and downregulated DEGs together. Compared to ORA, GSEA takes advantage of the log2 fold change, so the results are more informative and reliable.

2.9.3 Pathway visualisation

For the visualisation of the biological pathways, the Pathview package from Bioconductor within the R environment was used (Luo and Brouwer, 2013). This package is based on fully accessible biological pathway-based data from KEGG and can represent the pathways in a directly automated and readable way.

2.10 Code availability

The codes used for the analysis in this study are accessible through the following link: https://bitbucket.org/gsi-elephantproject/gsistudent/src/main/. Researchers and interested parties can explore the scripts implemented in the methodology to ensure transparency and reproducibility of the findings.

Chapter 3: Results

3.1 Introduction

The results are presented in three main sections. First, section 3.2 provides insights into sample inclusion and the quantification of the DEGs. Next, the identification of DEGs is shown is section 3.3. Hence, the results of the pathway enrichment analysis are represented in section 3.4, including both ORA and GSEA.

3.2 Sample inclusion and the quantification of differential expressed genes

Out of the 12 samples initially included in this study, nine demonstrated an acceptable level of variability, making them eligible for inclusion in the final results. The remaining three samples showed a too strong variability among the replicates. Consequently, it was decided to exclude the following three samples for further analysis: D1-02y (human), E2-0Gy (elephant), E2-2Gy (elephant). **Table 2** provides an overview of the samples included in the study and presents the differential gene expression numbers in humans and elephants under control conditions (0 Gy) and exposure to radiation at 2 Gy.

Table 2. *Comparison of number of differentially expressed genes in Homo sapiens and Loxodonta africana under control and 2 Gy radiation exposure. The included samples in the study are also represented.*

Samples included	Study group	Total DEGs	Significant DEGs (FDR ≤ 0.05)	Upregulated (+) / downregulated (-)
Homo sapiens				
D1, D2, D3	0 Gy	12320	7399	3871 (+)
D2, D3	2 Gy			3528 (-)
Loxodonta africana				
E1, E3	0 Gy	14039	2869	1968 (+)
E1, E3	2 Gy			901 (-)

A total of 12320 genes were identified as differential expressed in the human samples. 7388 DEGs emerged as significant (FDR ≤ 0.05), presenting 3871 upregulated genes and 3538 downregulated genes in response to 2 Gy radiation. On the contrary, 14039 genes were differentially expressed in the elephant samples, out of which 2869 were identified as significant (FDR ≤ 0.05). Under 2 Gy conditions, there were 1968 genes that were

upregulated, while 901 genes showed downregulation. This indicates that 60.05% of the DEGs identified in humans, were considered significant, while in elephants the percentage is 20.44%.

3.3 Identification of differential expressed genes

This section presents significant DEGs found in humans and elephants. The most highly expressed genes are highlighted to give insights into their potential roles in radiation responses. These results of the analysis provide valuable insights for subsequent pathway enrichment analyses, specified in the following section.

3.3.1 An overview of potentially interesting differential expressed genes

One frequently used approach to delve deeper into the identification and visualisation of DEGs involves the comparison of the FDR with the log fold-change, is a volcano plot representation (**Figure 1**). It represents the DEGs identified in humans and elephants and displays gene names which have a significant difference in expression levels between the irradiated (2 Gy) and control condition (0 Gy). For the analyses presented in **Figure 1**, a threshold of -1.5 and 1.5 for the base-2 log fold-change (log2FC) and FDR threshold of 0.05 was applied.

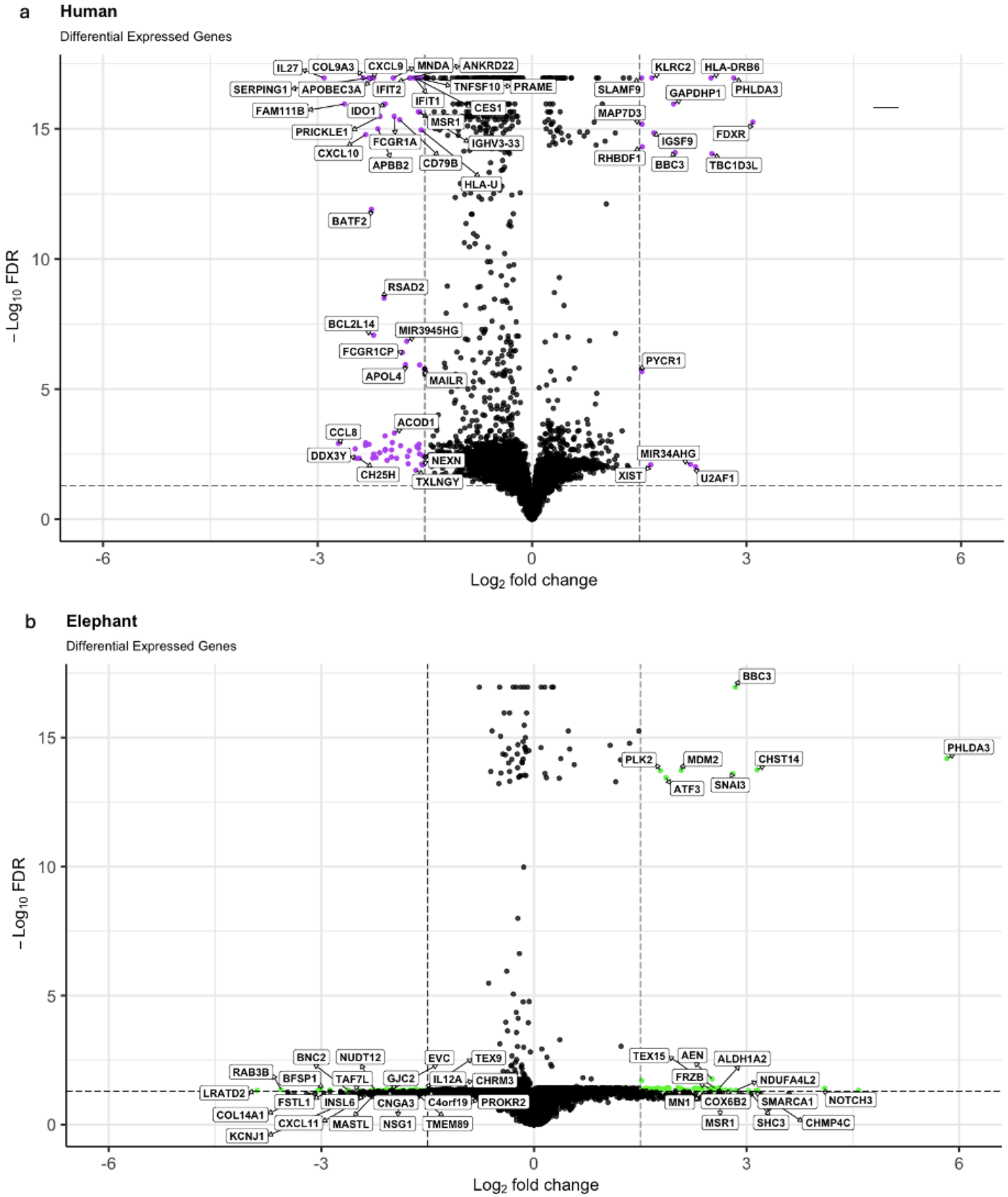
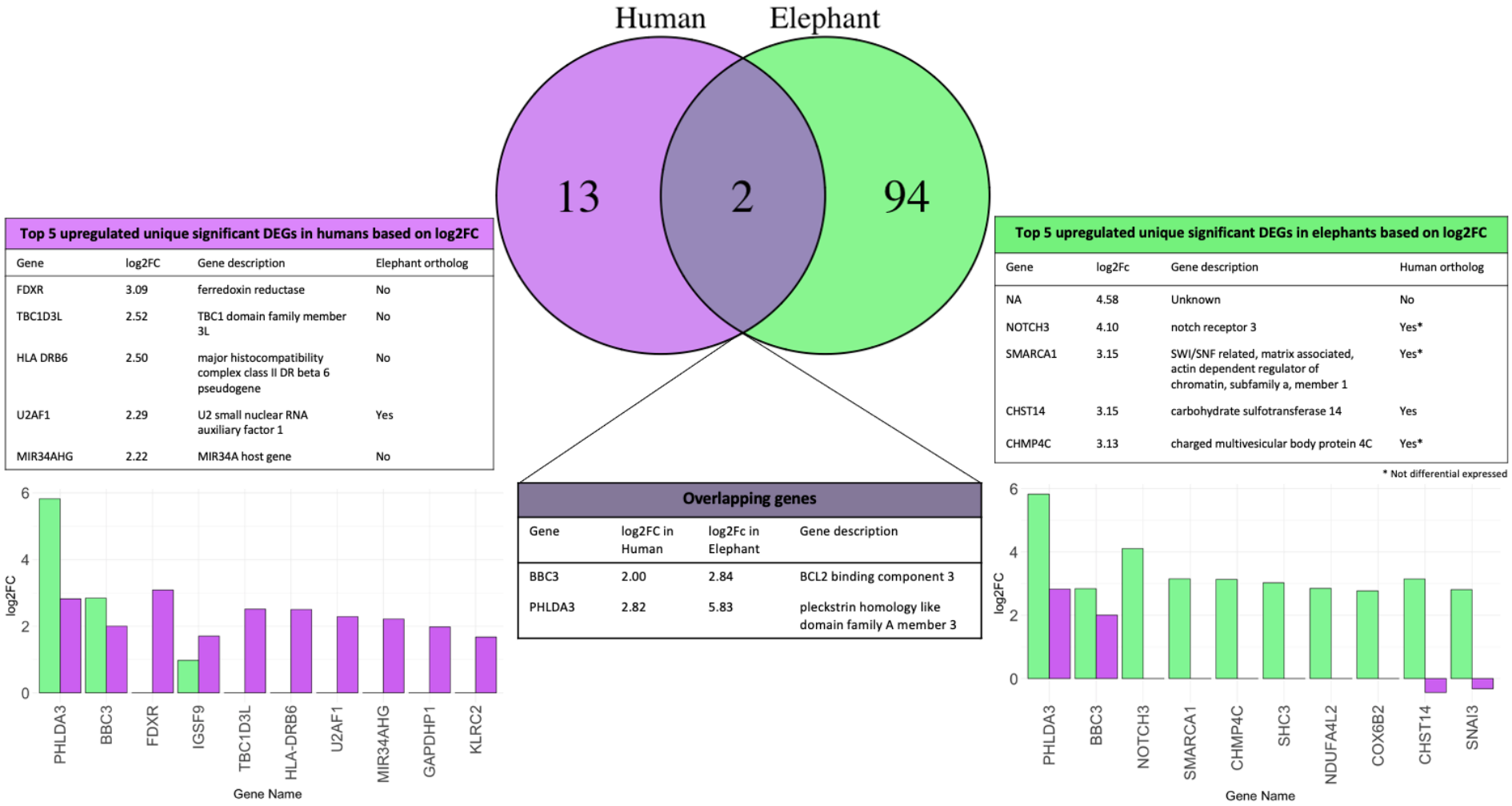


Figure 1. *Volcano plot of differential expressed genes in humans (in purple - a) and elephants (in green - b), with default log fold-change (x-axis) threshold of -1.5 and 1.5 and FDR (y-axis) threshold of 0.05. Single genes are depicted as dots. Genes exceeding the default threshold are labelled with their gene name.*

The gene expression levels of 79 human DEGs were greater than 1.5-fold change in the 2Gy conditions relative to the control condition (**Figure 1.a**). Respectively, 15 DEGs were upregulated, and 64 DEGs were downregulated. Seventy-eight out of 79 DEGs fall within the log2FC range of -3 to 3, except for the *FDXR* gene which shows a high expression with a log2FC > 3.

The DEGs identified in elephants are represented in **Figure 1.b**. A total of 144 DEGs showed a log fold-change exceeding 1.5 in response to the 2 Gy conditions when compared to the control conditions. Of the 144 DEGs, 92 DEGs were upregulated, and 52 were downregulated. The majority of the DEGs are less significant compared to humans, resulting in most DEGs close to the FDR threshold of 0.05. However, the DEGs are distributed over a broader log2FC scale compared to the human samples, with a notable number of genes above or below the lof2FC threshold of -3 to 3 respectively. While the DEGs in the Volcano plot of the human samples is well distributed in both horizontal and vertical direction, much more DEGs lay closer to zero in the vertical direction for the elephant samples, indicating weaker significant changes in gene expression compared to the control.

In addition to providing an overview of potentially interesting genes, **Figure 2** and **Figure 3** show two Venn diagrams, emphasising the overlap of up- and downregulated DEGs between humans and elephants. Similar to the Volcano plots, the log2FC cut off was set at -1.5 and 1.5, with an FDR lower than 0.05. The comparison shows two overlapping DEGs (*BBC3* and *PHLDA3*) in the upregulated gene figure (**Figure 2**). When looking for differences in the log2FC expression values for both genes, it is evident that both genes exhibit relatively higher upregulation in elephants compared to humans. Furthermore, in the comparison of downregulated DEGs (**Figure 3**), a single overlapping gene was identified: the *CXCL11* gene with a minor difference in expression level between the two species (human log2FC = 2.33, elephant log2FC = 2.24).



