RESEARCH PAPER

The evolutionary study of oncogenes and tumor suppressor genes in breast carcinoma among different vertebrate ancestors

Saima Shahzad Mirza[1], Mustafa Azizoglu[2], Farheen Aslam[3], Dibakar Roy[4], Hassan Raza Heral[5], Zainab Awais[6], Ayesha Aihetasham[7], Ayesha Siddiqa[8], Muhammad Waseem[9], Ayesha Masood[10], Shagufta Iram[11], Asma Rasheed[12]

[1]Microbial Bioengineering Laboratory, Department of Zoology, University of Education, Bank Road Campus, Lahore, Pakistan..
Email; saima.mirza@ue.edu.pk, ORCID: 0000-0002-3287-9879
[2]Department of Pediatric Surgery, Esenyurt Necmi Kadioglu State Hospital, Istanbul 34515, Turkey, Department of Stem Cell and Tissue Engineering & 3D Bioprinting, Istinye University, Istanbul 34460, Turkey.
Email: mdmazizoglu@gmail.com. ORCID: 0000-0002-8266-5203.
[3]Associate Professor, Department of Biotechnology, Lahore College for Women University, Pakistan.
Email: farheen.aslam@lcwu.edu.pk. ORCID: 0000-0003-4380-2401
[4]Department of Chemistry and Chemical Biology, Indiana University, Indianapolis, IN 46202.
Email: roydi@iu.edu: ORCID:https://orcid.org/0009-0003-5980-6076
[5]Assistant Professor, Department of Pathology, Faryal Dental College, Lahore, Pakistan.
Email: Hassanraza1245@gmail.com ORCID:0009-00039488-9188
[6]Department of Pathology, University college of Medicine and Dentistry, University of Lahore, Lahore, Pakistan.
Email: zainab_syed@hotmail.com. ORCID: 0009-0008-0745-418.
[7]Assistant Professor, Institute of Zoology, University of the Punjab, Quaid E Azam Campus, Lahore, Pakistan.
Email: ayesha.zool@pu.edu.pk. ORCID: 0000-0003-0533-6366
[8]Institute of Microbiology and Molecular Genetics, University of the Punjab Lahore, Pakistan. Email: Ayesha.mmg@pu.edu.pk. ORCID: 0000-0002-6046-809X
[9]Research Cell, University College of Medicine and Dentistry, university of Lahore, Pakistan.
Email: Waseem.mmg.pu@gmail.com. ORCID: 0000-0002-0819-129X.
[10]Professor in Department of Pathology, University College of Medicine and Dentistry, The University of Lahore, Lahore, Pakistan.
Email: ayesha.masood@ucm.uol.edu.pk. ORCID: 0000-0002-8386-506.
[11] Head of Department, Department of Pathology University College of Medicine and Dentistry, university of Lahore, University of Lahore.
ORCID: 0009-0008-3186-7068
[12]Professor Chemical Pathology, University College of Medicine and Dentistry, The University of Lahore, Pakistan.
Email: asma.rasheed@ucm.uol.edu.pk. ORCID: 0009-0003-8314-309X

ABSTRACT

Introduction: Cancer involves abnormal cell proliferation and migration throughout the body. Genetic changes occurring in hereditary breast cancer lead to the activation of oncogenes, resulting in gain-of-function effects. This study was designed to conduct a phylogenetic analysis of breast cancer-related oncogenes and tumor suppressor genes families in humans relevant to different species.

Methods: Data was retrieved from ensemble genome browser and NCBI. BLAST research executed for FASTA sequence of selected species. Phylogenetic analysis was conducted using MEGA X. Alignment was achieved through Clustal W, leading to the creation of neighbor-joining tree in circular form. Rigorous validation was performed using bootstrap method, employing 500 pseudo replicates.

Results: Among the eleven oncogenes, the CBLC and JUN gene has two paralogues, BCL3, DDX6, FEV, MYB, KARAS, LCK, NUP214 and MPL had 5, while TFG have no paralogues. In the 10 tumor suppressor genes, the CDH1, JAK2, PML, PTEN, GPC3, SOCS1, gene had 5 paralogues, EXT1 and WRN gene has 4, NPM1 has 2, VHL has 1 paralogue. Furthermore, phylogenetic trees were constructed for 21 gene families, encompassing 10 tumor suppressor genes and 11 oncogenes, utilizing vertebrate sequences. Discussion: The analysis revealed variable paralogue distribution among oncogenes and tumor suppressor genes, ranging from none to five. Phylogenetic trees of 21 gene families highlighted their evolutionary relationships across vertebrates.

*****Author for Correspondence: farheen.aslam@lcwu.edu.pk*

.

Conclusion: In conclusion, variability in paralogue numbers among oncogenes and tumor suppressor genes, with some genes having multiple paralogues and others, like TFG and VHL, having few or none was found. Phylogenetic trees highlighted evolutionary relationships**..**

Keywords: Breast cancer, Oncogenes, Tumor suppressor genes, Evolution, Mammals.
How to cite this article: Mirza SS, Azizoglu M, Aslam F, Roy D, Heral HR, Awais Z, Aihetasham A, Siddiqa A, Waseem M, Masood A, Iram S, Rasheed A, The evolutionary study of oncogenes and tumor suppressor genes in breast carcinoma among different vertebrate ancestors..Int J Drug Deliv Technol. 2026;16 (1s): 383-396; DOI: 10.25258/ijddt.16. 383-396
Source of support:. None
Conflict of interest: None

Graphical abstract

The graphical abstract is shown in the figure 1.

