Pakistan Veterinary Journal

ISSN: 0253-8318 (PRINT), 2074-7764 (ONLINE)
DOI: 10.29261/pakvetj/2025.240

RESEARCH ARTICLE

Comparative Genomics Analysis of Cation Channel Sperm-associated Proteins (CATSPERs) Associated with Sperm Motility in Livestock

Kun Li[1,2], Mubbashar Hassan[3], Hafiz Ishfaq Ahmad[4*], Abid Hussain Shahzad[3], Sanan Raza[3], Iram Qadeer[5], Sayyed Aun Muhammad[6], and Wentong Liu[7*]

[1]College of Veterinary Medicine, Nanjing Agricultural University, Nanjing 210095, China; [2]School of Basic Medicine, Hubei University of Arts and Science, 296 Longzhong Road, Xiangyang, 441053, China; [3]Department of Clinical Sciences, Theriogenology Section, College of Veterinary and Animal Sciences, Jhang, Sub Campus UVAS, Lahore, Pakistan; [4]Department of Animal Breeding and Genetics, Faculty of Veterinary and Animal Sciences, The Islamia University of Bahawalpur, Pakistan; [5]Department of Zoology, The Govt. Sadiq College Women University, Bahawalpur, Pakistan; [6]Department of Clinical Sciences, College of Veterinary and Animal Sciences, Jhang, Sub Campus UVAS, Lahore, Pakistan; [7]School of Basic Medicine, Hubei University of Arts and Science, 296 Longzhong Road, Xiangyang, 441053, China.
*Corresponding author: ishfaq.ahmad@iub.edu.pk; 11847@hbuas.edu.cn

ARTICLE HISTORY (25-478)

Received: May 06, 2025
Revised: August 02, 2025
Accepted: August 04, 2025
Published online: September 15, 2025

Key words:
CATSPER channels
Evolutionary genomics
Positive selection
Protein-protein interactions
Livestock fertility
Sperm motility

ABSTRACT

The CATSPER gene family encodes essential ion channels that govern sperm motility and male fertility across mammalian species. This study presents a comprehensive evolutionary and functional analysis of CATSPER genes (CATSPER 1-4) in cattle, sheep and pigs, integrating comparative genomics, analysis of selection pressure, and structural characterization. Using maximum likelihood models (M1a/M2a and M7/M8), strong signatures of positive selection in all four CATSPER genes were identified ($2\Delta\ell$=38.27-67.45, $P<0.05$), with Bayesian Empirical Bayes analysis revealing 11-17 positively selected sites per gene. These adaptive mutations were predominantly located in ion transport domains (Pfam) and showed lineage-specific patterns, particularly in primate CATSPER2/4. Concurrently, purifying selection maintained critical functional regions, as evidenced by high conservation scores (ConSurf) and structural integrity (ERRAT>74.32). Phylogenetic reconstruction demonstrated deep conservation of CATSPER proteins across Artiodactyla, with gene duplication events observed in rodent lineages. Functional characterization revealed 4-6 transmembrane domains per protein and identified key post-translational modification sites (phosphorylation, glycosylation). Protein-protein interaction networks (STRING) implicated CATSPER in a reproductive complex with AKAP3, SPAG6/16, and CABYR, regulating flagellar calcium signaling. Notably, nuclear hormone receptor binding motifs were predicted in promoter regions, suggesting endocrine regulation of CATSPER expression. These findings provide evolutionary and mechanistic insights into CATSPER-mediated fertility, highlighting the followings: 1) Adaptive evolution in ion transport domains correlates with species-specific sperm motility requirements, 2) Structural conservation maintains core channel function despite lineage-specific adaptations, and 3) Protein interaction networks integrate CATSPER into broader sperm motility pathways. The study establishes a framework for targeting CATSPER variants in livestock breeding programs and informs comparative models of male fertility across mammals.

To Cite This Article: Kun L, Hassan M, Ahmad HI, Shahzad AH, Raza S, Qadeer I, Muhammad SA and Liu W 2025. Comparative genomics analysis of cation channel sperm-associated proteins (CATSPERs) associated with sperm motility in livestock. Pak Vet J. http://dx.doi.org/10.29261/pakvetj/2025.240

INTRODUCTION

Mammalian spermatozoa become motile after ejaculation, but they need to acquire enhanced motility, known as hyperactivation, to penetrate the cumulus and zona pellucida before fertilizing the egg (Coy *et al.*, 2012). A voltage-sensitive Ca2+-selective current, regulated by the Cation Channel of Sperm (CATSPER),

allows Ca2+ to enter the sperm tail and trigger hyperactivation (Qi *et al.*, 2007). Thus, CATSPER regulates Ca2+ influx and facilitates hyperactivation, two processes that are crucial for male fertility. This enables sperm to get through the zona pellucida glycoprotein layer that surrounds the oocyte (Mina and Morel, 2020). Eutherian (placental) and monotreme (egg-laying) mammals both show sperm hyperactivation (Mina and Morel, 2020), which is different from non-mammalian species. Researchers have learned a lot about how sperm competition affects things we can see, but they still don't know exactly what genetic variables make sperm more competitive (Civetta and Ranz, 2019).

While recent breakthroughs in animal genomics have facilitated extensive studies on fertility-related genes, CATSPER genes have received comparatively little attention (Elango *et al.*, 2020). The availability of high-quality genome assemblies for key livestock species now gives us an unprecedented chance to look at the genetic structure of CATSPER genes in these animals (Park *et al.*, 2023). Such studies might show both functional features that have stayed the same and adaptations that are unique to each species and might have occurred during domestication or the establishment of a breed. Interestingly, early research shows that certain cattle breeds may display natural variation in CATSPER genes that may affect sperm motility characteristics (Montoto *et al.*, 2011). Such variations might be used as genetic tools in selective breeding of males to boost their reproductive performance in artificial insemination. These channels play a role in other parts of the cell function, apart from helping with calcium entry. The proteins connect different signaling pathways that control sperm capacitation, the process of chemotaxis, and how the acrosome responds (Yuan *et al.*, 2024).

The aim of this study was to explore the genomic characteristics of CATSPER genes among livestock species including cattle, sheep, and pigs, aiming at comparisons, as well as their possible influence on sperm motility. Employing various genomic and evolutionary tools, this study investigated the sequence variations, structural propertoes, and signals of selection in CATSPER genes.

MATERIALS AND METHODS

Data collection: The nucleotide and amino acid sequences of CATSPER gene of different mammalian species including cattle, sheep, and pigs, were analyzed in order to understand their evolutionary history and to identify regulatory regions within the genes. Sequence data were retrieved from publically available databases including the Ensembl Project for Automated Genome Annotation (Kersey *et al.*, 2016), and National Center for Biotechnology Information (NCBI) (Hancock *et al.*, 2014). The MEGA6 and ClustalW algorithms were used (Tamura *et al.*, 2013) to identify conserved regions and functional domains within the sequences. Phylogenetic trees were constructed to asses evolutionary changes in CATSPER proteins. Similarly, the study included analysis of genes regulatory regions responsible for the CATSPER expression.

Sequence analysis: The sequences were aligned using MUSCLE (Edgar, 2004) and MAFFT (Katoh and Standley, 2013), with manual modification to confirm reading frame correctness. Phylogenetic trees were rebuilt using IQ-TREE (Nguyen *et al.*, 2015) and MEGA X (Kumar *et al.*, 2018), with 1,000 bootstrap repeats. These trees were used as input file for selection analyses in the PAML (Yang, 2007). The Bayes Empirical Bayes (BEB) tests were employed on dataset to minimize false positive results. The positive selection results were further validated using analysis under model M8 in Selecton server and datamonkey webserver where various complementary approaches including Fast Unconstrained Bayesian AppRoximation (FUBAR) (Murrell *et al.*, 2013) and Mixed Effects Model of Evolution (MEME) (Murrell *et al.*, 2012) were used for episodic selection.