Top 5 upregulated unique significant DEGs in humans based on log2FC

Gene	log2FC	Gene description	Elephant ortholog
FDXR	3.09	ferredoxin reductase	No
TBC1D3L	2.52	TBC1 domain family member 3L	No
HLA DRB6	2.50	major histocompatibility complex class II DR beta 6 pseudogene	No
U2AF1	2.29	U2 small nuclear RNA auxiliary factor 1	Yes
MIR34AHG	2.22	MIR34A host gene	No

Overlapping genes

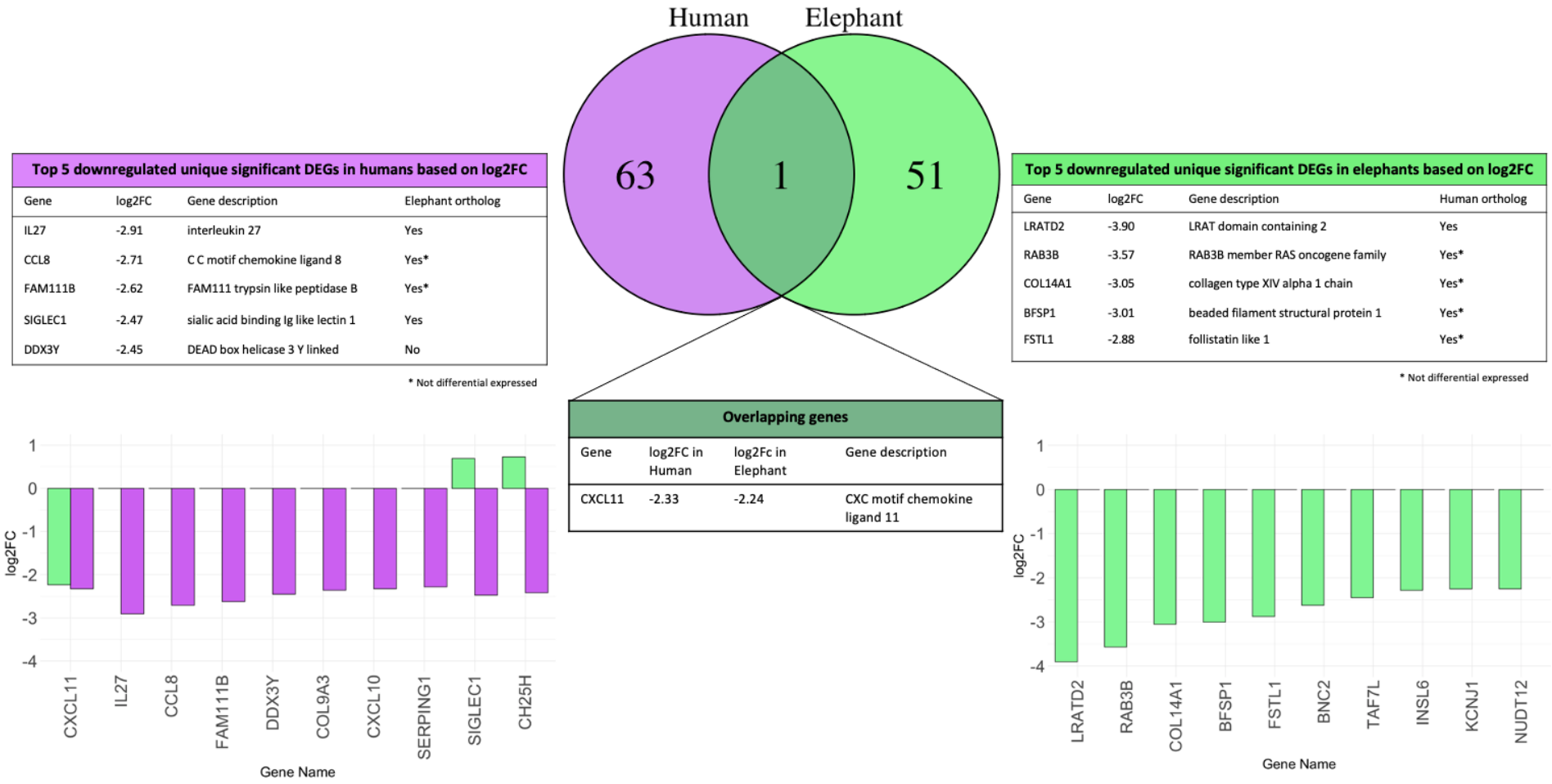
Gene	log2FC in Human	log2Fc in Elephant	Gene description
BBC3	2.00	2.84	BCL2 binding component 3
PHLDA3	2.82	5.83	pleckstrin homology like domain family A member 3

Top 5 upregulated unique significant DEGs in elephants based on log2FC

Gene	log2Fc	Gene description	Human ortholog
NA	4.58	Unknown	No
NOTCH3	4.10	notch receptor 3	Yes*
SMARCA1	3.15	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 1	Yes*
CHST14	3.15	carbohydrate sulfotransferase 14	Yes
CHMP4C	3.13	charged multivesicular body protein 4C	Yes*

* Not differential expressed

Figure 2. *Venn diagrams representing overlapping upregulated DEGs between humans and elephants, filtered for log 2-fold change threshold of -1.5 and 1.5, FDR lower than 0.05, and p-value lower than 0.05. The tables present the top 5 upregulated unique significant DEGs identified in both species, determined by the log 2-fold change. The bar plots show the variations in expression levels among the top 10 DEGs expressed in both species.*



Top 5 downregulated unique significant DEGs in humans based on log2FC

Gene	log2FC	Gene description	Elephant ortholog
IL27	-2.91	interleukin 27	Yes
CCL8	-2.71	C C motif chemokine ligand 8	Yes*
FAM111B	-2.62	FAM111 trypsin like peptidase B	Yes*
SIGLEC1	-2.47	sialic acid binding Ig like lectin 1	Yes
DDX3Y	-2.45	DEAD box helicase 3 Y linked	No

* Not differential expressed

Overlapping genes

Gene	log2FC in Human	log2Fc in Elephant	Gene description
CXCL11	-2.33	-2.24	CXC motif chemokine ligand 11

Top 5 downregulated unique significant DEGs in elephants based on log2FC

Gene	log2FC	Gene description	Human ortholog
LRATD2	-3.90	LRAT domain containing 2	Yes
RAB3B	-3.57	RAB3B member RAS oncogene family	Yes*
COL14A1	-3.05	collagen type XIV alpha 1 chain	Yes*
BFSP1	-3.01	beaded filament structural protein 1	Yes*
FSTL1	-2.88	follistatin like 1	Yes*

* Not differential expressed

Figure 3. *Venn diagrams representing overlapping downregulated DEGs between humans and elephants, filtered for log 2-fold change threshold of -1.5 and 1.5, FDR lower than 0.05, and p-value lower than 0.05. The tables present the top 5 downregulated unique significant DEGs identified in both species, determined by the log 2-fold change. The bar plots show the variations in expression levels among the top 10 DEGs expressed in both species.*

Regarding the upregulated DEGs in elephants (**Figure 2**), the top five unique DEGs which were most upregulated by radiation response included a gene with an unassigned name and function, *NOTCH3*, *SMARCA1*, *CHST14* and *CHMP4C*, with the unassigned gene having the second highest expression (log2FC = 4.5). The *PHLDA3* gene stood out as the most significantly upregulated gene in elephants (log2 fold-change = 5.83). Remarkably, all the identified genes of the top five unique elephant DEGs had human orthologs, but only the gene *CHST14* out of these 4 was differentially expressed (human log2FC = -0.44, elephant log2FC = 3.15). Considering the bar plots, six out of 10 DEGs were solely differentially expressed in elephants, except for the two overlapping genes, and for *CHST14* and *SNAI3*, where human gene expression was downregulated. Similarly, for human upregulated DEGs, four out of the top five unique genes lacked an elephant ortholog, indicating non-existence or are unidentified in elephants and have unique expression in humans. Among the top 10 showed in the bar plot, only the *IGSF9* gene exhibited similar expression in elephants, along with the two overlapping genes *BBC3* and *PHLDA3*.

Focussing on downregulated genes (**Figure 3**), the top five uniquely downregulated DEGs in elephants all had human orthologs. Among them, only *LRATD2* showed differential expression in humans; however, it did not reach statistical significance based on the chosen threshold (FDR ≤ 0.05). Within the top 10 downregulated genes in elephants, none were significantly expressed in humans. In humans, four of the top five uniquely downregulated DEGs had corresponding elephant orthologs, with two of them, *IL27* and *SIGLEC1*, also being differentially expressed in elephants. Among these, *SIGLEC1* is particularly significant. Examining the top 10 downregulated DEGs in humans, seven were exclusively expressed in humans. Notably, *SIGLEC1* and *CH25H* were significantly downregulated in humans but upregulated in elephants.

3.4 Pathway enrichment analysis

The pathway analysis involved ORA and GSEA using the GO and KEGG databases. In ORA, only significant genes (FDR ≤ 0.05) were taken into consideration for the analysis. In contrast, GSEA includes all DEGs and is based on log2FC values, providing a higher potential for identifying potential functional pathways.

3.4.1 Over-representation analysis

3.4.1.1 Over-representation analysis with the GO database

In humans, 15 pathway showed significant over-representation for both upregulated and downregulated genes (**Figure 4**). Pathways consisting of upregulated genes were associated with ribonucleoprotein and cytosolic ribosomal subunits. Additionally, the upregulated pathways were distributed under all three GO term categories (Biological Processes – BP, Cellular Components - CC and Molecular Functions - MF), while pathways associated with downregulated genes were linked to immune defence responses under the GO category BP. Gene ratios ranged from 0.00 to 0.10 for pathways with upregulated genes and 0.00 to 0.20 for pathways with downregulated genes. In elephants, no GO terms were observed for pathways with upregulated genes, and pathways with downregulated genes were solely related to immune responses (**Figure 5**).

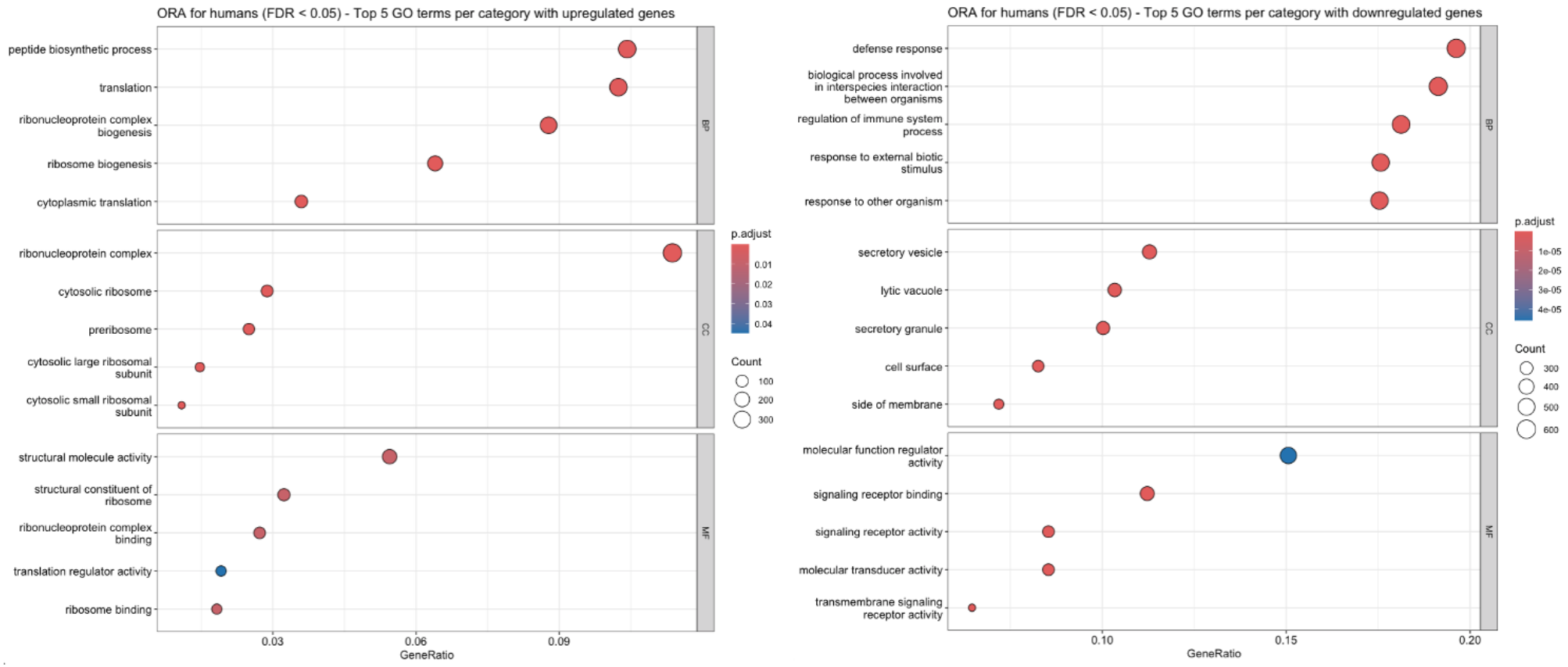


Figure 4. *Gene ontology (GO) - over-representation analysis (ORA). Pathway enrichment results in humans for differentially expressed genes with a false discovery rate (FDR) ≤ 0.05 and log2 fold change (log2FC) </> 0. The size of dots represents the number of genes in each pathway, while the colour of the dots corresponds to the adjusted p-value. The gene ratio of core enrichment genes to pathway genes provides insights into the specific impact of each pathway in the experimental conditions.*