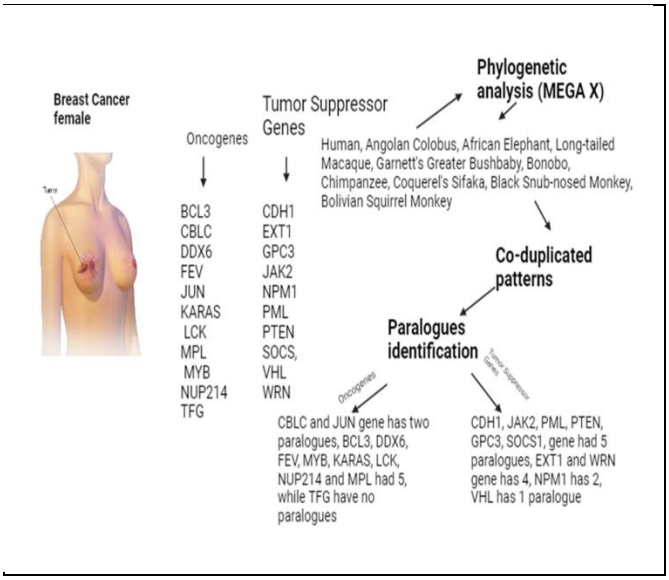

Fig 1: Indicates different oncogenes and tumor suppressor genes relationship and their paralogues among different vertebrates and humans

INTRODUCTION

Breast cancer (BC) refers to malignant neoplasms originating from breast tissue, most commonly arising from the epithelial lining of the milk ducts. In the United States, BC is the most frequently diagnosed cancer among women and ranks as the second leading cause of cancer-related mortality [1]. The TNM classification system is widely employed for staging BC, assessing tumor size (T), regional lymph node involvement (N), and distant metastasis (M). Stage 0 represents carcinoma in situ, including both ductal carcinoma in situ (DCIS) and lobular carcinoma in situ (LCIS). Stages 1–3 indicate increasing local and regional spread, including lymph node involvement, whereas Stage 4 denotes metastatic disease, generally associated with a poor prognosis [2]. Globally, BC accounts for approximately 25% of all cancer cases in women, with an estimated 1.67 million new diagnoses reported in 2012. Incidence rates are notably higher in less developed countries compared to more developed regions [3]. According to the World Health Organization, approximately 2.26 million new BC cases were diagnosed in 2020, representing one in every eight cancer diagnoses worldwide, making BC one of the most aggressive and prevalent malignancies globally [4]. Standard therapeutic approaches i.e., chemotherapy, radiotherapy, surgery, and immunotherapy—remain limited in efficacy and often yield suboptimal long-term outcomes [5]. On a global scale, BC leads all cancers in women, with approximately 2 million new cases and 600,000 deaths annually. In Middle East and North Africa (MENA) region, the highest incidence rates have been reported in Algeria and Iraq (≥60/100,000), while the lowest occur in Saudi Arabia and Yemen (<30/100,000). Mortality rates are highest in Iraq, the Syrian Arab Republic, Algeria, and Sudan (>20/100,000), and lowest in Saudi Arabia (7.6/100,000). Although incidence in MENA remains lower than in most world regions, mortality rates (16.9/100,000) surpass those of all other regions except Sub-Saharan Africa. Among women younger than 50 years, incidence rates are 5.5 times lower than in women aged ≥50 years and remain below those observed in similarly aged populations in Western countries. Projections indicate that by 2050, the BC burden in the MENA region will reach approximately 219,000 new cases and 88,900 deaths, reflecting increases of 86% and 117%, respectively [6].

Molecular-level investigations of cancer cells revealed wide spectrum of genetic alterations involving various combinations of oncogenes (OGs) and tumor suppressor genes (TSGs) [7]. In hereditary breast cancer (HBC), genetic mutations often activate OGs, leading to gain-of-function effects, while concomitantly inactivating TSGs through loss-of-function mutations. These alterations disrupt normal cell cycle regulation, inhibit proper cell cycle checkpoints, and impair DNA repair mechanisms. Epigenetic dysregulation further contributes to the aberrant expression and function of these genes. Historically, TSGs were referred to as recessive oncogenes due to their characteristic inheritance patterns [8]. Examples include H-RAS as a prototypical oncogene and the retinoblastoma protein (RB) as a key TSG, both identified through approaches such as functional assays, linkage analysis, positional cloning, and studies of genetically predisposed individuals [9]. Comparative genomic analyses have identified numerous genes subject to duplication or deletion in cancer, with major contributions arising from the Human Genome Project [10]. Mutations in *BRCA1* and *BRCA2* substantially increase the risk of both breast and ovarian cancers, while rare hereditary variants in *TP53* (encoding the p53 protein) and *PTEN* are associated with elevated susceptibility to multiple cancer types, including BC [11, 12].

Additional germline mutations in *ATM, CHK2, NBS1, RAD50, PALB2*, and *BRIP1* have been linked to heightened BC risk. Among these, *ATM* and *CHK2* encode kinases central to the DNA damage response, *RAD50* is essential for DNA repair, *PALB2* encodes a protein that interacts with *BRCA2*, and *BRIP1* encodes a *BRCA1*-interacting DNA helicase. These genes are also implicated in the pathogenesis of other malignancies, including pancreatic and prostate cancers [13]. The figure 1 describes the details in graphical form.

