Zhao et al. *Zool. Res.* 2026, 47(2): 519−529
https://doi.org/10.24272/j.issn.2095-8137.2025.315
https://cstr.cn/32248.14.zr.2025.315

ZR | Zoological Research

Article Open Access

Regulatory diversification of conserved *Runx* genes drives morphological innovation in soft-shell turtles

En-Hao Zhao[1,2,#], Fang Li[2,3,#], Mao-Lei Wei[1,2], Meng-Rao Chen[2,3], Qi-Ran Chen[2,3], Wei Sun[2,3], Chu-Tian Ge[2,3,*], Zong-Ji Wang[2,3,*]

[1] *College of Fisheries and Life Sciences*, *Shanghai Ocean University*, *Shanghai* 201306, *China*

[2] *College of Biological and Environmental Sciences*, *Zhejiang Wanli University*, *Ningbo*, *Zhejiang* 315100, *China*

[3] *Institute of Animal Sex and Development*, *Zhejiang Wanli University*, *Ningbo*, *Zhejiang* 315100, *China*

ABSTRACT

The Runt-related transcription factor (*Runx*) family governs a range of developmental processes through sequence-specific DNA recognition and transcriptional regulation. To investigate the evolutionary dynamics and functional roles of *Runx* genes in turtles, genomic and transcriptomic datasets spanning 29 Testudines species were analyzed using homology prediction, sequence alignment, phylogenetic reconstruction, evolutionary rate estimation, and spatiotemporal expression profiling. Comparative genomic analyses revealed strong *Runx* gene conservation across Testudines, while lineage-specific divergence was evident in Trionychidae (soft-shelled turtles), characterized by expanded intronic regions, unique single-nucleotide variants, and accelerated evolutionary rates. Notably, *Runx2* exhibited clade-specific variation in polyglutamine/polyalanine (QA) repeat length and the Q:A ratio, which significantly correlated with morphological traits, including plastron length and cervical scute patterning. Transcriptomic data revealed stage- and tissue-specific expression of *Runx* genes in embryonic gonads and carapacial ridges (CR), with divergent upstream regulatory inputs and downstream targets implicated in carapace morphogenesis and sex determination pathways. These findings identify QA repeat modulation in *Runx2* as a previously unrecognized molecular mechanism driving morphological innovation while maintaining core function in turtles. This work advances understanding of the conserved and lineage-specific roles of *Runx* genes and clarifies the molecular basis of morphological and reproductive trait evolution in Testudines.

Keywords: *Runx* genes; Turtles; *Runx2* QA repeat; Skeletal development; Sex determination

INTRODUCTION

The Runt-related transcription factor family encodes a set of evolutionarily conserved regulators that orchestrate gene expression through sequence-specific DNA binding, playing essential roles across diverse development contexts (Rennert et al., 2003). Central to all Runx proteins is a highly conserved 128-amino acid Runt domain that, in vertebrates, forms a heterodimer with core-binding factor β (*CBF-β*), thereby enhancing transcriptional activation of downstream effectors (Golling et al., 1996). This domain functions across phyla, mediating critical developmental processes such as neurogenesis and sexual differentiation in *Drosophila*, and governing gonadal development, hematopoiesis, and skeletal patterning in vertebrates (Crute et al., 1996; Westendorf & Hiebert, 1999; Wheeler et al., 2000).

While invertebrate chordates typically possess a single *Runx* ortholog, two rounds of whole-genome duplication in early vertebrate evolution produced three paralogs (*Runx1*, *Runx2*, and *Runx3*) in jawed vertebrates (Akiyama et al., 2005; Dehal & Boore, 2006; Hecht et al., 2008; Levanon & Groner, 2004; Nah et al., 2014). These paralogs have undergone extensive subfunctionalization, exhibiting tissue-specific expression and divergent roles in development (Yano-Sakamoto et al., 2024). *Runx1* is indispensable for hematopoietic stem cell specification during endothelial-to-hematopoietic transition within the aorta-gonad-mesonephros region (Aikawa et al., 2006; Okuda et al., 2001). *Runx2* functions as a master regulator of osteogenic differentiation,

Received: 03 October 2025; Accepted: 04 November 2025; Online: 05 November 2025

Foundation items: This work was supported by the Ningbo Natural Science Foundation (2023J043), Ningbo Youth Science and Technology Innovation Leading Talent Program (2025QL076), National Natural Science Foundation of China (32102772, 32303000, 32403022, 32325049, U22A20529), Zhejiang Provincial Natural Science Foundation (Q24C190009), Zhejiang Provincial Top Discipline of Biological Engineering (Level A), Zhejiang Wanli University (ZS2022008, ZS2024001), and Zhejiang Province Natural Science Foundation (LZ25C190001)

[#]Authors contributed equally to this work

*Corresponding authors, E-mail: cge@zwu.edu.cn; zongji.wang@zwu.edu.cn

directing skeletal formation, craniofacial patterning, and chondrocyte lineage progression (Adhami et al., 2015; Otto et al., 1997; Stricker et al., 2002). In zebrafish (*Danio rerio*), genome duplication led to retention of two *Runx2* paralogs, *Runx2a* and *Runx2b* (Flores et al., 2004), with loss of *Runx2b* eliminating intermuscular bone formation (Nie et al., 2022). *Runx3* has been implicated in immune regulation, neural development (Inoue et al., 2008; Li et al., 2002), and chondrocyte maturation during osteogenesis (Stricker et al., 2002; Yoshida et al., 2004). Beyond skeletal and hematopoietic systems, *Runx* genes also participate in gonadal differentiation. Loss of *Runx1* or *Runx2* in murine models results in granulosa cell disruption and predisposition to ovarian tumorigenesis (Bridges et al., 2022). Additionally, *Runx1* interacts complementarily and redundantly with forkhead box L2 (*Foxl2*) (Nicol et al., 2019), a key female pathway determinant, to sustain granulosa cell identity (Nicol et al., 2019). In temperature-dependent sex determination systems, such as those of turtles, *Foxl2* activation is modulated by signal transducer and activator of transcription 3 (*STAT3*) phosphorylation (Wu et al., 2024), highlighting a potential regulatory intersection. Moreover, both *Runx1* and *Runx2* are necessary for proper differentiation of the epididymal epithelium (Toriseva et al., 2024), underscoring their broader involvement in reproductive tract development.

Among the three vertebrate *Runx* paralogs, *Runx2* exhibits functional specialization conferred by its polyglutamine (polyQ)/polyalanine (polyA) (QA) repeat domain, located between the conserved N-terminal MSDVS and C-terminal VPRLR motifs. This domain is known to enhance transactivation activity (Newton & Pask, 2020; Thirunavukkarasu et al., 1998) and plays a pivotal role in skeletal and cartilage development, including the modulation of craniofacial length in mammals (Newton & Pask, 2020; Sears et al., 2007). Due to their tandem-repeat nature, QA domains are highly susceptible to replication slippage, leading to substantial interspecies polymorphism. In mammals, QA repeats typically range from 13–31 glutamines (Q) to 19–22 alanines (A), while amphibians possess shorter sequences (6–10Q:2A) (Yano-Sakamoto et al., 2024). Variation in the Q:A ratio directly influences α-helical stability, thereby modulating transactivation potency (Rose & Meier, 2004).

