Hindawi
Journal of Zoological Systematics and Evolutionary Research
Volume 2023, Article ID 6693488, 15 pages
https://doi.org/10.1155/2023/6693488

Research Article

Positive Selection in Zinc Finger Protein Reveals Genetic Signatures of Adaptive Evolution in Undifferentiated Stem Cells during Evolution in Mammals

Mahboob Ali,[1] Palwasha Khan,[2] Mahreen Mahmood,[3] Jilong Han,[4] Gulnaz Afzal,[5] Iram Qadeer,[6] Syed Ali Azmal,[7] Nadia R. A. El-Mouhty,[8] Samy F. Mahmoud,[9] Ibrahim Jafri,[9] Muhammad Ali,[10] and Nighat Hashmi[11]

[1]*Rural Health Center, Akhtarabad, District Okara, Pakistan*
[2]*District Headquarter Hospital, Rajanpur, Pakistan*
[3]*Dental Surgeon Rural Health Centre, Akhtarabad, Punjab, Pakistan*
[4]*College of Animal Science and Technology, Shihezi University, China*
[5]*Department of Zoology, The Islamia University of Bahawalpur, Pakistan*
[6]*Government Sadiq College Women University, Bahawalpur, Pakistan*
[7]*Department of Animal Genetics and Breeding, Sirajganj Government Veterinary College, Sirajganj under the Department of Livestock Services, Bangladesh*
[8]*Department of Chemistry, College of Science, Taif University, P.O. Box 11099, Taif 21944, Saudi Arabia*
[9]*Department of Biotechnology, College of Science, Taif University, P.O. Box 11099, Taif 21944, Saudi Arabia*
[10]*Basic Health Unit, Mangrotha, Tehsil Taunsa Sharif, District DG Khan, Pakistan*
[11]*Rural Health Center Garh Maharaja, Tehsil Ahmadpur Sial, District Jhang, Pakistan*

Correspondence should be addressed to Jilong Han; hanjilong10000@126.com and Syed Ali Azmal; syed.azmal@yahoo.com

Received 20 April 2023; Revised 10 May 2023; Accepted 15 May 2023; Published 29 May 2023

Academic Editor: Amjad Bashir

Positive selection refers to the process by which certain genetic variations are more likely to be passed on to future generations because they confer some advantage in terms of survival or reproduction. Zinc finger proteins are a type of transcription factor that plays a role in regulating gene expression, particularly in undifferentiated stem cells. Suppose it has been found that certain zinc finger proteins show genetic signatures of positive selection in mammals. In that case, it suggests that these proteins may have played a role in adaptive evolution in undifferentiated stem cells. This could mean that the specific genetic changes in these zinc finger proteins gave an advantage to the organisms that possessed them, helping them survive and reproduce more effectively. These genetic changes may have allowed the organisms to adapt to changing environments or to develop new abilities, such as increased resistance to disease or faster growth. Undifferentiated stem cells that underwent adaptive evolution during the evolution of mammals can be identified genetically by the outcomes of positive selection on zinc finger proteins. Because of selection pressures like environmental shifts or the introduction of novel pathogens, it is plausible that some zinc finger proteins have experienced fast evolution. The emergence of novel activities or higher expression levels of these proteins as a result of this quick evolution may have given the creatures that possessed them a survival edge. Another possible outcome of positive selection in zinc finger proteins is the emergence of new genetic variations that allow for increased diversity and plasticity in stem cells. This increased diversity and plasticity could have allowed for more efficient adaptation to changing environments and could have played a role in the evolution of new organisms or new characteristics in existing organisms. Overall, the results of positive selection in zinc finger proteins can provide insight into how adaptive evolution occurred in undifferentiated stem cells during the evolution of mammals and how this evolution may have contributed to the development of new organisms and new characteristics and adaptations to changing environments.

1. Introduction

Undifferentiated cells that display an unlimited capacity for self-renewal and pluripotency are called pluripotent stem cells (PSCs) [1]. Even though PSCs have the potential to provide significant resources for the study and preservation of rare animal species, ethical and technical challenges have prevented researchers from attempting very many equivalent derivations with these species. PSCs can now be obtained from a wider variety of mammals, including some that are threatened with extinction, thanks to recent developments in somatic cell reprogramming into induced pluripotent stem cells (iPSCs) [2]. Many taxonomic groups have successfully used iPSC technology, including primates and Carnivora. We reserve all of our rights. There is a strict ban on recycling without prior permission [3]. However, there is an ongoing debate regarding the possibility of species-specific differences in the characteristics of the PSCs, including but not limited to the pluripotent state, the degree of reprogramming efficiency, and the optimal culture conditions. Variations in gene regulatory mechanisms can help explain some of the phenotypic differences between individuals and species [4]. By analyzing the similarity and differences between networks, comparative methods can shed light on how gene regulatory processes have developed throughout evolution [5]. The observation of purifying selection and positive selection enables one to infer evolutionary conservation and adaptations. The development of phenotypic adaptations in cetacean lipid metabolism has genetic signatures that we recently described. Numerous studies have demonstrated the role of natural selection in the evolution of adaptive traits in animals and plants [6]. Importantly, how changes are positioned in the regulatory network's hierarchy will affect how a regulatory process is regulated. However, how natural selection affects the PGRN genes and how the mammalian PGRN's evolutionary patterns function are still unclear. Adaptive evolution refers to the process by which species evolve in response to changing environmental conditions. This process occurs when certain genetic variations conferring a survival advantage become more common in a population, leading to new traits and adaptations [7]. Studying adaptive evolution in proteins can provide important insights into the mechanisms underlying the evolution of species and the molecular basis of biological adaptation. Proteins are the workhorses of the cell, performing various functions such as catalyzing chemical reactions, transmitting signals, and providing structure [8]. By analyzing the genetic signatures of proteins in different species of mammals, researchers can gain insights into the evolutionary pressures that shaped their function and role in biological processes. For example, they can identify amino acid substitutions that have arisen in response to changing environments and led to the development of new traits and adaptations [9].

Phylogenetic analyses have been used to reconstruct the evolutionary history of ZFPs, revealing patterns of gene duplication, gene loss, and sequence change that have contributed to the increased functional diversity of this important protein family [10]. Molecular evolution approaches, such as site-specific positive selection tests, have revealed amino acid changes that are expected to have functional ramifications, providing insight into the molecular underpinnings of ZFP development. Several mammalian lineages have been studied using comparative genomics to learn more about the evolution of ZFPs. By comparing the sequences and architectures of ZFPs across species, researchers have found lineage-specific modifications in gene expression and function that have contributed to the evolution of complex traits and developmental processes [11].

The study of adaptive evolution in proteins is crucial because it sheds light on the molecular foundation of biological adaptability and, by extension, on the mechanisms underpinning the evolution of species. This data can be utilized to create innovative solutions to problems in human health and the environment, as well as to better understand the processes that drive the development of proteins and other biological systems. The purpose of this research was to determine how stable the mammalian PGRN has been throughout the course of evolution [12]. (i) We did this by calculating the ratio of synonymous to nonsynonymous site substitution rates (dN/dS) to understand how strong purifying selection is throughout time and how quickly evolution is proceeding. (ii) To better comprehend the evolution of the PGRN design, we examined the estimated conservation profiles throughout the PGRN subcircuits. (iii) We found genes undergoing positive selection for lineages and looked at the concordance in the functional areas of the proteins to discover the phylogenetic inference of the differences in the PGRN.

2. Materials and Methods

The nucleotide and amino acid sequences of the ZFP42 gene can be downloaded from the Ensembl database using the gene ID or gene name. The sequences can be downloaded in FASTA format, which is a standard format for biological sequence data. The nucleotide and amino acid sequences of the ZFP42 gene were aligned with the sequences from other species using alignment software such as Clustal Omega, MUSCLE, or MAFFT [13]. Alignment is necessary to ensure that the sequences are in the same reading frame and to identify the conserved regions of the gene [14]. The maximum likelihood method was utilized during the phylogenetic investigation, which was carried out with the MEGA 6 software. The bootstrap test was carried out with a maximum likelihood approach, and a value of 1000 was used for the advanced log-likelihood values [15].

2.1. Sequence Analysis. Alignment programs like Clustal Omega, MUSCLE, and MAFFT were used to compare and align the ZFP42 gene's nucleotide and amino acid sequences with those of other species. Making sure the sequences are all in the same reading frame and finding the conserved parts of the gene both require alignment [16]. Several statistical techniques, including the codon-based maximum likelihood method included in the PAML (Phylogenetic Analysis by Maximum Likelihood) software program, were used to conduct the positive selection analysis. For each codon in

the gene, this technique calculates the dN/dS ratio and checks to see if it is significantly higher than 1, which denotes positive selection. Positive selection operating on particular branches of the phylogenetic tree can also be found using additional techniques, such as the branch-site model [17]. Phylogenetic trees were created using a variety of techniques, including neighbor joining, Bayesian inference, and maximum likelihood. Based on the degree of similarity or difference between the ZFP42 gene sequences from other species, these methods use algorithms to infer the evolutionary relationships between the sequences. It is possible to examine and evaluate the tree's topology and branch lengths in order to deduce evolutionary links and identify positive selection. Software like MEGA (Molecular Evolutionary Genetics Analysis) or FigTree was used to illustrate the findings of positive selection and phylogenetic analysis [18]. With the use of these tools, users can see the phylogenetic tree, spot branches that have undergone positive selection, and view the dN/dS ratio and other important statistics.