Throughout Earth's evolutionary history, there has been a clear progression from simple to increasingly complex life forms. For instance, life evolved from unicellular organisms, such as bacteria, to multicellular placozoans, and from diploblastic organisms to bilaterians possessing a third germ layer [14]. Likewise, chordates gave rise to vertebrates [15]. These major evolutionary transitions were often accompanied by the emergence of novel structures and functions, driven in part by genetic innovations such as gene duplications [16]. It has been proposed that extensive gene duplication events occurred at the origin of the vertebrate lineage, contributing to the widespread presence of gene families in modern vertebrates [17-20]. The expansion in gene number has been linked to the increased morphological and anatomical complexity observed in vertebrates compared to non-vertebrate chordates, such as cephalochordates and tunicates. The presence of paralogons in human and other vertebrate genomes has led to the widely discussed two-round (2R) whole-genome duplication hypothesis, which proposes that two successive whole-genome duplications occurred early in vertebrate evolution after divergence from amphioxus-like invertebrate ancestors [21-25]. In contrast, an alternative model suggests that the high number of paralogs in vertebrates could result from continuous, small-scale duplications of individual genes rather than large block duplications [26-30]. If all mammalian cells are equally susceptible to oncogenic transformation, the risk of cancer would theoretically increase proportionally with body mass and lifespan. However, Peto's paradox describes the observation that cancer incidence does not scale as predicted in large, long-lived animals [31]. In humans, a single copy of *TP53* (two alleles) is essential for preventing tumor development. This study aims to investigate the evolutionary history of oncogenes and tumor suppressor genes within human gene families in relation to other species, with a particular focus on *Peto's paradox* and its implications for breast cancer susceptibility in large-bodied animals. An overview of this relationship is presented in Figure 2.

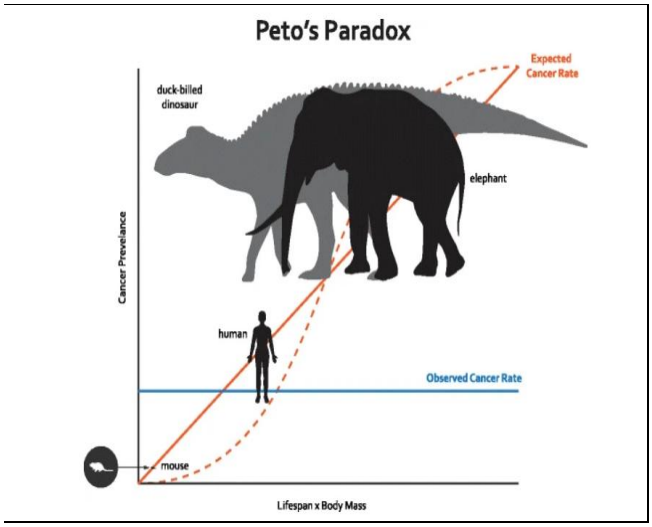

Figure 2: Describes the Peto's paradox among vertebrates sourced from the study [17]

MATERIAL AND METHODS

Species selection

About ten species were selected during this study. These included *Homo sapiens* (Humans), *Colobus angolensis palliates* (Angolan Colobus), *Loxodonta africana* (African Elephant), *Macaca fasciculari* (Long-tailed Macaque), *Otolemur garnettii* (Garnett's Greater Bushbaby), *Pan paniscus* (Bonobo), *Pan troglodyte* (Chimpanzee), *Propithecus coquereli* (Coquerel's Sifaka), *Rhinopithecus bieti* (Black Snub-nosed Monkey), and *Saimiri boliviensis boliviensis* (Bolivian Squirrel Monkey).

Oncogenes and tumor suppressor genes under study

Twenty gene families of OGs and TSGs families were selected, out of which. Eleven were OGs, including *BCL3, CBLC, DDX6, FEV, JUN, KARAS, LCK, MPL, MYB, NUP214* and *TFG* and Ten were tumor suppressor genes including *CDH1, EXT1, GPC3, JAK2, NPM1, PML, PTEN, SOCS1, VHL* and *WRN.*

Study limitations

This study is limited to the oncogenes and tumor suppressor genes for breast canacer in different vertebrates selected and humans to study the effect of gene duplications.

Paralogue sequence collection

The paralogue sequences were completed through the utilization of BLASTP within the Ensemble genome browser, accessible at this URL: https://www.ensembl.org/index.html?redirect=no [32]. In order to augment genes families with sequences from species for which sequences was unavailable in Ensemble, a BLASTP analysis was conducted. The information search was conducted using the protein database, which could be accessed via NCBI (https://www.ncbi.nlm.nih.gov/) and Joint Genome Institute (http://www.jgi.doe.gov/).

Sequences alignment and phylogenetic analysis

For phylogenetic analysis of each genes family, we employed MEGA X version 5 [33]. Additionally, we utilized CLUSTAL W, a multiple alignment tool, with default parameters as described by, [34] to align the amino acid

sequences. Amino acid sequence alignments were manually refined when necessary. For the pairwise alignment, the gap opening penality was kept 10, and gap extension penality was kept 0.1. For the multiple alignments, the gap opening penality was kept 10, and gap extension penality was kept 0.2. Moreover, the gap separation distance was kept 4 and all other parameters were kept as default and run the alignment by CLUSTAL W.

Neighbor-joining method

Phylogenetic trees for each gene family were reconstructed using the Neighbor-joining (NJ) method (NJ) as outlined by Russo [35], Subsequently, the total deletion option was applied to eliminate gaps in the amino acid. For amino acid substitution, we employed the False proportion (p) method along with the passion-corrected amino acid distance, as our chosen approach. As both methods yielded identical results, only the outcomes derived from (NJ) tree, utilizing uncorrected p-distance values, are presented here. Protein sequences that exhibited deviations causing disruptions in the alignment were excluded. The accuracy of the analysis was affirmed through the use of the bootstrap method with 500 pseudo-replicates, which generated the bootstrap probability for every internal branch within the tree [36, 37], using the MEGA X version 5.