At the molecular level, *Runx* genes function as nodal regulators that integrate upstream signaling and drive downstream transcriptional networks in a tissue-specific manner (Otto et al., 2003). These transcription factors operate as key effectors of the transforming growth factor-β (*TGF-β*) superfamily (Ito & Miyazono, 2003) and are differentially activated through tissue-specific cues (Lee et al., 2000). During osteogenesis, bone morphogenetic proteins (*Bmp4* and *Bmp7*) activate *Runx2* via SMAD (mothers against decapentaplegic homolog)-mediated pathways (Phimphilai et al., 2006), with *Bmp7* demonstrating conserved osteogenic functions across vertebrates (Dong et al., 2022). In turn, *Runx2* controls transcription of major skeletal genes, including bone gamma-carboxyglutamate protein (*BGLAP*, osteocalcin) in osteoblasts, and secreted phosphoprotein 1 (*SPP1*, osteopontin) and collagen type I alpha 1 chain (*COL1A1*, collagen I) in chondrocytes (Komori, 2009). In gonads, *Runx1* is activated by Wnt family member 4 (*Wnt4*) (Naillat et al., 2015) and regulates steroidogenic genes such as cytochrome P450 family 11 subfamily A member 1 (*Cyp11a1*), cytochrome P450 family 19 subfamily A member 1 (*Cyp19a1*), and follicle stimulating hormone receptor (*FSHR*) (Ojima et al., 2019; Zhong et al., 2020), establishing *Runx1* as a central regulator of granulosa cell differentiation, folliculogenesis, and steroidogenesis.

Despite extensive characterization in mammals and classical model systems, the evolutionary dynamics, structural diversification, and functional specialization of *Runx* genes remain largely unresolved in reptiles, particularly within Testudines. Extant turtles are divided into two major suborders: Cryptodira and Pleurodira, with Cryptodira encompassing several derived clades including Trionychia and Durocryptodira (Shaffer et al., 2017). Turtles possess a unique body plan defined by the presence of a bony shell comprising the plastron and carapace (Zangerl, 1969) and, like other reptiles and select teleosts, exhibit temperature-dependent sex determination (Hui et al., 2021; Trukhina et al., 2013). These features raise fundamental questions regarding the evolutionary innovations and regulatory adaptations of *Runx* genes that support the emergence of such derived morphological and reproductive traits.

To address this, the present study conducted a comprehensive comparative analysis of *Runx* gene evolution and function across 29 turtle species. Through integrated genomic and transcriptomic approaches, the analysis revealed both conserved and lineage-specific innovations in *Runx* gene architecture, expression dynamics, and regulatory connectivity. Overall, the findings obtained in this study establish *Runx* genes as critical components of turtle developmental programs and provide mechanistic insights into the molecular architecture underlying carapace formation and sex determination in Testudines.

MATERIALS AND METHODS

Sequence retrieval and *Runx* gene homology prediction

The longest isoforms of Runx protein sequences from chicken (*Gallus gallus*), mouse (*Mus musculus*), African clawed frog (*Xenopus laevis*), zebrafish (*D. rerio*), and American alligator (*Alligator mississippiensis*) were retrieved from the NCBI Protein Database (https://www.ncbi.nlm.nih.gov/) and served as reference sequences (Supplementary Table S1). These references were used to predict homologous *Runx* genes in the genomes of 29 Testudines species (Supplementary Table S1).

Homology searches were performed using BLAST (v2.2.26) (Altschul et al., 1990) with an *E*-value threshold of $<1\times10^{-3}$. BLAST hits with alignment coverage or sequence similarity below 30% were discarded. Fragmented BLAST hits were merged using Solar (v0.9.6) (https://sourceforge.net/p/treesoft/code/HEAD/tree/branches/lh3/solar/) (Li et al., 2006; Yu et al., 2006), incorporating flanking extensions of 400 kb. Gene structures were annotated using GENEWISE (v2.4.1) (Birney et al., 2004), and the predicted sequences were validated through BLASTP searches against the NCBI non-redundant (nr) database.

Phylogenetic and evolutionary analyses

Coding sequences of *Runx* genes from 29 turtle species were aligned using MAFFT (v.7.525) (Katoh et al., 2002). Phylogenetic relationships were reconstructed under a maximum-likelihood framework using IQ-TREE (v.2.2.2.7) (Minh et al., 2020), with branch support evaluated via 1 000 bootstrap replicates and SH-aLRT tests. The best-fit

nucleotide substitution model (TPM3u+F+I+R3) was selected using ModelFinder (Kalyaanamoorthy et al., 2017).

Lineage-specific *Runx* motifs were identified using MEME (Bailey & Elkan, 1994) with a third-order Markov background model. Ancestral sequence reconstruction was inferred using the codeml module in PAML (v.4.9) (Yang, 2007), based on a published Testudines phylogeny (Huang et al., 2023). Taxon-specific insertions, deletions, and substitutions were identified by aligning *Runx* sequences using MAFFT (v.7.525) (Katoh et al., 2002) and visualizing aligned regions in Jalview (v.2.11.3.2) (Waterhouse et al., 2009).

QA repeat lengths in *Runx2* were quantified across taxa. Ancestral states of Q:A ratios were inferred under a Brownian motion model using Phytools (Revell, 2024) and visualized with contMap (Revell et al., 2018). Phylogenetic generalized least squares (PGLS) models were implemented using the phylolm package (Tung Ho & Ané, 2014) to test associations between Q:A ratios and phenotypes obtained from the CheloniansTraits Database (Wang et al., 2025). Pearson correlation analyses were applied to phylogenetically corrected residuals to validate trait associations.

Three-dimensional (3D) structures of *Runx* proteins were predicted using AlphaFold3 (Abramson et al., 2024), with domain annotation performed via InterPro (Paysan-Lafosse et al., 2023). Mutation mapping and structural visualization were performed in PyMOL (v2.6) (Delano, 2002). Protein complexes between *Runx2* and *CBF-β* or *SMAD1* from the Chinese soft-shelled turtle (*Pelodiscus sinensis*) and yellow pond turtle (*Mauremys mutica*) were modeled using AlphaFold3. Binding interfaces were identified in PyMOL, and complex stability was evaluated by generating *Runx2* variants with different Q:A ratios (1, 3, 5, 8, 12). Interaction energies between each variant and its partner (*CBF-β* or *SMAD1*) were calculated using the AnalyzeComplex command in FoldX (v.5.1) (Delgado et al., 2019; Guerois et al., 2002; Schymkowitz et al., 2005).