2.2. Tests for Selection. Positive selection led to the evolution of PGRP genes or genomic regions, which meant that they underwent alterations to increase their fitness and environmental adaption [19]. Positive selection was discovered in protein-coding genes using the PAML (Phylogenetic Analysis by Maximum Likelihood) software tool. The data are analyzed by PAML using the codeml application [17]. In order to identify positive selection, the codeml program examines the rates of synonymous and nonsynonymous changes. The proportion of synonymous to nonsynonymous changes in protein-coding genes is measured here. Positive selection is indicated by a dN/dS ratio greater than one, whereas purifying selection is indicated by a ratio lower than one [20]. By computing the likelihood log ratio (2lnL) using the 2 distributions, PAML's output file indicates a positive selection [21]. The likelihood ratio test (LRT) and the Bayes empirical Bayes (BEB) approach, which contrast a null model that assumes no positive selection with a different model that enables positive selection [22], are the most often employed statistics. To test for positive selection, the LRT evaluates how well the alternative (selection) model fits the data compared to the null (no selection) model. In order to calculate the posterior probability that a given codon site has been subject to positive selection, the BEB technique is used [23]. Models based on codon use predict the frequencies of synonymous and nonsynonymous changes in protein-coding genes. Using these models, scientists have been able to identify protein-coding genes that have been positively selected. The branch-site and site-specific codon-based models were employed to identify positive selection. Different from the site-specific model, which examines positive selection at particular codon sites, the branch-site model examines positive selection along a single branch of the phylogenetic tree [24].

Protein-coding sequences that have undergone positive selection were also analyzed using the Datamonkey website [25, 26]. Methods for detecting positive selection were found on Datamonkey, and they included the fixed effects likelihood (FEL), the random effects likelihood (REL), the fast unconstrained Bayesian approximation (FUBAR), and the mixed effects model of evolution (MEME) [26, 27]. The likelihood-based approaches FEL, REL, and MEME all employ a phylogenetic tree to identify areas of positive selection. The ongoing selective forces on the sequence across the entire tree or at each site are identified by FEL and REL. However, MEME permits various selective pressures to act on various sequence locations [28].

On the other hand, FUBAR is a Bayesian approach that estimates the posterior probability of positive selection at each site in the sequence with the help of a Markov chain Monte Carlo (MCMC) algorithm [29]. Protein-coding sequence data showed evidence of positive selection, which was further verified using the Selecton server. To calculate the posterior probabilities of various selection patterns at each site in the protein-coding sequences, the MCMC model utilized in the Selecton server employed a Bayesian methodology. In addition to positive and negative selection, it also provides for the option of no selection at any given site. Using the posterior distribution as a sample, the MCMC method can estimate the likelihood of each candidate region [30].

2.3. Phylogenetic Analysis. Phylogenetic trees will be built using ZFP42 sequence data from several species in order to examine evolutionary links and changes in the gene. Maximum likelihood was used to create the phylogenetic trees in MEGA (Molecular Evolutionary Genetics Analysis) 10.0.5 [15]. We first built a tree with the neighbor-joining method and then assessed its topology with the maximum likelihood technique and the Whelan and Goldman (WAG) substitution model [31]. We did 1000 bootstrap runs to assess the robustness of the tree structure further. TreeBeST's resulting species tree serves as a standard against which gene trees and other phylogenetic trees can be judged. Phylogenetic network analysis revealed reticulation events in the evolutionary history of the proteins [32].

Positive selection can then be determined by establishing which specific amino acid positions in a protein sequence have been subject to positive selection. By definition, positive selection causes beneficial mutations to become more common in a population over time [33]. Positive selection modifies particular amino acid residues in proteins to improve their function or alter their structure. By comparing synonymous (silent) and nonsynonymous (amino acid altering) substitution rates at each amino acid site, we applied a widely used method for detecting positive selection. Positive selection can be inferred from the observation of a site-specific increase in the frequency of nonsynonymous substitutions relative to synonymous substitutions [34]. Further, convergent evolution at specific places was investigated by comparing the amino acid sequences of homologous proteins from various species. Positive selection may be inferred when two distantly related species independently develop identical amino acid changes at the same place [35, 36].

2.4. Conservation Analyses. ConSurf [37] was used to conduct the conservation study of ZFP42. The ZFP42 amino acid sequences utilized in this investigation were retrieved

from public databases like GenBank and UniProt. The ConSurf server is a bioinformatics platform for analyzing protein sequences for conservation. Conserved areas in protein structures are often utilized to locate functional domains and residues [38, 39]. Each residue in the protein sequence is assigned a conservation score from 1 (most varied) to 9 (most conserved) by the ConSurf service. When visualizing a protein's structure, it might be useful to highlight conserved residues by giving them a distinct color or form. We looked into the synteny conservation of genomic regions surrounding the ZFP42 gene across mammalian species. Using the ConSurf library (https://consurf.tau.ac.il/) [40], the evolutionary conservation of amino acid residues in the mammalian ZFP42 protein was studied. These amino acids are more common in protein-protein interactions and enzymatic crannies. Changes in conserved amino acids are more harmful than polymorphisms in more flexible areas of a protein because they affect the protein's function and structure [41].

2.5. Protein Modeling and Structural Analysis. Crystal structural comparison of human ZFP42 protein and human compound C2H2 zinc finger domains [42] reveals high homology between the two. As a result, it served as a starting point for creating a 3D model of the human ZFP42 protein. Software packages like SWISS-MODEL [43], I-TESSAR [44], and Phyre2 (https://www.sbg.bio.ic.ac.uk/phyre2/html) [45] were used for protein modeling.

3. Results

The zinc finger protein 42 (ZFP42) is a DNA-binding transcription factor involved in regulating gene expression. The protein contains one or more zinc finger domains, which are structural motifs found in various DNA-binding proteins. Zinc finger domains consist of repeating sequences of amino acids that chelate zinc ions and are responsible for the DNA-binding activity of the protein. ZFP42 is a C2H2-type zinc finger protein, meaning that it contains two cysteine and two histidine residues in each zinc finger domain that coordinate the zinc ion. ZFP42 interacts with particular DNA sequences via its C2H2 zinc finger domains, which enables it to connect to target genes and control their expression. In this study, we looked at the evolution of the zinc finger protein ZNF91 subfamily and at the role it may play in the formation of mammalian undifferentiated stem cells. The functional effects of positively selected sites in the ZNF42 subfamily were discovered utilizing a combination of comparative genomics, phylogenetic analysis, and molecular modeling. Several positively selected sites within the ZNF42 subfamily have been identified, and many of these are located inside the proteins' DNA-binding domains. Signs of convergent evolution, when different lineages share adaptive changes, were also found at some of these sites. We used molecular modeling to estimate the effect of amino acid modifications on protein structure and DNA binding affinity in order to evaluate the potential functional consequences of these favorably selected sites. We found that several strategically placed motifs were predicted to affect the selectivity of DNA binding, which may allow the proteins to regulate distinct target genes or signaling cascades.

3.1. Gene Structure and Annotation Analysis. Annotations to the genome are often shown by colored blocks along the chromosome and transcriptional arrows. You should first locate the genes, exons, introns, promoters, and untranslated regions (UTRs) in the region of interest. Enhancers, promoters, and transcription factor binding sites, among other regulatory elements, had their positions mapped out thanks to genome annotations. These components can be crucial in regulating gene expression and shedding light on the genes' putative roles. Annotations of the genome uncovered evolutionarily conserved areas, which are significant for maintaining genomic structure and function. These areas can help you deduce whether or not your study area is biologically significant. Both a Kruppel-type zinc finger domain (involved in DNA binding) and a BTB/POZ domain (engaged in protein-protein interactions) are present in the ZFP42 protein. The protein functions as a transcriptional regulator, controlling processes including embryonic stem cell self-renewal and differentiation. In humans, there are three isoforms of the ZFP42 gene that result from alternative splicing: ZFP42 isoform 1, ZFP42 isoform 2, and ZFP42 isoform 3. The first isoform, 1, is the longest and most well-studied. In contrast, the N-terminal area is absent in isoform 2, and both the N-terminal and C-terminal regions are absent in isoform 3 (Figure 1). Similar to its human counterpart, the mouse ZFP42 gene consists of eight exons covering roughly 49 kilobases. The rodent ZFP42 protein, which regulates self-renewal and differentiation in embryonic stem cells, has both a Kruppel-type zinc finger domain and a BTB/POZ domain.

The term "gene synteny" describes the preservation of a gene's physical relationship from one species to another. Several mammalian species' synteny of the zinc finger protein 42 (ZFP42) gene has been analyzed. The ZFP42 gene is situated in close proximity to C11orf98 and CEP295NL on human chromosome 4. These adjacent genes are preserved in a wide range of mammalian species (Figure 2). ZFP42 is a gene found on chromosome 7 in mice, and its close neighbors include the Dctn2 and Rad23b genes. DNA repair is mediated by the Rad23b gene, while intracellular transport is regulated by the Dctn2 gene. Humans and numerous other mammals share these genes.