Selection tests: The signature of selection on each codon was identified using maximum likelihood approaches employed in PAML (Phylogenetic Analysis by Maximum Likelihood) package (Yang, 2007) to find out the codons under selection in the CATSPER gene sequences. Using likelihood ratio tests (LRTs), opposing models were compared to identify the best-fitting model among M8 vs. M8a, M2 vs. M2a (Yang, 2007). The accurate branch lengths of the phylogenetic tree were calculated using model M0 with fixed length of 0 for selection analysis at each codon site. The Bonferroni corrections and the Benjamini-Hochberg approaches were used to reduce the probabilities of false-positive results. The BEB approach was used to identify the codons under selection pressure during molecular evolution with posterior probabilities at 0.9 and 0.95 (Yang, 2007). To further validate the PAML results, the dataset was analyzed with selecton server that detected the codons under selection pressure using the MCMC model and the results were visualized in colored figures that displayed the results with neutral, purifying and positive selection during the course of evolution in mammalian lineages (Yang, 2007).

Phylogenetic analysis: Phylogenetic reconstruction was conducted to assess the evolutionary relationships among CATSPER gene sequences across multiple species. First, high-quality coding sequences were retrieved from public databases (NCBI GenBank/Ensembl) and aligned using MUSCLE v3.8 with codon-aware settings to maintain reading frame integrity. The alignment was manually curated to remove poorly aligned regions and gaps. The best-fit nucleotide substitution model (GTR+Γ+I) was selected using ModelFinder in IQ-TREE (Wong *et al.*, 2025) based on the Bayesian Information Criterion (BIC). Maximum likelihood (ML) trees were then reconstructed using IQ-TREE v3.0 with 1000 ultrafast bootstrap replicates to assess branch support. For Bayesian inference, MrBayes v3.2.7 (Wang *et al.*, 2025) was employed running two independent Markov Chain Monte Carlo (MCMC) (Barido-Sottani *et al.*, 2024) analyses for 2 million generations (sampling every 1000 generations) with a 25% burn-in. Convergence was confirmed by ensuring the average standard deviation of split frequencies fell below 0.01 and by examining trace plots in Tracer v1.7. The resulting phylogenetic trees were visualized and annotated using FigTree v1.4.4 (http://tree.bio.ed.ac.uk/software/figtree/). To further validate tree topology, gene trees were compared against a trusted species tree using TreeBeST,

reconciling discrepancies due to potential incomplete lineage sorting or gene duplication events. Branch-specific tests for selection were performed using the branch-site model in PAML to detect episodic diversifying selection along particular lineages.

Conservation and gene enrichment analysis: This study examined synteny around CATSPER genes to study functional conservation. ConSurf assessed amino acid conservation in human CATSPER proteins (Ashkenazy *et al.*, 2016), focusing on critical residues in active sites and protein interactions (Islam *et al.*, 2021). EnrichNet was used to analyze gene-pathway associations in syntenic regions (Liu *et al.,* 2017).

3D protein modeling: Homology modeling was used to predict CATSPER structures, using Swiss Model (Waterhouse *et al.*, 2018), I-TASSER (Zhang, 2008), and Phyre2 (Kelley *et al.*, 2015). Models were refined via energy minimization (Amber force field, UCSF Chimera; Huang *et al.*, 2014) and validated with ProSA (Wiederstein and Sippl, 2007). Selecton (v2.2) mapped selection pressures on codon alignments (Stern *et al.*, 2007).

RESULTS

Evidence of adaptive evolution in CATSPER genes: This study demonstrated compelling evidence of positive selection acting on CATSPER genes across diverse animal species. To identify signatures of selection, likelihood ratio tests were conducted comparing site-based evolutionary models—specifically, the nearly neutral model (M1a) versus the positive selection model (M2a) and the neutral model (M7) versus the positive selection model (M8)-using reconstructed phylogenetic trees as input. These analyses revealed statistically significant evidence of diversifying selection ($P<0.05$), suggesting that CATSPER genes have undergone adaptive evolution in certain lineages. The results strongly supported adaptive evolution in CATSPER1–4, as indicated by significant likelihood ratio test statistics ($2\Delta\ln L=38.27$, 13.93, 67.45 and 46.98, respectively; $P<0.05$), as presented in Table 1. Further validation was obtained through analyses using the Selection server and Bayesian Empirical Bayes (BEB) methods, which identified multiple sites under positive selection (Fig. 1). Conservation analysis conducted with the ConSurf server revealed that these positively selected sites were highly conserved across mammalian clades, underscoring their critical role in CATSPER function throughout mammalian evolution. Additionally, regions exhibiting high amino acid conservation were identified in both solvent-exposed and buried domains of the proteins, highlighting the evolutionary significance of these residues.

Sites under positive selection: The Mechanistic-Empirical Combination model implemented in the Selecton server identified specific sites under positive selection pressure in CATSPER genes. These sites are likely functionally significant, given their high conservation and localization within regions critical for protein activity. Analyses in the present study revealed a pattern of selective amino acid retention in regions exhibiting positive selection signals. These findings aligned with the essential role of CATSPER channels in sperm motility and fertilization, as illustrated in Fig. 2.