Evolutionary rate estimation for *Runx* genes was performed using the codeml branch model in PAML (v.4.9) (Yang, 2007), with a published Testudines phylogeny as the phylogenetic framework (Thomson et al., 2021). Both one-ratio and two-ratio models were applied to detect branches with significantly accelerated evolution. Likelihood ratio tests (LRTs) were conducted to compare the models. The free-ratio model was employed to estimate evolutionary rates for each branch independently.

Expression analysis

Transcriptomic datasets were retrieved from the NCBI Sequence Read Archive, including carapacial ridge tissues from the pond turtle (*Mauremys reevesii*, Geoemydidae, PRJNA713512) (Yang et al., 2022) and *P. sinensis* (Trionychidae, PRJNA742492) (Zhang et al., 2021), as well as gonadal tissues of the red-eared slider (*Trachemys scripta,* Emydidae, PRJNA1187197) spanning developmental stages 14 to 25 (Fan et al., 2025). These species collectively represent major Testudines clades. Carapacial ridge samples were selected due to their key role in shell morphogenesis (Kuraku et al., 2005), while gonadal samples were used for their relevance to sex determination, particularly in *T. scripta*, which exhibits temperature-dependent sex determination and co-expression of *Runx1* and *Foxl2* (Nicol et al., 2019).

Raw RNA sequencing (RNA-seq) reads were quality-checked and filtered using fastp (v.0.23.2) (Chen et al., 2018). Expression quantification was conducted using the new Tuxedo pipeline (Pertea et al., 2016): reads were aligned to the corresponding reference genomes using HISAT2 (v.2.2.1) (Kim et al., 2015), and transcript quantification was performed using StringTie (v.2.2.1) (Pertea et al., 2015). Gene-level expression values were calculated using tximport (v.1.28.0) (Soneson et al., 2016). To account for variations across samples, Fragments Per Kilobase of transcript per Million mapped reads (FPKM) were normalized using a scaling method based on the trimmed mean of M values (TMM), under the assumption that most genes maintain consistent expression across samples.

Gonadal tissues from *T. scripta* embryos at defined developmental stages were dissected, and total RNA was isolated using TRIzol® reagent (Invitrogen, USA). First-strand cDNA synthesis was performed using the RevertAid First Strand cDNA Synthesis Kit (K1622, Thermo Scientific, USA). Gene-specific primers targeting the predicted *T. scripta Runx1* sequence were designed for quantitative real-time polymerase chain reaction. Reactions were conducted with SYBR® Premix Ex Taq™ II (TaKaRa, Japan) using synthesized cDNA as the template. Glyceraldehyde-3-phosphate dehydrogenase (*Gapdh)* served as the internal reference gene, and relative expression levels of *Runx1* were calculated using the $2^{-\Delta\Delta Ct}$ method. Primer sequences are provided in Supplementary Table S2.

RESULTS

Runx gene conservation and intronic expansion in Testudines

To investigate the evolutionary conservation and structural diversification of *Runx* genes in turtles, genome assemblies from 29 Testudines species representing 10 of 14 recognized families were analyzed alongside representative Runx protein sequences from five major vertebrate clades (see Materials and Methods). Homology-based predictions identified all three canonical *Runx* paralogs (*Runx1*, *Runx2*, and *Runx3*) in every turtle species examined (Figure 1A), indicating a high degree of retention and conservation of this gene family across the lineage. Gene duplication events were infrequent, observed only in the eastern box turtle (*Terrapene carolina*) and common snapping turtle (*Chelydra serpentina*) (each harboring two *Runx1* copies) and in the pig-nosed turtle (*Carettochelys insculpta*) (harboring two *Runx3* copies), suggesting these are likely species-specific duplications.

Phylogenetic reconstruction based on coding sequences revealed three well-supported, paralog-specific clades (*Runx1*, *Runx2*, and *Runx3*) (Supplementary Figure S1), each of which further subdivided into subclades corresponding to the major Testudines radiations—Durocryptodira, Trionychia, and Pleurodira. This topological congruence between *Runx* gene trees and species phylogeny reinforces the evolutionary stability of the *Runx* repertoire in turtles. Moreover, the species-specific duplications observed in *T. carolina*, *C. serpentina*, and *C. insculpta* were each phylogenetically nested within their respective clades, forming discrete sister branches indicative of recent, post-speciation duplication events.

Although exon structure and conserved *Runx* motifs were largely uniform across Testudines (Figure 1B; Supplementary Figures S2, S3), substantial variation was detected in intronic regions. In particular, intron 1 (false discovery rate (FDR)=3.66×

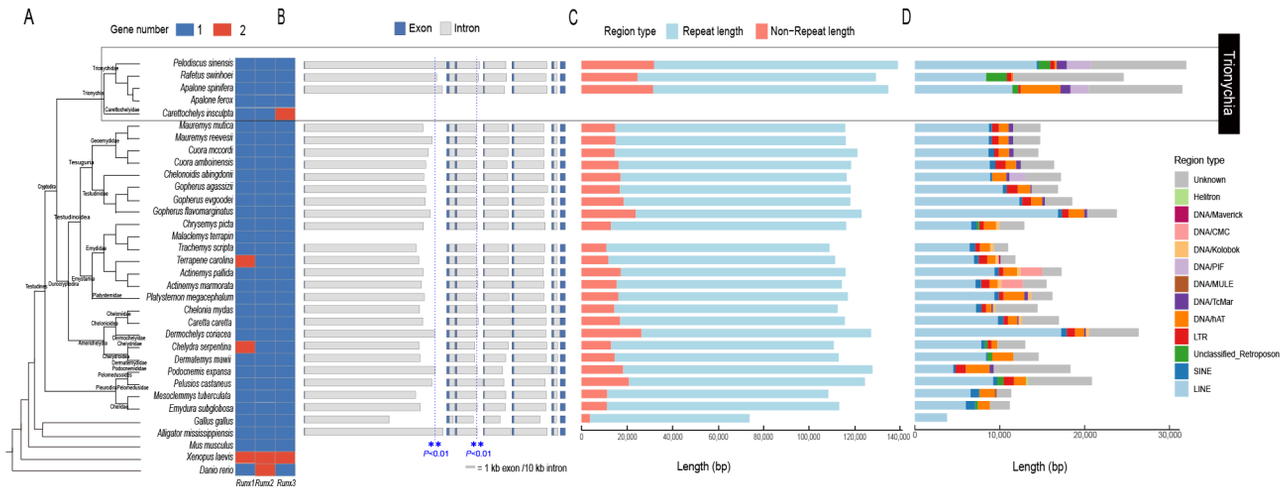

Figure 1 ***Runx*** **gene prediction and intronic divergence in Testudines**

A: Left panel: Phylogenetic relationships and Latin names of 29 Testudines species. Right panel: Heatmap showing copy numbers of three *Runx* genes (*Runx1*, *Runx2*, and *Runx3*) across species. B: Exon (blue) and intron (gray) lengths of *Runx1* in Testudines. Intron lengths are scaled to one-tenth of actual size. Wilcoxon tests revealed significantly longer first and third introns in Trionychidae compared to other turtles (**: P<0.01). C: Proportional composition of repetitive and non-repetitive sequences in the first intron of *Runx1* across species. D: Type and length distribution of repetitive elements within the first intron of *Runx1*. Three species (*Apalone spinifera*, *Apalone ferox*, and *Malaclemys terrapin*) lacked complete intron predictions and are excluded in panels B–D.