Although the genes immediately next to ZFP42 are conserved between humans and mice, the order and orientation of these genes with respect to ZFP42 are not. This points to the possibility of chromosomal rearrangements involving the ZFP42 locus throughout mammalian evolution. Zinc finger protein 42 (ZFP42) in mammals has a complex evolutionary history, and its gene tree might shed light on its ancestry and connections to other genes. Many different kinds of mammals share a copy of the gene ZFP42. This includes primates, rodents, carnivores, and ungulates. ZFP42 sequences from several mammalian species were used to create a gene tree, which demonstrates the high degree of conservation of this gene and the evolutionary relationship among various mammalian species. Several

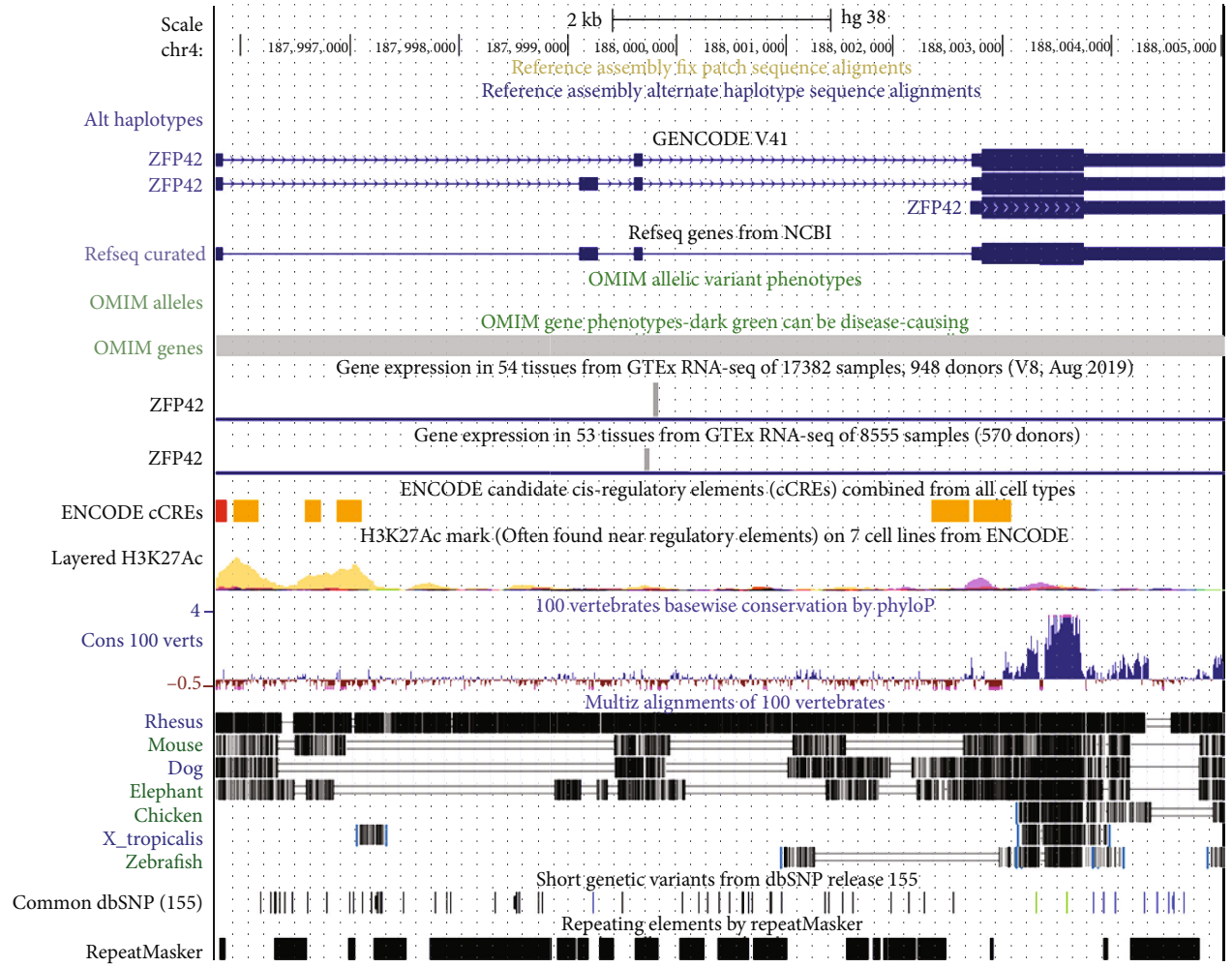

Figure 1: Genome annotation displays a genome location of zinc finger protein 42. The introns are shown as horizontal lines connecting the blocks representing the coding exons. Thinner blocks at the beginning and conclusion of the aligning regions represent the 5′ and 3′ untranslated regions (UTRs). The aligning sections (often exons) are displayed as black blocks when the query is a cDNA. Darkness in dense display mode is proportional to the amount of features that align with the region or the quality of the match.

evolutionary processes, including gene duplications, speciation, and loss, are depicted in the gene tree for ZFP42 in mammals. The ZFP42 gene, for instance, has three recognized isoforms in humans but only two in mice. These variants are thought to have resulted via alternative splicing following the split between humans and mice (Figure 2). In addition, the gene tree shows that some species have lost the ZFP42 gene, while others have multiple copies. For example, in the primate lineage, at least two gene duplication events have given rise to ZFP42 paralogs, such as ZFP42L1 and ZFP42L2. Overall, the gene tree of ZFP42 in mammals suggests that the gene has undergone a complex evolutionary history, with some species retaining the gene, others losing it, and others undergoing gene duplication events. Despite these events, the conservation of the gene across a wide range of mammalian species indicates that it likely plays an important biological role.

3.2. Positive Selection. Adaptive selection is a process in which certain alleles, or versions of a gene, become more common in a population over time due to the frequency of favorable traits associated with those alleles. The concept of adaptive selection is central to the theory of evolution and is believed to play a key role in the development of new species and the evolution of complex traits. Evidence of adaptive selection in the evolution of zinc finger protein 42 (ZFP42) would suggest that this gene has undergone adaptive evolution, in which changes in its genetic makeup have occurred due to natural selection acting on traits that improve the chances of survival and reproduction. Observing adaptive selection in ZFP42 would provide important insights into the evolution of this gene and its role in the development and function of various tissues and cell types in mammals. The amount and locations of histidine and cysteine residues involved in zinc atom coordination are used to classify zinc fingers. The C2H2 class is the first to be characterized, and it is distinguished by the presence of cysteines as the first pair of zinc-coordinating residues and histidines as the second pair. Some members of this family have been shown to bind zinc dependently to DNA or RNA in various

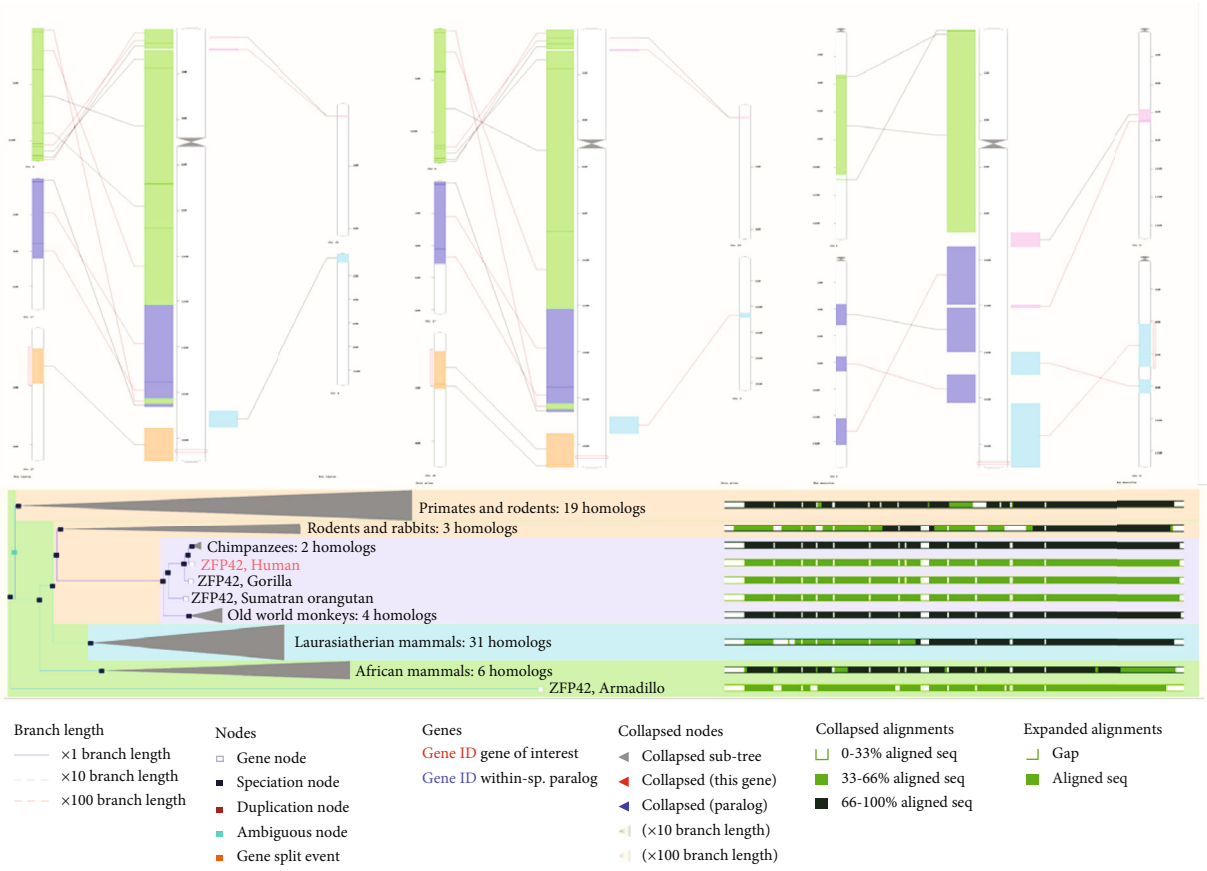

Figure 2: Gene synteny refers to the conservation of the physical linkage of genes across different species. Zinc finger protein 42 (ZFP42) gene synteny has been studied in different mammalian species. Gene tree shows the evolutionary relationships among genes across different species. The gene tree of ZFP42 in mammals suggests that the gene has undergone a complex evolutionary history, with some species retaining the gene, others losing it, and others undergoing gene duplication events. Despite these events, the conservation of the gene across a wide range of mammalian species indicates that it likely plays an important biological role.

experimental experiments. The structural stability of the C2H2 zinc fingers depends not only on the conserved zinc ligand residues but also on several other locations. Four residues following the second cysteine are the most conserved spots, and it is usually an aromatic or aliphatic residue. In addition, a comprehensive domain-wide profile was created (Figure 3).