Co-duplicated genes families study

Among the topology of 20 genes families, the trees of 20 gene families, including 10 genes *CDH1, EXT1, GPC3, JAK2, NPM1, PML, PTEN, SOCS1, VHL* and *WRN* from TSGs and 10 genes including *BCL3, CBLC, DDX6, FEV, JUN, KARAS, LCK, MPL, MYB, NUP214* and *TFG* from OGs were studied with vertebrates sequences.

RESULTS

This work was conducted in response to previous work by [17], by targeting other OGs and TSGs to explore the functions and evolutionary processes of other breast cancer associated genes.

Oncogenes and tumor suppressor genes families

The table 1 comprises of all OGs and table 2 consists of TSGs under study. The table shows gene members, location on chromosomes, proteins accession number, and molecular functions.

Table 1: The oncogenes with the molecular functions

Genes	Member	Chromosomal Location	Proteins Accession number	Molecular Functions
BCL3				
	NFKBIE	6	O00221	Inhibits NF-kappa-B Inhibits DNA-binding
	POTEJ	2	P0CG39	pk binding
				structural component of postsynaptic actin cytoskeleton
	POTEM	14	A6NI47	Virus receptor activity
				dna binding rna binding
	POTEG	14	Q6S5H5	-------
	POTEB2	15	H3BUK9	-------
CBLC				
	CBLB	3	Q13191	Phosphotyrosine residue binding calcium ion binding Zinc ion binding
	CBL	11	P2268	calcium ion binding ephrin receptor binding

DDX6				
	DDX54	12	Q8TDD1	ATP binding nuclear estrogen receptor binding RNA binding
	EIF4A2	3	Q14240	ATP hydrolysis activity RNA binding
	DDX53	X	Q86TM3	RNA binding RNA helicase activity
	DDX4	5	Q9NQI0	ATP hydrolysis activity molecular condensate scaffold activity
	DDX49	19	Q9Y6V7	RNA binding RNA helicase activity
FEV				
	ERFL	19	A0A1W2PQ73	DNA-binding Tf activity sequence-specific DNA binding
	ELK1	X	P19419	chromatin binding DNA-binding transcription factor activity
	SPI1	11	P17947	chromatin binding protein sequestering activity
	ETV1	7	P50549	DNA-binding Tf activity sequence-specific double-stranded DNA binding
	ETV6	12	P41212	DNA-binding Tf activity protein domain specific binding
MYB				

	SNAPC4	9	Q5SXM2	DNA binding specific DNA binding
	MYBL2	20	P10244	sequence-specific DNA binding
	MYBL1	8	P10243	RNA polymerase II-specific sequence-specific DNA binding
	DMTF1	7	Q9Y222	RNA polymerase II-specific DNA-binding Tf activity
	CDC5L	6	Q99459	DNA binding identical protein binding RNA binding
MPL				
	EPOR	19	P19235	erythropoietin receptor activity identical protein binding
	IL21R	16	Q9HBE5	cytokine receptors activity trans membrane signaling receptors activity
	IL2RB	22	P14784	coreceptor activity interleukin-2 receptor activity
	CSF2RB	22	P32927	coreceptor activity signaling receptor activity
	IL4R	16	P24394	cytokine receptor activity interleukin-4 receptor activity
NUP214				
	NPAP1	15	Q9NZP6	nuclear localization sequence binding structural

				constituent of nuclear pore
	POM121	7	Q96HA1	------
	POM121L12	7	Q8N7R1	------
	POM121L2	6	Q96KW2	------
	POM121C	7	A8CG34	------
TFG				No paralogues
JUN				
	JUNB	19	P17275	DNA binding
	JUND	19	P17535	enzyme binding nuclear receptor binding
LCK				
	JAK1	1	P23458	ATP binding growth hormone receptor binding
	PTK2	8	Q05397	actin bindingSource molecular function activator activity
	FER	5	P16591	actin binding lipid binding
	SRMS	20	Q9H3Y6	ATP binding signaling receptor binding
	TNK1	17	: Q13470	ATP binding protein tyrosine kinase activity
KRAS				
	DIRAS3	1	O95661	GDP binding

				GTPase activity
	ERAS	X	Q7Z444	G protein activity GDP binding
	RASL12	15	Q9NYN1	G protein activity GDP binding
	HRAS	11	P01112	G protein activity GDP binding
	RAP1A	1	P62834	G protein activity
				guanyl-nucleotide exchange factor activity

Table 2: The tumor suppressor gene families with the molecular functions

Genes	**Member**	**Chromosomal Location**	**Proteins Accession numbers.**	**Molecular Functions**
CDH1				
		16	P55287	calcium ion binding
				adherens junction organization
	CDH8	16	P55286	identical protein binding calcium ion binding
	CDH17	8	Q12864	integrin binding calcium ion binding
	PCDH7	4	O60245	cyclase activity antioxidant activity oxidoreductase activity
	CDH20	18	Q9HBT6	ligase activity protein tag activity
JAK2				
	JAK1	1	P23458	ATP binding
				growth hormone receptor binding
	LCK	1	P06239	identical protein binding protein kinase binding
	PTK2	8	Q05397	integrin binding
				protein phosphatase binding
	FER	5	P16591	lipid binding
				protein tyrosine kinase binding