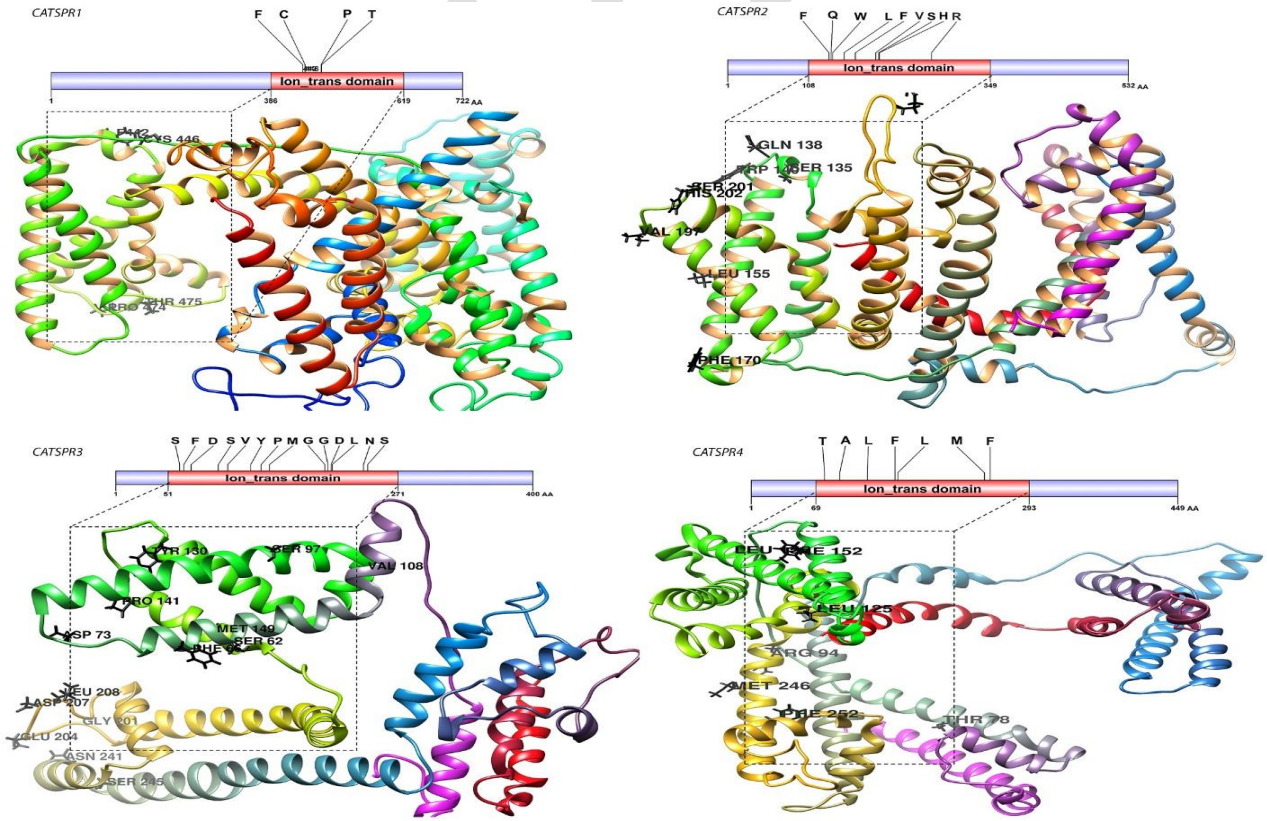

Fig. 1: Positive selection in the mammalian CATSPER gene family. This figure illustrates areas of the CATSPER gene family where positive selection has occurred across several mammalian species.

Fig. 2: Positive selection analysis of CATSPER genes using the selecton server. Yellow and magenta emphasize areas that have undergone positive selection, which may suggest functional significance and evolutionary adaptability. Blue and green zones of purifying selection—which point to the evolutionary preservation of function—showcase the dynamic character of evolutionary forces operating upon the CATSPER genes. The study reveals certain areas that are probably involved in reproductive benefits in mammals.

Table 1: Positive selection analysis of CATSPER genes using maximum likelihood to estimate the influence of selection on codons

Gene	Category	Model	Parameter estimates	PAML	MEME	FUBAR
CATSPER1	Nearly Neutral	M1a	P1=0.40609; P2=0.59391 ω1=0.07676; ω2=1.00	0	79, 80, 112, 131	112, 131, 207, 223
	Positive Selection	M2a	P1=0.37590; P2=0.37290; P=30.25121 ω1=0.08002; ω2=1.00; ω3=2.345	207, 223, 325, 327,372		
	Beta	M7	p=0.29797; q=0.25118	0		
	Beta & w>1	M8	p0=0.67913; p=0.37807; q=0.69928 P1=0.32087; ω1=1.75	79, 80, 114, 133		
CATSPER2	Nearly Neutral	M1a	P1=0.61207; P2=0.38793 ω1=0.10277; ω2=1.00000	0	94, 135, 271	135, 271, 578, 581
	Positive Selection	M2a	P1=0.59433; P2=00.26864; P=0.13703 ω1=0.04937; ω2= 0.36373; ω3=1.93294	88, 94, 135, 271		
	Beta	M7	p=0.35626; q=0.53763	0		
	Beta & w>1	M8	p0=0.80699; p=0.53277; q=1.56821 P1=0.19301; ω1=2.30951	88, 94, 175, 176		
CATSPER3	Nearly Neutral	M1a	P1=0.61242; P2=0.38758 ω1=0.14497; ω2=1.00000	0	41, 45, 46, 81, 93	46, 81, 93, 220, 257
	Positive Selection	M2a	P1=0.59434; P2=0.33754; P=0.06812	22, 27, 149,		
	Beta	M7	p=0.55182; q=0.84969	0		
	Beta & w>1	M8	p0=0.83797; p=0.83938; q=2.21098 P1=0.16203; ω1=1.45169	21, 27, 41, 45, 81, 220, 257		
CATSPER4	Nearly Neutral	M1a	P1=0.61783; P2=0.38217 ω1=0.07313; ω2=1.00000	0	17, 22, 34, 37, 40	25, 34, 40, 41, 64, 69
	Positive Selection	M2a	P1=0.59998; P2=0.29960; P=0.10042 ω1=0.04185; ω2=0.46523; ω3=1.74114	23, 34, 37, 40, 41, 357, 453		
	Beta	M7	p=0.26978; q=0.49522	0		
	Beta & w>1	M8	p0=0.85615; p=0.36397; q=1.12431 P1=0.14385; ω1=1.93677	17, 22, 34, 37, 40, 41, 64, 69		

Purifying selection and functional integrity: The present study also indicated purifying selection at work, maintaining the integrity of CATSPER genes by removing deleterious mutations. This was supported by the conserved nature of regions under purifying selection, as shown in Fig. 2.

Comprehensive selection analysis: In this study, the HyPhy package, including MEME, and FUBAR, was used to identify evolutionary signals revealing of positive selection. The convergence of results from these varied methods provided strong support to the recognized sites undergoing positive selection (Fig. 3), underscoring the adaptive significance of these genes in vertebrate evolution.

Phylogenetic conservation and evolution of CATSPERs: The molecular phylogenetic research into CATSPER proteins across mammalian species revealed a high degree of sequence and structural conservation, indicative of their critical role in sperm motility and male fertility. The phylogenetic analysis underscored the evolutionary pressure on these genes to maintain their function, with CATSPER gene sequences exhibiting substantial conservation across the studied species (Fig. 4).

The phylogenetic tree revealed evolutionary relationships among CATSPER1–4 genes in livestock species (cattle, sheep, pigs, goats, and related species), constructed using nucleotide/amino acid alignments (MEGA6, ClustalW). Each paralog (CATSPER1–4) formed distinct clusters, confirming independent gene lineages. Orthologs cluster by species (*Bos taurus, Capra hircus, Ovis aries*) reflected conserved sequences, while divergent CATSPER2 positioning suggested functional specialization. Species like *Sus scrofa* and *Equus caballus* exhibited unique evolutionary trajectories (Fig. 5).