3.3. Conservation Analysis. The zinc finger protein 42, also known as ZFP42, is a transcription factor that is present in all mammalian species. The protein's conservation across species shows it serves a job that has been conserved throughout evolution, specifically in the regulation of embryonic stem cell self-renewal and differentiation. Sequence similarity and preserved expression patterns are two ways to quantify gene conservation. The term "sequence similarity" is used to describe the degree to which the DNA and amino acid sequences of a gene are similar from one species to another. Conversely, when we talk about whether or not a gene's expression is conserved across species, we are referring to whether or not that gene is expressed in a consistent manner. According to the amino acid sequence similarity study, there is a significant degree of conservation of the ZFP42 protein across mammalian species, with a high degree of similarity between the human, mouse, and other mammalian ZFP42 proteins (Figure 4). The ZFP42 gene's DNA sequence is similarly well conserved, showing almost no variation between species. ZFP42 serves as a transcription factor, and its DNA-binding specificity and transcriptional activity are both dependent on its domain structure. ZFP42 is essential for the development and function of many mammalian organs and cell types, although our knowledge of its domain structure and involvement in gene regulation is limited. Proteins with several zinc finger domains are encoded by the ZFP42 (zinc finger protein 42) gene found in mammals. Transcription factors, or zinc finger proteins, attach to DNA to control gene expression. ZFP42 is a protein that has been found to be functionally conserved across multiple mammalian species, including humans, chimps, mice, rats, and dogs. Some parts of the protein show more divergence between species than others, but overall, there is a high degree of conservation.

3.4. Adaptive Selection. Zinc finger protein 42 (ZFP42) has undergone an intricate process of evolution that reflects the shifting genetic and environmental influences that have changed the gene over time. Several mammals, including humans, include the DNA-binding transcription factor

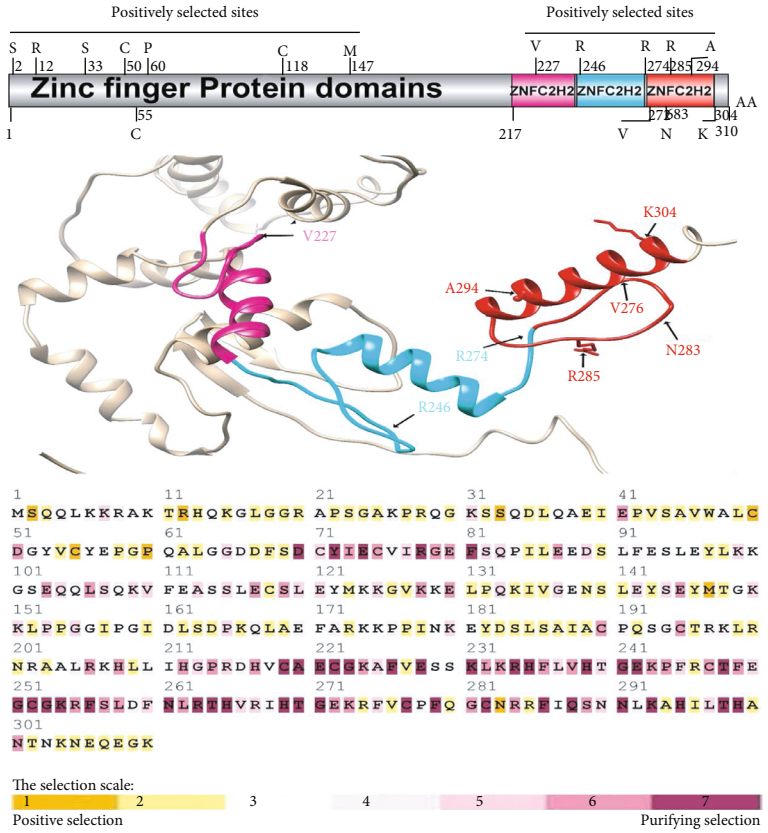

FIGURE 3: Zinc finger protein 3D structure, domain alignment, and positively chosen locations. The ZNFC2H2 domain structure of a zinc finger protein. Positive selection analysis of the ZFP42 gene using the Selecton server. A plot of the dN/dS ratio along the length of the ZFP42 protein sequence. The red line indicates the threshold for positive selection (dN/dS > 1). The Bayes empirical Bayes (BEB) analysis of the ZFP42 protein sequence identified four amino acid sites (positions 47, 156, 207, and 258) with a high probability of being under positive selection (posterior probability > 0.95). The amino acid substitutions at these sites are shown in red. Structural representation of the ZFP42 protein showing the location of the positively selected amino acid sites (red spheres) in the context of the protein structure. The protein is shown in cartoon representation, with the selected amino acids highlighted in space-filling representation. Overall, these results suggest that the ZFP42 gene has undergone positive selection, with specific amino acid sites playing a key role in the adaptive evolution of the protein.

ZFP42, which is involved in controlling the expression of certain genes. Natural selection, gene duplication and loss, recombination, mutations, and gene duplication are likely to have all had an impact on the evolution of ZFP42. ZFP42's structure and function may vary as a result of mutations, which are the main source of genetic variation in populations. Gene duplication and loss can also play a role in the evolution of ZFP42, as these events can result in multiple copies of the gene in a genome and create new functional units. When a species' genes change due to selective forces in its environment, this is called adaptive evolution. Several studies have shown evidence of positive selection operating on the gene for zinc finger protein 42 (ZFP42) during its adaptive evolution in different mammalian species (Figure 5). Natural selection promotes variants in a population's genetic makeup that also give a selective advantage, which is known as positive selection. ZFP42 is a protein shown to have experienced positive selection in both humans and other primates and animals. These areas may have undergone adaptive modifications in response to environmental forces, as they are considered critical for regulating gene expression. An aBSREL analysis discovered evidence of episodic diversifying selection on eight of your phylogeny's forty-four branches. Formal tests for diversifying selection were conducted on 44 different branches. After making adjustments for the effects of repeated tests, the significance of the findings was determined by applying the likelihood ratio test at a significance level of p 0.05. In the table that presents the detailed data, you will see the significance and the number of rate categories inferred at each branch. A statistical summary of the models that best fit the data is presented in the following table. Baseline MG94xREV is shorthand for the MG94xREV baseline model, which assumes that each branch has a single rate category. The full adaptive model is shorthand for the adaptive aBSREL model, which deduces the optimal number of rate categories for each branch.

The aBSREL (adaptive Bayesian selection and recombination estimation with local likelihoods) method is a statistical framework that can detect evidence of episodic diversifying selection in a given gene. Diversifying selection, also known as positive selection, refers to instances in which certain alleles, or gene versions, become more common in a

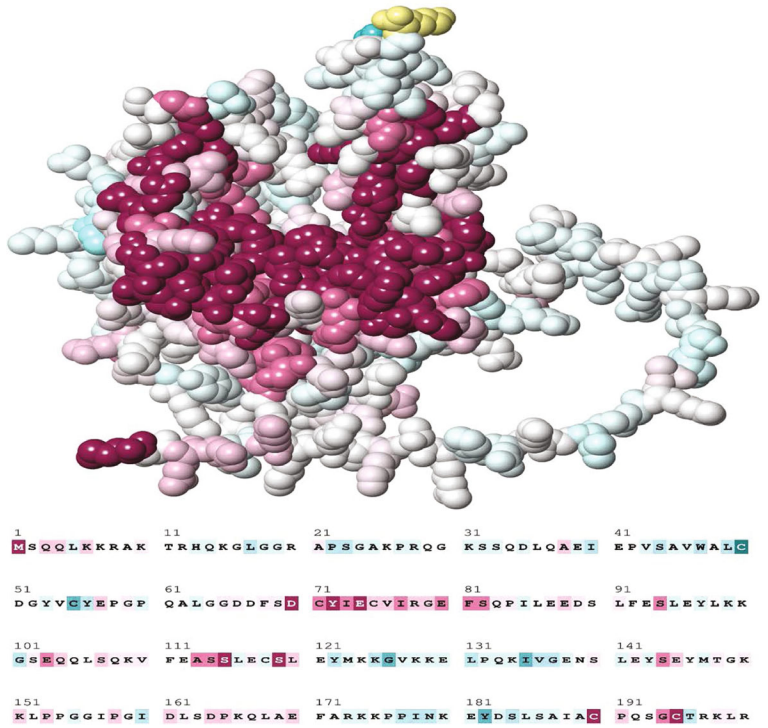

Figure 4: High sequence similarity and conserved ZFP42 gene expression patterns across several mammalian species imply that the protein has an essential and evolutionarily conserved function. It is also a useful model for researching stem cell biology and early development due to the gene's conservation across species.