	SRMS	20	Q9H3Y6	ATP binding signaling receptor binding
PML				
	TRIM66	11	O15016	zinc ion binding nucleoplasm
	TRIM5	11	Q9C035	protein homodimerization activity pk binding
	TRIM10	6	Q9UDY6	ubiquitin protein ligase activity zinc ion binding
	TRIM43	2	Q96BQ3	------
	TRIM43B	2	A6NCK2	oxidoreductase activity transferase activity
PTEN				
	TRIM66	11	O15016	zinc ion binding cyclase activity
	TRIM5	11	: Q9C035	transcription coactivator activity ubiquitin protein ligase activity
	TRIM10	6	Q9UDY6	zinc ion binding lyase activity
	TRIM43B	2	A6NCK2	structural molecule activity transporter activity
EXT1				
	EXTL3	8	O43909	magnesium ion binding protein-hormone receptor activity
	EXT2	11	Q93063	metal ion binding protein heterodimerization activity
	EXTL1	1	Q92935	histone binding protein folding chaperone
	EXTL2	1	Q9UBQ6	manganese ion binding lyase activity
NPM1				

	NPM2	8	Q86SE8	histone binding RNA binding
	NPM3	10	O75607	chromatin binding histone binding
VHL	No paralogues			
	VHLL	1	Q6RSH7	dominant-negative VHL to serve as a protector of HIFalpha
WRN				
	RECQL4	8	O94761	ATP binding bubble DNA binding
	RECQL5	17	O94762	ATP binding enzyme activator activity
	BLM	15	P54132	DNA/DNA annealing activity DNA helicase activity
	RECQL	12	P46063	DNA helicase activity ATP hydrolysis activity
GPC3				
	GPC5	3	P78333	dna binding rna binding
	GPC2	7	Q8N158	ligase activity protein tag activity
	GPC4	X	O75487	catalytic activity gtpase activity
	GPC1	2	P35052	copper ion binding fibroblast growth factor binding
	GPC6	13	Q9Y625	isomerase activity ligase activity
SOCS1				
	GPC5	13	P78333	cyclase activity antioxidant activity
	GPC2	7	Q8N158	rna metabolic process protein folding
	GPC4	X	O75487	hydrolase activity lyase activity
	GPC1	2	P35052	copper ion binding fibroblast growth factor binding

	GPC6	13	Q9Y625	ligase activity protein tag activity

Phylogenetic study of genes

The evolutionary trees of 21 genes families, i.e., 10 genes *CDH1, EXT1, GPC3, JAK2, NPM1, PML, PTEN, SOCS1, VHL* and *WRN* from TSGs and 11 genes including B*CL3, CBLC, DDX6, FEV, JUN, KARAS, LCK, MPL, MYB, NUP214* and *TFG* from OGs were embedded with vertebrates and invertebrates.

Oncogenes

The trees for oncogenes under study with different species have been designed and the paralogues were also described. The gene families are listed in the table 1 for oncogenes as these are the paralogues and are duplicated among vertebrates and humans.

A comprehensive comparative genomic analysis revealed the distribution of multiple oncogenes across ten studied species. The *BCL3* gene (Figure 1, Supplementary File S1) encodes a protein implicated in breast cancer, with five paralogues, and is present in all studied species. The *CBLC* gene (Figure 2, Supplementary File S2) has two paralogues (*CBLB* and *CBL*), functions in phosphotyrosine, calcium, and zinc ion binding, and is present in all species. The *DDX6* gene (Figure 3, Supplementary File S3) possesses five paralogues (*DDX5*, *EIF4A2*, *DDX53*, *DDX4*, *DDX49*), functions in ATP and RNA binding, and is present in all species. The *FEV* gene (Figure 4, Supplementary File S4) has five paralogues, plays a role in evolutionary adaptation, and is absent in *Loxodonta africana* and *Otolemur garnettii*. The *JUN* gene (Figure 5, Supplementary File S5) has two paralogues (*JUND* and *JUNB*), regulates transcription, and is present in all species. The *KARAS* gene (Figure 6, Supplementary File S6) has five paralogues (*DIRAS3*, *ERAS*, *RASL12*, *HRAS*, *RAP1A*) and is absent in *Otolemur garnettii*. The *LCK* gene (Figure 7, Supplementary File S7) has five paralogues, is involved in T-cell signaling, and is absent in *Rhinopithecus bieti*. The *MPL* gene (Figure 8, Supplementary File S8) has five paralogues, participates in cytokine receptor-mediated signaling, and is absent in *Macaca fascicularis* and *Saimiri boliviensis boliviensis*. The *MYB* oncogene (Figure 9, Supplementary File S9) has five paralogues, regulates cell proliferation and apoptosis, and is present in all species. The *NUP214* gene (Figure 10, Supplementary File S10) has five paralogues, encodes a nuclear pore complex protein, and is present in all species. Finally, the *TFG* gene (Figure 11, Supplementary File S11) has no known paralogues and is absent in *Loxodonta africana*, *Rhinopithecus bieti*, and *Saimiri boliviensis boliviensis*.

Tumor suppressor genes

The trees for tumor suppressor genes under study with different species have been designed and the paralogues were also described. The gene families are listed in the table 2 for tumor suppressor genes as these are the paralogues and are duplicated among vertebrates and humans.

