Research Proposal by Aaria Rahman - Aaria Rahman.pdf

by Sanaul Haque

Submission date: 13-Oct-2025 11:23PM (UTC+0700)

Submission ID: 2779975595

File name: Research_Proposal_by_Aaria_Rahman_-_Aaria_Rahman.pdf (211.98K)

Word count: 1729

Character count: 10004

IARCO 2025 RESEARCH PROPOSAL

Beyond TP53 Gene: Exploring Gene Network Robustness to Solve Peto's Paradox

Full Name: Aaria Rahman

Institution: Sunnydale

Category: Junior

Class: Class 12 (A Level)

Country: Bangladesh

Submission Date: September 27, 2025

Registered Email Address: aaria.rahman.college@gmail.com

Research Topic: Beyond TP53 Gene: Exploring Gene Network Robustness to Solve Peto's
Paradox

Research Problem

Uncontrolled cell growth brought on by accumulated mutations in DNA repair, genes, oncogenes and tumour suppressors causes cancer. Based on probability, cancer incidence should be higher in species with more cells and longer lifespans. Peto's paradox is the contradiction that big mammals like whales and elephants do not develop cancer [1].

Multiple current explanations are single-gene based, such as the expansion of TP53 in elephants [4] or the unique extracellular matrix molecules that provide naked mole rats their resistance to cancer [3]. However, whole molecular systems, not just individual genes, regulate cancer suppression. This leaves open the question of whether strong gene interaction networks, rather than single mutations, better explain why certain species are resistant to cancer.

This study addresses the need to shift from single-gene models to systems-level analyses to understand cancer resistance across mammals better.

Existing Literature

Peto [1] first presented the cancer incidence paradox, which was later formalised into evolutionary theory by Caulin and Maley in 2011, who proposed that natural selection produces diverse cancer-suppression mechanisms across lineages [2].

Since then, lineage -specific adaptations have been identified. Naked mole rats are cancer-resistant because they accumulate high-molecular-weight hyaluronan, which inhibits uncontrolled proliferation [3]. Elephants carry approximately 20 copies of TP53 gene, which promotes apoptosis [4]. Later assessment showed that not all copies are equally functional,

pushing for a comprehensive explanation [9]. In contrast, whales have developed enhanced DNA repair pathways and tumour suppression mechanisms [5].

Comparative genomic research indicates extensive mechanisms are involved. A recent analysis revealed that the prevalence of cancer in mammals is correlated to the variation in gene copy numbers in pathways linked to cancer [6]. Large-scale cross-species analyses have also discovered genes linked to cancer resistance that may mediate human cancer risk [7]. Rapid molecular molecular changes across lineages is shown by evolutionary studies of the p53 network, which is in line with cancer-related selection pressures [8].

Reassesments of the elephant TP53 story argue that concentrating only on single-gene expansions runs the risk of exaggerating their significance [9]. Even widely circulated public narratives, like the NIH summary *How Elephants Beat Cancer* [10], highlights how much of the field's discourse has centered on TP53, leaving other systemic explanations less explored.

Expanding beyond single genes has been emphasised in recent work. Such as, studies on the evolutionary role of TP53 indicate that it might serve a variety of purposes [11]. While Bgee [14] allows for cross-species expression comparisons, databases like STRING [12] and BioGRID [13] provide the tools to investigate entire interaction networks at the systems level. The significance of employing a variety of resources to guarantee coverage and accuracy is further highlighted by a comparative analysis of interaction databases [15].

Research Question

Is resistance in large, long-lived mammals better explained by the robustness of cancer-related gene networks than by single-gene adaptations?

3

Research Topic Justification

This study is novel because it reframes Peto's paradox as a network problem. It suggests that strong, fail-safe systems evolved to suppress cancer rather than concentrating on distinct single-gene adaptations. This has direct biomedical relevance: enhancing human pathway redundancy rather than concentrating on specific targets may be a more effective way to prevent cancer.

Methodology

Study Design

This study will adopt a comparative computational biology approach, including genomic, transcriptomic and protein interaction data across a diverse set of mammalian species.