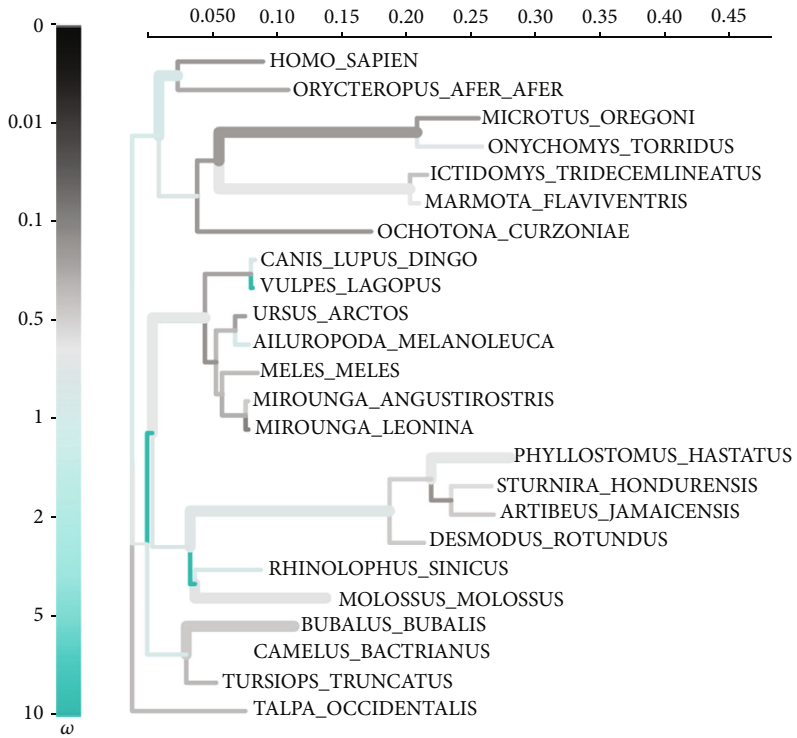

Figure 5: The aBSREL (adaptive Bayesian selection and recombination estimation with local likelihoods) method is a statistical framework that can be used to detect evidence of episodic diversifying selection in a ZFP42 gene.

population over time due to the frequency of favorable traits associated with those alleles. If aBSREL has found evidence of episodic diversifying selection in zinc finger protein 42 (ZFP42), this would suggest that this gene has undergone adaptive evolution, in which changes in its genetic makeup have occurred as a result of natural selection acting on traits that improve the chances of survival and reproduction.

3.5. Recombination Analysis. The mechanism of recombination also plays a significant role in the genesis of ZFP42. During meiosis, recombination occurs when genetic material is transferred between homologous chromosomes. This might result in the development of novel genetic combinations. ZFP42's evolution is thought to have been guided in part by natural selection. The frequency with which a population exhibits a change in its genetic makeup that increases its prospects of survival and reproduction rises with time. Proof of adaptive evolution can be found in the observation of positive selection, or the enrichment of specific alleles in a population. It has the potential to shed light on the history of ZFP42 and its function in mammalian tissue and cell development. The evolution of ZFP42 is a multifaceted and ever-changing process that is a product of the interaction of numerous forces. The evolution of ZFP42 holds key information about the historical significance of gene regulation in shaping mammalian development and physiology. Understanding the evolution of ZFP42 and its role in the development and function of different tissues and cell types in mammals would benefit greatly from the observation of episodic diversifying selection in this gene utilizing aBSREL.

More study is required to determine how ZFP42 and other genes have shaped mammalian evolution and the emergence of fundamental biological processes. Evidence of episodic diversifying selection in a gene can be found using the FEL analysis, a statistical technique (Table 1). Alleles, or variants of genes, that are associated with more desirable qualities tend to become more widespread over time; this process is known as "diversifying selection" (Figure 6). Zinc finger protein 42 (ZFP42) may have undergone adaptive evolution, in which changes to its genetic makeup have occurred as a result of natural selection acting on traits that increase the likelihood of survival and reproduction, if a FEL analysis has found evidence of episodic diversifying selection in this gene. Observing episodic diversifying selection in ZFP42 using FEL analysis would provide important insights into the evolution of this gene and its role in the development and function of various tissues and cell types in mammals. Further research is needed to understand the precise mechanisms by which ZFP42 and other genes have influenced the evolution of mammals and the development of key biological processes.

3.6. Branch-Site Unrestricted Statistical Test for Episodic Diversification. Based on the likelihood ratio test, there is evidence of episodic diversifying selection in this dataset ($p = 1.036e - 11$). BUSTED analysis (v4.1) was performed on the alignment from home/datamonkey/datamonkey-js-server/production/app/busted/output/63e61790f4c29257da21727f using HyPhy v2.5.46. This analysis included site-to-site synonymous rate variation. BUSTED is a statistical method that can detect evidence of episodic diversifying selection in a given gene. Diversifying selection, also known as positive selection, refers to instances in which certain alleles, or gene versions, become more common in a population over time due to the frequency of favorable traits associated with those alleles (Figure 7). If a BUSTED analysis has found evidence of episodic diversifying selection in zinc finger protein 42 (ZFP42), this would suggest that this gene has undergone adaptive evolution, in which changes in its genetic makeup have occurred as a result of natural selection acting on traits that improve the chances of survival and reproduction.

Observing episodic diversifying selection in ZFP42 using BUSTED analysis would provide important insights into the evolution of this gene and its role in the development and function of various tissues and cell types in mammals. Further research is needed to understand the precise mechanisms by which ZFP42 and other genes have influenced the evolution of mammals and the development of key biological processes.

3.7. Genetic Algorithm for Recombination Detection. GARD discovered clear recombination bottlenecks. As fast as 27.00 models per second, GARD analyzed 11826 models. With 942 possible breakpoints in the alignment, a search space of 6148571947736 models with up to 5 breakpoints was available, of which the genetic algorithm investigated 0%. Zinc finger protein 42 (ZFP42) may have experienced genetic recombination, the swapping of bits of DNA between pairs of identical chromosomes, if recombination breakpoints can be found in it. Recombination breakpoints can occur when two chromosomes exchange genetic material, leading to the formation of new combinations of genes and the rearrangement of the genetic material. The genetic analysis for recombination detection (GARD) is a bioinformatics tool that can identify recombination breakpoints in a given gene or genomic region. If GARD has found evidence of recombination breakpoints in ZFP42, this would suggest that the gene has undergone rearrangements due to genetic recombination (Figure 8). Observing recombination breakpoints in ZFP42 may provide important insights into the evolution of this gene and its role in the development and function of various tissues and cell types in mammals. Further research is needed to understand the precise mechanisms by which recombination breakpoints have influenced the evolution of ZFP42 and the development of key biological processes.

4. Discussion

The zinc finger protein 42 (ZFP42) is a transcription factor that plays a role in regulating gene expression in undifferentiated stem cells. It is involved in the maintenance of stem cell pluripotency and self-renewal, which are critical properties for the proper functioning of these cells [46]. Studies have revealed that ZFP42 has undergone adaptive evolution in mammals during evolution. This adaptive evolution has been associated with genetic signatures that indicate changes in the structure and function of the protein. These changes

Table 1: Detailed site-by-site results from the FEL analysis.

Partition	Codon	Alpha	Beta	Alpha = beta	LRT	*p*-value	Total branch length	*p*-asmp	Class
1	85	0.000	3.045	1.861	4.038	0.0445	7.453	0.0000	Diversifying
1	133	0.000	4.170	2.596	6.446	0.0111	10.396	0.0000	Diversifying
1	137	0.000	0.665	0.469	2.753	0.0971	1.876	0.0000	Diversifying
1	150	0.475	2.555	1.973	2.873	0.0901	7.903	0.0000	Diversifying
1	155	0.000	2.270	1.692	7.144	0.0075	6.777	0.0000	Diversifying
1	245	0.000	1.896	1.425	5.115	0.0237	5.706	0.0000	Diversifying
1	258	0.000	0.637	0.448	2.746	0.0975	1.794	0.0000	Diversifying
1	281	0.000	0.907	0.703	2.994	0.0836	2.816	0.0000	Diversifying
1	305	0.000	0.665	0.471	3.325	0.0682	1.887	0.0000	Diversifying
1	354	0.000	0.885	0.592	4.568	0.0326	2.372	0.0000	Diversifying
1	400	0.000	0.456	0.239	3.837	0.0501	0.956	0.0000	Diversifying
1	404	0.000	0.779	0.506	5.059	0.0245	2.025	0.0000	Diversifying

