This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0). To view a copy of the license, visit **https://creativecommons.org/licenses/by-nc/4.0/**

Annals of Clinical and Analytical Medicine
Original Research

Evolutionary tinkering of a proto-oncogene: bootstrap-validated selection pressure analysis across MYC phylogeny

Evolutionary analysis of the MYC protein

Ahmed Azharuddin[1] Ehtesham Ahmed Shariff[1] Raiyan Ehtesham Ahmed Sharieff[2] Farahnaz Muddebihal[3]
Mohammed Daham Alanazi[4] Saleh Hamdan Alanazi[5]

[1]Department of Emergency Medical Services, Faculty of Emergency Medical Services, King Saud University, Riyadh, Saudi Arabia.
[2]Department of Medicine, Faculty of Medicine, Alfaisal University, Riyadh, Saudi Arabia.
[3]Department of Preventive Dentistry, Faculty of Dentistry, Jouf University, Sakaka, Saudi Arabia.
[4]Department of Emergency Medical Services, Prince Sultan bin Abdulaziz College for Emergency Medical Services, King Saud University, Riyadh, Saudi Arabia.
[5]Department of Medical Supply, King Saud University Medical City, Riyadh, King Saud University, Saudi Arabia.

Abstract

Aim: This study seeks to elucidate the evolutionary trajectory of MYC proto-oncogenes across vertebrate lineages, with particular emphasis on the role of gene duplication events in shaping their diversification. By examining conserved functional domains, the research aims to generate insights that may advance our understanding of oncogenic mechanisms and inform future directions in cancer biology.

Methods: Phylogenetic analyses were conducted on MYC protein sequences derived from 37 representative vertebrate species to investigate evolutionary relationships. Multiple sequence alignment was performed using the MUSCLE program with default parameters to ensure accurate positional homology across taxa. Neighbor Joining (NJ) and Maximum Likelihood (ML) methods were employed to infer phylogenetic trees. Statistical robustness of the inferred clades was evaluated by calculating bootstrap support values based on 1,000 replicates, thereby providing confidence estimates for major branching patterns.

Results: Comparative phylogenetic analyses revealed that gene duplication events in early vertebrate evolution gave rise to the c-MYC lineage, establishing a foundation for subsequent diversification. Strong sequence conservation was observed among hominids, with human, bonobo, and chimpanzee MYC proteins exhibiting 98.7% similarity. Human MYC sequences shared 87% identity with gorilla homologs and 83% with orangutan homologs, reflecting gradual divergence among great apes.

Conclusion: Elevated evolutionary rates in key functional domains indicate positive selection driving adaptive changes in oncogenic regulation. The ability of MYC to change and adapt across species (its evolutionary plasticity) may help explain why it can act as an oncogene. The resulting phylogenetic patterns provide essential context for interpreting MYC-driven tumorigenesis.

Keywords

myc, phylogeny, cancer, oncoproteins, evolutionary analysis, bootstrapping

DOI: 10.4328/ACAM.22986 Ann Clin Anal Med 2026;17(10):1094-1098
Received: 12.11.2025 Accepted: 22.12.2025 Published Online: 26.01.2026 Printed: 01.10.2026
Corresponding Author: Ahmed Azharuddin, Department of Emergency Medical Services, Faculty of Emergency Medical Services, King Saud University, Riyadh, Saudi Arabia. **Email:** aazharuddin@ksu.edu.sa **P.:** +966 58 036 3785
Corresponding Author ORCID: https://orcid.org/0000-0001-8797-3523
Other Authors ORCID: Ehtesham Ahmed Shariff, https://orcid.org/0000-0001-6847-3961 • Raiyan Ehtesham Ahmed Sharieff, https://orcid.org/0009-0009-8188-6731 • Farahnaz Muddebihal, https://orcid.org/0000-0002-9296-7686 • Mohammed Daham Alanazi, https://orcid.org/0009-0003-9688-0478 • Saleh Hamdan Alanazi, https://orcid.org/0009-0006-8879-2568