No GO terms found for upregulated genes (q-value < 0.05)

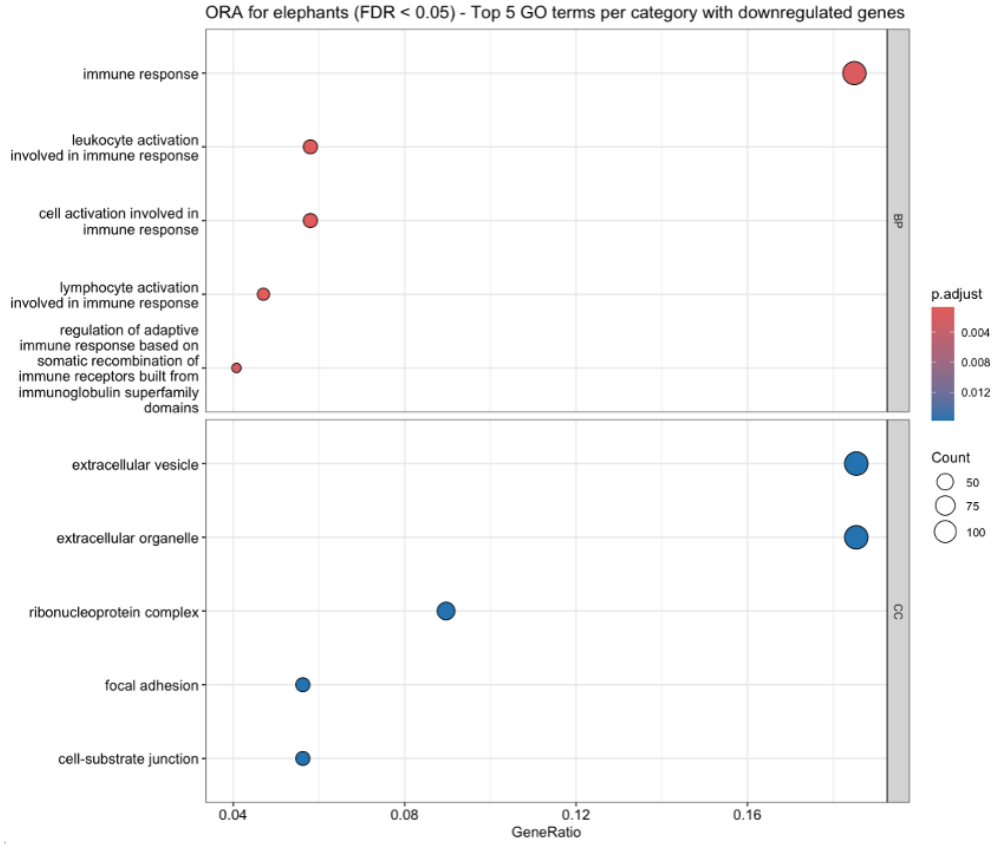


Figure 5. *Gene ontology (GO) - over-representation analysis (ORA). Pathway enrichment results in elephants for differentially expressed genes with a false discovery rate (FDR) ≤ 0.05 and log2 fold change (log2FC) </> 0. The size of dots represents the number of genes in each pathway, while the colour of the dots corresponds to the adjusted p-value. The gene ratio of core enrichment genes to pathway genes provides insights into the specific impact of each pathway in the experimental conditions.*

3.4.1.2 Over-representation analysis with the KEGG database

The ORA analysis based on the KEGG database identified two significant pathways with upregulated genes in humans related to the ribosome and several pathways associated with downregulated genes, notably the cytokine-cytokine receptor interaction (**Supplementary figure 1**). Despite the low gene ratios of these pathways, these pathways were still considered significant. No over-representation of KEGG pathways was found in elephants (**Supplementary figure 2**).

3.4.2 Gene set enrichment analysis

3.4.2.1 Gene set enrichment analysis with the GO database

As represented in **Supplementary figure 3**, the most activated biological processes in GO terms were associated with ribosomal large and small subunit biogenesis, with notably high gene ratios. In addition, the top six suppressed GO terms were linked to positive regulation and production of (beta) interleukin-1, with gene ratios ranging from 0.60 to 0.75. All identified GO terms in humans were significant, although the gene counts were lower compared to ORA. Nonetheless, no significant GO terms were found in elephants (**Supplementary figure 4**).

3.4.2.2 Gene set enrichment analysis with the KEGG database

In humans, 17 significant pathways were identified, including activated pathways related to ribosomal functions and suppressed pathways related to infections and diseases, implying a correlation with immune response related functions (**Supplementary figure 5**). Remarkably, the p53 signaling pathway was significantly activated, suggesting a potential link to DNA damage response in humans. One significant pathway was found in elephants, namely the butanoate metabolism pathway with an adjusted p-value of 0.023 and gene ratio of 0.235 (**Supplementary figure 6**).

3.4.3 Pathview based on KEGG database

An alternative approach to get more insight into KEGG pathways involves the use of the visualisation tool Pathview. This tool visualises the pathways of interest, presenting the respective log2FC values for the included DEGs. In **Figure 6** the p53 signalling pathway is represented for both species. In humans (**Figure 6.a**), *MDM2* (log2FC = 1.02) and *TP53* (log2FC = 0.23) are both upregulated, where *MDM2* induces the negative feedback loop to reduce p53 protein activity. The high upregulation of p48 and *GADD45A* (log2FC = 0.79) indicate roles in DNA repair and damage prevention, while the downregulation of

rrM2B (log2FC = -0.16), which codes for the p53R2 protein may weaken these DNA repair mechanisms. Moreover, the elevated levels of *BBC3* (log2FC = 2.00), coding for the PUMA protein, and *BAX* (log2FC = 0.54) will enhance a higher apoptosis rate. The downregulation of *cdc2* (log2FC = -0.11) and *ZNF385A* (log2FC = -0.49) implies a p53-dependent cell cycle arrest upon DNA damage. Conversely, elephant cells show no *TP53* expression at this four-hour time point, but a unique upregulation of *CDKN1A* (log2FC = 1.07), coding for the cyclin-dependent kinase inhibitor p21 protein, which promotes cell cycle arrest in response to DNA damage. There is a strong upregulation of *BAX* (log2FC = 1.22) and *BBC3* (log2FC = 2.85), which are key mediators of p53-induced apoptosis. The upregulation of pro-apoptotic genes such as *BBC3*, which encodes a direct activator protein (PUMA) which induces mitochondrial outer membrane permeabilization and apoptosis, as well as *ZMAT3*, also known as *PAG608* (log2FC = 1.00) and a key splicing regulator in the TP53 pathway, *SENS1* coding for sestrin (log2FC = 0.06) and *CCNG1* coding for cyclin G (log2FC = 1.23), further support the fast and effective apoptotic response in elephants cells upon radiation-induced DNA damage. The upregulation of *MDM2* (log2FC = 2.08) and *CDKN1A* in elephants could indicate a balance between cell cycle arrest and apoptosis.

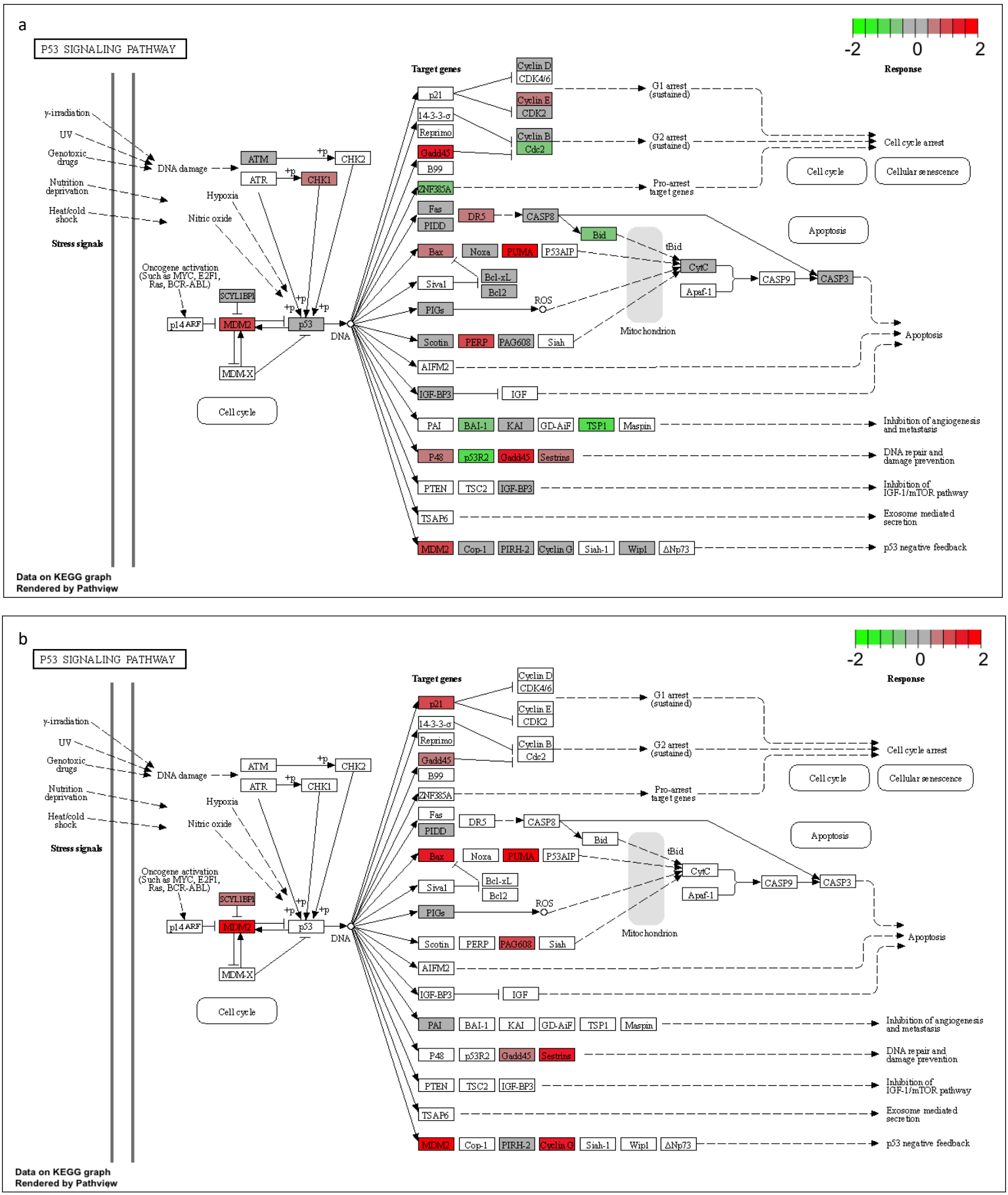


Figure 6. *p53 signalling pathway for humans (a) and for elephants (b) based on KEGG database and generated by Pathview. Upregulated proteins in red, downregulated proteins in green and moderately expressed protein in gray.*

Looking at the apoptosis pathway different patterns are seen between humans (**Supplementary figure 7.a**) and elephants (**Supplementary figure 7.b**). In humans, many downregulated DEGs are pro-apoptotic genes within the apoptosis pathway. For example, *TNF* (log2FC = -0.71), *TNFSF10* (log2FC = -1.61), *TRADD* (log2FC = -0.54) are all playing a crucial role in the TNF-induced apoptosis pathway. The downregulation of IL3RA (log2FC = -0.49) may impact cell survival. Additionally, the downregulation of *CTSB* (log2FC = -0.56), *CASP10* (log2FC = -0.41), and *CASP7* (log2FC = -0.49) refer to a reduced apoptotic response. Conversely, the upregulation of *BBC3* (PUMA), *GADD45* and *BAX*, and *cdc2* enhance the pro-apoptotic response. Elephants have a different pattern within the apoptosis pathway. Fewer DEGs were expressed compared to humans, and the DEGs show higher expression levels. The increased expression of *TRKA* (log2FC = 0.91), *PI3K* (log2FC = 0.22) and PUMA suggest the activation of apoptosis induction. *BAX* and *GADD45* further contribute to the pro-apoptotic response. Additionally, the downregulation of *FLIP* (log2FC = -0.47) and *LAPXIAP* (log2FC = -0.02) indicate a suppression of pro-survival mechanisms.