10^{-3}, Wilcoxon test) and intron 3 (FDR=7.33×10^{-3}, Wilcoxon test) of *Runx* were significantly elongated in Trionychidae relative to other families (Figure 1B). Comparable patterns of intronic expansion were detected for *Runx2* and *Runx3* (Supplementary Figure S2), indicating lineage-specific expansion of non-coding regions within the *Runx* gene family.

These intronic expansions were closely associated with the accumulation of repetitive elements, particularly within *Runx1* intron 1. Relative to non-Testudines outgroups, turtles exhibited higher densities of repetitive sequences in this region, with Trionychidae showing the most pronounced enrichment (Figure 1C). Repeat classification revealed that this enrichment was driven primarily by two categories—Unclassified Retroposons and Unknown Repeats, both of which were significantly overrepresented in Trionychidae compared to other families (Figure 1D). This atypical accumulation of repeats suggests that evolutionary constraints on intronic architecture were relaxed in this lineage, allowing repetitive elements to proliferate. Such expansions may alter chromatin dynamics, RNA splicing, or epigenetic regulation, thereby influencing *Runx1* expression and function in lineage-specific developmental contexts.

Lineage-specific Runx2 QA divergence shapes phenotypic evolution in turtles

To investigate the evolutionary dynamics of the *Runx2* QA repeat domain in turtles, polyQ and polyA tract lengths were quantified across 29 Testudines species. The QA region, located between the conserved MSDVS and VPRLR motifs upstream of the RUNT domain (Supplementary Figure S4), is known to modulate transcriptional activity and has been implicated in vertebrate craniofacial evolution (Newton & Pask, 2020; Sears et al., 2007). Comparative analyses revealed pronounced lineage-specific differences in both QA repeat length and the Q:A ratio, consistent with clade-specific constraints and potentially reflecting morphological divergence.

Among Trionychia (softshell) turtles, the QA domain consistently exhibited longer polyQ tracts, predominantly spanning 8–9 residues (Figure 2A; Supplementary Figure S4), with corresponding Q:A ratios exceeding 1.8 in all species examined (Figure 2B)—significantly higher than those observed in other lineages. In contrast, non-trionychian cryptodires showed intermediate repeat lengths (6–8 residues) and moderate Q:A ratios (Supplementary Figure S4). Within Emysternia, QA repeats were remarkably conserved: all seven species possessed identical polyQ lengths of 7 and a stable Q:A ratio of 1.75 (Supplementary Figure S4), consistent with strong purifying selection. In contrast, basal pleurodires (side-necked turtles) exhibited the shortest polyQ tracts, ranging from 5–6 residues, and the lowest Q:A ratios among all taxa examined (Figure 2A, B; Supplementary Figure S4).

Notably, the Q:A ratio exhibited greater evolutionary variability than either polyQ or polyA length alone. Ancestral state reconstruction using *G. gallus* as an outgroup estimated a Q:A ratio of 2.38 for the Testudines ancestor (Supplementary Figure S5), exceeding that of all extant lineages, thus suggesting gradual contraction of the QA domain during turtle evolution accompanied by lineage-specific diversification.

To assess functional relevance, phylogenetic linear modeling (phylolm) was used to test associations between the Q:A ratio and turtle phenotypes. Significant phylogenetic correlations (P<0.05, Wald test) were detected for three traits: maximum cervical scute count, plastron length, and juvenile survival rate (Figure 2C). These associations were also supported by Pearson correlation analysis of trait residuals derived from univariate phylogenetic models (P<0.05, t-test) (Figure 2D, E), confirming independence from phylogenetic signal.

Among these traits, cervical scute count exhibited the strongest association with the Q:A ratio. Trionychids, which lack this anterior carapacial element, are known to exhibit premature ossification of the carapacial ridge, disrupting marginal bone plate formation. These findings suggest that *Runx* genes, particularly *Runx2*, modulate carapacial ridge differentiation through this developmental mechanism. Associations with plastron length (a key skeletal component)

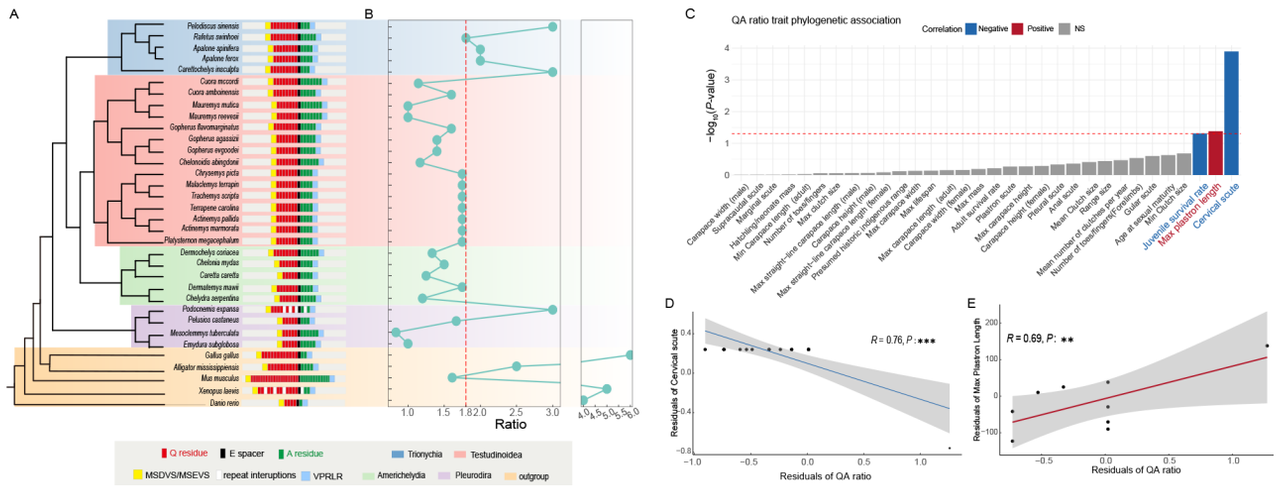

Figure 2 Lineage-specific variation in *Runx2* QA tandem repeat domain

A: QA repeat structures and phylogenetic relationships among Testudines. Color shading denotes major taxonomic groups. QA domain components: yellow squares, MSDVS/MSEVS initiation motif; red, glutamine residues (Q); black, glutamic acid spacers (E); green, alanine residues (A); blue, VPRLR terminus. B: Q:A ratios for each species. Red dashed line denotes lowest ratio (1.8) observed in *Rafetus swinhoei* (Trionychidae). C: Phylogenetic linear modeling of associations between Q:A ratio and phenotypic traits. Traits are plotted on the x-axis; association strength ($-\log_{10}$(*P*-value) on the y-axis. Bar colors indicate significance: gray (non-significant), blue (significant negative correlation), and red (significant positive correlation). D, E: Pearson correlations between phylogenetically corrected Q:A ratio residuals, juvenile survival rate (D), and maximum plastron length (E). *R* values and *P* values are provided.

and juvenile survival rate (a physiological trait) further indicate pleiotropic effects of QA repeat divergence. Together, these results suggest that lineage-specific divergence in *Runx2* QA repeats may have broad functional consequences, contributing to both morphological innovations (e.g., scute patterning, shell morphology) and physiological adaptation (offspring viability) in turtles.