Introduction

The MYC proto-oncogene encodes a nuclear phosphoprotein that regulates apoptosis, cell cycle, and transformation, and is frequently activated in human cancers. MYC is a 439–amino–acid, 48.8 kDa oncogenic protein that forms dimers with basic helix–loop–helix (bHLH) partners to bind DNA, specifically recognizing the core sequence 5'-CAC.[GA]TG-3'.[1] The MYC protein dimerizes with MAX, binds E-box DNA, and regulates target gene transcription.[2] Amplification of this gene is frequently observed in numerous human cancers. Translocations of the MYC gene are associated with Burkitt lymphoma and multiple myeloma in human patients.[3] MYC amplifies transcription via cofactors and chromatin modifiers; its level-dependent deregulation drives cancer.[4] MYC is translated from CUG/AUG into multiple isoforms; its activity depends on regulatory elements and factors. MYC deficiency impairs ribosome biogenesis and hinders the expansion of natural killer (NK) cells, thereby compromising the body's anticancer immune response.[5]

c-MYC is rapidly induced to sustain growth, driving cells into the cell cycle; its inhibition blocks mitogenic signals and promotes differentiation.[6] c-MYC drives proliferation but triggers apoptosis if survival signals from cytokines or adhesion receptors are lost.[7] Targeted histone acetylation genes may depend on the initial binding of c-MYC. This binding facilitates the subsequent recruitment of other transcription factors necessary to complete transcriptional activation.[8] MYC regulates transcription through both activation and repression. In contrast, Mad/Max complexes compete with MYC/Max at E-box elements, thereby repressing gene expression and suppressing cellular differentiation.[9] MYC expression is influenced by signaling mutations and indels. Previous studies have shown that, due to its nuclear localization and lack of a druggable active site, direct targeting of MYC is challenging; therefore, indirect therapeutic strategies have been explored.[10]

It activates transcription via TRRAP and CBP/p300, while Miz-1 and Sp1 repress tumor suppressors.[11] MYC stability is regulated by Fbxw7 and Aurora-A, impacting tumor growth. It binds the VEGFA promoter to drive angiogenesis and acts as a universal transcription amplifier through flexible, unstructured regions.[12] MYC also forms a "topoisome" with topoisomerases to relieve transcriptional stress, underscoring its central role in cancer biology and therapeutic targeting.[13] MYC's evolutionary role has been extensively studied, with key research on C-MYC expression, cancer involvement, and cell transformation revealing its broad biological functions. Yet, sequence-based phylogenetic insights into MYC homologs remain limited. This study analyzes MYC homologs across diverse organisms to clarify their phylogenetic relationships and evolutionary dynamics. By exploring functional conservation and divergence, we aim to deepen our understanding of MYC's role in cellular transformation, adaptation, and its relevance to physiology and disease.

Materials and Methods

Sequence Retrieval and Multiple Sequence Alignment

MYC protein sequences (n = 37) representing different species were downloaded from the Swiss Uniprot (Uniprot ID-P01106.1) database.[14] We used the PSI-Blast program to search for similar sequences across species in the non-redundant sequence database.[15] The BLOSUM62 matrix was used to identify functionally more conserved sequences.[16] PSI-BLAST threshold and expected value were set to 0.005 and 10, respectively. A BLAST-generated similar sequence was used for multiple sequence alignment using the MUSCLE program.[17] We used the BLOSUM62 matrix with a gap-open penalty of -12.0 and a gap-extension penalty of -1.0. Sequence weighting scheme weight1 was used in iterations 1 and 2, and weight2 was used for tree-dependent refinement.[16]