The *CDH1* gene (Figure 12, Supplementary File S12), with five paralogues, is involved in calcium ion binding and adherens junction organization, and is present in all studied species. The *EXT1* gene (Figure 13, Supplementary File S13) has four paralogues (*EXTL3*, *EXT2*, *EXTL1*, *EXTL2*) and is conserved across all species. The *GPC3* gene (Figure 14, Supplementary File S14) has five paralogues and is also present in all species. The *JAK2* gene (Figure 15, Supplementary File S15), with five paralogues, shows evolutionary conservation and key roles in cellular signaling, present in all organisms examined. The *NPM1* gene (Figure 16, Supplementary File S16) has two paralogues (*NPM2*, *NPM3*) and is present in all species. The *PML* gene (Figure 17, Supplementary File S17) possesses five paralogues and is present in all species. The *PTEN* gene (Figure 18, Supplementary File S18) has five paralogues and is conserved in all studied organisms. The *SOCS1* gene (Figure 19, Supplementary File S19) has five paralogues (*GPC5*, *GPC2*, *GPC4*, *GPC1*, *GPC6*) and is present in all species. The *VHL* gene (Figure 20, Supplementary File S20) has one paralogue and is present in all species. Finally, the *WRN* gene (Figure 21, Supplementary File S21) has four paralogues and is present in all species. For all these genes, no absence was detected across the ten species analyzed, showing their evolutionary conservation and potential functional indispensability.

DISCUSSION

This study was conducted by forwarding the work of [17] to explore the remaining genes study to explore the evolution of genes duplication patterns among different species by targeting the species as previously described by [17]. TSGs and OGs can be studied across vertebrate to examine how these genes have evolved. Similar to other mammals, cats and dogs exhibit comparable cancer incidence, suggesting that these genes follow a similar evolutionary pattern across mammals. In contrast, elephants possess more paralogous copies of these genes, which may contribute to their increased resistance to cancer [38].

They present an in-depth analysis of cancer gene duplications across mammals, examining 548 recognized human genes and 63 mammalian genes. The findings revealed that 87% of human genes are replicated in about one mammalian genome, with many lineage-specific expansions of cancer-related genes. The researchers concluded the naked mole rat exhibits highest number of cancer gene replications, related with studies, suggest naked mole rats possess highly effective cancer defense mechanisms and unique anti-aging traits [39-41] and currently [42], no cases of cancer have been observed in this species. In contrast, cancer has been noted in murine rodents, i.e., brown rats and mice [43]. The researchers found that these have more copies of cancer-related genes compared to most other mammals, with the exception of two-toed sloths. The sloths had the highest number of caretakers, somatic (TSGs), and both TSGs and oncogenes. This aligns with previous findings suggesting that rodent lineages exhibit more gene family expansions than other mammals.

Additionally, both two-toed sloths and nine-banded armadillos xenarthran mammals that diverged from other eutherians early in the Cenozoic also possess large numbers of cancer gene duplications. Despite being distantly related to rodents and primates, these species show significant gene replications. The study also noted that while sloths have not been reported to develop cancer, their fur contains an anticancer fungus [44-46]. Interestingly, these species do not have exceptionally large body sizes or long lifespans.

During this study, different species have their relation with breast cancer related genes i.e., the *BCL3* gene is present in various species, including *Colobus angolensis*, *Homo sapiens*, *Loxodonta africana*, *Macaca fascicularis*, *Otolemur garnettii*, *Pan paniscus*, *Pan troglodytes*, *Propithecus coquereli*, *Rhinopithecus bieti*, and *Saimiri boliviensis*, and has five paralogues. It is linked to breast cancer across these species. The *CBLC* gene, found in the same species, has two paralogues (*CBLB* and *CBL*) and functions in binding phosphotyrosine residues, calcium ions, and zinc ions. Bats have great metabolic rate and long life [47]. They found small brown bat were related with other laurasiatherians with many TSGs. The *DDX6* gene, with five paralogues (*DDX5, EIF4A2, DDX53, DDX4*, and *DDX49*), is involved in ATP and RNA binding and is present in the same organisms. The *FEV* gene, also with five paralogues, provides insights into evolutionary history and genetic diversity in these species. The *JUN* gene, with two paralogues (*JUND* and *JUNB*), is similarly widespread. The *KARAS* gene, with five paralogues (*DIRAS3*, *ERAS, RASL12, HRAS, RAP1A*), is found across the same species. The *LCK* gene, with five paralogues, and the *MPL* gene, with five paralogues, play roles in cellular signaling and are also present in these species. The *MYB* and *NUP214* genes, each with five paralogues, are identified across the same organisms, contributing to biological processes (i.e., cell cycle regulation). Finally, the *TFG* gene, which lacks paralogues, has a unique evolutionary history in these species.

For the tumor suppressor genes, The *CDH1* gene, found in species like *Colobus angolensis*, *Homo sapiens*, *Loxodonta africana*, *Macaca fascicularis*, *Otolemur garnettii*, *Pan paniscus*, *Pan troglodytes*, *Propithecus coquereli*, *Rhinopithecus bieti*, and *Saimiri boliviensis*, plays a role in calcium ion binding and adherens junction organization. It has five paralogues. The *EXT1* gene, also present in these organisms, has four paralogues (*EXTL3, EXT2, EXTL1, EXTL2*). The *GPC3* gene has five paralogues and is identified in the same species. *JAK2*, a gene important in signaling pathways, has five paralogues and is found across these species. The *NPM1* gene, with two paralogues (*NPM2, NPM3*), is involved in histone and RNA binding. The *PML* gene has five paralogues, as does the P*TEN* gene, in these species. The *SOCS1* gene has five paralogues (*GPC5, GPC2, GPC4, GPC1, GPC6*). The *VHL* gene, with one paralogue, shows unique evolutionary adaptations across these species. Finally, the WRN gene, with four paralogues, reflects evolutionary changes that have contributed to species survival.