Adaptive evolution and lineage-specific mutations in *Runx* genes across Testudines

To evaluate patterns of molecular evolution in turtle *Runx* genes, lineage-specific substitution rates were estimated using *G. gallus* as an outgroup (Figure 3). This analysis revealed significantly accelerated substitution rates within the Trionychidae lineage relative to other turtle families ($P<1\times10^{-5}$, *t*-test; Supplementary Figure S6). In contrast, *Runx2* exhibited markedly reduced rates within Emysternia (Figure 3B), consistent with strong evolutionary constraint on this paralog. Despite these lineage-specific rate differences, branch model comparisons detected no significant heterogeneity in selection pressure among lineages ($P>0.05$, LRT; Supplementary Table S3), indicating broad evolutionary conservation of *Runx* genes in Testudines.

To investigate the consequences of accelerated evolution in Trionychidae, coding sequence divergence within this lineage was examined in detail. A total of 27 Trionychidae-specific single-nucleotide substitutions were identified: 12 in *Runx1*, 10 in *Runx2*, and five in *Runx3* (Figure 4A, D; Supplementary Figure S7). Most substitutions were synonymous and confined to the third codon position, consistent with neutral drift. Only two nonsynonymous mutations were detected: a Ser-to-Pro substitution in *Runx1* (TCT→CCT at position 934) and a Thr-to-Ala change in *Runx3* (ACT→GCT at position 712) (Figure 4B, E). Domain prediction and 3D structural modeling confirmed that both amino acid substitutions occurred outside critical functional regions, including the conserved RUNT domain, and did not alter predicted tertiary structure (Figure 4C, F).

Collectively, these results suggest that while *Runx* genes in Trionychidae have accumulated lineage-specific mutations under relaxed sequence constraint, core regulatory functions appear preserved through purifying selection.

Stage- and tissue-specific expression dynamics of *Runx* genes in Testudines

To resolve the spatial and temporal regulation of *Runx* genes during embryogenesis, transcriptomic data were analyzed from developing gonads and carapacial ridges in representative Testudines species. The carapacial ridge, a transient embryonic structure central to shell formation, offers a developmental window to compare gene expression programs between hard-shelled turtles bearing cervical scutes and soft-shelled turtles lacking carapace margins. Known upstream regulators and downstream targets were also integrated to reconstruct preliminary *Runx* regulatory networks.

In the carapacial ridge of *P. sinensis* and *M. reevesii*, all three *Runx* paralogs (*Runx1*, *Runx2*, and *Runx3*) were expressed during stages 15–16, corresponding to mid-to-late phases of ridge development. Expression levels increased progressively, with *M. reevesii* exhibiting earlier activation, as low *Runx* expression levels were already detectable at stage 15 (Figure 5A, B). Expression peaks of upstream components of the *TGF-β* (*TGFβ1/2/3*, *TGFβR1/2/3*) and *BMP* (*BMP4/7*) signaling pathways preceded *Runx* activation, implicating these pathways as candidate upstream inducers of *Runx* transcription. In contrast, downstream targets, including *BGLAP* and *Col1a1*, were transcriptionally up-regulated following *Runx* activation, consistent with *Runx*-mediated activation of genes essential for dermal bone formation and collagen matrix deposition during carapacial ridge development.

In embryonic gonads, which transition through bipotential, sex determination, and sexual differentiation phases, *Runx* expression peaked predominantly during the sex determination phase, with paralog-specific and species-

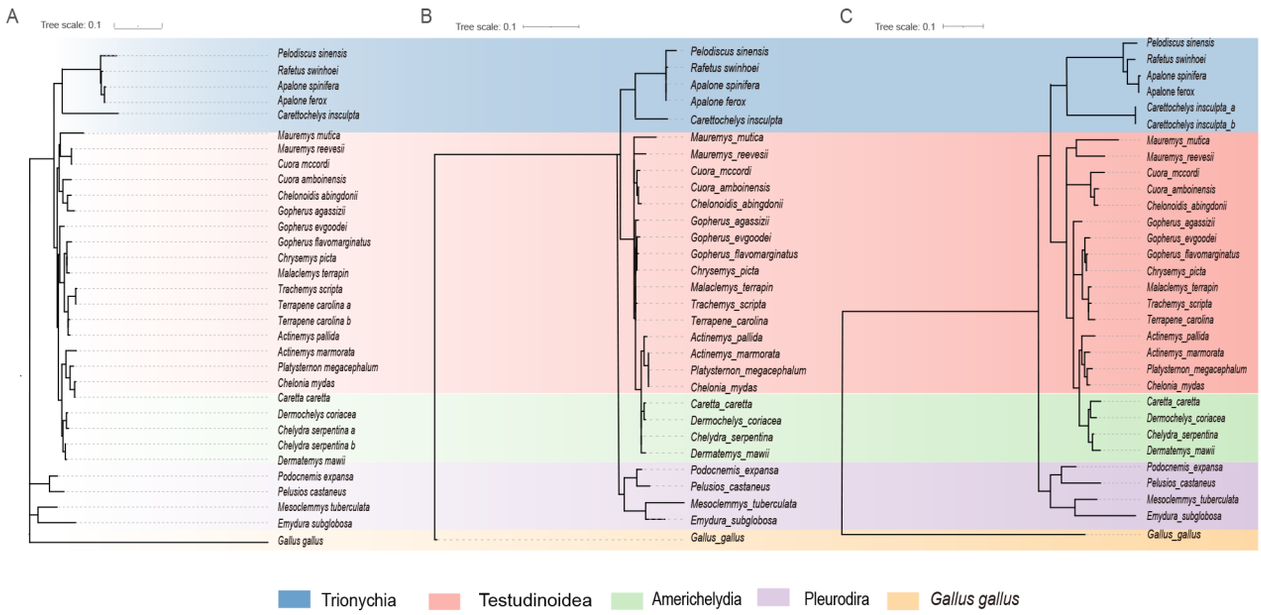

Figure 3 Accelerated evolution of *Runx* genes in Testudines

A–C: Synonymous substitution rates (*dS*) for *Runx1* (A), *Runx2* (B), and *Runx3* (C) under the free-ratio model across Testudines lineages.