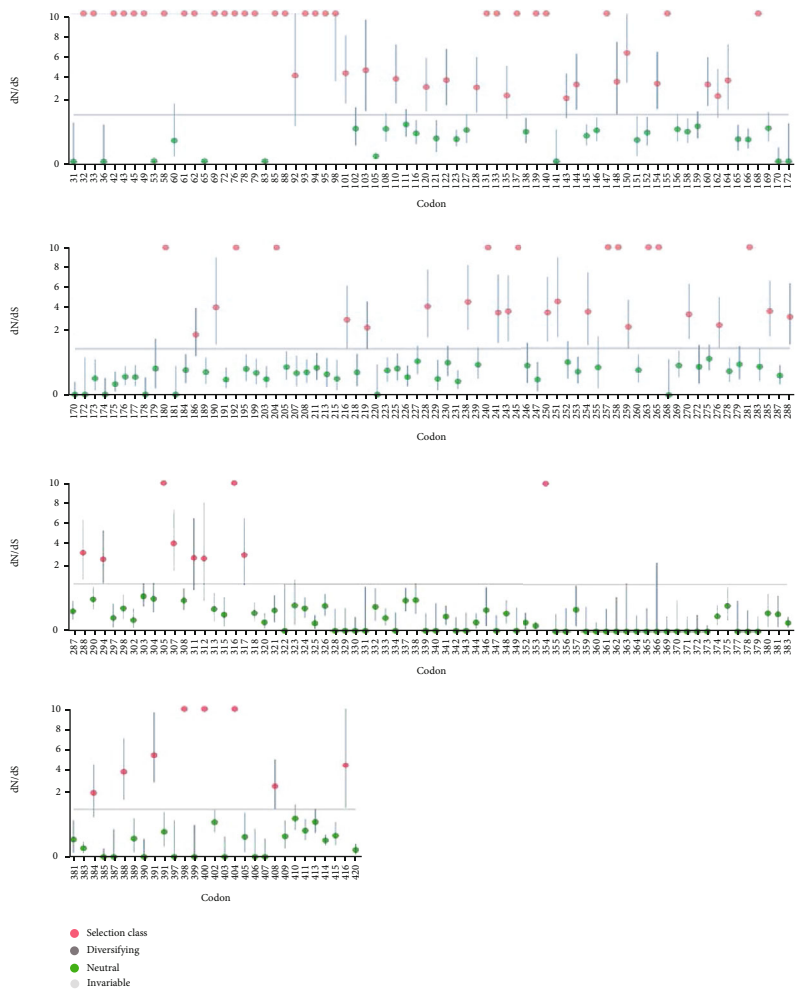

Figure 6: FEL analysis has found evidence of episodic diversifying selection in zinc finger protein 42 (ZFP42), and this would suggest that this gene has undergone adaptive evolution.

are thought to have allowed ZFP42 to better regulate the expression of specific genes in response to changing environmental conditions, which has been critical for the survival and success of mammals over time [47]. The genetic signatures of adaptive evolution in ZFP42 provide insight into the molecular mechanisms underlying the evolution of stem cells in mammals and how these cells have responded to changing environmental conditions over time. Understanding the role of ZFP42 in the regulation of gene expression in stem cells may also have implications for the development

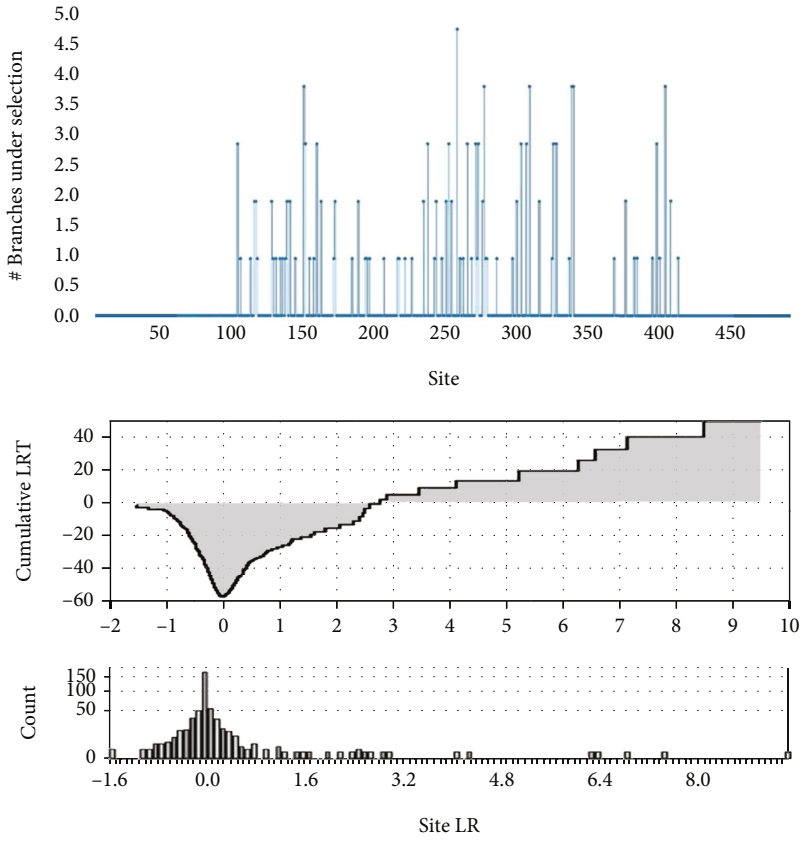

Figure 7: Cumulative distribution of the likelihood ratio test for the BUSTED test broken down by the contributions of individual sites.

of regenerative medicine and the treatment of diseases that involve stem cells. Zinc finger protein 42 (ZFP42) has been shown to have undergone adaptive evolution in mammals, revealing genetic signatures of this evolution [48]. As a transcription factor, ZFP42 plays an important role in regulating gene expression. The changes in its structure and function that have occurred over time have allowed it to better regulate the expression of specific genes in response to changing environmental conditions. This adaptive evolution has likely been critical for the survival and success of mammals during evolution. Studies investigating the genetic signatures of ZFP42's adaptive evolution have provided insight into the molecular mechanisms underlying the evolution of gene regulation in mammals [49]. They may have implications for the development of regenerative medicine and the treatment of diseases involving gene expression changes [50].

Studies have identified genetic signatures of ZFP42's adaptive evolution, providing insight into the molecular mechanisms underlying the evolution of gene regulation in mammals. These studies suggest that ZFP42 has evolved to better respond to environmental changes and maintain cellular functions, such as pluripotency and self-renewal, in stem cells. The adaptive evolution of ZFP42 highlights the importance of transcription factors in gene regulation and the role that evolution plays in shaping their function [51, 52]. Understanding the molecular mechanisms underlying the evolution of ZFP42 may have important implications for developing regenerative medicine and treating diseases that involve gene expression changes. Positive selection is a process of natural selection in which certain alleles, or versions of a gene, become more common in a population over time due to an increase in the frequency of favorable traits associated with those alleles. In the zinc finger protein 42 (ZFP42) case, researchers have identified instances of positive selection in this gene, suggesting that it has undergone evolutionary changes that have allowed it to better adapt to its environment. ZFP42 is a gene involved in regulating gene expression, and it is believed to play a role in the development and function of various tissues and cell types [53]. The observation of positive selection in this gene may indicate that changes in ZFP42 have played a role in the evolution of mammals by contributing to the development and function of these tissues and cells (Figure 3). Further research is needed to understand the precise mechanisms by which ZFP42 and other genes have influenced the evolution of mammals and the development of various tissues and cell types. However, the identification of positive selection in ZFP42 highlights the importance of this gene in the evolutionary process and its potential role in the development of key biological processes.

The analysis of zinc finger protein 42 (ZFP42) can reveal important insights into the process of adaptive evolution, which refers to changes in the genetic makeup of a species that have occurred due to natural selection, acting on traits that improve the chances of survival and reproduction. In the case of ZFP42, researchers may analyze the gene to