Chapter 4: Discussion

4.1 Introduction

This lack of correlation between body size, longevity and cancer risk across mammalian species was already described in the seventies as Peto's paradox. Especially African elephants, with a large body size and long lifespan, have developed protective mechanisms against cancer, and therefore do have one of the lowest cancer incidence rates across the animal kingdom (Seluanov et al., 2018). The first evidence for this evolutionary selection of cancer suppression mechanisms was discovered by two independent research groups, who identified an expanded number of *TP53* gene copies in the elephant genome, followed by a study reporting 7 to 11 extra copies of LIF (Abegglen et al., 2015; Vazquez et al., 2018). Besides looking into these protective tumour suppressor genes and their up- and downstream signalling genes, this study also aimed to identify pathway or genes which could play a crucial role in the balancing act between cancer prevention and avoiding stem cell exhaustion and premature ageing in African elephants.

4.2 Differential expressed genes

Following differential expression analysis, a considerable quantity of DEGs were identified in both elephants and humans. The identification of the high numbers of genes in human (12320) and elephants (14039) raises the questions about their biological significance, how these findings align with other studies and the fraction related proportionally to the total number of genes in each species. The suggested protein-coding gene count in the human genome is approximately 19000 (Amaral et al., 2023), making the 12320 identified DEGs a significant fraction (approximately 65.00%) of the estimated total. Although the total gene count in African elephants is not thoroughly studied, the total number of protein-coding genes in African elephants is estimated to be 22142 (Shi et al., 2024), and 14039 of the 22142 represents a comparable fraction of approximately 63.00%. The high number of DEGs in both species could be related to the high radiosensitivity of cell type under investigation, namely peripheral blood mononuclear cells (Fang et al., 2021). Another reason could be the increased sensitivity of NOISeq in detecting DEGs compared to more common identifier tools (Tarazona et al., 2011; 2015). Despite this, elephants have a low percentage (20.44%) of significant DEGs over the total identified DEGs, in contrast to humans

(60.05%). This substantial difference demonstrates the need for careful consideration regarding sample variability and the impact of the relatively low number of donors, together with a critical examination of the statistical thresholds used in this study.

The identification of DEGs by the use of Volcano plots (**Figure 1**) and Venn diagrams (**Figure 2** and **3**) provided insights into the differences of gene expression responses after radiation exposure. The wide dispersion of log2FC values and lower significance of elephant DEGs could indicate a higher level of difference in gene expressions between to two treatment groups (control vs radiation-exposed). Upregulated elephant DEGs, including the overlapping *BBC3* and *PHLDA3* which play a pro-apoptotic role (Han et al., 2001; Kawase et al., 2009), show higher log2FC values compared to their human counterparts. While it is important to be cautious in comparing these values without considering significance, this observation could indicate a more substantial upregulation of DNA damage response in elephants. A similar pattern is observed in downregulated DEGs, where the top two in elephants display relatively higher log2FC values. In addition, the overlapping downregulation of *CXCL11* could suggest common immune response aspects in both species. This is unsurprising given that peripheral blood mononuclear cells are immune cells (Kleiveland, 2015). Moreover, chemokines are critical in directing immune cell migration, a process which is necessary for an effective anti-tumour immune response. Downregulation of the *CXCL11* gene may reduce immune cell migration to cells with DNA damage, affecting the genomic stability (Tokunaga et al., 2018; Ozga, Chow and Luster, 2021). The top 10 up/downregulated genes present distinct expression values between the two species what could indicate specie specific responses to DNA damage.

Particularly in the humans, a majority of the highest upregulated genes are related to immune responses, such as *HLA DRB6*, *IL27*, *CCL8* and *SIGLEC1 (McCarthy, 2011; Yoshida and Hunter, 2015; Zheng et al., 2015; Butler-Laporte et al., 2020)*. The *FDXR* and *FAM111B* gene identified in humans are directly associated with DNA damage response (O'Brien et al., 2018; Arowolo et al., 2022). In contrast, elephants have only one gene directly related to immune responses, which is *FSTL1* (Mattiotti et al., 2018). Moreover, elephant genes are more involved in less familiar processes, such as endosomal sorting (*CHMP4C*) and glycosaminoglycan biosynthesis (*CHST14*) (Zhang et al., 2010; Carlton et

al., 2012). *NOTCH3*, *SMARCA1* and *FSTL1* play a role in cell differentiation and gene expression regulation. However, assigning and dividing genes according to their function is complex and their role may vary depending on specific molecular conditions. Therefore, the next step of this project was the identification of pathways, which can provide more insight into the connection between the DEGs in both species.

Looking into the potential role of tumour suppressor genes in the cancer suppression mechanisms in elephants, the absence of the *TP53* retrogenes were notable. Significant *TP53* expression was only observed in humans, indicating no apparent involvement of the TP53 gene in elephants four hours post-irradiation. Additionally, there was no association identified with the *LIF6* pseudogene in elephants. Moreover, a highly upregulated gene (ENSLAFG00000028505), unique to elephants without an annotated gene name or function description, should be considered for further investigation for its possible role in DNA damage response.

4.3 Pathways based on GO and KEGG database

The pathway enrichment analysis used ORA and GSEA based on GO and KEGG databases to provide insights into the responses of humans and elephants to induced DNA damage. The analysis of GO terms identified different patterns between humans and elephants. Both species showed a relation to immune response pathways associated with downregulated genes. ORA with the KEGG database showed less significant pathways compared to the GO database in humans, and no pathways in elephants. GSEA with the GO database revealed highly activated biological processes related to ribosomal subunit biogenesis in humans, indicating an upregulation of the synthesis and processing of ribosomal RNAs (Sloan et al., 2016). Additionally, suppressed human GO terms were linked to interleukin-1 production. Notably, no significant GO terms were found in elephants, indicating potential differences in DNA damage response between both species at four hours post-irradiation. GSEA with the KEGG database again showed activated ribosomal functions, which confirms the GO results in human samples. Additionally, the p53 signaling pathway was significantly activated, suggesting a potential link to DNA damage response in humans. In elephants, the butanoate metabolism pathway was the only significant finding, indicating specific metabolic responses. This may suggest a direction toward tighter regulation of their metabolism, possibly maintaining it at a consistently low level. Such regulation could potentially serve

as a cancer suppression mechanism, postulated as one of the hypotheses to solve Peto's paradox (Nagy, Victor and Cropper, 2007; Dang, 2012; 2015).

The difference in interpreting results between ORA and GSEA include that, in ORA, the presence of a group of up-regulated genes in a pathway does not necessarily mean that the pathway has an increased activity. It could also indicate reduced activity due to increased expression of inhibitory elements. In contrast, GSEA provides insights into both activated and suppressed pathways, considering the expression value within the analysis.

Moreover, the gene ratio used in this study shows the proportion of DEGs within a given pathway. However, other studies often prefer the use of the fold enrichment score or the enrichment score. These alternatives are considered as more reliable indicators to show pathway significance because they consider both the number of genes in the pathway and the overall distribution of genes in the gene list. For example, the fold enrichment score calculates how many times the pathway is more likely to be enriched with DEGs compared to what would be expected by chance. This helps to minimise biases associated with the gene sets.

Lastly, this study considered the q-value cut-off as a determinant of significant pathways. Looking at the results, the general q-value cut-off of 0.05 does affect the number of significant pathways in elephants. Therefore, it could be worth to consider higher q-value cut-offs in further analysis. It is essential to recognise that the q-value cut-off is used as a threshold to control the FDR (or adjusted p-value), and therefore the likelihood of false positives (Storey, 2003). The use of a higher q-value cut-off will allow a higher proportion of false positives and a lower confidence of pathways being truly significant. Given the high variability in the samples and the sensitivity of the results, the decision was made to keep the q-value threshold at 0.05. This choice controls the FDR and keeps the reliability of the identified significant pathways.

4.4 Pathview: the p53 signalling and apoptosis pathway

The analysis of the p53 signalling pathway and apoptosis pathway in humans and elephants show different patterns and functions of the involved DEGs. In the p53 signalling pathway, *MDM2* and *TP53* are upregulated in humans, indicating the activation of the negative feedback loop between *MDM2* and p53 activity (Mendoza,

Mandani and Momand, 2014). However, elephant cells do not show *TP53* expression at four hours, which raises the question about the cellular processes independent of the typical p53-mediated pathway. The observed upregulation of *CDKN1A (*p21 protein*)*, known to allow DNA damage repair while inhibiting apoptosis, could suggest are crucial role of this protein in elephants (Abbas and Dutta, 2009). However, the upregulation of *MDM2* in elephants together with *CDKN1A* and the absence of *TP53* may indicate a different response in elephants compared to humans. Further insights from the Abegglen paper show an upregulation of p21 and p53 proteins in elephants at five hours post-irradiation, indicating that transcription process was already completed (Abegglen et al., 2015). The variation in our findings where p21 was observed but not p53 four hours post-irradiation outlines potential differences in variation in gene expression patterns observed at different time points after irradiation or the difference in data analysis between both studies.

Turning to the apoptosis pathway in humans, many downregulated pro-apoptotic DEGs were found, despite the upregulation of apoptotic genes *BBC3* (*PUMA)* and *GADD45*. Despite the fact that fewer DEGs were significantly expressed in elephants, a clear upregulated pro-apoptosis pattern was observed, referring to *TRKA*, *PI3K*, *PUMA*, *BAX*, *AIP*, and *GADD45*. This could contribute to the fact that elephants blood cells undergo more and faster apoptosis when treated with the same doses of radiation compared to human blood cells, and therefore show are stronger apoptotic response at this same time point ((Preston et al., 2023). The downregulation of *FLIP* (also called c-FLIP) and *LAPXIAP*, both apoptosis suppression factors ((Safa, 2013; Chaudhary et al., 2016), indicate that there is also a contribution to the pro-survival mechanisms in elephants, which could point to the balancing act.

4.5 Limitations

4.5.1 Sample inclusion

Unfortunately, certain samples showed exhibited excessively high variability, precluding their inclusion in this study. As a consequence, each subgroup for elephants and the irradiated group for humans ended up with only two samples. Due to the inclusion of less than three sample in each subgroup the study needs careful considerations. The low number of samples not only imply insufficient statistical power, but also influences the

generalisability of the conclusions. A more extensive sample size is essential to ensure the reliability of the assumptions drawn regarding the effect of irradiation and DNA damage response in both species.

4.5.2 Applied cut-offs

This study included various cut-offs in different parts of the study. First, considerations should be given to the exclusion of DEGs with log2FC values under 1.5 or above -1.5 for identifying potential important DEGs, as these genes may represent subtle but meaningful changes in gene expression that could contribute to biological processes or pathways.

4.5.3 Pathway enrichment analysis tool & functional annotation

Although clusterProfiler is a widely used R package for functional annotation and pathways analysis, certain limitations should be acknowledged. Challenges become clear due to the requirement for a functional annotation through an organism database (OrgDb). The absence of such a database for elephants (*Loxodonta africana*) results in the exclusion of around 3000 potentially impactful genes. This exclusion may lead to the potential loss of crucial information and problems associated with transferring annotations across species.

4.6 Future directions

4.6.1 Elephant functional annotation

The consideration of elephant-specific annotation databases has the potential to provide more accurate information pathways and correlated genes. Therefore, finding a way to use the elephant database from Ensembl directly for functional annotation instead of relying on the human database is crucial. Otherwise, alternative approaches, such as the non-model-based approach for clusterProfiler or developing custom scripts could improve functional annotation. However, since inappropriate background set selection heavily influences enrichment results, careful consideration is necessary.

4.6.2 Extra samples

Increasing the sample size of this study is crucial to further look into potential mechanisms involved in DNA damage responses induced by radiation. Incorporating fibroblast samples could offer insight into additional mechanisms in DNA damage

response, as they respond sensitively to radiation-induced stress compared to humans (Preston et al., 2023).

Moreover, including additional analysis on irradiated samples and studying response at 12- or 24-hours post-irradiation should give insights into the DNA damage response mechanisms over a broader frame of time. This allows for more observation of temporal variations, which could reveal how the cellular response evolve over different stages post-irradiation.

4.6.3 Clustering

Next, clustering of the genes could be performed to detect subgroups of correlated genes. Here, you perform pathway enrichment analysis on each of these cluster alone. Using tools like PathfindR package for clustering may give hidden pattern or pathways in the data, offering a better perspective on the biological interpretation of the response to radiation.

Chapter 5: Conclusion

The data presented here show differences in significant DEGs, with overlap observed in only a few genes; *PHLDA3*, *BBC3*, and *CXCL11*, mainly associated with pro-apoptotic and immune responses. ORA using the GO database revealed immune response patterns associated with downregulated genes in both humans and elephants. Humans showed increased ribosomal subunit biogenesis and activated p53 signaling, indicative of a DNA damage response, while elephants showed a butanoate metabolism pathway, potentially serving as a cancer suppression mechanism. Moreover, species-specific differences in p53 signaling and apoptosis pathways were found.

It is important to take into consideration that this study only provides a view at a very narrow time point of four hours post-irradiation. This specific time point implies that certain responses may have already occurred, while others are yet to be revealed. Despite the significant impact of this limitation and the high variability between the replicates resulting in limited sample size, this study presents potential suggestions for future investigation into species specific DNA damage responses. As future directions, incorporating different time points and more samples could add crucial findings to this study.