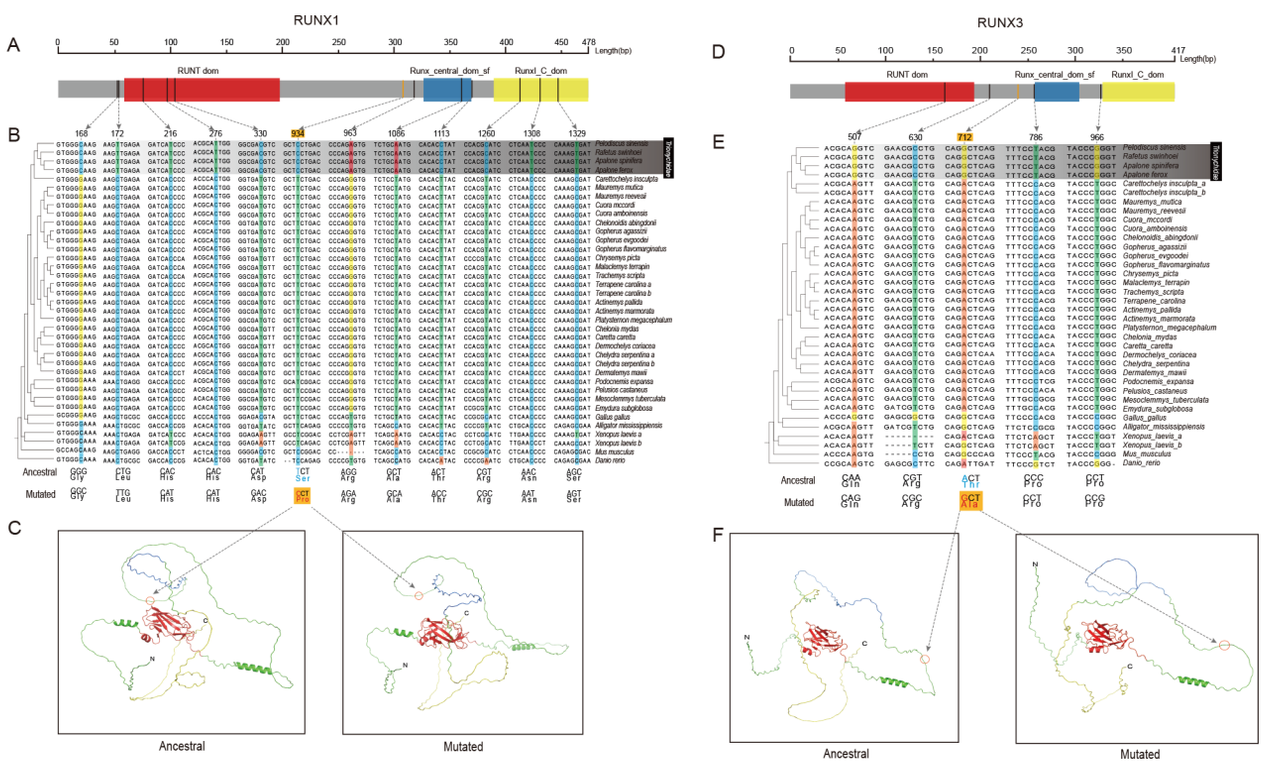

Figure 4 Sequence divergence and structural modeling of *Runx1* and *Runx3*

A, D: Schematic of protein length, structural domains, and positions of lineage-specific mutations in *Runx1* and *Runx3*. Colored rectangles represent structural domains; black lines mark mutation sites. B, E: Phylogenetic relationships and aligned *Runx* coding sequences. Species within black boxes belong to Trionychidae. Colored block s denote nucleotide bases at mutation sites; ancestral and derived codons are shown below. Nonsynonymous mutations are highlighted in orange. C, F: Predicted 3D protein structures of *Runx1* and *Runx3* before and after amino acid substitutions, with mutation sites indicated.

specific patterns (Figure 5C, D). In *T. scripta*, *Runx1* exhibited female-biased expression, aligning with its proposed role in ovarian development. Quantitative PCR confirmed the transcriptomic trends, supporting its involvement in temperature-dependent sex determination. Network reconstruction revealed species-specific regulatory architectures. In *T. scripta*, *TGF-β* family genes peaked during the bipotential phase, followed by *Runx1* up-regulation during sex determination. Notably, *Cyp19a1*, a key ovarian differentiation gene, was up-regulated downstream of *Runx1*, supporting its involvement within the female differentiation pathway. In contrast, other canonical *Runx* targets, such as *FSHR*, lacked consistent regulatory association.

Together, these results demonstrate that *Runx* paralogs are

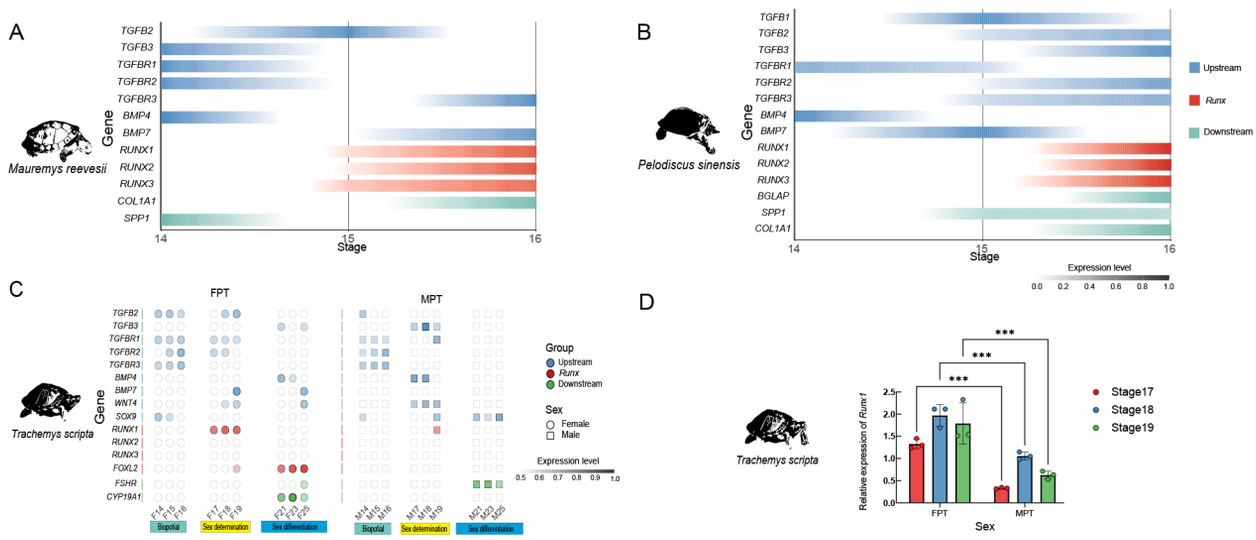

Figure 5 Spatiotemporal expression patterns of *Runx* genes in Testudines

A, B: Expression of *Runx* genes, their upstream regulators, and downstream targets in the carapacial ridge of *M. reevesii* (A) and *P. sinensis* (B). Gene expression (RPKM) was normalized per gene after filtering low-expression genes (below median expression). C: Expression of *Runx* genes and associated regulatory factors in gonads of *T. scripta*. Normalization and filtering were performed as described above. D: Female-biased expression of *Runx1* in *T. scripta* embryonic gonads. Bar plots show relative expression levels (qPCR) at stages 17–19. Expression was significantly higher at the female-producing temperature (FPT) compared with male-producing temperature (MPT). Error bars represent mean±standard deviation (SD); ***: P<0.001.