Distance Disparity Index Estimation and Phylogeny

All taxa were selected, and the neighbor-joining method was used. We applied the Jones–Taylor–Thornton (JTT) amino acid substitution model to construct phylogenetic trees.[18] To assess rate constancy among lineages, a molecular clock test was performed by comparing the Maximum Likelihood (ML) values for the given topology with and without molecular clock constraints under the JTT model (+G). Using a substitution- rate matrix (Q), the matrix (F), which consists of the observed proportions of amino acid pairs between a pair of sequences with their divergence time t, is given by the following equation F (t) = Ae2tQ where A denotes the diagonal matrix of the equilibrium amino acid frequencies for Q. From this equation, the evolutionary distance d = 2tQ can be iteratively computed by a maximum- likelihood method. Pairwise deletion of gaps and missing data in sequences was performed to infer sequence phylogeny. A phylogenetic tree was generated using neighbor-joining with 500 bootstrap replications.[19] We used the amino acid substitution model with the Jones-Taylor-Thorton (JTT) method to model substitutions across species.[16] To analyze evolutionary rates among taxa, we have used the discrete gamma distribution method.[20] Evolutionary analyses were conducted in MEGA 12.[21] All taxa were selected, and the ordinary least squares method was used.[22] To assess rate variation among sites, a rate test was run using the maximum-likelihood method.[23] Time trees were computed to estimate divergence times for all branching points in the phylogenetic tree. These estimates were obtained using the RelTime method, which does not require assumptions about lineage rate variations.[24] Tajima's relative test function tests the hypothesis of a molecular evolutionary clock (i.e., a constant rate of molecular evolution) between two samples using an out-group sample. The molecular clock performs a Maximum Likelihood test of the molecular clock hypothesis for a given tree topology and sequence alignment.[23] The molecular clock hypothesis assumes that all tips of the tree are equidistant from the root.[18] In this study, the molecular clock was applied to estimate divergence times among lineages and to assess whether evolutionary rates were consistent across taxa.

Ethical Approval

All analyses were conducted using in silico data, thus not requiring formal ethical approval from a research ethics committee.

Statistical Analysis

Phylogenetic analyses were performed using MEGA 12 software. Bootstrap support values were calculated using 500 or 1,000 replicates as specified for each analysis. Evolutionary distances, molecular clock tests, Tajima's neutrality test, Tajima's relative rate test, and rate variation analyses were conducted using the software's statistical algorithms. Bootstrap values were used to assess the robustness of phylogenetic relationships.

Reporting Guidelines

No specific reporting guideline was applicable to this in silico phylogenetic analysis.

Results

Distance Matrices and Disparity Index

MEGA 12 computes and presents the disparity index for each site, and it is more powerful than a chi-square test for testing the equality of base frequencies between sequences.[21,25] Disparity Index >0 shows sequence pairs with base composition biases

exceeding expectations from divergence or chance.

Phylogenetic trees show short, conserved branches in mammals, and MYC factors suggest that oncoproteins are ancient or rapidly evolving. Across species, MYC oncoproteins are essential for core cellular functions, suggesting broadly similar evolutionary rates. Indeed, a significant rate difference is only evident between the more closely related organisms for which all MYC oncoprotein sequences are available. But the composition distances for MYC oncoproteins in more divergent groups seem to approach similar values (Supplementary Figure S1). Furthermore, the overall number of substitutions between MYC oncoproteins and their presumed ancestor appears to be essentially the same. MYC oncoproteins appear to show a lower substitution rate among closely related species. The mean distances within subfamilies of these proteins are in the range of 0.00 to 0.278.

Neutrality & Molecular Clock Tests

Tajima's neutrality test compares segregating sites with nucleotide diversity; differences in 4Nv estimates indicate non- neutral evolution (where N is the effective population size and v is the mutation rate per site) (Table 1).

Mean evolutionary rates in these categories were 0.00, 0.01, 0.09, 0.60, and 4.30 substitutions per site. The amino acid frequencies are 7.69% (A), 5.11% (R), 4.25% (N), 5.13% (D), 2.03% (C), 4.11% (Q), 6.18% (E), 7.47% (G), 2.30% (H), 5.26% (I), 9.11% (L), 5.95% (K), 2.34% (M), 4.05% (F), 5.05% (P), 6.82% (S), 5.85% (T), 1.43% (W), 3.23% (Y), and 6.64% (V). For estimating ML values, a tree topology was automatically computed. The maximum log-likelihood for this computation was -4221.694. The molecular clock hypothesis was tested by comparing log-likelihood values with and without the clock. The non-clock model consistently showed a higher likelihood, and significance was assessed using a chi-squared test with 2 degrees of freedom (Table 1).