Data Collection

The initial step involves selecting around 15-20 mammalian species of a broad spectrum of body sizes and lifespans. Such include small, ephemeral species like mice and rats, medium-sized species like dogs, humans as intermediate and large, long-lived species like elephants and whales. Additionally, cancer-resistant diminutive species like bats and naked mole rats will be included for comparative analysis. Life-history traits, inclusively body mass and lifespan, will be retrieved from databases such as AnAge and PanTHERIA, whereas cancer incidence data will be obtained from published comparative oncology research [6], [7].

Genomic assemblies and annotations for each species will be sourced from Ensembl and NCBI Genome, with orthologous genes identified through resources like OrthoDB or Ensembl Compara. This data will provide the basis for identifying essential cancer-related genes

implicated in tumour suppression, DNA repair and apoptosis. STRING-db [12] and BioGRID [13] will be utilised to reconstruct protein interaction networks.

Transcriptomic datasets will be integrated to capture the baseline expression of these oncogenes. RNA-seq data from normal tissues, including liver and blood, will be sourced from databases like Bgee [14] and the NCBI SRA. Gene expression values will be standardised to transcripts per million (TPM) to facilitate cross-species comparability.

Network Construction

Protein-protein interaction networks will be established for each species, concentrating on pathways relevant to cancer suppression, such as DNA repair, apoptosis and immune surveillance. Robustness within each network will be assessed based on redundancy, characterised by the existence of multiple alternative pathways for a biological process, connectivity, measured as the average number of interactions per gene and fault tolerance, defined as the network's resilience when individual nodes are computationally eliminated [15]. Superimposing normalised gene expression data will facilitate the identification of hub genes that are both structurally central and highly active in large, long-lived species.

Statistical Analysis

The robustness metrics will be compared to body size, lifespan and cancer incidence data to test whether network-level characteristics explain cancer resistance more effectively than single-gene measures. Statistical models will be implemented using Phylogenetic Generalised Least Squares (PGLS) [8] to account for non-independence of species resulting from shared evolutionary history. Model comparisons will be made between network robustness, TP53 copy number and combined predictors [4], [9], with performance assessed via explained variance (R^2) and the Akaike Information Criterion (AIC). Visualisations, including scatterplots, heatmaps and phylogenetic tree, will depict the correlations between robustness and species characteristics.

Ethical considerations

Only publicly available genomic, transcriptomic and interaction databases (Ensembl, NCBI, STRING and Bgee) are used in this project. No new human or animal data will be collected and no laboratory experiments will be conducted. This way, neither the welfare of animals nor human participants is at risk. All sources will be appropriately referenced, guaranteeing ethical and open research procedures.

Expected Outcomes

It is anticipated that this study will demonstrate that the cancer-related gene networks of large, long-lived mammals are more redundant, connected and resilient to disturbance than those of smaller, cancer-prone species. Compared to single-gene adaptations like TP53 copy number expansion, these robust systems are likely better able to explain cancer resistance. Transcriptomic data integration should also show that resistant species exhibit strong expression of highly connected hub genes, underscoring evolution's dependence on robust molecular systems rather than discrete genetic alterations.

Limitations

Data on cancer incidence in different species is still limited and inconsistent, frequently found only in zoo records rather than in systematic surveys [15]. In rare species, interactions may only depend on interference because protein interaction networks are skewed toward well-studied organisms. The tissue type and coverage of transcriptomic data also differ. Lastly, as this is a computational study, the findings will show correlations but not confirmation without wet-lab experimental validation.

References

[1] R. Peto, "Epidemiology, multistage models, and short-term mutagenicity tests," *International Journal of Epidemiology*, vol. 45, no. 3, pp. 621–637, Jun. 2016, doi: 10.1093/ije/dyv199.

[2] A. F. Caulin and C. C. Maley, "Peto's Paradox: evolution's prescription for cancer prevention," *Trends in Ecology & Evolution*, vol. 26, no. 4, pp. 175–182, Feb. 2011, doi: 10.1016/j.tree.2011.01.002.

[3] X. Tian *et al.*, "High-molecular-mass hyaluronan mediates the cancer resistance of the naked mole rat," *Nature*, vol. 499, no. 7458, pp. 346–349, Jun. 2013, doi: 10.1038/nature12234.

[4] M. Sulak *et al.*, "TP53 copy number expansion is associated with the evolution of increased body size and an enhanced DNA damage response in elephants," *eLife*, vol. 5, Sep. 2016, doi: 10.7554/elife.11994.