embedded within temporally and spatially restricted gene networks in Testudines. They respond to distinct upstream signals, particularly the *TGF-β* and *BMP* pathways, and contribute to lineage-specific gene expression programs governing both carapacial morphogenesis and gonadal differentiation.

DISCUSSION

Turtles constitute a deeply divergent tetrapod lineage, distinguished by a unique body plan that integrates an enclosing shell with highly derived craniofacial architecture. The protective shell, composed of the carapace and plastron, exhibits pronounced developmental divergence among lineages. In hard-shelled turtles, lateral expansion of the axial ribs contributes to formation of the carapace margin, whereas in soft-shelled turtles, developmental constraints within the carapacial ridge are associated with premature ossification and subsequent loss of marginal bony plates (Rice et al., 2015). Such structural disparities point to lineage-specific genetic mechanisms shaping skeletal evolution, a view supported by recent evidence for clade-specific amino acid compositions in components of the simple epidermal differentiation complex (*SEDC*) and S100-fused type proteins (*SFTPs*) across hard-shelled, soft-shelled, and leatherback turtles (Chen et al., 2025). Despite these advances, the molecular determinants driving this phenotypic diversification remain largely unresolved.

Comparative genomic analyses revealed that *Runx* genes display exceptional conservation across Testudines, with no lineage-specific variation in copy number, motif composition, exon organization, or protein length detected over an evolutionary timescale exceeding 260 million years (Lyson et al., 2010). This degree of stability contrasts sharply with patterns observed in other vertebrates, where lineage-specific duplication events have expanded the *Runx* repertoire. For instance, teleost fish such as zebrafish retain two *Runx2* paralogs (*Runx2a* and *Runx2b*) as a consequence of whole-genome duplication (Flores et al., 2004; Sato et al., 2023), while allotetraploidy in African clawed frog resulted in duplication of the entire *Runx* gene set (Session et al., 2016). The absence of comparable expansions in turtles underscores an unusually strong evolutionary constraint acting on *Runx* gene architecture within this lineage, suggesting that maintenance of a stable *Runx* complement may be critical for turtle developmental or physiological programs. Notably, even in the limited cases where individual turtle species retain duplicated *Runx* loci, these events lack consistent phylogenetic structure and do not define broader clade-level trends. This pattern indicates that although episodic relaxation of selective pressure may permit sporadic duplication or divergence (Figure 3), such changes are rarely retained. Instead, pervasive purifying selection appears to have preserved *Runx* gene structure and dosage across Testudines. Collectively, these observations establish *Runx* genes as among the most evolutionarily constrained transcription factor families in turtles and highlight their exceptional structural and functional stability within amniotes.

Despite pronounced conservation of *Runx* coding architecture, intronic regions emerged as critical regulators of *Runx* expression dynamics. Marked elongation of the first intron in *Runx1* was detected in Trionychidae, a structural feature predicted to delay transcriptional initiation (Heyn et al., 2015). These expanded introns harbor lineage-specific accumulations of retrotransposons and complex repeat elements, which are known to influence local regulatory landscapes (Feschotte & Pritham, 2007). Notably, this architectural divergence correlates with delayed *Runx1* activation during carapacial ridge development in *P. sinensis* compared to *M. reevesii* (Figure 5A, B), suggesting that intronic repeat expansion contributes to regulatory divergence. Although this association remains speculative, targeted functional validation—including chromatin conformation mapping (Hi-C or 4C), epigenomic profiling, and spatiotemporal expression confirmation via *in situ*

hybridization or immunohistochemistry—will be required to establish whether retrotransposon insertions causally affect *Runx1* transcriptional timing. The current analyses provide a comparative genomic framework that highlights candidate regulatory mechanisms and provides a foundation for future experimental resolution.

Coding sequence analyses further identified Trionychidae-specific substitutions in *Runx* genes, the majority of which were synonymous and localized to third codon positions (Figure 4), consistent with pervasive purifying selection on protein sequence. The few nonsynonymous variants occurred outside functional domains, including the RUNT domain, underscoring strong selective constraint on *Runx* structural integrity. Although largely constrained, such substitutions may still contribute to functional divergence. Synonymous chances can influence mRNA splicing efficiency or translational control (Pagani et al., 2005), while missense substitutions outside core domains may alter protein stability or proteasomal degradation (Hovland et al., 2023). Together, these findings suggest that even subtle sequence variation, acting primarily through regulatory rather than structural mechanisms, may fine-tune *Runx* activity and contribute to lineage-specific regulatory divergence without compromising core domain integrity.

Interestingly, the polyQ tract embedded within the *Runx2* QA repeat domain exhibited stepwise reduction across turtle lineages, decreasing from Trionychidae through Testudinoidea and Chelonioidea to Pleurodira (Figure 2B). This progressive contraction likely contributes to the diversification of skeletal traits among Testudines. Prior research has linked coordinated expression of Msh homeobox 2 (*Msx2*) and *Runx2* to dermal bone reduction in turtles (Sato et al., 2023). Our findings expand this framework by identifying elevated Q:A ratios uniquely in Trionychidae. These ratios are strongly associated with key morphological adaptations such as the absence of cervical scutes and abbreviated plastron length (Figure 2C–E), suggesting that QA domain variation serves as a regulatory modulator of *Runx2*-driven developmental programs. This model posits that QA repeat architecture fine-tunes ossification dynamics—promoting marginal bone formation in hard-shelled turtles while constraining carapacial ossification in Trionychidae.

Structural modeling using AlphaFold3 indicated that variation in QA repeat length indirectly modulated transcriptional activity by altering the stability of *Runx2* interactions with cofactors such as *CBF-β* and *SMAD1*. Moderate QA lengths (Q:A ratio of 3) conferred the highest complex stability, whereas shorter or excessively long repeats reduced stability, with very long repeats potentially prone to aggregation (Pelassa et al., 2014) (Supplementary Figure S8). These findings suggest that variation in QA architecture may fine-tune *Runx2* transcriptional output without affecting the conserved RUNT domain, providing a mechanistic basis for functional divergence among turtle lineages.