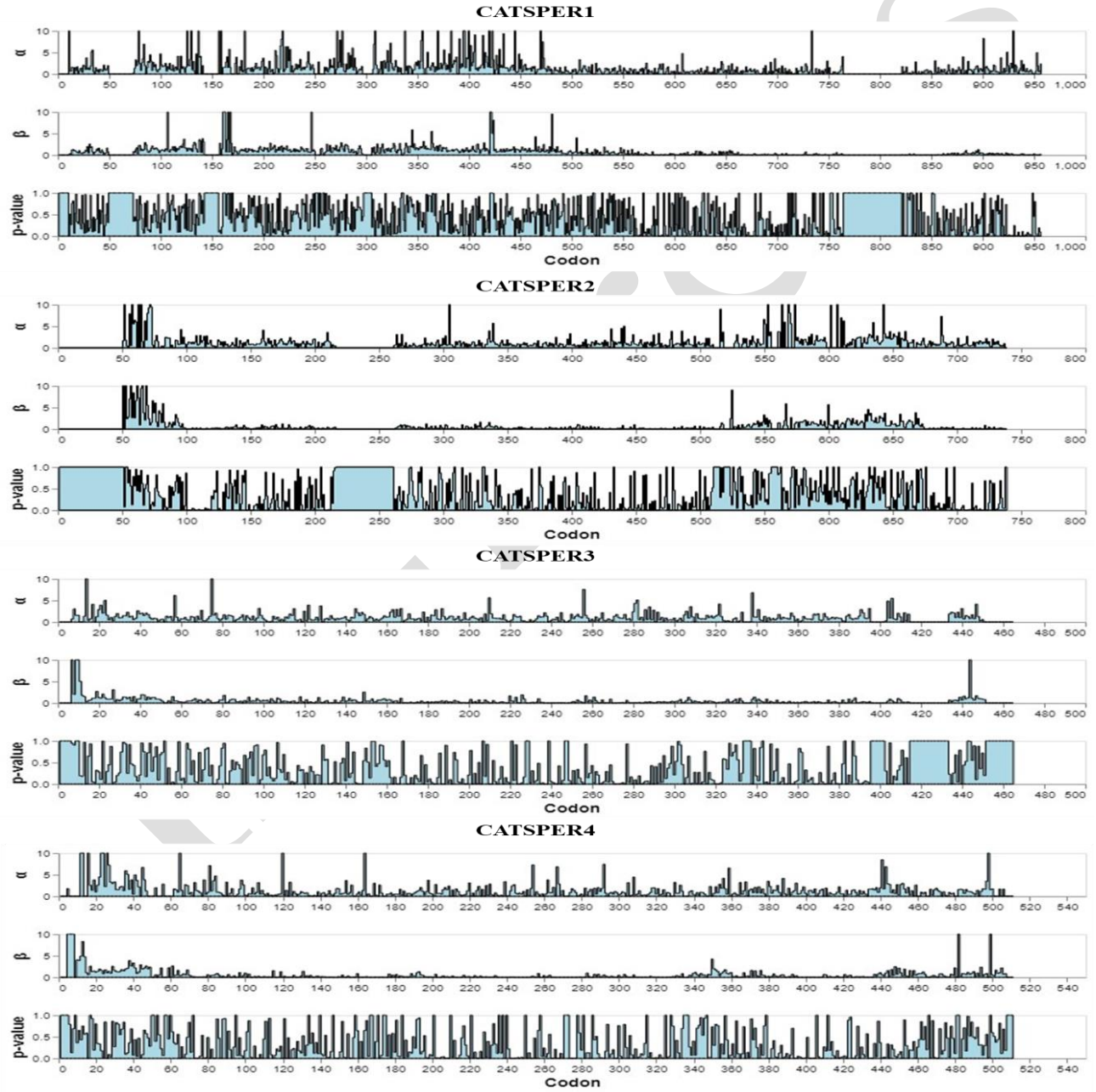

Fig. 3: Site-specific positive selection detected by FEL Analysis. Fixed Effects Likelihood (FEL) analysis drives the maximum likelihood estimates of synonymous (α) and non-synonymous (β) substitution rates for every site within the CATSPER genes displayed here. Sites with non-synonymous rates (β) above synonymous rates (α) indicate positive selection; estimates over 10 are limited at this level for clarity.

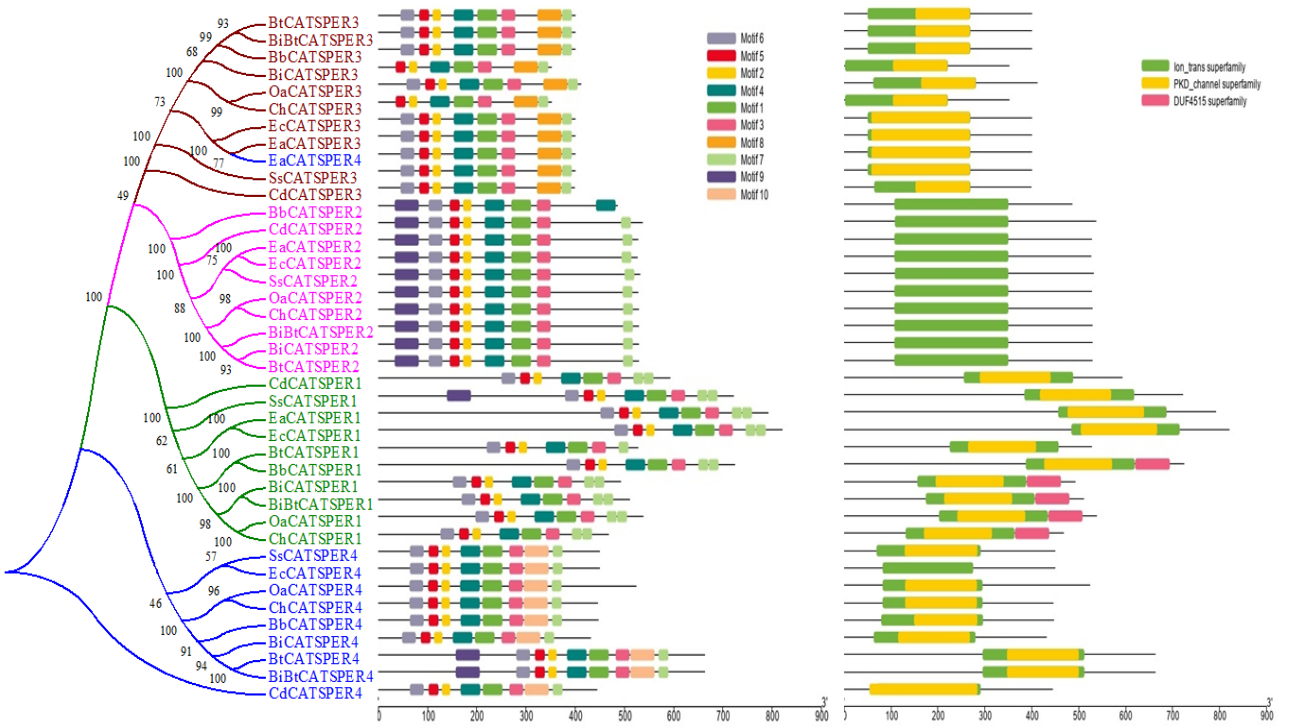

Fig. 4: Phylogenetic relationships, conserved motifs, and protein domain architectures of CATSPER isoforms across livestock species. The left panel displays a maximum likelihood phylogenetic tree of CATSPER isoforms (CATSPER1–4) from cattle (Bt), sheep (Oa), pigs (Ss), buffalo (Bb), bison (Bi), goat (Ch), donkey (Cd), and horse (Ea), with bootstrap values indicated at branch nodes. The middle panel shows the distribution of conserved motifs (Motif 1–10) identified using the MEME Suite, with each motif represented by a distinct colored box. The right panel presents predicted functional domains from the NCBI Conserved Domain Database (CDD), including the Ion_trans superfamily (green), PKD_channel superfamily (yellow), and DUF4515 superfamily (pink). The alignment reveals conserved structural organization among CATSPER proteins, reflecting both functional conservation and evolutionary divergence across livestock species.