[5] M. Tollis, J. D. Schiffman, and A. M. Boddy, "Evolution of cancer suppression as revealed by mammalian comparative genomics," *Current Opinion in Genetics & Development*, vol. 42, pp. 40–47, Feb. 2017, doi: 10.1016/j.gde.2016.12.004.

[6] S. Matthews, V. N. Fard, M. Tollis, and C. Seoighe, "Variable gene copy number in cancer-related pathways is associated with cancer prevalence across mammals," *Molecular Biology and Evolution*, Mar. 2025, doi: 10.1093/molbev/msaf056.

[7] N. U. Nair *et al.*, "Cross-species identification of cancer resistance–associated genes that may mediate human cancer risk," *Science Advances*, vol. 8, no. 31, Aug. 2022, doi: 10.1126/sciadv.abj7176.

[8] C. N. Passow, A. M. Bronikowski, H. Blackmon, S. Parsai, T. S. Schwartz, and S. E. McGaugh, "Contrasting Patterns of Rapid Molecular Evolution within thep53Network across Mammal and Sauropsid Lineages," *Genome Biology and Evolution*, vol. 11, no. 3, pp. 629–643, Jan. 2019, doi: 10.1093/gbe/evy273.

[9] L. Nunney, "Cancer suppression and the evolution of multiple retrogene copies of TP53 in elephants: A re‑evaluation," *Evolutionary Applications*, vol. 15, no. 5, pp. 891–901, Apr. 2022, doi: 10.1111/eva.13383.

[10] S. J. Gaughran, E. Pless, and S. C. Stearns, "How elephants beat cancer," *eLife*, vol. 5, Oct. 2016, doi: 10.7554/elife.21864.

[11] K. Voskarides and N. Giannopoulou, "The role of TP53 in adaptation and evolution," *Cells*, vol. 12, no. 3, p. 512, Feb. 2023, doi: 10.3390/cells12030512.

[12] STRING, "STRING: functional protein association networks," *STRING*, Jul. 26, 2023. https://string-db.org/

[13] M. T. Lab, "BioGRID Interaction Database," *Copyright © 2015 Mike Tyers, All Rights Reserved.* https://thebiogrid.org/

[14] Bgee - Bring Gene Expression Expertise, "Bgee: gene expression data in animals," *Bgee*. https://bgee.org/

[15] A. K. Bajpai *et al.*, "Systematic comparison of the protein-protein interaction databases from a user's perspective," *Journal of Biomedical Informatics*, vol. 103, p. 103380, Jan. 2020, doi: 10.1016/j.jbi.2020.103380.

Research Proposal by Aaria Rahman - Aaria Rahman.pdf

9 Zahra Ziaastani, Behjat Kalantari-Khandani, Mohammad-Javad Niazi, Ali Kazemipour. "Identification of critical genes and metabolic pathways in rheumatoid arthritis and osteoporosis toward drug repurposing", Computers in Biology and Medicine, 2024
Publication

1%

10 Sophie Matthews, Vahid Nikoonejad Fard, Marc Tollis, Cathal Seoighe. "Variable gene copy number in cancer-related pathways is associated with cancer prevalence across mammals", Molecular Biology and Evolution, 2025
Publication

1%

11 www.science.org
Internet Source

1%

12 Emma Palefsky, Himani K. Patel, Trey A. Doss, Emilye C. Eischeid et al. "Characterizing Elephant MDM2's Role in p53 Regulation", The FASEB Journal, 2025
Publication

<1%

13 Linxia Sun, Zhikang Xu, Mengqi Shuai, Chengxu Li, Guang Yang, Shixia Xu. "Natural resistance to cancers in long-lived mammals: genomic mechanisms and experimental evidence to explain Peto's paradox", Science China Life Sciences, 2025
Publication

<1%

14 experts.umn.edu
Internet Source

<1%

15 Faheem Ahmed, Anupama Samantasinghar, Afaque Manzoor Soomro, Sejong Kim, Kyung Hyun Choi. "A systematic review of computational approaches to understand

cancer biology for informed drug repurposing", Journal of Biomedical Informatics, 2023

Publication

16 discovery.researcher.life
    Internet Source
                                                        <1%

17 accounts.public.ce.basespace.illumina.com
    Internet Source
                                                        <1%

Exclude quotes   On   Exclude matches   Off
Exclude bibliography   Off