References

Abbas, T. and Dutta, A., 2009. p21 in cancer: intricate networks and multiple activities. *Nature reviews. Cancer*, [online] 9(6), pp.400–414. https://doi.org/10.1038/nrc2657.

Abegglen, L.M., Caulin, A.F., Chan, A., Lee, K., Robinson, R., Campbell, M.S., Kiso, W.K., Schmitt, D.L., Waddell, P.J., Bhaskara, S., Jensen, S.T., Maley, C.C. and Schiffman, J.D., 2015. Potential Mechanisms for Cancer Resistance in Elephants and Comparative Cellular Response to DNA Damage in Humans. *JAMA*, [online] 314(17), pp.1850–1860. https://doi.org/10.1001/jama.2015.13134.

Amaral, P., Carbonell-Sala, S., De La Vega, F.M., Faial, T., Frankish, A., Gingeras, T., Guigo, R., Harrow, J.L., Hatzigeorgiou, A.G., Johnson, R., Murphy, T.D., Pertea, M., Pruitt, K.D., Pujar, S., Takahashi, H., Ulitsky, I., Varabyou, A., Wells, C.A., Yandell, M., Carninci, P. and Salzberg, S.L., 2023. The status of the human gene catalogue. *Nature*, [online] 622(7981), pp.41–47. https://doi.org/10.1038/s41586-023-06490-x.

Andrews, S., 2010. *Babraham Bioinformatics - FastQC A Quality Control tool for High Throughput Sequence Data*. [online] Available at: <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/> [Accessed 7 January 2024].

Arowolo, A., Malebana, M., Sunda, F. and Rhoda, C., 2022. Proposed Cellular Function of the Human FAM111B Protein and Dysregulation in Fibrosis and Cancer. *Frontiers in Oncology*, [online] 12, p.932167. https://doi.org/10.3389/fonc.2022.932167.

Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., Harris, M.A., Hill, D.P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J.C., Richardson, J.E., Ringwald, M., Rubin, G.M. and Sherlock, G., 2000. Gene Ontology: tool for the unification of biology. *Nature genetics*, [online] 25(1), pp.25–29. https://doi.org/10.1038/75556.

Bartas, M., Brázda, V., Volná, A., Červeň, J., Pečinka, P. and Zawacka-Pankau, J.E., 2021. The Changes in the p53 Protein across the Animal Kingdom Point to Its Involvement in Longevity. *International Journal of Molecular Sciences*, [online] 22(16), p.8512. https://doi.org/10.3390/ijms22168512.

Blighe, K., Rana, S. and Lewis, M., 2023. *EnhancedVolcano: publication-ready volcano plots with enhanced colouring and labeling*. [online] Available at: <https://bioconductor.org/packages/devel/bioc/vignettes/EnhancedVolcano/inst/doc/EnhancedVolcano.html> [Accessed 8 January 2024].

Bolger, A.M., Lohse, M. and Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, [online] 30(15), pp.2114–2120. https://doi.org/10.1093/bioinformatics/btu170.

Bray, N.L., Pimentel, H., Melsted, P. and Pachter, L., 2016. Near-optimal probabilistic RNA-seq quantification. *Nature Biotechnology*, [online] 34(5), pp.525–527. https://doi.org/10.1038/nbt.3519.

Butler-Laporte, G., Kreuzer, D., Nakanishi, T., Harroud, A., Forgetta, V. and Richards, J.B., 2020. Genetic Determinants of Antibody-Mediated Immune Responses to Infectious Diseases Agents: A Genome-Wide and HLA Association Study. *Open Forum Infectious Diseases*, [online] 7(11), p.ofaa450. https://doi.org/10.1093/ofid/ofaa450.

Carlson, M., 2019. *org.Hs.eg.db: Genome wide annotation for Human.*

Carlton, J.G., Caballe, A., Agromayor, M., Kloc, M. and Martin-Serrano, J., 2012. ESCRT-III governs the Aurora B-mediated abscission checkpoint through CHMP4C. *Science (New York, N.Y.)*, 336(6078), pp.220–225. https://doi.org/10.1126/science.1217180.

Caulin, A.F., Graham, T.A., Wang, L.-S. and Maley, C.C., 2015. Solutions to Peto's paradox revealed by mathematical modelling and cross-species cancer gene analysis. *Philosophical Transactions of the Royal Society B: Biological Sciences*, [online] 370(1673), p.20140222. https://doi.org/10.1098/rstb.2014.0222.

Chaudhary, A.K., Yadav, N., Bhat, T.A., O'Malley, J., Kumar, S. and Chandra, D., 2016. A potential role of X-linked inhibitor of apoptosis protein in mitochondrial membrane permeabilization and its implication in cancer therapy. *Drug discovery today*, [online] 21(1), pp.38–47. https://doi.org/10.1016/j.drudis.2015.07.014.

Chen, H. and Boutros, P.C., 2011. VennDiagram: a package for the generation of highly-customizable Venn and Euler diagrams in R. *BMC Bioinformatics*, [online] 12(1), p.35. https://doi.org/10.1186/1471-2105-12-35.

Dang, C.V., 2012. Links between metabolism and cancer. *Genes & Development*, [online] 26(9), pp.877–890. https://doi.org/10.1101/gad.189365.112.

Dang, C.V., 2015. A metabolic perspective of Peto's paradox and cancer. *Philosophical Transactions of the Royal Society B: Biological Sciences*, [online] 370(1673), p.20140223. https://doi.org/10.1098/rstb.2014.0223.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M. and Gingeras, T.R., 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics (Oxford, England)*, 29(1), pp.15–21. https://doi.org/10.1093/bioinformatics/bts635.

Fang, F., Yu, X., Wang, X., Zhu, X., Liu, L., Rong, L., Niu, D. and Li, J., 2021. Transcriptomic profiling reveals gene expression in human peripheral blood after exposure to low-dose ionizing radiation. *Journal of Radiation Research*, [online] 63(1), pp.8–18. https://doi.org/10.1093/jrr/rrab091.

Fleming, J. m., Creevy, K. e. and Promislow, D. e. l., 2011. Mortality in North American Dogs from 1984 to 2004: An Investigation into Age-, Size-, and Breed-Related Causes of Death. *Journal of Veterinary Internal Medicine*, [online] 25(2), pp.187–198. https://doi.org/10.1111/j.1939-1676.2011.0695.x.

García-Cao, I., García-Cao, M., Martín-Caballero, J., Criado, L.M., Klatt, P., Flores, J.M., Weill, J.-C., Blasco, M.A. and Serrano, M., 2002. 'Super p53' mice exhibit enhanced

DNA damage response, are tumor resistant and age normally. *The EMBO journal*, 21(22), pp.6225–6235. https://doi.org/10.1093/emboj/cdf595.

Han, J., Flemington, C., Houghton, A.B., Gu, Z., Zambetti, G.P., Lutz, R.J., Zhu, L. and Chittenden, T., 2001. Expression of bbc3, a pro-apoptotic BH3-only gene, is regulated by diverse cell death and survival signals. *Proceedings of the National Academy of Sciences*, [online] 98(20), pp.11318–11323. https://doi.org/10.1073/pnas.201208798.

Kanehisa, M., 2019. Toward understanding the origin and evolution of cellular organisms. *Protein Science*, [online] 28(11), pp.1947–1951. https://doi.org/10.1002/pro.3715.

Kanehisa, M., Furumichi, M., Sato, Y., Kawashima, M. and Ishiguro-Watanabe, M., 2023. KEGG for taxonomy-based analysis of pathways and genomes. *Nucleic Acids Research*, [online] 51(D1), pp.D587–D592. https://doi.org/10.1093/nar/gkac963.

Kanehisa, M. and Goto, S., 2000. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Research*, [online] 28(1), pp.27–30. https://doi.org/10.1093/nar/28.1.27.

Kawase, T., Ohki, R., Shibata, T., Tsutsumi, S., Kamimura, N., Inazawa, J., Ohta, T., Ichikawa, H., Aburatani, H., Tashiro, F. and Taya, Y., 2009. PH Domain-Only Protein PHLDA3 Is a p53-Regulated Repressor of Akt. *Cell*, [online] 136(3), pp.535–550. https://doi.org/10.1016/j.cell.2008.12.002.

Kleiveland, C.R., 2015. Peripheral Blood Mononuclear Cells. In: K. Verhoeckx, P. Cotter, I. López-Expósito, C. Kleiveland, T. Lea, A. Mackie, T. Requena, D. Swiatecka and H. Wichers, eds. *The Impact of Food Bioactives on Health: in vitro and ex vivo models*. [online] Cham (CH): Springer. Available at: <http://www.ncbi.nlm.nih.gov/books/NBK500157/> [Accessed 24 January 2024].

Liao, Y., Smyth, G.K. and Shi, W., 2014. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*, [online] 30(7), pp.923–930. https://doi.org/10.1093/bioinformatics/btt656.

Luo, W. and Brouwer, C., 2013. Pathview: an R/Bioconductor package for pathway-based data integration and visualization. *Bioinformatics*, [online] 29(14), pp.1830–1831. https://doi.org/10.1093/bioinformatics/btt285.

Mattiotti, A., Prakash, S., Barnett, P. and van den Hoff, M.J.B., 2018. Follistatin-like 1 in development and human diseases. *Cellular and Molecular Life Sciences*, [online] 75(13), pp.2339–2354. https://doi.org/10.1007/s00018-018-2805-0.

McCarthy, N., 2011. Early exposure is inflammatory. *Nature Reviews Immunology*, [online] 11(3), pp.157–157. https://doi.org/10.1038/nri2945.

Mendoza, M., Mandani, G. and Momand, J., 2014. The MDM2 gene family. *Biomolecular concepts*, [online] 5(1), pp.9–19. https://doi.org/10.1515/bmc-2013-0027.

Nagy, J.D., Victor, E.M. and Cropper, J.H., 2007. Why don't all whales have cancer? A novel hypothesis resolving Peto's paradox. *Integrative and Comparative Biology*, [online] 47(2), pp.317–328. https://doi.org/10.1093/icb/icm062.

Noble, R., Kaltz, O. and Hochberg, M.E., 2015. Peto's paradox and human cancers. *Philosophical Transactions of the Royal Society B: Biological Sciences*, [online] 370(1673), p.20150104. https://doi.org/10.1098/rstb.2015.0104.

Nueda, M. j., Ferrer, A. and Conesa, A., 2012. ARSyN: a method for the identification and removal of systematic noise in multifactorial time course microarray experiments. *Biostatistics*, [online] 13(3), pp.553–566. https://doi.org/10.1093/biostatistics/kxr042.

Nunney, L., 2013. The real war on cancer: the evolutionary dynamics of cancer suppression. *Evolutionary Applications*, [online] 6(1), pp.11–19. https://doi.org/10.1111/eva.12018.

O'Brien, G., Cruz-Garcia, L., Majewski, M., Grepl, J., Abend, M., Port, M., Tichý, A., Sirak, I., Malkova, A., Donovan, E., Gothard, L., Boyle, S., Somaiah, N., Ainsbury, E., Ponge, L., Slosarek, K., Miszczyk, L., Widlak, P., Green, E., Patel, N., Kudari, M., Gleeson, F., Vinnikov, V., Starenkiy, V., Artiukh, S., Vasyliev, L., Zaman, A. and Badie, C., 2018. FDXR is a biomarker of radiation exposure in vivo. *Scientific Reports*, [online] 8, p.684. https://doi.org/10.1038/s41598-017-19043-w.

Ozga, A.J., Chow, M.T. and Luster, A.D., 2021. Chemokines and the immune response to cancer. *Immunity*, [online] 54(5), pp.859–874. https://doi.org/10.1016/j.immuni.2021.01.012.

Peto, R., Roe, F.J., Lee, P.N., Levy, L. and Clack, J., 1975. Cancer and ageing in mice and men. *British Journal of Cancer*, [online] 32(4), pp.411–426. Available at: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2024769/> [Accessed 21 December 2023].

Preston, A.J., Rogers, A., Sharp, M., Mitchell, G., Toruno, C., Barney, B.B., Donovan, L.N., Bly, J., Kennington, R., Payne, E., Iovino, A., Furukawa, G., Robinson, R., Shamloo, B., Buccilli, M., Anders, R., Eckstein, S., Fedak, E.A., Wright, T., Maley, C.C., Kiso, W.K., Schmitt, D., Malkin, D., Schiffman, J.D. and Abegglen, L.M., 2023. Elephant TP53-RETROGENE 9 induces transcription-independent apoptosis at the mitochondria. *Cell Death Discovery*, [online] 9(1), pp.1–11. https://doi.org/10.1038/s41420-023-01348-7.

Rainer, J., 2017. *EnsDb.Hsapiens.v79: Ensembl based annotation package.*

Rau, A., Gallopin, M., Celeux, G. and Jaffrézic, F., 2013. Data-based filtering for replicated high-throughput transcriptome sequencing experiments. *Bioinformatics (Oxford, England)*, 29(17), pp.2146–2152. https://doi.org/10.1093/bioinformatics/btt350.

RStudio Team, 2020. *RStudio: Integrated Development for R. RStudio, PBC, Boston, MA*. Available at: <http://www.rstudio.com/.>.