The molecular clock test was performed by comparing the ML value for the given topology with and without the molecular clock constraints under the Jones-Taylor-Thornton model (+G).[18] Differences in evolutionary rates among sites were modeled using a discrete Gamma (G) distribution (with the shape parameter shown). The null hypothesis of equal evolutionary rate throughout the tree was rejected at a 5% significance level (*P* = 7.973E-061).

Tajima's relative rate test indicates equal evolutionary rates between sequences A (Homo sapiens) and B (Pan paniscus), with sequence C (Callorhinus ursinus) as the outgroup (Table 2). The x2 test statistic was 0.33 (*P* = .56370 with 1 degree of freedom) *P* value less than 0.05 is often used to reject the null hypothesis of equal rates between lineages.

Rate Correlation Test and Phylogenetic Analysis

The null hypothesis of rate independence of evolutionary rates among lineages in the user-supplied phylogeny was not rejected (*P* value > 0.05). The log-likelihood of the evaluated tree is (-5557.47). A discrete Gamma distribution was used to model evolutionary rate differences among sites (5 categories (+G, parameter = 0.1679)). This analysis involved 37 amino acid sequences. There were 714 positions in the final dataset. Phylogenetic trees that include both vertebrate and non-vertebrate species show a coherent evolutionary relationship. Within these trees, species relationships are consistently inferred, with Homo sapiens, Pan paniscus, and Pan troglodytes clustering together alongside Gorilla gorilla and Pongo abelii. The remaining primates are grouped in another cluster. The Minimum Evolution method corrects for multiple hits and selects the tree with the smallest branch sum (S). However, construction of an ME tree is time-consuming because, in principle, the S values for all topologies must be evaluated. Because the number of possible topologies (unrooted trees) rapidly increases with the number of taxa, it becomes very difficult to examine all topologies. The NJ method produces an unrooted tree because it does not require the assumption of a constant rate of evolution.

Table 2. Results of Tajima's Relative Rate Test Applied to Three Sequence Comparisons

Configuration	Count
Identical sites in all three sequences	408
Divergent sites in all three sequences	0
Unique differences in Sequence A	1
Unique differences in Sequence B	2
Unique differences in Sequence C	27

Finding the root requires an out-group taxon. In the absence of out-group taxa, the root is sometimes given at the midpoint of the longest distance connecting two taxa in the tree, which is referred to as midpoint rooting. To test the tree's reliability, we have used the Bootstrap method for a neighbor-joining tree (Supplementary Figure S2). The user-specified tree topology was analyzed using the least-squares method—the sum of branch lengths was 1.33665460 for the displayed tree. The evolutionary distances were computed using the JTT matrix-based method and are expressed in units of amino acid substitutions per site. The Homo sapiens taxon is colored in red in the light green cluster.

Ordinary Least Squares (OLS) based rate variation among sites was modeled with a gamma distribution (shape parameter = 1). All ambiguous positions were removed for each sequence pair. There were 714 positions in the final dataset. We have also calculated relative and absolute divergence times for all branching points in the tree. The time tree (Supplementary Figure S3) shows the same topology as the active tree, with local clock rates and divergence times estimated for all branching points. All divergence time estimates are based solely on the branch lengths in the active tree and were calculated based on the maximum likelihood-based method. RelTime analysis applied multiple divergence-time calibrations, with the Homo sapiens node constrained by the MRCA taxa Pan paniscus and Pan troglodytes.