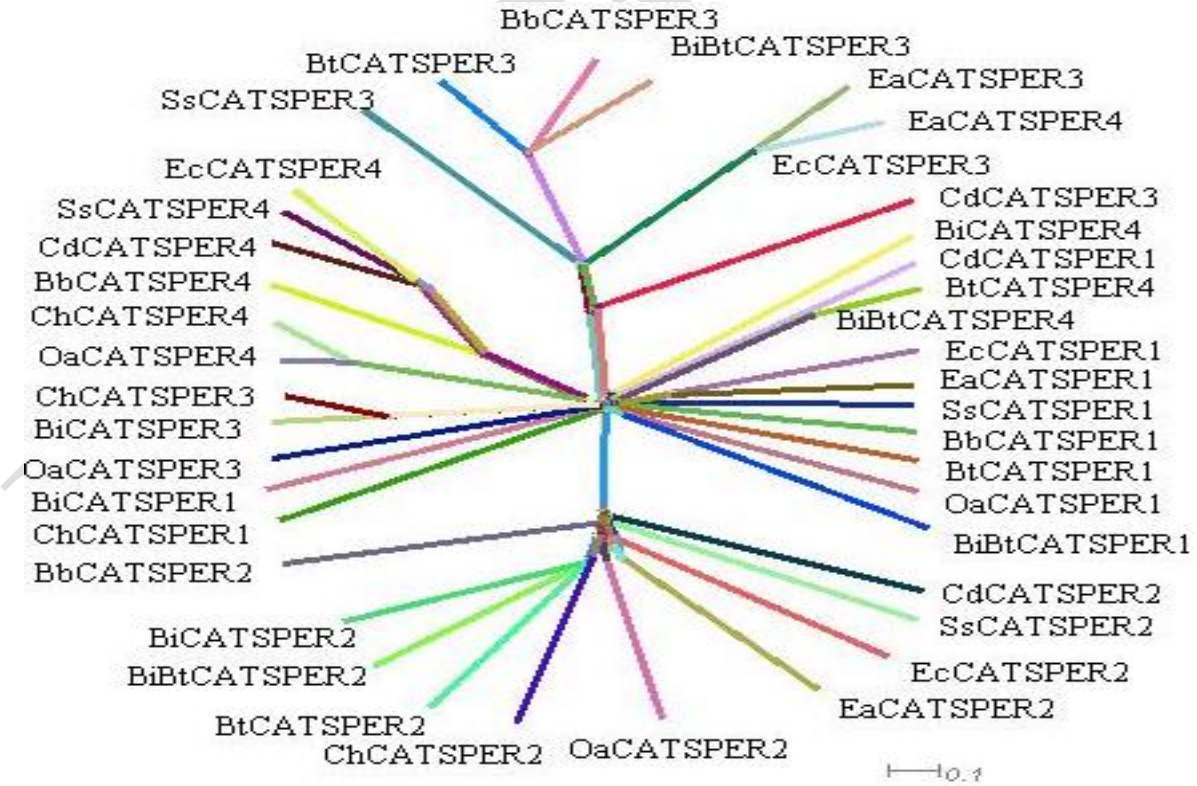

Fig. 5: Unrooted phylogenetic tree of CATSPER isoforms across livestock species. The unrooted tree depicts evolutionary relationships among CATSPER1–4 isoforms from cattle (Bt), buffalo (Bb), bison (Bi), sheep (Oa), goat (Ch), pig (Ss), donkey (Cd), horse (Ea), and their hybrids (e.g., BiBt). Constructed using the Neighbor-Joining method from amino acid sequences, the tree highlights distinct isoforms through color-coded branches, revealing clustering patterns that reflect sequence divergence and functional diversification within the CATSPER gene family. The scale bar indicates evolutionary distance.

Functional analysis of CATSPER proteins: The functional analysis of CATSPER proteins investigated the biological roles of these proteins, revealing multiple transmembrane domains integral to their function as ion channels in the plasma membrane of the sperm cell (Fig. 6). These domains were found to be essential for calcium ion transport, a key process in sperm hyperactivation and fertilization (Table 2). The ERRAT quality factor and GRAVY hydropathy index assessed protein structure quality and hydrophobicity, respectively, with results indicating high structural integrity and distinct hydrophobic characteristics of CATSPER proteins (Table 3).

Table 2: Conserved motifs detected in CATSPER proteins (CATSPER1, CATSPER2, CATSPER3, CATSPER4) gene family

No.	Sequence	width	Pfam Description
	TLFDLSDPMNFQNLLVAIFTLFILAT		
1	LDGWTDJYQDLDARN	41	Ion transport protein
2	YWKDGWNVLDFVIIVJLL	18	Ion transport protein
3	IIFILJGSFIFLNLFVAVMVTNFZNSLKK RILKLLTYSRGJRTIITALGRSLPSMA	29	Ion transport protein
4	SILILLFILMYIFA	41	Ion transport protein
5	LFEVSDWIFLSIYISEFLLKW	21	No Description
6	KKLJESPAFKNFIIFLIFLNTIVLVLETE	29	No Description
7	EFVHVLEKMQEBLPZKKQLQE ELVEKFKKTLRHTDPMVLDDFGTS	21	No Description
8	LPFIDIYLSTLDNQDATIYKLQE QGLSQAVPRHTIREILDSSRQKKLM	50	No Description
9	LGDQHQLVRFSIKPRHVERITH QGQQHQIAFSEVDDKKGNGNN	50	No Description
10	KLPLVHCAVARSEMSGVPQEPFMG	50	No Description

Post-translational modifications, such as phosphorylation and glycosylation, were also found to be significant ($P<0.05$) for CATSPER function. Additionally, nuclear hormone receptors (NHRs) appeared to regulate CATSPER gene expression, as evidenced by potential NHR binding sites in CATSPER gene promoters and the observed effects of androgen therapy on CATSPER mRNA expression levels in sperm motility (Fig. 7).

Protein-protein interaction analysis of CATSPER complex: Utilizing the STRING database, the computational analysis identified several key proteins (SPAG6, PPP1R12A, PPP1CC, PPP1CB, SPA17, ROPN1, AKAP3, FSCB) that potentially interacted with CATSPER and played a role in ion transport and signaling pathways crucial for sperm function. Sperm-associated antigens 6 and 16 (SPAG6 and SPAG16), integral components of the axoneme structure, were shown to interact with CATSPER. SPAG16, in particular, was important for flagellar movement and, by extension, sperm motility. Additionally, the calcium-binding tyrosine phosphorylation-regulated protein (CABYR), in cooperation with CATSPER and AKAP3, was linked to the control of sperm motility. Moreover, the control of sperm motility by the protein phosphatase 1 catalytic subunit gamma (PPP1CC) was identified through its interaction with the CATSPER channel (Fig. 8).