Safa, A.R., 2013. Roles of c-FLIP in Apoptosis, Necroptosis, and Autophagy. *Journal of Carcinogenesis & Mutagenesis*, Suppl 6, p.003. https://doi.org/10.4172/2157-2518.S6-003.

Seluanov, A., Gladyshev, V.N., Vijg, J. and Gorbunova, V., 2018. Mechanisms of cancer resistance in long-lived mammals. *Nature reviews. Cancer*, [online] 18(7), pp.433–441. https://doi.org/10.1038/s41568-018-0004-9.

Shi, M., Chen, F., Sahu, S.K., Wang, Q., Yang, S., Wang, Z., Chen, J., Liu, H., Hou, Z., Fang, S.-G. and Lan, T., 2024. Haplotype-resolved chromosome-scale genomes of the Asian and African Savannah Elephants. *Scientific Data*, [online] 11(1), p.63. https://doi.org/10.1038/s41597-023-02729-4.

Sloan, K.E., Warda, A.S., Sharma, S., Entian, K.-D., Lafontaine, D.L.J. and Bohnsack, M.T., 2016. Tuning the ribosome: The influence of rRNA modification on eukaryotic ribosome biogenesis and function. *RNA Biology*, [online] 14(9), pp.1138–1152. https://doi.org/10.1080/15476286.2016.1259781.

Storey, J.D., 2003. The positive false discovery rate: a Bayesian interpretation and the q-value. *The Annals of Statistics*, [online] 31(6), pp.2013–2035. https://doi.org/10.1214/aos/1074290335.

Sulak, M., Fong, L., Mika, K., Chigurupati, S., Yon, L., Mongan, N.P., Emes, R.D. and Lynch, V.J., 2016. TP53 copy number expansion is associated with the evolution of increased body size and an enhanced DNA damage response in elephants. *eLife*, [online] 5, p.e11994. https://doi.org/10.7554/eLife.11994.

Tarazona, S., Furió-Tarí, P., Turrà, D., Pietro, A.D., Nueda, M.J., Ferrer, A. and Conesa, A., 2015. Data quality aware analysis of differential expression in RNA-seq with NOISeq R/Bioc package. *Nucleic Acids Research*, 43(21), p.e140. https://doi.org/10.1093/nar/gkv711.

Tarazona, S., García-Alcalde, F., Dopazo, J., Ferrer, A. and Conesa, A., 2011. Differential expression in RNA-seq: A matter of depth. *Genome Research*, [online] 21(12), pp.2213–2223. https://doi.org/10.1101/gr.124321.111.

The Gene Ontology Consortium, Aleksander, S.A., Balhoff, J., Carbon, S., Cherry, J.M., Drabkin, H.J., Ebert, D., Feuermann, M., Gaudet, P., Harris, N.L., Hill, D.P., Lee, R., Mi, H., Moxon, S., Mungall, C.J., Muruganugan, A., Mushayahama, T., Sternberg, P.W., Thomas, P.D., Van Auken, K., Ramsey, J., Siegele, D.A., Chisholm, R.L., Fey, P., Aspromonte, M.C., Nugnes, M.V., Quaglia, F., Tosatto, S., Giglio, M., Nadendla, S., Antonazzo, G., Attrill, H., dos Santos, G., Marygold, S., Strelets, V., Tabone, C.J., Thurmond, J., Zhou, P., Ahmed, S.H., Asanitthong, P., Luna Buitrago, D., Erdol, M.N., Gage, M.C., Ali Kadhum, M., Li, K.Y.C., Long, M., Michalak, A., Pesala, A., Pritazahra, A., Saverimuttu, S.C.C., Su, R., Thurlow, K.E., Lovering, R.C., Logie, C., Oliferenko, S., Blake, J., Christie, K., Corbani, L., Dolan, M.E., Drabkin, H.J., Hill, D.P., Ni, L., Sitnikov, D., Smith, C., Cuzick, A., Seager, J., Cooper, L., Elser, J., Jaiswal, P., Gupta, P., Jaiswal, P., Naithani, S., Lera-Ramirez, M., Rutherford, K., Wood, V., De Pons, J.L., Dwinell, M.R., Hayman, G.T., Kaldunski, M.L., Kwitek, A.E., Laulederkind, S.J.F., Tutaj, M.A., Vedi, M., Wang, S.-

J., D'Eustachio, P., Aimo, L., Axelsen, K., Bridge, A., Hyka-Nouspikel, N., Morgat, A., Aleksander, S.A., Cherry, J.M., Engel, S.R., Karra, K., Miyasato, S.R., Nash, R.S., Skrzypek, M.S., Weng, S., Wong, E.D., Bakker, E., Berardini, T.Z., Reiser, L., Auchincloss, A., Axelsen, K., Argoud-Puy, G., Blatter, M.-C., Boutet, E., Breuza, L., Bridge, A., Casals-Casas, C., Coudert, E., Estreicher, A., Livia Famiglietti, M., Feuermann, M., Gos, A., Gruaz-Gumowski, N., Hulo, C., Hyka-Nouspikel, N., Jungo, F., Le Mercier, P., Lieberherr, D., Masson, P., Morgat, A., Pedruzzi, I., Pourcel, L., Poux, S., Rivoire, C., Sundaram, S., Bateman, A., Bowler-Barnett, E., Bye-A-Jee, H., Denny, P., Ignatchenko, A., Ishtiaq, R., Lock, A., Lussi, Y., Magrane, M., Martin, M.J., Orchard, S., Raposo, P., Speretta, E., Tyagi, N., Warner, K., Zaru, R., Diehl, A.D., Lee, R., Chan, J., Diamantakis, S., Raciti, D., Zarowiecki, M., Fisher, M., James-Zorn, C., Ponferrada, V., Zorn, A., Ramachandran, S., Ruzicka, L. and Westerfield, M., 2023. The Gene Ontology knowledgebase in 2023. *Genetics*, [online] 224(1), p.iyad031. https://doi.org/10.1093/genetics/iyad031.

Thomas, P.D., 2017. The Gene Ontology and the meaning of biological function. *Methods in molecular biology (Clifton, N.J.)*, [online] 1446, pp.15–24. https://doi.org/10.1007/978-1-4939-3743-1_2.

Tokunaga, R., Zhang, W., Naseem, M., Puccini, A., Berger, M.D., Soni, S., McSkane, M., Baba, H. and Lenz, H.-J., 2018. CXCL9, CXCL10, CXCL11/CXCR3 axis for immune activation - a target for novel cancer therapy. *Cancer treatment reviews*, [online] 63, pp.40–47. https://doi.org/10.1016/j.ctrv.2017.11.007.

Tyner, S.D., Venkatachalam, S., Choi, J., Jones, S., Ghebranious, N., Igelmann, H., Lu, X., Soron, G., Cooper, B., Brayton, C., Hee Park, S., Thompson, T., Karsenty, G., Bradley, A. and Donehower, L.A., 2002. p53 mutant mice that display early ageing-associated phenotypes. *Nature*, [online] 415(6867), pp.45–53. https://doi.org/10.1038/415045a.

Vazquez, J.M., Sulak, M., Chigurupati, S. and Lynch, V.J., 2018. A Zombie LIF Gene in Elephants Is Upregulated by TP53 to Induce Apoptosis in Response to DNA Damage. *Cell Reports*, [online] 24(7), pp.1765–1776. https://doi.org/10.1016/j.celrep.2018.07.042.

Vincze, O., Colchero, F., Lemaître, J.-F., Conde, D.A., Pavard, S., Bieuville, M., Urrutia, A.O., Ujvari, B., Boddy, A.M., Maley, C.C., Thomas, F. and Giraudeau, M., 2022. Cancer risk across mammals. *Nature*, [online] 601(7892), pp.263–267. https://doi.org/10.1038/s41586-021-04224-5.

Wijesooriya, K., Jadaan, S.A., Perera, K.L., Kaur, T. and Ziemann, M., 2022. Urgent need for consistent standards in functional enrichment analysis. *PLOS Computational Biology*, [online] 18(3), p.e1009935. https://doi.org/10.1371/journal.pcbi.1009935.

Wirén, S., Häggström, C., Ulmer, H., Manjer, J., Bjørge, T., Nagel, G., Johansen, D., Hallmans, G., Engeland, A., Concin, H., Jonsson, H., Selmer, R., Tretli, S., Stocks, T. and Stattin, P., 2014. Pooled cohort study on height and risk of cancer and cancer death. *Cancer Causes & Control*, [online] 25(2), pp.151–159. https://doi.org/10.1007/s10552-013-0317-7.

Wu, T., Hu, E., Xu, S., Chen, M., Guo, P., Dai, Z., Feng, T., Zhou, L., Tang, W., Zhan, L., Fu, X., Liu, S., Bo, X. and Yu, G., 2021. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *The Innovation*, [online] 2(3), p.100141. https://doi.org/10.1016/j.xinn.2021.100141.

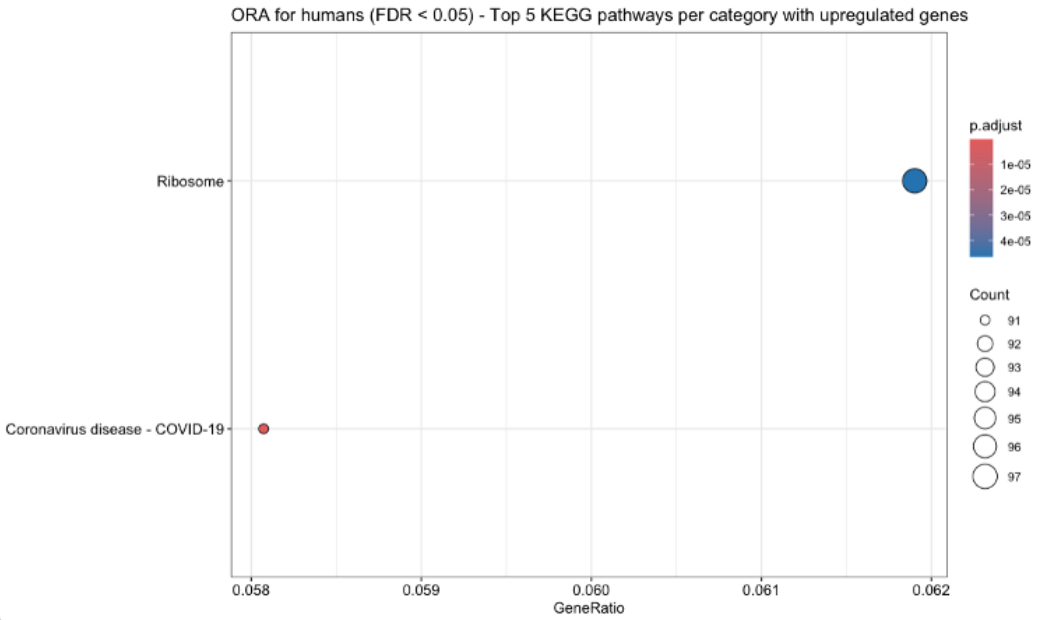
Yoshida, H. and Hunter, C.A., 2015. The Immunobiology of Interleukin-27. *Annual Review of Immunology*, [online] 33(1), pp.417–443. https://doi.org/10.1146/annurev-immunol-032414-112134.

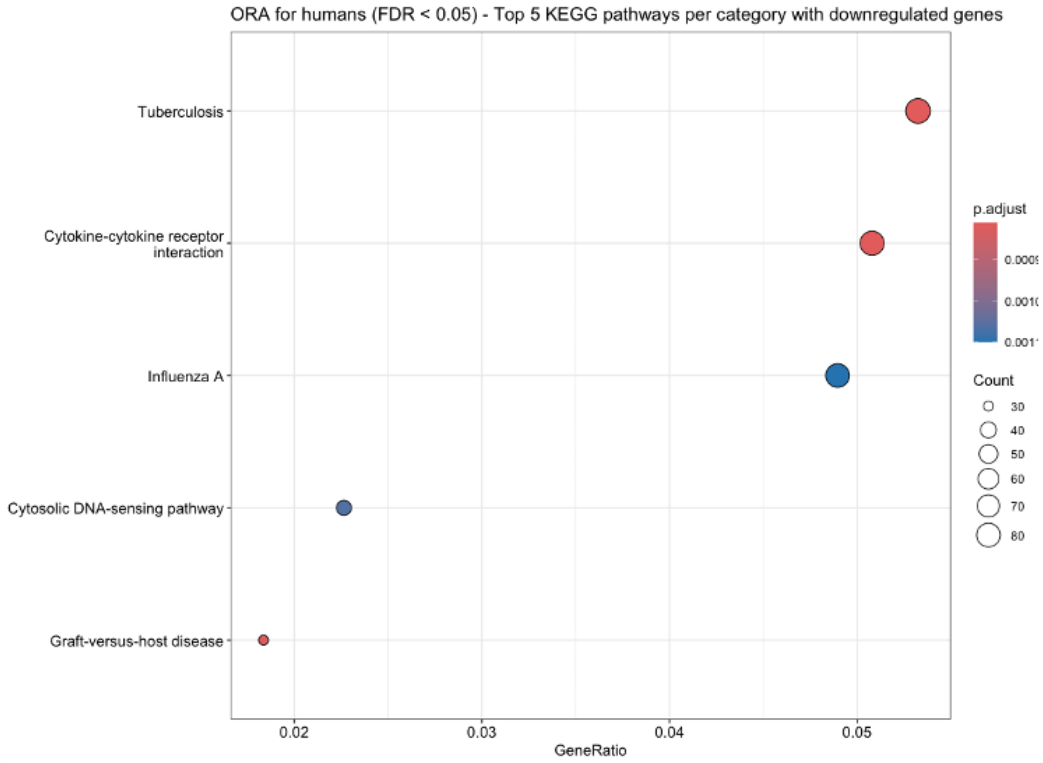
Yu, G., Wang, L.-G., Han, Y. and He, Q.-Y., 2012. clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. *OMICS : a Journal of Integrative Biology*, [online] 16(5), pp.284–287. https://doi.org/10.1089/omi.2011.0118.