Table 1. Results of Evolutionary Analyses: Tajima's Neutrality Test and Molecular Clock Hypothesis Evaluation

M	S	Ps	Θ	Π	D
Results from Tajima's Neutrality Test					
37	214	0.299720	0.071797	0.035216	-1.914071
		lnL	**Parameters**	**(+G)**	**(+I)**
Results from a test of molecular clocks:					
1.	With Clock	-4425.006	37	0.220	n/a
2.	Without Clock	-4231.761	72	0.15	n/a

Discussion

The phylogenetic analysis reveals distinct evolutionary patterns in the MYC protein across mammals. Primates (Homo sapiens, Pan troglodytes, Gorilla) show tight clustering (disparity index ~0.1-0.15), indicating strong functional conservation, particularly among hominids. The trend may indicate faster evolution of the MYC oncoprotein than suggested by the disparity index. Canids display greater divergence (Canis lupus familiaris-dingo pairwise distance 0.35 vs 0.2 in felids), suggesting dietary adaptation-driven evolution. Artiodactyls (Camelus, Bos taurus) cluster with rumen-specific convergence (distance 0.25-0.3), while Vicugna shows unique positioning (0.4 disparity) potentially linked to altitudinal adaptation. Chiropterans split between Pteropus (flight metabolism) and Myotis (echolocation systems), with Rhinolophus ferrumequinum showing intermediate divergence (0.2-0.25). Rodent polyphyly (Mus vs Rattus at 0.45 distance) highlights rapid compensatory evolution (Supplementary Figure S1). Branch lengths correlate with the intensity of ecological specialization, ranging from conserved primate profiles to diversified carnivore patterns.

The bootstrap-supported topology reveals critical insights into MYC protein evolution. Strong nodal support ([> 85%]) at key mammalian divergence points confirms conserved functional domains under purifying selection. However, moderate support ([65-75%]) in Laurasiatherian lineages suggests accelerated evolutionary rates, potentially linked to metabolic adaptations. The unresolved Cetartiodactyla node implies incomplete lineage sorting or convergent evolution. These patterns underscore MYC's dual role as a conserved transcriptional regulator and a driver of lineage-specific adaptation.

MYC's phylogeny reflects core evolutionary constraints and adaptive shifts across mammals. Conserved primate nodes highlight stable transcriptional regulation, while carnivoran clustering suggests niche-specific cell cycle control—the cetartiodactyl polytomy points to either rapid diversification or enduring pleiotropic limits (Supplementary Figure S2). Rodent- specific branching patterns highlight potential compensatory evolution in rapidly proliferating tissues. This mosaic evolutionary landscape positions MYC as both a conserved developmental gatekeeper and a molecular substrate for lineage-specific physiological adaptations.

The phylogenetic clustering reveals distinct MYC evolutionary trajectories across mammalian orders. Close primate (Pan/Macaca) grouping underscores strong functional conservation in higher cognitive species, while Canis' separate branching suggests carnivore-specific adaptive pressures. The Rattus-ovis-Camelus association indicates potential convergence in MYC regulation among herbivorous/desert-adapted species. Unresolved nodes near Rhinolephus and Flerepas may reflect either rapid radiation events or persistent functional constraints in insectivorous lineages. These patterns highlight MYC's dual role as a conserved developmental regulator and as a substrate for niche-specific molecular tweaking, particularly in metabolic and environmental stress-response pathways.

Phylogenetic analysis of MYC across mammals reveals conserved brain-related functions in primates, metabolic adaptations in arboreal apes, and dietary-driven divergence in canids. Artiodactyls show rumen-specific MYC regulation and hypoxia-linked variants in vicuñas. Chiropterans split between flight-optimized MYC in megabats and auditory adaptations in microbats. Rodents display polyphyletic patterns, possibly due to rapid evolution or lab strain artifacts (Supplementary Figure S3). Across evolutionary lineages, MYC serves as a conserved anchor that maintains core cell-cycle regulation. At the same time, it serves as a molecular driver that enables ecological specialization. Variation in terminal branch lengths reflects this dual role, ranging from the constrained profiles of large-brained primates to the diversified patterns seen in dietary specialists and environmental extremophiles.

Limitations

The analysis was constrained by limited taxonomic sampling (37 species), restricting comprehensive evolutionary comparisons. Cetartiodactyla phylogeny remained unresolved due to insufficient nodal support, and the findings relied solely on sequence data, lacking experimental validation of MYC protein functions. Potential lab strain artifacts in rodent lineages may have skewed evolutionary interpretations, and the unaddressed role of epigenetic regulation in MYC evolution leaves key mechanistic questions unanswered.