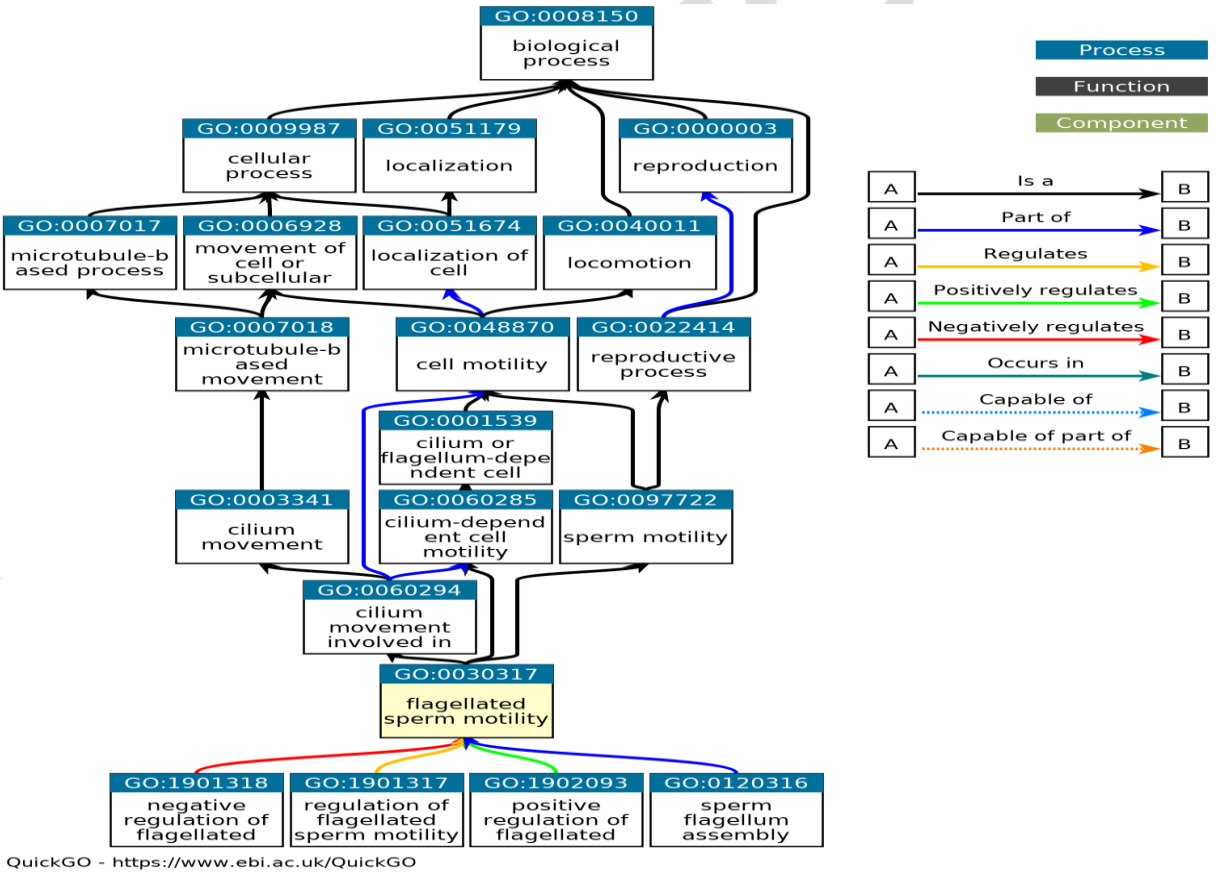

Fig. 6: Gene Ontology (GO) hierarchy of biological processes associated with flagellated sperm motility. This diagram depicts the hierarchical organization of GO terms linked to flagellated sperm motility (GO:0030317) and its regulatory pathways. The top-level term biological process (GO:0008150) branches into key subcategories, including reproduction (GO:0000003), localization (GO:0051179), and locomotion (GO:0040011), which further specify processes such as cell motility (GO:0048870), sperm motility (GO:0097722), and cilium-dependent movement (GO:0060285). The central term, flagellated sperm motility (GO:0030317), is influenced by both negative (GO:1901318) and positive (GO:1902093) regulation and is functionally associated with sperm flagellum assembly (GO:0120316). Color-coded arrows denote relationship types (e.g., is a, part of, regulates), clarifying functional and hierarchical connections. Generated using QuickGO (https://www.ebi.ac.uk/QuickGO).

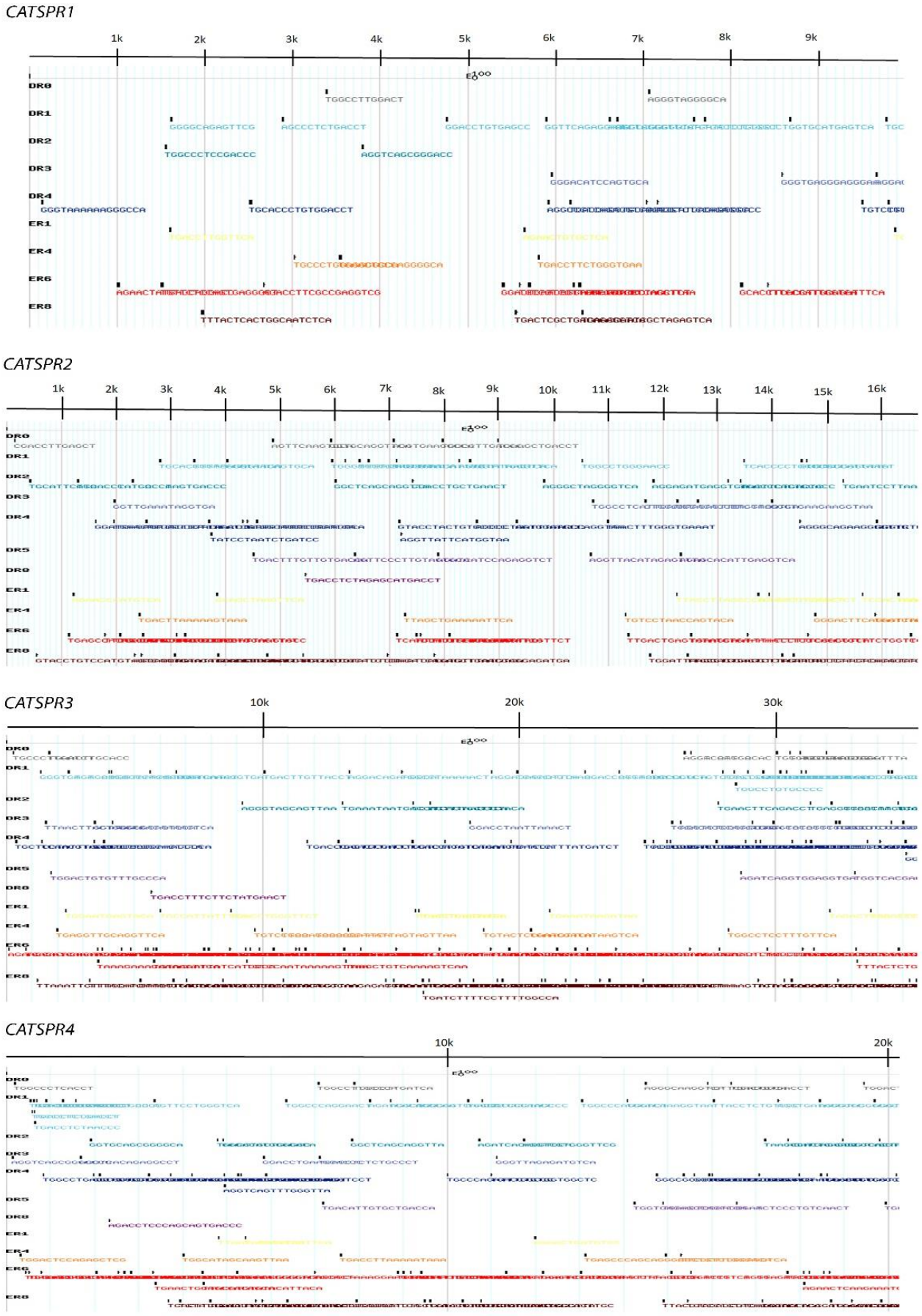

Fig. 7: The regulatory influence of nuclear hormone receptors (NHRs) on CATSPER gene expression. This figure highlights the possible regulatory routes regulated by different ligands, showing the connections between nuclear hormone receptors (NHRs) and the promoter regions of CATSPER genes. This illustrates gene expression profiles upon ligand binding, or other relevant measures demonstrating the regulatory influence of NHRs on CATSPER expression. Emphasizing the complex network of gene control, this graphic overview illustrates how NHRs can regulate CATSPER activity and support sperm function.