Zhang, L., Müller, T., Baenziger, J.U. and Janecke, A.R., 2010. Congenital Disorders of Glycosylation with Emphasis on loss of Dermatan-4-Sulfotransferase. In: L. Zhang, ed. *Progress in Molecular Biology and Translational Science*, Glycosaminoglycans in Development, Health and Disease. [online] Academic Press. pp.289–307. https://doi.org/10.1016/S1877-1173(10)93012-3.

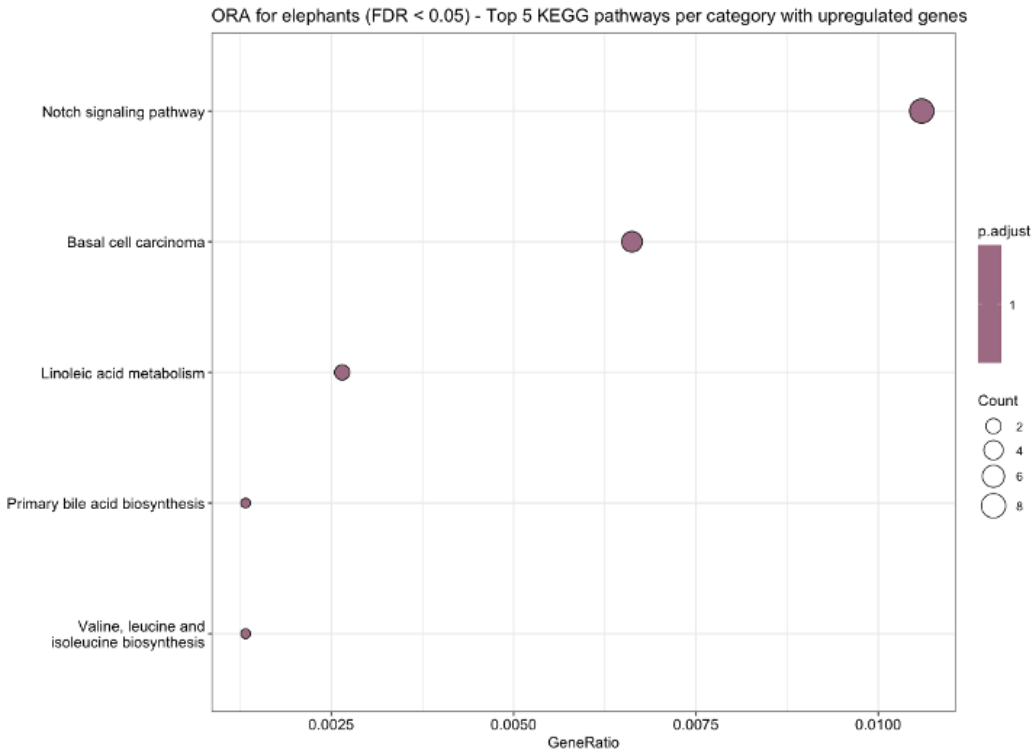
Zheng, Q., Hou, J., Zhou, Y., Yang, Y., Xie, B. and Cao, X., 2015. Siglec1 suppresses antiviral innate immune response by inducing TBK1 degradation via the ubiquitin ligase TRIM27. *Cell Research*, [online] 25(10), pp.1121–1136. https://doi.org/10.1038/cr.2015.108.

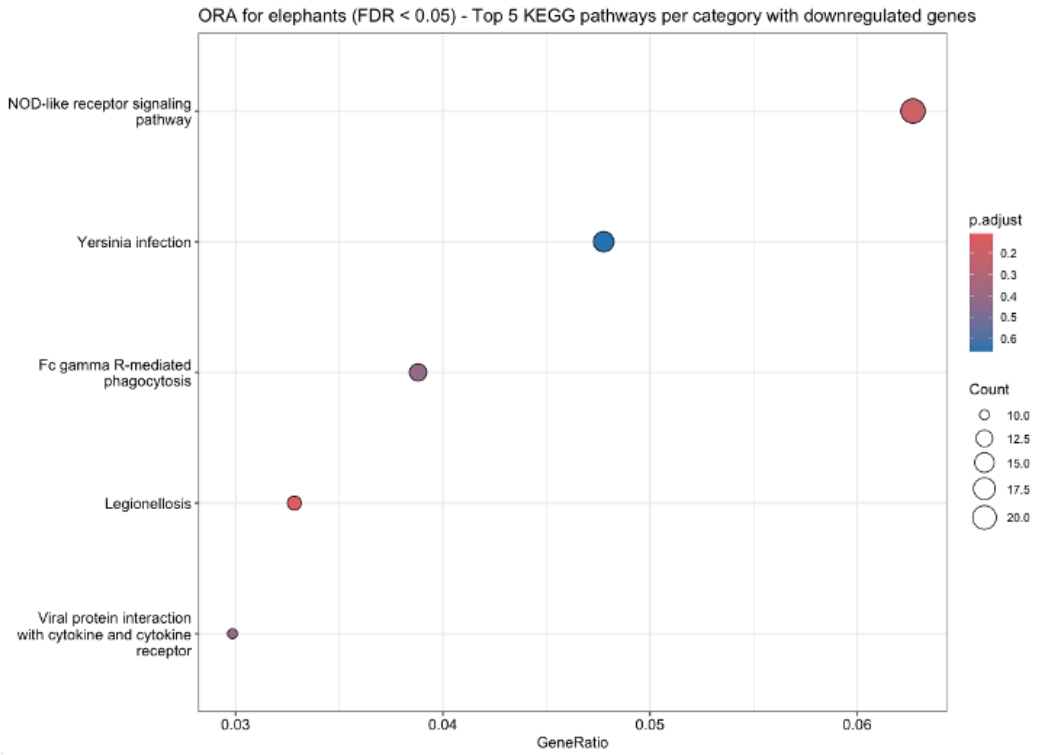
Appendix A: Supplementary figures



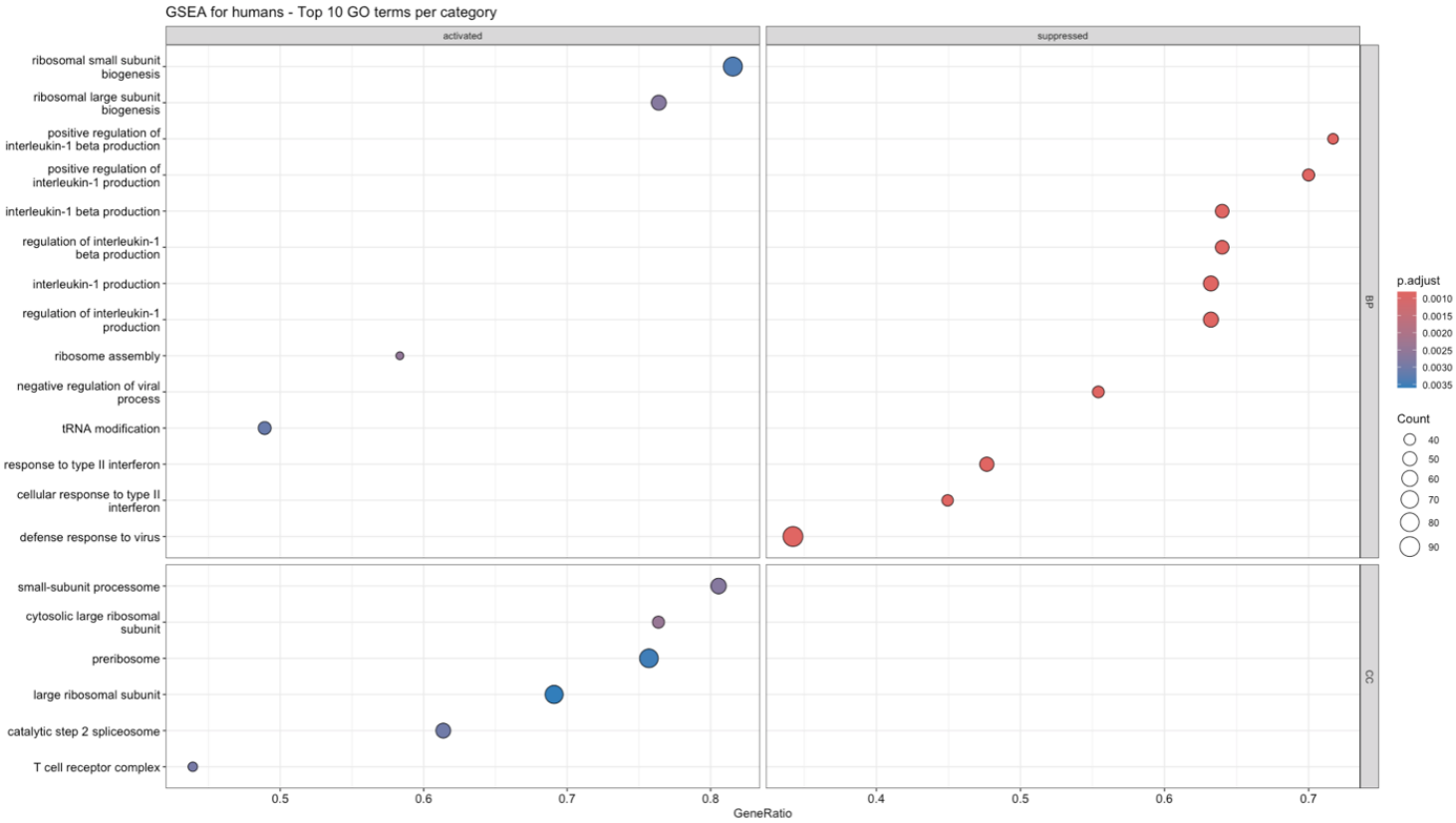


Supplementary figure 1. *Kyoto Encyclopedia of genes and genomes (KEGG) – over-representation analysis (ORA). Pathway enrichment results in humans for differentially expressed genes with a false discovery rate (FDR) ≤ 0.05 and log2 fold change (log2FC) </> 0. The size of dots represents the number of genes in each pathway, while the colour of the dots corresponds to the adjusted p-value. The gene ratio of core enrichment genes to pathway genes provides insights into the specific impact of each pathway in the experimental conditions.*