CONCLUSION

In conclusion, the analysis of oncogenes and tumor suppressor genes reveals significant variability in the number of paralogues across different gene families, with some genes exhibiting multiple paralogues while others, such as *TFG* and *VHL*, have none or a single paralogue. All 10 tumor suppressor genes have paralogues, but out of 11, oncogenes genes include *BCL3, CBLC, DDX6, FEV, JUN, KARAS, LCK, MPL, MYB, NUP214*, have paralogues, and 1 gene have not paralogue include *TFG*. The construction of phylogenetic trees based on vertebrate sequences further signifies the evolutionary relationships among these genes. Furthermore, further research could include a more in-depth investigation into the functional implications of these paralogues, as well as exploring how the paralogue-based diversification might influence the development of cancer. Comparative studies across different species could provide additional insights into the evolutionary conservation or divergence of these gene families. More studies are needed to explore the work completely on other related genes.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

Not applicable.

HUMAN AND ANIMAL RIGHTS

No animals/humans were used for studies that are the basis of this research.

CONSENT FOR PUBLICATION

Not applicable.

AVAILABILITY OF DATA AND MATERIALS

Not applicable.

FUNDING

None.

CONFLICT OF INTEREST

The authors declare no conflict of interest, financial or otherwise.

REFERENCE

1. Sharma GN, Dave R, Sanadya J, Sharma P, Sharma K. Various types and management of breast cancer: an overview. Journal of advanced pharmaceutical technology & research. 2010;1(2):109.
2. Ji H, Hu C, Yang X, Liu Y, Ji G, Ge S, et al. Lymph node metastasis in cancer progression: molecular mechanisms, clinical significance and therapeutic interventions. Signal Transduction and Targeted Therapy. 2023;8(1):367.
3. Reyes-Monasterio A, Lozada-Martinez ID, Cabrera-Vargas LF, Narvaez-Rojas AR. Breast cancer care in Latin America: The ghost burden of a pandemic outbreak. International Journal of Surgery. 2022;104:106784.
4. Lv L, Zhao B, Kang J, Li S, Wu H. Trend of disease burden and risk factors of breast cancer in developing countries and territories, from 1990 to 2019: Results from the Global Burden of Disease Study 2019. Frontiers in public health. 2023;10:1078191.

5. Wang J, Wu S-G. Breast cancer: an overview of current therapeutic strategies, challenge, and perspectives. Breast Cancer: Targets and Therapy. 2023:721-730.
6. Zahwe M, Bendahhou K, Eser S, Mukherji D, Fouad H, Fadhil I, et al. Current and future burden of female breast cancer in the Middle East and North Africa region using estimates from GLOBOCAN 2022. International Journal of Cancer. 2025;156(12):2320-2329.
7. Dakal TC, Dhabhai B, Pant A, Moar K, Chaudhary K, Yadav V, et al. Oncogenes and tumor suppressor genes: functions and roles in cancers. MedComm. 2024;5(6):e582.
8. Hansen M. Genetic predisposition and cancer. Developmental Biology and Cancer: CRC Press; 2024. p. 277-312.
9. Peta C, Tsirimonaki E, Fedonidis C, Koliou X, Sakellaridis N, Mangoura D. Two tails for neurofibromin: a tale of two microtubule-associated proteins. Clinical and Basic Aspects of Neurofibromatosis Type 1: IntechOpen; 2021.
10. Bell DW. Our changing view of the genomic landscape of cancer. The Journal of Pathology: A Journal of the Pathological Society of Great Britain and Ireland. 2010;220(2):231-243.
11. Jamwal A, Chatterjee M, Hooda K, Jan N, Mir MA. Insights from Genetic Studies: The Role of p53 in Hereditary Breast Cancer Syndrome. p53 in Breast Cancer: CRC Press; 2025. p. 291-300.
12. Murawski M, Jagodziński A, Bielawska-Pohl A, Klimczak A. Complexity of the genetic background of oncogenesis in ovarian cancer—Genetic instability and clinical implications. Cells. 2024;13(4):345.
13. Fernandez A, Artola M, Leon S, Otegui N, Jimeno A, Serrano D, et al. Cancer Vulnerabilities Through Targeting the ATR/Chk1 and ATM/Chk2 Axes in the Context of DNA Damage. Cells. 2025;14(10):748.
14. Lamża Ł. Diversity of 'simple'multicellular eukaryotes: 45 independent cases and six types of multicellularity. Biological Reviews. 2023;98(6):2188-2209.
15. Malakhov V, Gantsevich M. The origin and main trends in the evolution of bilaterally symmetrical animals. Paleontological Journal. 2022;56(8):887-937.
16. Shlyakhtina Y, Moran KL, Portal MM. Genetic and non-genetic mechanisms underlying cancer evolution. Cancers. 2021;13(6):1380.
17. Nawaz Y, Zaib Khan A, Tanvir F, Ambreen S, Zafar J, Umar A, et al. Evolutionary History of Oncogenes and Tumor Suppressor Genes Families: The Ancestral Origin in Mammals and Duplication Patterns. Letters in Drug Design & Discovery. 2024;21(7):1226-1239.
18. Dornburg A, Mallik R, Wang Z, Bernal MA, Thompson B, Bruford EA, et al. Placing human gene families into their evolutionary context. Human Genomics. 2022;16(1):56.
19. Baines C, Meitern R, Kreitsberg R, Sepp T. Comparative study of the evolution of cancer gene duplications across fish. Evolutionary Applications. 2022;15(11):1834-1845.
20. Comaills V, Castellano-Pozo M. Chromosomal instability in genome evolution: From cancer to macroevolution. Biology. 2023;12(5):671.
21. Lamb TD. Analysis of paralogons, origin of the vertebrate karyotype, and ancient chromosomes retained in extant species. Genome Biology and Evolution. 2021;13(4):evab044.
22. Ray L, Medeiros D. Linking vertebrate gene duplications to the new head hypothesis. Biology. 2023;12(9):1213.
23. Peterson KJ, Beavan A, Chabot PJ, McPeek MA, Pisani D, Fromm B, et al. MicroRNAs as indicators into the causes and consequences of whole-genome duplication events. Molecular Biology and Evolution. 2022;39(1):msab344.
24. Gout J-F, Hao Y, Johri P, Arnaiz O, Doak TG, Bhullar S, et al. Dynamics of gene loss following ancient whole-genome duplication in the cryptic Paramecium complex. Molecular biology and evolution. 2023;40(5):msad107.
25. Morini M, Bergqvist CA, Asturiano JF, Larhammar D, Dufour S. Dynamic evolution of transient receptor potential vanilloid (TRPV) ion channel family with numerous gene duplications and losses. Frontiers in Endocrinology. 2022;13:1013868.
26. Tasnim M, Wahlquist P, Hill JT. Zebrafish: unraveling genetic complexity through duplicated genes. Development Genes and Evolution. 2024;234(2):99-116.
27. Brasó-Vives M, Marlétaz F, Echchiki A, Mantica F, Acemel RD, Gómez-Skarmeta JL, et al. Parallel evolution of amphioxus and vertebrate small-scale gene duplications. Genome Biology. 2022;23(1):243.
28. Mottes F, Villa C, Osella M, Caselle M. The impact of whole genome duplications on the human gene regulatory networks. PLoS Computational Biology. 2021;17(12):e1009638.
29. Assis R, Conant G, Holland B, Liberles DA, O'Reilly MM, Wilson AE. Models for the retention of duplicate genes and their biological underpinnings. F1000Research. 2024;12:1400.
30. Begum T, Duchen P, Bucao CF, Robinson-Rechavi M. Phylogenetic modeling provides evidence for sudden shifts in expression after small-scale duplication in vertebrates and strong support for the ortholog conjecture. bioRxiv. 2023.
31. Caulin AF, Maley CC. Peto's Paradox: evolution's