Table 3: Physiochemical characteristic of the Sus scrofa (reference species) CATSPER gene family

Gene	Chromosome	Exons	AA	MW (D)	ERRAT	VERIFY	pI	EC	AI	II	GRAVY
CATSPER1	2	12	722	82819.72	76.35	82.61	8.17	69330	76.07	49.47	-0.531
CATSPER2	1	15	*532*	62121.46	74.32	80.39	8.66	89380	50.95	105.70	0.079
CATSPER3	2	8	*400*	46529.16	78.43	84.12	6.80	48360	102.97	34.38	0.117
CATSPER4	6	11	*449*	51807.37	80.49	81.42	5.30	65890	107.66	42.26	0.187

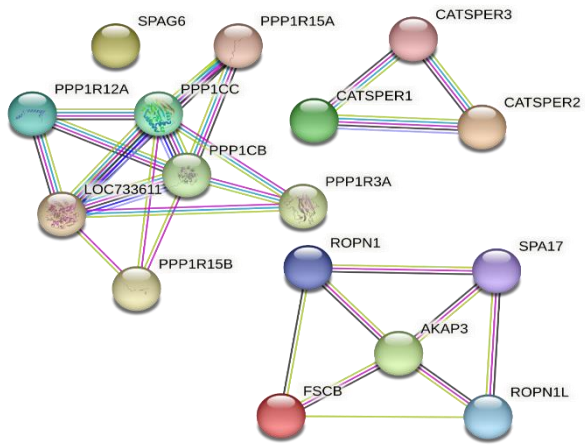

Fig. 8: Protein-protein interaction of CATSPER1, CATSPER2, CATSPER3, CATSPER4 gene family. The figure displays a panel of key genes involved in sperm function, including CATSPER1, CATSPER2, and CATSPER3, which encode ion channels critical for sperm hyperactivation and motility. Other genes, such as SPAG6, ROPN1, and AKAP3, are implicated in sperm flagellar structure and motility regulation. Protein phosphatase regulatory subunits (PPP1R12A, PPP1R15A, PPP1R15B, PPP1CB, PPP1CC, PPP1R3A) and sperm adhesion proteins (SPA17, FSCB, ROPN11) further highlight pathways essential for sperm signaling, motility, and fertilization competence. The inclusion of LOC733611 (uncharacterized locus) suggests potential novel roles in sperm biology.

DISCUSSION

This study revealed that CATSPER genes have experienced positive selection during adaptive evolution, emphasizing their role in sperm function and male fertility. The maximum likelihood ratio approaches compared the positive selection and neutral evolution models (M1a vs. M2a and M7 vs. M8), revealed signs of positive selection in CATSPER genes. These results were further validated by Bayesian Empirical Bayes (BEB) methods, which recognized the specific amino acid sites found under selective pressure during molecular evolution. The conservation analysis was performed using the ConSurf server and found that these sites are conserved across the mammalian lineages, suggesting that these are functionally essential, preserving the functioning and structural integrity of CATSPER channels during evolution (Yang *et al.*, 2024). Remarkably, these positively selected sites were found in both solvent-exposed and buried regions of the proteins, supporting the subtle variations in these domains may affect fertilization and sperm motility (Darszon *et al.*, 2011). The selecton server used the MCMC model and identified the key sites that likely impact ion channel activity. The convergence of results from these independent methods strengthens the case for adaptive evolution in CATSPER genes. Structural assessments using ERRAT and GRAVY indicated that these proteins maintained high integrity and possessed distinct hydrophobic properties, which are likely crucial for their membrane localization and ion conductance. Additionally, post-translational modifications, such as phosphorylation and glycosylation, appear to modulate CATSPER activity, potentially in response to physiological cues during sperm maturation. The presence of nuclear hormone receptor (NHR) binding sites in CATSPER gene promoters suggests that their expression may be hormonally regulated, offering a possible explanation for observed variations in sperm function under different endocrine conditions (Ali *et al.*, 2023).

Selected evolutionary factors influence proteins in the reproductive system and frequently experience rapid evolution and functional divergence. Mammalian sperm-specific proteins, which mostly consist of cell surface binding proteins and enzymes, undergo fast evolutionary changes. Studies on sperm-specific channels and transporters are limited. However, there is evidence of positive selection on indel substitutions in the first exon of the CATSPER1 gene in mammals.

The detection of evolutionary rates often relies on the use of non-synonymous (dN) and synonymous (dS) nucleotide substitution values, along with the dN/dS ratio (Jeffares *et al.*, 2014). We computed the dN and dS values for cattle, sheep and pig orthologous gene pairs of CATSPER and CATSPERβ genes using the codeml program, which is integrated into the PAML package. The dN value of CatSper and CATSPERβ genes was 5.8-fold and 3.2-fold greater, respectively, than housekeeping genes and non-sperm tissue-specific genes, and 1.5-fold greater than other sperm-specific genes (Cai and Clapham, 2008). Both CatSpers and CatSperβ seem to be influenced by strong selective pressures that encourage variation in amino acids. Different domains of CatSpers have shown evidence of rapid evolution in all four CatSper proteins. For example, CATSPER1 exhibited significant positive selection in the N-terminal cytoplasmic domain but showed evolutionary constraint in the 6-TMS domains (dN/dS=0.062). The TMS domains of CATSPER4 exhibited comparable limitations, as indicated by a dN/dS value of 0.077. Conversely, the TMS domains of CATSPER2 and CATSPER3 experienced stronger selective pressures, with dN/dS values of 0.435 and 0.172, respectively.

The unique evolutionary patterns observed in CatSper channel domains may enhance the diversity of channels in their ability to regulate protein or second-messenger interactions (Lobley *et al.*, 2003). It is possible that CatSperβ, being the only auxiliary protein, has developed at a quicker rate to adapt to the varied evolutionary locations of CatSper proteins, leading to a high dN/dS ratio. The availability of several CATSPER sequences, particularly from invertebrates, enables us to quantify the extent of functional differentiation between CATSPER1-4, as demonstrated in earlier studies. The functional divergence coefficients between pairs of CATSPER1-4 groups were calculated and found to be statistically significant ($P<0.05$).

Therefore, it seems that each CATSPER group has undergone modified evolutionary limitations following its

separation from a hypothetical ancestral duplication (Rivera, 2022). Consistently, the analysis of functional distance also reveals that all four CATSPERS have similarly lengthy functional branch lengths. This suggests that all four CatSpers diverged from the putative primordial CATSPER protein at comparable overall evolutionary rates. However, even though the general evolutionary rates are similar, the evolutionary processes may have differing effects on specific sites that are distinct in each CATSPER gene. Our data indicate that CATSPER1-4 has undergone significant changes in functional limitations following potential ancestral replication, resulting in evolutionary novelties. Ultimately, the swift development of CATSPERS continues to have a significant impact on regulating CATSPER functions in mammalian lineages. In light of the crucial role CATSPER proteins play in sperm motility and male fertility, mutations in these proteins can have profound implications for male infertility (Ren and Xia, 2010). This study has identified various isoforms of the CATSPER family, each potentially exerting distinct influences on sperm motility and fertilization processes.

Conclusions: The present study suggests that CATSPER protein plays a vital role in mammalian sperm motility. The identification of many conserved regulatory domains in the promoters of CATSPER genes demonstrates the importance of this protein in gene expression control. This study on the genomes of different species indicated that CATSPER genes are linked coevolutionary in mammalian lineages. This suggests that these proteins control the beating of sperm flagella. The loss of these proteins might be due to sperm signaling pathways picking up new components by coincidence or due to the sperm adaptation to a new fertilization environment, which may have developed separately in different taxonomic groupings. This study also reveals that the way CATSPER is controlled in non-metazoan is different from the way it is controlled in livestock species.