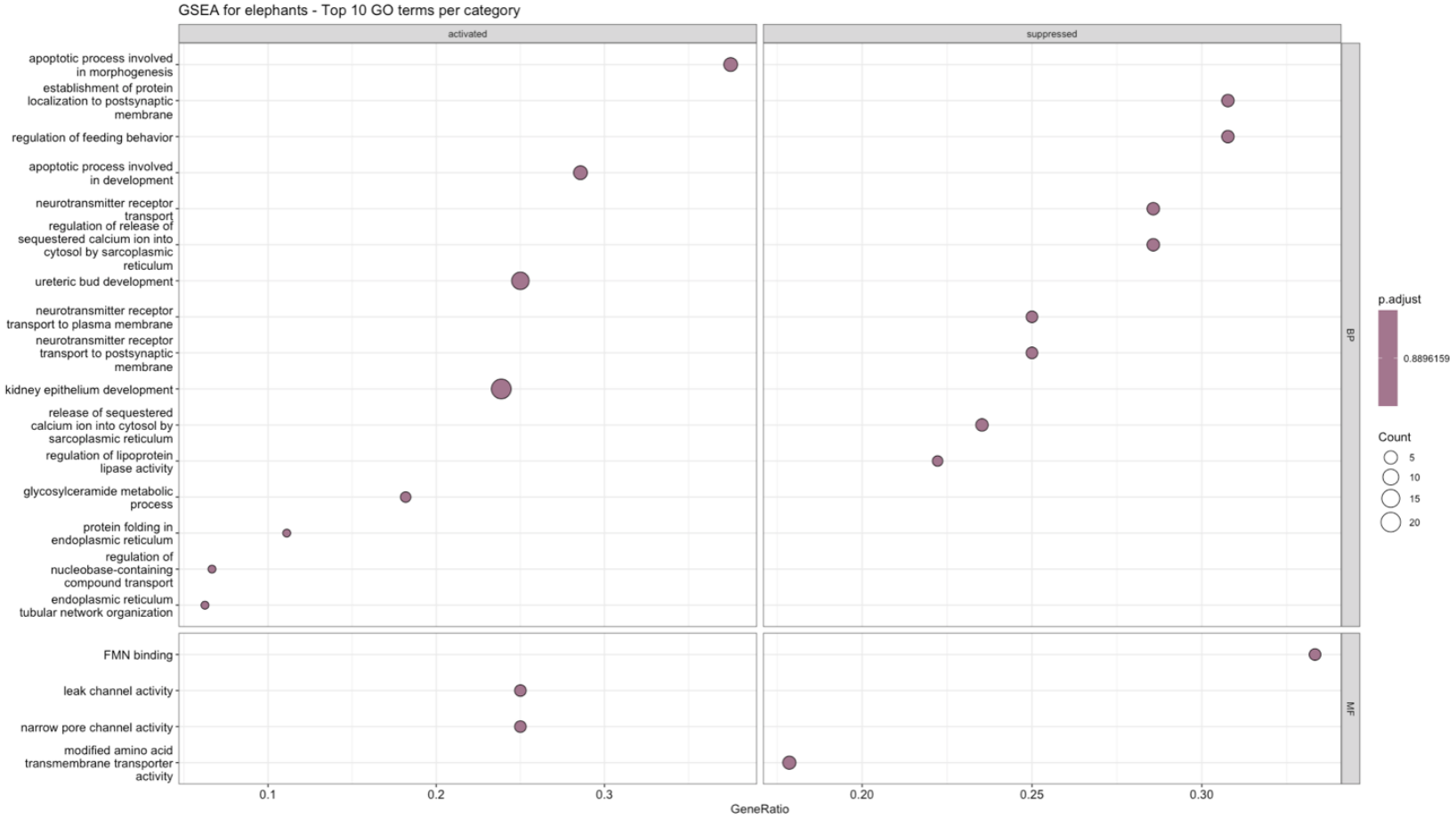




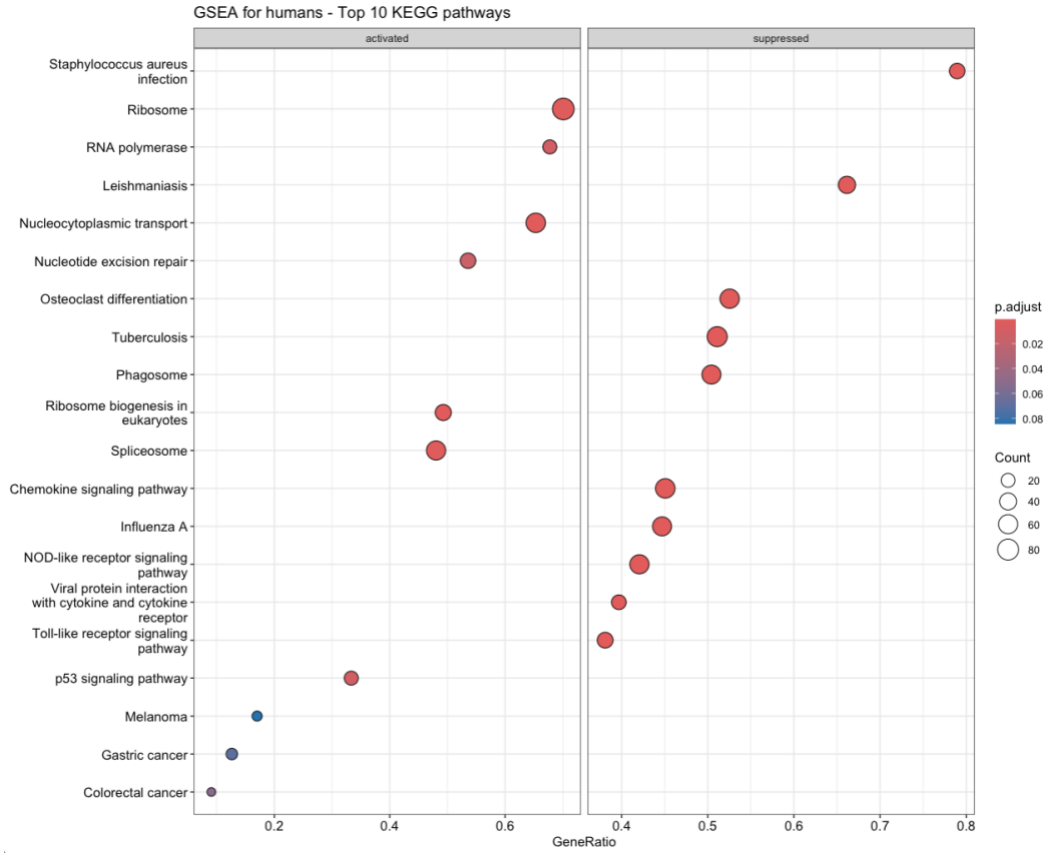
Supplementary figure 2. *Kyoto Encyclopedia of genes and genomes (KEGG) – over-representation analysis (ORA). Pathway enrichment results in elephants for differentially expressed genes with a false discovery rate (FDR) ≤ 0.05 and log2 fold change (log2FC) </> 0. The size of dots represents the number of genes in each pathway, while the colour of the dots corresponds to the adjusted p-value. The gene ratio of core enrichment genes to pathway genes provides insights into the specific impact of each pathway in the experimental conditions.*



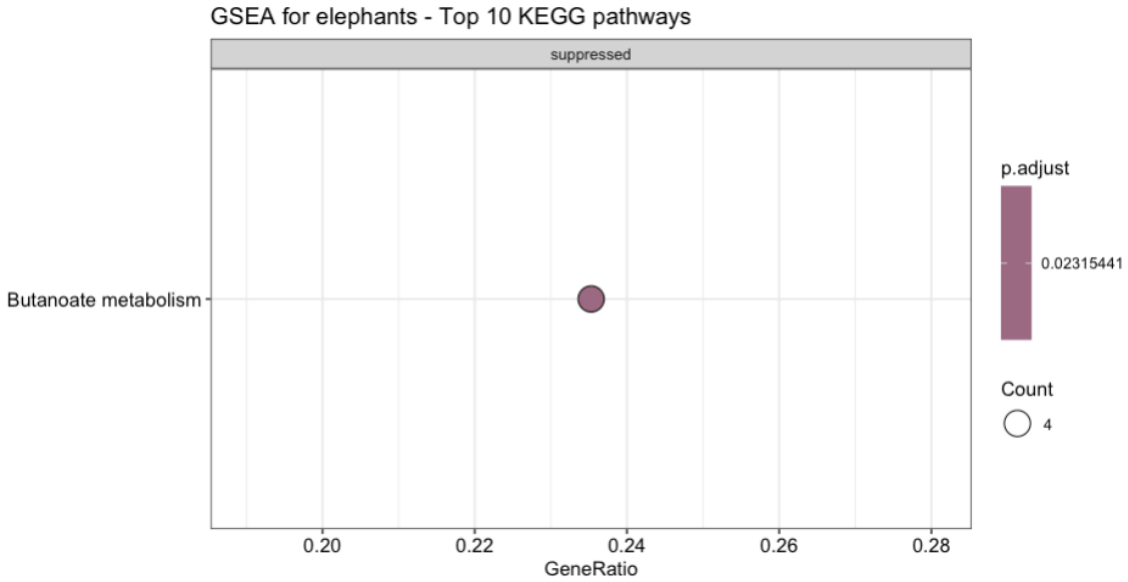
Supplementary figure 3. *Gene ontology (GO) – gene set enrichment analysis (GSEA). Pathway enrichment results in humans for differentially expressed genes without a false discovery rate (FDR) threshold and log2 fold change (log2FC) </> 0. The size of dots represents the number of genes in each pathway, while the colour of the dots corresponds to the adjusted p-value. The gene ratio of core enrichment genes to pathway genes provides insights into the specific impact of each pathway in the experimental conditions.*



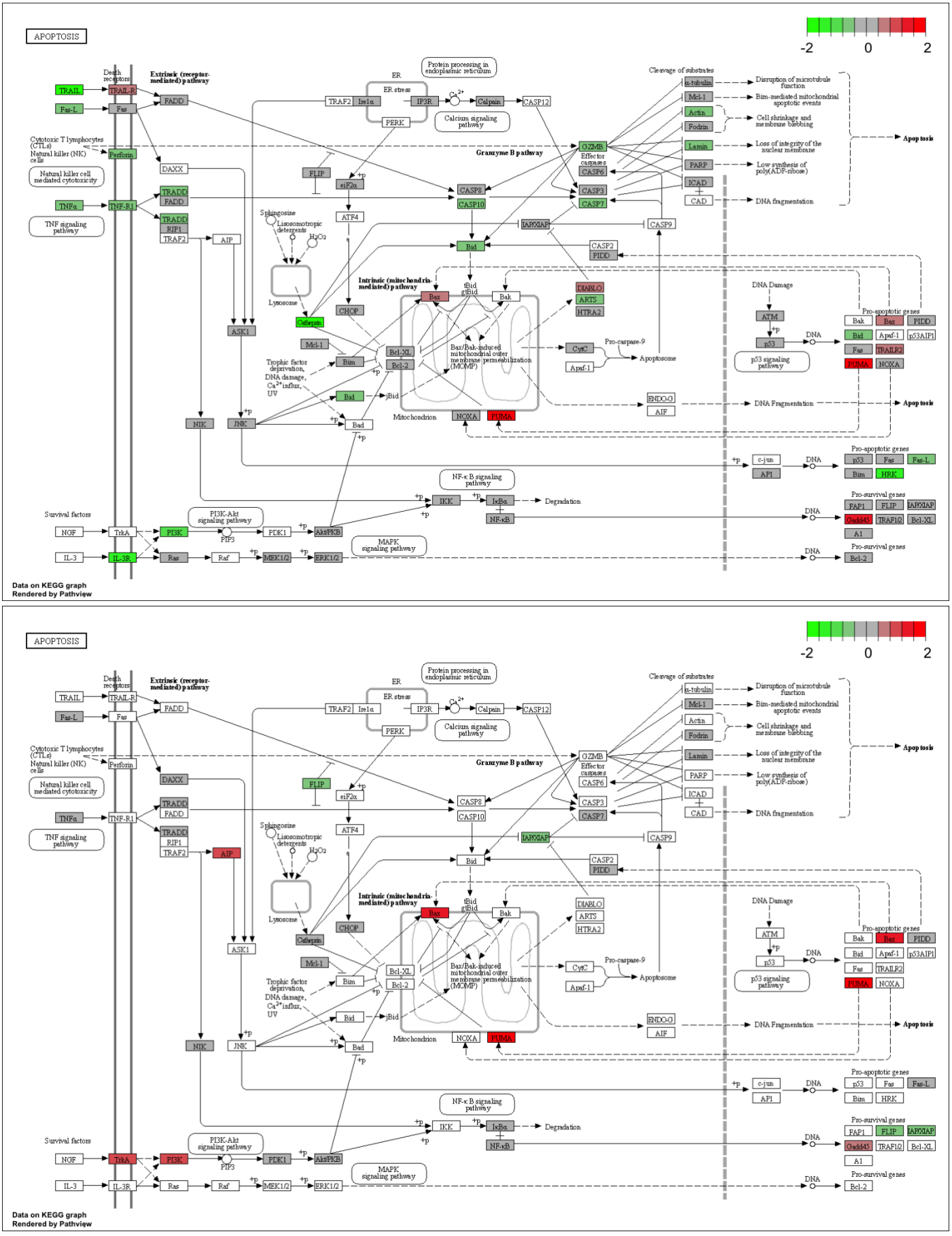
Supplementary figure 4. *Gene ontology (GO) – gene set enrichment analysis (GSEA). Pathway enrichment results in elephants for differentially expressed genes without a false discovery rate (FDR) threshold and log2 fold change (log2FC) </> 0. The size of dots represents the number of genes in each pathway, while the colour of the dots corresponds to the adjusted p-value. The gene ratio of core enrichment genes to pathway genes provides insights into the specific impact of each pathway in the experimental conditions.*



Supplementary figure 5. *Kyoto Encyclopedia of genes and genomes (KEGG) – gene set enrichment analysis (GSEA). Pathway enrichment results in humans for differentially expressed genes without a false discovery rate (FDR) threshold and log2 fold change (log2FC) </> 0. The size of dots represents the number of genes in each pathway, while the colour of the dots corresponds to the adjusted p-value. The gene ratio of core enrichment genes to pathway genes provides insights into the specific impact of each pathway in the experimental conditions.*



Supplementary figure 6. *Kyoto Encyclopedia of genes and genomes (KEGG) – gene set enrichment analysis (GSEA). Pathway enrichment results in elephants for differentially expressed genes without a false discovery rate (FDR) threshold and log2 fold change (log2FC) </> 0. The size of dots represents the number of genes in each pathway, while the colour of the dots corresponds to the adjusted p-value. The gene ratio of core enrichment genes to pathway genes provides insights into the specific impact of each pathway in the experimental conditions.*



Supplementary figure 7. *Apoptosis pathway for humans (a) and for elephants (b) based on KEGG database and generated by Pathview. Upregulated proteins in red, downregulated proteins in green and moderately expressed protein in gray.*