Conclusion

This study demonstrates the widespread presence and rapid evolution of MYC oncoproteins across species, reflecting their involvement in key cellular processes, including growth, apoptosis, metabolism, and mitochondrial biogenesis. Phylogenetic analysis of 37 MYC protein sequences reveals complex evolutionary relationships shaped by gene duplication events early in vertebrate history, giving rise to the c-MYC lineage. Human MYC shows strong conservation with Pan paniscus and Pan troglodytes, followed by Gorilla gorilla and Pongo abelii, mirroring the established vertebrate branching order. These findings provide a robust framework for future functional studies on MYC's role in cancer and cellular regulation. Future work should focus on experimentally validating these in silico findings through functional assays and comparative studies in model organisms. Such approaches, including gene expression profiling and protein interaction analyses, will be essential for confirming evolutionary insights and clarifying the mechanistic role of MYC in oncogenesis.

Declarations

Animal and Human Rights Statement
Not applicable.

Informed Consent
Not applicable.

Data Availability
All sequence data used in this study were obtained from publicly available databases.

Conflict of Interest
The authors declare no conflict of interest.

Funding
None.

Author Contributions (CRediT Taxonomy)
Conceptualization: A.A., E.A.S.
Methodology: A.A., E.A.S.
Software: A.A., R.E.A.S.
Formal Analysis: A.A., E.A.S., R.E.A.S.
Investigation: A.A., F.M.
Data Curation: R.E.A.S., M.D.A.
Writing – Original Draft: A.A.
Writing – Review & Editing: E.A.S., F.M., M.D.A., S.H.A.
Supervision: A.A., S.H.A.

AI Usage Disclosure
The authors declare that no AI-assisted technologies were used.

Abbreviations

BLAST: Basic local alignment search tool
JTT: Jones–taylor–thornton
MEGA: Molecular evolutionary genetics analysis
ML: Maximum likelihood
MUSCLE: Multiple sequence comparison by log-expectation
MYC: Myelocytomatosis proto-oncogene
NJ: Neighbor joining
NK: Natural killer
OLS: Ordinary least squares
PSI-BLAST: Position-specific iterated basic local alignment search tool
TRRAP: Transformation/transcription domain-associated protein