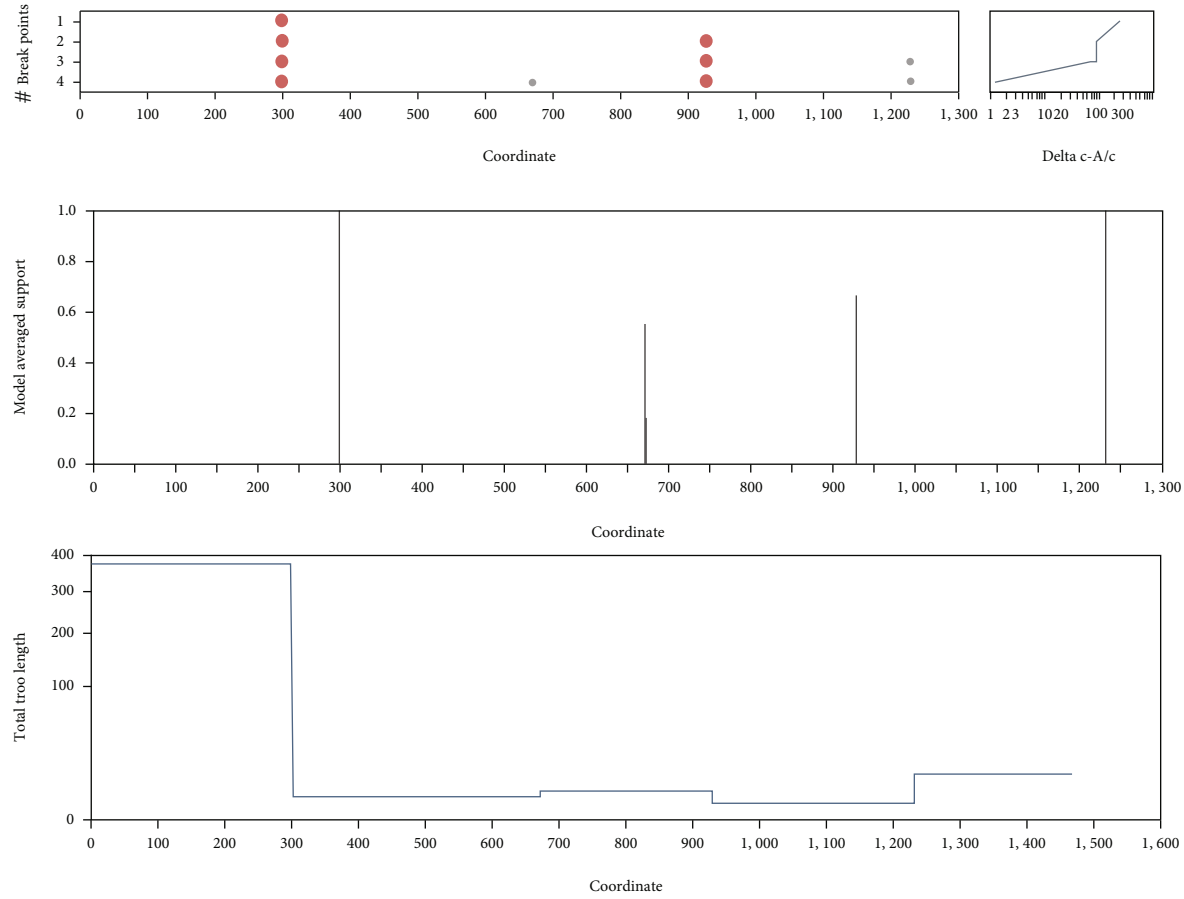

Figure 8: GARD has found evidence of recombination breakpoints in ZFP42, and this would suggest that the gene has undergone rearrangements due to genetic recombination.

identify instances of positive selection, a process in which certain alleles become more common in a population over time due to an increase in the frequency of favorable traits associated with those alleles [54].

Such analysis can provide evidence for the role of ZFP42 in the evolution of mammals, as changes in this gene may have played a role in the development and function of various tissues and cell types. In particular, ZFP42 is believed to be involved in regulating gene expression, and its analysis may reveal how changes in this gene have influenced the regulation of genes in various tissues and cell types, leading to their evolution [55]. Overall, the analysis of ZFP42 can provide valuable insights into the process of adaptive evolution, highlighting the importance of this gene in the evolutionary process and its potential role in the development of key biological processes in mammals. Episodic positive selection, also known as diversifying selection, refers to instances in which certain alleles, or versions of a gene, become more common over time due to the frequency of favorable traits associated with those alleles (Figure 6). Evidence of episodic positive selection in zinc finger protein 42 (ZFP42) would suggest that this gene has undergone adaptive evolution, in which changes in its genetic makeup have occurred due to natural selection acting on traits that improve the chances of survival and reproduction [56].

There have been several studies that have identified evidence of episodic positive selection in ZFP42. For example, researchers have compared the sequences of ZFP42 from different species or populations and found that certain gene regions have evolved faster than others, suggesting that these regions may have undergone positive selection [57, 58]. In addition, phylogenetic analysis and comparative genomics have provided further evidence for episodic positive selection in ZFP42, as gene changes have occurred at different times in different species. Overall, the evidence of episodic positive selection in ZFP42 highlights the importance of this gene in the evolutionary process and its potential role in the development and function of various tissues and cell types in mammals. Further research is needed to understand the precise mechanisms by which ZFP42 and other genes have influenced the evolution of mammals and the development of key biological processes.

5. Conclusions

Adaptive evolution refers to changes in the genetic makeup of a species that have occurred due to natural selection, which acts on traits that improve the chances of survival and reproduction. ZFP42 is a gene that is involved in the regulation of undifferentiated stem cells, which are cells that

have the potential to differentiate into many different cell types. The observation of positive selection in this gene may indicate that changes in ZFP42 have played a role in the evolution of mammals by contributing to the development and function of undifferentiated stem cells. In conclusion, ZFP42's adaptive evolution in mammals reveals genetic signatures that shed light on the evolution of gene regulation in mammals. It highlights the importance of transcription factors in gene regulation and the role of evolution in shaping their function. Further research in this area has the potential to lead to new therapeutic strategies for diseases that involve changes in gene expression. Overall, the findings of positive selection in ZFP42 reveal important genetic signatures of adaptive evolution in undifferentiated stem cells, which have likely played a crucial role in the evolution of mammals over time. Further research is needed to understand the precise mechanisms by which ZFP42 and other genes have influenced the evolution of stem cells and other key biological processes in mammals.

Data Availability

All data relevant to this paper shall be openly available to the readers.

Conflicts of Interest

The authors declared no conflict of interest.

Acknowledgments

The researchers would like to acknowledge the Deanship of Scientific, Research, Taif University for funding this work.

References

[1] R. Abu-Dawud, N. Graffmann, S. Ferber, W. Wruck, and J. Adjaye, "Pluripotent stem cells: induction and self-renewal," *Philosophical Transactions of the Royal Society, B: Biological Sciences*, vol. 373, no. 1750, 2018.

[2] G. Liu, B. T. David, M. Trawczynski, and R. G. Fessler, "Advances in pluripotent stem cells: history, mechanisms, technologies, and applications," *Stem Cell Reviews and Reports*, vol. 16, no. 1, pp. 3–32, 2020.

[3] M. Punetha, K. K. Bajwa, S. Dua et al., "Pluripotent stem cells for livestock health and production," *Current Stem Cell Research & Therapy*, vol. 17, no. 3, pp. 252–266, 2022.

[4] M. Mossahebi-Mohammadi, M. Quan, J.-S. Zhang, and X. Li, "FGF signaling pathway: a key regulator of stem cell pluripotency," *Frontiers in Cell and Development Biology*, vol. 8, p. 79, 2020.

[5] F. Martínez Sosa and M. Pilot, "Molecular mechanisms underlying vertebrate adaptive evolution: a systematic review," *Genes*, vol. 14, no. 2, p. 416, 2023.

[6] J. Merilä and A. P. Hendry, "Climate change, adaptation, and phenotypic plasticity: the problem and the evidence," *Evolutionary Applications*, vol. 7, no. 1, pp. 1–14, 2014.

[7] A. C. Vinton, S. J. Gascoigne, I. Sepil, and R. Salguero-Gómez, "Plasticity's role in adaptive evolution depends on environmental change components," *Trends in Ecology & Evolution*, vol. 37, no. 12, pp. 1067–1078, 2022.

[8] K. M. Poluri, K. Gulati, S. Sarkar, K. M. Poluri, K. Gulati, and S. Sarkar, "Structural and functional properties of proteins," in *Protein-Protein Interactions: Principles and Techniques*, pp. 1–60, Springer, 2021.

[9] M. Camps, A. Herman, E. Loh, and L. A. Loeb, "Genetic constraints on protein evolution," *Critical Reviews in Biochemistry and Molecular Biology*, vol. 42, no. 5, pp. 313–326, 2007.

[10] N. D. Temperley, S. Berlin, I. R. Paton, D. K. Griffin, and D. W. Burt, "Evolution of the chicken Toll-like receptor gene family: a story of gene gain and gene loss," *BMC Genomics*, vol. 9, no. 1, pp. 1–12, 2008.

[11] D. C. King, J. Taylor, L. Elnitski, F. Chiaromonte, W. Miller, and R. C. Hardison, "Evaluation of regulatory potential and conservation scores for detecting cis-regulatory modules in aligned mammalian genome sequences," *Genome Research*, vol. 15, no. 8, pp. 1051–1060, 2005.

[12] H. Ellegren, "Comparative genomics and the study of evolution by natural selection," *Molecular Ecology*, vol. 17, no. 21, pp. 4586–4596, 2008.

[13] F. Sievers and D. G. Higgins, "Clustal omega," *Current Protocols in Bioinformatics*, vol. 48, no. 1, 2014.

[14] F. Sievers and D. G. Higgins, "The Clustal Omega Multiple Alignment Package," in *Multiple Sequence Alignment: Methods and protocols*, pp. 3–16, Springer, 2021.

[15] S. Kumar, G. Stecher, M. Li, C. Knyaz, and K. Tamura, "MEGA X: molecular evolutionary genetics analysis across computing platforms," *Molecular Biology and Evolution*, vol. 35, no. 6, pp. 1547–1549, 2018.

[16] F. Sievers and D. G. Higgins, "Clustal omega for making accurate alignments of many protein sequences," *Protein Science*, vol. 27, no. 1, pp. 135–145, 2018.

[17] Z. Yang, "PAML 4: phylogenetic analysis by maximum likelihood," *Molecular Biology and Evolution*, vol. 24, no. 8, pp. 1586–1591, 2007.

[18] K. Tamura, J. Dudley, M. Nei, and S. Kumar, "MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0," *Molecular Biology and Evolution*, vol. 24, no. 8, pp. 1596–1599, 2007.

[19] T. Little and N. Cobbe, "The evolution of immune-related genes from disease carrying mosquitoes: diversity in a peptidoglycan- and a thioester-recognizing protein," *Insect Molecular Biology*, vol. 14, no. 6, pp. 599–605, 2005.