Conflict of interest: The authors declare that there is no conflict of interest.

Acknowledgment: This work was supported by the Natural Science Foundation of China (Grant No. W2433064).

Authors' contributions: LK and HIA were responsible for conceptualization, investigation, methodology, data curation, project administration, visualization and writing up the original draft. LK, MH, HIA, WL, AHS, SR, IQ and SAM carried out review and editing of the manuscript.

REFERENCES

Ali H, Mahmood M, Adnan M, *et al.*, 2023. Comparative evolutionary and structural analyses of the TYRP1 gene reveal molecular mechanisms of biological functions in mammals. J King Saud Univ Sci 35(6): 102772.

Ashkenazy H, Abadi S, Martz E, *et al.,* 2016. ConSurf 2016: An improved methodology to estimate and visualize evolutionary conservation in macromolecules. Nucleic Acids Res 44:344-50.

Barido-Sottani J, Schwery O, Warnock RC, *et al.,* 2024. Practical guidelines for Bayesian phylogenetic inference using Markov chain Monte Carlo (MCMC). Open Res Europe 3:204.

Cai X and Clapham DE, 2008. Evolutionary genomics reveals lineage-specific gene loss and rapid evolution of a sperm-specific ion channel complex: CatSpers and CatSperβ. PloS One 3(10):3569.

Civetta A and Ranz JM, 2019. Genetic factors influencing sperm competition. Front Genet 10:820.

Coy P, García-Vázquez FA, Visconti PE, *et al.,* 2012. Roles of the oviduct in mammalian fertilization. Reproduction 144(6):649-60.

Darszon A, Nishigaki T, Beltran C, *et al.*, 2011. Calcium channels in the development, maturation, and function of spermatozoa. Physiol Rev 91(4): 1305-55.

Edgar RC, 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics 5(1):113; doi: 10.1186/1471-2105-5-113.

Elango K, Kumaresan A, Sharma A, *et al.*, 2020. Sub-fertility in crossbred bulls: Deciphering testicular level transcriptomic alterations between zebu *(Bos indicus)* and crossbred *(Bos taurus x Bos indicus)* bulls. BMC Genomics 21(1):502; doi: 10.1186/s12864-020-06907-1.

Hancock JM, Zvelebil MJ, Stoesser G, *et al.*, 2014. Dictionary of Bioinformatics and Computational Biology. Eds JM Hancock and Zvelebil MJ; National Center for Biotechnology Information (NCBI): doi.org/10.1002/9780471650126.dob0477.pub2.

Huang CC, Meng EC, Morris JH, *et al.*, 2014. Enhancing UCSF Chimera through web services. Nucleic Acids Res 42(W1):W478-W484.

Islam R, Rahaman M, Hoque H, *et al.*, 2021. Computational and structural based approach to identify malignant nonsynonymous single nucleotide polymorphisms associated with CDK4 gene. Plos One 16(11):0259691.

Jeffares DC, Tomiczek B, Sojo V, *et al.*, 2014. A beginners guide to estimating the non-synonymous to synonymous rate ratio of all protein-coding genes in a genome. In "Parasite Genomics Protocols". Springer, New York NY, USA, pp: 65-90.

Katoh K and Standley DM, 2013. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. Mol Biol Evol 30(4):772-80.

Kelley LA, Mezulis S, Yates CM, *et al.*, 2015. The Phyre2 web portal for protein modeling, prediction and analysis. Nat Protoc 10(6):845-58.

Kersey P, Allen JJE, Armean I, *et al.*, 2016. Ensembl Genomes 2016: more genomes, more complexity. Nucleic Acids Res 44(D1):D574-D580.

Kumar S, Stecher G, Li M, Knyaz C, *et al.*, 2018. MEGA X: Molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol 35(6):1547-49.

Liu L, Wei J and Ruan J, 2017. Pathway enrichment analysis with networks. Genes 8(10):246.

Lobley A, Pierron V, Reynolds L, *et al.*, 2003. Identification of human and mouse CatSper3 and CatSper4 genes: Characterisation of a common interaction domain and evidence for expression in testis. Reprod Biol Endocrinol 1:53: doi: 10.1186/1477-7827-1-53.

Mina CG and Morel D, 2020. Equine Reproductive Physiology, Breeding and Stud Management. 5th Ed, CABI books.

Montoto LG, Magana C and Tourmente M, 2011. Sperm competition, sperm numbers and sperm quality in muroid rodents. PLoS One 6(3):18173.

Murrell B, Wertheim JO, Moola S, *et al.*, 2012. Detecting individual sites subject to episodic diversifying selection. PLoS Genet 8(7):e1002764.

Murrell B, Moola S, Mabona A, *et al.*, 2013. FUBAR: a fast, unconstrained bayesian approximation for inferring selection. Mol Biol Evol 30(5):1196-1205.

Nguyen LT, Schmidt HA, Von Haeseler A, *et al.*, 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol 32(1):268-74.

Park YJ, Pang WK and Pang MG, 2023. Integration of omics studies indicates that species-dependent molecular mechanisms govern male fertility. J Anim Sci Biotechnol 14(1):28: doi.org/10.1186/s40104-023-00836-1.

Qi H, Moran MM, Navarro B, *et al.*, 2007. All four CatSper ion channel proteins are required for male fertility and sperm cell hyperactivated motility. Proc Nat Acad Sci USA 104(4):1219-23.

Ren D and Xia J, 2010. Calcium signaling through CatSper channels in mammalian fertilization. Physiology (Bethesda) 25(3):165-75.

Rivera AM, 2022. Investigating the duplication and evolution of essential fertilization proteins. PhD Thesis, University of Washington, USA.

Stern A, Doron-Faigenboim A, Erez E, *et al.*, 2007. Selecton 2007: advanced models for detecting positive and purifying selection using a Bayesian inference approach. Nucleic Acids Res 35:W506-W511.

Tamura K, Stecher G, Peterson D, *et al.*, 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol Biol Evol 30(12):2725-29.

Wang J, Chen F, Xiao X, *et al.,* 2025. A comprehensive protocol for Bayesian phylogenetic analysis using MrBayes: From sequence

alignment to model selection and phylogenetic inference. Bio-protocol 15(8):5276.

Waterhouse A, Bertoni M, Bienert S, *et al*., 2018. SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res 46(W1):W296-W303.

Wiederstein M and Sippl MJ, 2007. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Res 35:W407-W410.

Wong TK, Ly-Trong N, Ren H, *et al.,* 2025. IQ-TREE 3: Phylogenomic Inference Software using Complex Evolutionary Models. DOI:10.32942/X2P62N.

Yang Y, Yang L, Han X, *et al.,* 2024. The regulation role of calcium channels in mammalian sperm function: a narrative review with a focus on humans and mice. Peer J 12:e18429: doi: 10.7717/peerj.18429.

Yang Z, 2007. PAML 4: phylogenetic analysis by maximum likelihood. Mol Biol Evol 24(8):1586-91.

Yuan S, Zhang Y, Dong PY, *et al.,* 2024. A comprehensive review on potential role of selenium, selenoproteins and selenium nanoparticles in male fertility. Heliyon 10(15):e34975: doi: 10.1016/j.heliyon.2024.e34975.

Zhang Y, 2008. I-TASSER server for protein 3D structure prediction. BMC Bioinformatics 9(1):40.