prescription for cancer prevention. Trends in ecology & evolution. 2011;26(4):175-182.
32. Clamp M, Andrews D, Barker D, Bevan P, Cameron G, Chen Y, et al. Ensembl 2002: accommodating comparative genomics. Nucleic Acids Research. 2003;31(1):38-42.
33. Kumar S, Nei M, Dudley J, Tamura K. MEGA: a biologist-centric software for evolutionary analysis of DNA and protein sequences. Briefings in bioinformatics. 2008;9(4):299-306.
34. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic acids research. 1994;22(22):4673-4680.
35. Russo C, Takezaki N, Nei M. Efficiencies of different genes and different tree-building methods in recovering a known vertebrate phylogeny. Molecular biology and evolution. 1996;13(3):525-536.
36. Felsenstein J. Confidence limits on phylogenies: an approach using the bootstrap. Evolution. 1985;39(4):783-791.
37. Whelan S, Goldman N. A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. Molecular biology and evolution. 2001;18(5):691-699.
38. Abegglen LM, Caulin AF, Chan A, Lee K, Robinson R, Campbell MS, et al. Potential mechanisms for cancer resistance in elephants and comparative cellular response to DNA damage in humans. Jama. 2015;314(17):1850-1860.
39. Buffenstein R. The naked mole-rat: a new long-living model for human aging research. The Journals of Gerontology Series A: Biological Sciences
40. Medical Sciences. 2005;60(11):1369-1377.
41. Seluanov A, Hine C, Azpurua J, Feigenson M, Bozzella M, Mao Z, et al. Hypersensitivity to contact inhibition provides a clue to cancer resistance of naked mole-rat. Proceedings of the National Academy of Sciences. 2009;106(46):19352-19357.
42. Tian X, Azpurua J, Hine C, Vaidya A, Myakishev-Rempel M, Ablaeva J, et al. High-molecular-mass hyaluronan mediates the cancer resistance of the naked mole rat. Nature. 2013;499(7458):346-349.
43. Delaney M, Ward J, Walsh TF, Chinnadurai S, Kerns K, Kinsel M, et al. Initial case reports of cancer in naked mole-rats (Heterocephalus glaber). Veterinary pathology. 2016;53(3):691-696.
44. Andervont HB, Dunn TB. Occurrence of tumors in wild house mice. Journal of the National Cancer Institute. 1962;28(5):1153-1163.
45. Higginbotham S, Wong WR, Linington RG, Spadafora C, Iturrado L, Arnold AE. Sloth hair as a novel source of fungi with potent anti-parasitic, anti-cancer and anti-bacterial bioactivity. PloS one. 2014;9(1):e84549.
46. Lee B-r, Oh S, Lee S-h, Kim Y, Youn S, Kim Y, et al. Squamous cell carcinoma in a nine-banded armadillo (Dasypus novemcinctus). Journal of Zoo Wildlife Medicine. 2015;46(2):333-334.
47. Thomas F, Asselin K, MacDonald N, Brazier L, Meliani J, Ujvari B, et al. Oncogenic processes: a neglected parameter in the evolutionary ecology of animals. Comptes Rendus Biologies. 2024;347(G1):137-157.
48. Munshi-South J, Wilkinson GS. Bats and birds: exceptional longevity despite high metabolic rates. Ageing research reviews. 2010;9(1):12-19