[20] J. Van den Eynden and E. Larsson, "Mutational signatures are critical for proper estimation of purifying selection pressures in cancer somatic mutation data when using the dN/dS metric," *Frontiers in Genetics*, vol. 8, p. 74, 2017.

[21] X. Xia and X. Xia, "Maximum Likelihood in Molecular Phylogenetics," in *Bioinformatics and the Cell: Modern Computational Approaches in Genomics, Proteomics and Transcriptomics*, pp. 381–395, Spriger, 2018.

[22] Z. Yang, W. S. Wong, and R. Nielsen, "Bayes empirical Bayes inference of amino acid sites under positive selection," *Molecular Biology and Evolution*, vol. 22, no. 4, pp. 1107–1118, 2005.

[23] K. Scheffler and C. Seoighe, "A Bayesian model comparison approach to inferring positive selection," *Molecular Biology and Evolution*, vol. 22, no. 12, pp. 2531–2540, 2005.

[24] Z. Yang and R. Nielsen, "Codon-substitution models for detecting molecular adaptation at individual sites along specific lineages," *Molecular Biology and Evolution*, vol. 19, no. 6, pp. 908–917, 2002.

[25] M. J. Ahmad, H. I. Ahmad, M. M. Adeel et al., "Evolutionary analysis of makorin ring finger protein 3 reveals positive selection in mammals," *Evolutionary Bioinformatics*, vol. 15, 2019.

[26] S. Weaver, S. D. Shank, S. J. Spielman, M. Li, S. V. Muse, and S. L. Kosakovsky Pond, "Datamonkey 2.0: a modern web application for characterizing selective and other evolutionary processes," *Molecular Biology and Evolution*, vol. 35, no. 3, pp. 773–777, 2018.

[27] B. Murrell, S. Moola, A. Mabona et al., "FUBAR: a fast, unconstrained Bayesian approximation for inferring selection," *Molecular Biology and Evolution*, vol. 30, no. 5, pp. 1196–1205, 2013.

[28] S. L. K. Pond and S. D. Frost, "Datamonkey: rapid detection of selective pressure on individual sites of codon alignments," *Bioinformatics*, vol. 21, no. 10, pp. 2531–2533, 2005.

[29] B. P. Carlin and S. Chib, "Bayesian model choice via Markov chain Monte Carlo methods," *Journal of the Royal Statistical Society: Series B: Methodological*, vol. 57, no. 3, pp. 473–484, 1995.

[30] J. P. Huelsenbeck, B. Larget, and M. E. Alfaro, "Bayesian phylogenetic model selection using reversible jump Markov chain Monte Carlo," *Molecular Biology and Evolution*, vol. 21, no. 6, pp. 1123–1133, 2004.

[31] S. Whelan and N. Goldman, "A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach," *Molecular Biology and Evolution*, vol. 18, no. 5, pp. 691–699, 2001.

[32] W. C. Wheeler, "Phylogenetic network analysis as a parsimony optimization problem," *BMC Bioinformatics*, vol. 16, no. 1, pp. 296–299, 2015.

[33] M. M. Desai and D. S. Fisher, "Beneficial mutation–selection balance and the effect of linkage on positive selection," *Genetics*, vol. 176, no. 3, pp. 1759–1798, 2007.

[34] J. A. Tennessen, "Positive selection drives a correlation between non-synonymous/synonymous divergence and functional divergence," *Bioinformatics*, vol. 24, no. 12, pp. 1421–1425, 2008.

[35] H. I. Ahmad, J. Zhou, M. J. Ahmad et al., "Adaptive selection in the evolution of programmed cell death-1 and its ligands in vertebrates," *Aging*, vol. 12, no. 4, pp. 3516–3557, 2020.

[36] Y.-D. Li, Z.-Y. Xie, Y.-L. Du et al., "The rapid evolution of signal peptides is mainly caused by relaxed selection on non-synonymous and synonymous sites," *Gene*, vol. 436, no. 1-2, pp. 8–11, 2009.

[37] G. Celniker, G. Nimrod, H. Ashkenazy et al., "ConSurf: using evolutionary data to raise testable hypotheses about protein function," *Israel Journal of Chemistry*, vol. 53, no. 3-4, pp. 199–206, 2013.

[38] H. I. Ahmad, A. Iqbal, N. Ijaz et al., "Molecular evolution of the activating transcription factors shapes the adaptive cellular responses to oxidative stress," *Oxidative Medicine and Cellular Longevity*, vol. 2022, Article ID 2153996, 13 pages, 2022.

[39] F. Glaser, T. Pupko, I. Paz et al., "ConSurf: identification of functional regions in proteins by surface-mapping of phylogenetic information," *Bioinformatics*, vol. 19, no. 1, pp. 163-164, 2003.

[40] O. Goldenberg, E. Erez, G. Nimrod, and N. Ben-Tal, "The ConSurf-DB: pre-calculated evolutionary conservation profiles of protein structures," *Nucleic Acids Research*, vol. 37, Supplement 1, pp. D323–D327, 2009.

[41] M. P. Miller and S. Kumar, "Understanding human disease mutations through the use of interspecific genetic variation," *Human Molecular Genetics*, vol. 10, no. 21, pp. 2319–2328, 2001.

[42] R. Mackeh, A. K. Marr, A. Fadda, and T. Kino, "C2H2-type zinc finger proteins: evolutionarily old and new partners of the nuclear hormone receptors," *Nuclear Receptor Signaling*, vol. 15, 2018.

[43] A. Waterhouse, M. Bertoni, S. Bienert et al., "SWISS-MODEL: homology modelling of protein structures and complexes," *Nucleic Acids Research*, vol. 46, no. W1, pp. W296–W303, 2018.

[44] J. Yang and Y. Zhang, "I-TASSER server: new development for protein structure and function predictions," *Nucleic Acids Research*, vol. 43, no. W1, pp. W174–W181, 2015.

[45] C. Manjusha, A. Santhiagu, S. Soumiya, V. Adarsh, and S. J. Prakash, "Phyre 2 and I-TASSER web portal for protein modeling, prediction and validation of gel Q and gel K genes from gellan gum producing bacterial strainSphingomonas paucimobilisATCC 31461," *Research Journal of Pharmacy and Technology*, vol. 12, no. 1, pp. 27–36, 2019.

[46] D. Pal and M. Rao, "Long noncoding RNAs in pluripotency of stem cells and cell fate specification," in *Long Non Coding RNA Biology*, pp. 223–252, Springer, 2017.

[47] D. Guallar, R. Pérez-Palacios, M. Climent et al., "Expression of endogenous retroviruses is negatively regulated by the pluripotency marker Rex1/Zfp42," *Nucleic Acids Research*, vol. 40, no. 18, pp. 8993–9007, 2012.

[48] H. D. Tadepally, "Evolution of C2H2-Zinc Finger Genes in Mammalian Genomes, [M.S. thesis]," UniversitédeMontréal, 2007.

[49] H. I. Ahmad, F. A. Khan, M. A. Khan et al., "Molecular evolution of the bactericidal/permeability-increasing protein (BPIFA1) regulating the innate immune responses in mammals," *Genes*, vol. 14, no. 1, p. 15, 2022.

[50] D. Gökbuget and R. Blelloch, "Epigenetic control of transcriptional regulation in pluripotency and early differentiation," *Development*, vol. 146, no. 19, 2019.

[51] H. I. Ahmad, G. Afzal, M. N. Iqbal et al., "Positive selection drives the adaptive evolution of mitochondrial antiviral signaling (MAVS) proteins-mediating innate immunity in mammals," *Frontiers in Veterinary Science*, vol. 8, p. 1680, 2022.

[52] M. A. Missinato, S. Murphy, M. Lynott et al., "Conserved transcription factors promote cell fate stability and restrict reprogramming potential in differentiated cells," *Nature Communications*, vol. 14, no. 1, 2023.

[53] W. D. Gifford, S. L. Pfaff, and T. S. Macfarlan, "Transposable elements as genetic regulatory substrates in early development," *Trends in Cell Biology*, vol. 23, no. 5, pp. 218–226, 2013.

[54] K. Wang, D. Liu, J. Hernandez-Sanchez et al., "Genome wide association analysis reveals new production trait genes in a male Duroc population," *PLoS One*, vol. 10, no. 9, article e0139207, 2015.

[55] P. Khaitovich, W. Enard, M. Lachmann, and S. Pääbo, "Evolution of primate gene expression," *Nature Reviews Genetics*, vol. 7, no. 9, pp. 693–702, 2006.

[56] H. I. Ahmad, G. Afzal, S. Sadia et al., "Structural and evolutionary adaptations of Nei-like DNA glycosylases proteins involved in base excision repair of oxidative DNA damage in vertebrates," *Oxidative Medicine and Cellular Longevity*, vol. 2022, Article ID 1144387, 20 pages, 2022.

[57] H. I. Ahmad, A. R. Asif, M. J. Ahmad et al., "Adaptive evolution of peptidoglycan recognition protein family regulates the innate signaling against microbial pathogens in vertebrates," *Microbial Pathogenesis*, vol. 147, article 104361, 2020.

[58] G. Shuler and T. Hagai, "Rapidly evolving viral motifs mostly target biophysically constrained binding pockets of host proteins," *Cell Reports*, vol. 40, no. 7, article 111212, 2022.