References

1. *Gaballa A, Krenz B, Uhl L. MYC: the guardian of its own chaos. Bioessays. 2025;47(7):e70010. doi:10.1002/bies.70010*
2. *Speltz TE, Qiao Z, Swenson CS, et al. Targeting MYC with modular synthetic transcriptional repressors derived from bHLH DNA-binding domains. Nat Biotechnol. 2023;41:541-551. doi:10.1038/s41587-022-01504-x*
3. *Shaw T, Cockrell H, Panchal R, Abraham A, Sawaya D. Burkitt lymphoma presenting as perforated appendicitis. Am Surg. 2022;88(3):547-548. doi:10.1177/00031348211029854*
4. *Jha RK, Kouzine F, Levens D. MYC function and regulation in physiological perspective. Front Cell Dev Biol. 2023;11:1268275. doi:10.3389/fcell.2023.1268275*
5. *Khameneh HJ, Fonta N, Zenobi A, et al. Myc controls NK cell development, IL-15-driven expansion, and translational machinery. Life Sci Alliance. 2023;6(7):e202302069. doi:10.26508/lsa.202302069*
6. *Nussinov R, Zhang W, Liu Y, Jang H. Mitogen signaling strength and duration can control cell cycle decisions. Sci Adv. 2024;10(27):eadm9211. doi:10.1126/sciadv.adm9211*
7. *Uroz M, Wistorf S, Serra-Picamal X, et al. Regulation of cell cycle progression by cell-cell and cell-matrix forces. Nat Cell Biol. 2018;20:646-654. doi:10.1038/s41556-018-0107-2*
8. *Llombart V, Mansour MR. Therapeutic targeting of "undruggable" MYC. eBioMedicine. 2022;75:103756. doi:10.1016/j.ebiom.2021.103756*
9. *Zielke N, Vähärautio A, Liu J, Kivioja T, Taipale J. Upregulation of ribosome biogenesis via canonical E-boxes is required for Myc-driven proliferation. Dev Cell. 2022;57(8):1024-1036. doi:10.1016/j.devcel.2022.03.018*
10. *Steinberger J, Robert F, Hallé M, et al. Tracing MYC expression for small molecule discovery. Cell Chem Biol. 2019;26(5):699-710. doi:10.1016/j.chembiol.2019.02.007*
11. *Wang C, Ma X. The role of acetylation and deacetylation in cancer metabolism. Clin Transl Med. 2025;15(1):e70145. doi:10.1002/ctm2.70145*
12. *Yu J, Liu D, Yuan Y, Sun C, Su Z. Rethinking MYC inhibition: a multi- dimensional approach to overcome cancer's master regulator. Front Cell Dev Biol. 2025;13:1601975. doi:10.3389/fcell.2025.1601975*
13. *Das SK, Lewis BA, Levens D. MYC: a complex problem. Trends Cell Biol. 2023;23(3):235-246. doi:10.1016/j.tcb.2022.07.006*
14. *The UniProt Consortium. UniProt: a worldwide hub of protein knowledge. Nucleic Acids Res. 2019;47:D506-515. doi:10.1093/nar/gky1049*
15. *Altschul SSF, Madden TTL, Schäffer AA, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 1997;25(17):3389-3402. doi:10.1093/nar/25.17.3389*
16. *Henikoff S, Henikoff JG. Amino acid substitution matrices from protein blocks. Proc Natl Acad Sci U S A. 1992;89(22):10915-1099. doi:10.1073/pnas.89.22.10915*
17. *Edgar RC. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 2004;32(5):1792-1797. doi:10.1093/nar/gkh340*
18. *Jones DT, Taylor WR, Thornton JM. The rapid generation of mutation data matrices from protein sequences. Bioinformatics. 1992;8(3):275-282. doi:10.1093/ bioinformatics/8.3.275*
19. *Hillis DM, Bull JJ. An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. Syst Biol. 1993;42(2):182-192. doi:10.1093/ sysbio/42.2.182*
20. *Yang Z. Among-site rate variation and its impact on phylogenetic analyses. Trends Ecol Evol. 1996;11(9):367-372. doi:10.1016/0169-5347(96)10041-0*
21. *Kumar S, Stecher G, Suleski M, Sanderford M, Sharma S, Tamura K. MEGA12: molecular evolutionary genetic analysis version 12 for adaptive and green computing. Mol Biol Evol. 2024;41(12):1-9. doi:10.1093/molbev/msae263*
22. *Xia X, Yang Q. A distance-based least-square method for dating speciation events. Mol Phylogenet Evol. 2011;59(2):342-353. doi:10.1016/j.ympev.2011.01.017*
23. *Yang Z. Maximum-likelihood models for combined analyses of multiple sequence data. J Mol Evol. 1996;42:587-596. doi:10.1007/BF02352289*
24. *Tamura K, Battistuzzi FU, Billing-Ross P, Murillo O, Filipski A, Kumar S. Estimating divergence times in large molecular phylogenies. Proc Natl Acad Sci U S A. 2012;109(47):19333-19338. doi:10.1073/pnas.1213199109*
25. *Kumar S, Gadagkar SR. Disparity index: a simple statistic to measure and test the homogeneity of substitution patterns between molecular sequences. Genetics. 2001;158(3):1321-1327. doi:10.1093/genetics/158.3.1321*

How to Cite

Ahmed Azharuddin, Ehtesham Ahmed Shariff, Raiyan Ehtesham Ahmed Sharieff, Farahnaz Muddebihal, Mohammed Daham Alanazi, Saleh Hamdan Alanazi. Evolutionary tinkering of a proto-oncogene: bootstrap-validated selection pressure analysis across MYC phylogeny. Ann Clin Anal Med 2026;17(10):1094-1098. doi:10.4328/ACAM.22986