Comparative analyses further demonstrated that turtles maintained a short and stable polyQ tract (7–9 Q), in contrast to the expanded and more variable repeats observed in squamates (12–21 Q) and other reptiles (Newton & Pask, 2020; Thirunavukkarasu et al., 1998). Crocodilians also exhibit remarkably conserved repeat length (10 Q, Q:A ratio 2.5), closely approximating the reconstructed ancestral state for Testudines (2.36) (Supplementary Figure S5). This cross-lineage conservation underscores a strong evolutionary constraint on QA repeat length in turtles, likely reflecting the need to balance transcriptional fidelity with lineage-specific plasticity.

Although associations between Q:A ratios and morphological traits remain correlative, functional validation will require reporter assays, such as luciferase reporter experiments, to test the effects of Trionychidae-specific QA variants on downstream targets such as *BGLAP* and *SPP1*. Nonetheless, the present comparative genomic and structural data provide a robust mechanistic framework for understanding lineage-specific adaptations and highlight promising directions for future functional studies.

While no evidence of positive selection was detected, branch-specific rate analyses identified significantly elevated synonymous substitution rates (*dS*) in Trionychidae relative to other Testudines (Figure 3; Supplementary Figure S6). Given that synonymous substitutions can modulate mRNA stability, splicing, and translational kinetics (Kimchi-Sarfaty et al., 2007), elevated *dS* may contribute to *Runx* expression levels, and thus influence shell patterning in this lineage.

The carapacial ridge, an embryonic structure essential for rib-derived carapace formation (Nagashima et al., 2007), displayed marked interspecific differences between Trionychidae and other turtles (Session et al., 2016). Transcriptomic profiling confirmed stage 14–16 up-regulation of all three *Runx* paralogs in both *P. sinensis* and *M. reevesii* (Figure 5A,B), consistent with developmental roles during carapacial ridge morphogenesis. However, comparative regulatory analyses revealed divergence in downstream targets: *SPP1* expression was down-regulated in *M. reevesii* but up-regulated in *P. sinensis*, implicating lineage-specific *Runx-SPP1* dynamics in shaping carapacial ridge architecture.

Runx genes also contributed to gonadal development under temperature-dependent sex determination. In *T. scripta*, *Runx1* exhibited female-biased expression during the thermosensitive period (Figure 5C) (Chen et al., 2024), consistent with sex-specific regulatory roles. While *Foxl2* maintains conserved female-biased expression across species (Georges et al., 2014), *Runx* genes demonstrated regulatory plasticity. Furthermore, while *TGF-β/BMP* pathways consistently regulate *Runx* genes (Ito & Miyazono, 2003; Phimphilai et al., 2006), the *Wnt4* pathway showed species-specific expression, absent during the thermosensitive period in *T. scripta*, highlighting lineage-restricted roles in sex determination.

CONCLUSION

Integrative genomic and transcriptomic analyses demonstrated that, despite structural conservation of *Runx* genes across turtles, regulatory sequence divergence—including QA domain variation, intronic expansion, and elevated synonymous substitution rates—has contributed to morphological and reproductive diversification within Testudines. These findings delineate a molecular framework through which conserved developmental regulators such as *Runx* genes are co-opted and modified to mediate lineage-specific evolutionary innovations, exemplified by distinctive shell morphologies and temperature-dependent sex determination in turtles.

SUPPLEMENTARY DATA

Supplementary data to this article can be found online.

COMPETING INTERESTS

The authors declare that they have no competing interests.

AUTHORS' CONTRIBUTIONS

Z.J.W, conceived and designed the study. E.H.Z., F.L., M.L.W., M.R.C., Q.R.C., and W.S. collected the data and conducted bioinformatics analyses. E.H.Z., Z.J.W., F.L., and C.T.G. jointly drafted, revised, and edited the manuscript. All authors read and approved the final version of the manuscript.

ACKNOWLEDGMENTS

Assistance with experimental work was provided by Xin-Qi Han, Rui Li, and Jia-Le Mo (Institute of Animal Sex and Development, Zhejiang Wanli University).

REFERENCES

Abramson J, Adler J, Dunger J, et al. 2024. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature*, **630**(8016): 493–500.

Adhami MD, Rashid H, Chen HY, et al. 2015. Loss of Runx2 in committed osteoblasts impairs postnatal skeletogenesis. *Journal of Bone and Mineral Research*, **30**(1): 71–82.

Aikawa Y, Nguyen LA, Isono K, et al. 2006. Roles of HIPK1 and HIPK2 in AML1‐ and p300‐dependent transcription, hematopoiesis and blood vessel formation. *The European Molecular Biology Organization Journal*, **25**(17): 3955–3965.

Akiyama H, Kim JE, Nakashima K, et al. 2005. Osteo-chondroprogenitor cells are derived from Sox9 expressing precursors. *Proceedings of the National Academy of Sciences of the United States of America*, **102**(41): 14665–14670.

Altschul SF, Gish W, Miller W, et al. 1990. Basic local alignment search tool. *Journal of Molecular Biology*, **215**(3): 403–410.

Bailey TL, Elkan C. 1994. Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proceedings*. *International Conference on Intelligent Systems for Molecular Biology*, **2**: 28–36.

Birney E, Clamp M, Durbin R. 2004. GeneWise and genomewise. *Genome Research*, **14**(5): 988–995.

Bridges K, Yao HHC, Nicol B. 2022. Loss of *Runx1* induces granulosa cell defects and development of ovarian tumors in the mouse. *International Journal of Molecular Sciences*, **23**(22): 14442.

Chen MR, Li F, Chen QR, et al. 2025. Comparative genomic identification and analyses of epider-mal differentiation complex genes in turtles. *Acta Hydrobiologica Sinica*, **49**(5): 052510. (in Chinese)

Chen QR, Sun W, Jin L, et al. 2024. Overexpression of Kdm6b induces testicular differentiation in a temperature-dependent sex determination system. *Zoological Research*, **45**(5): 1108–1115.

Chen SF, Zhou YQ, Chen Y, et al. 2018. Fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*, **34**(17): i884–i890.

Crute BE, Lewis AF, Wu ZN, et al. 1996. Biochemical and biophysical properties of the core-binding factor α2 (*AML1*) DNA-binding domain. *Journal of Biological Chemistry*, **271**(42): 26251–26260.

Dehal P, Boore JL. 2006. Two rounds of whole genome duplication in the ancestral vertebrate. *Public Library of Science Biology*, **3**(10): e314.

Delano WL. 2002. PyMOL: an open-source molecular graphics tool. *CCP4 Newsl*. *Protein Crystallogr*, **40**(1): 82–92.

Delgado J, Radusky LG, Cianferoni D, et al. 2019. FoldX 5.0: working with RNA, small molecules and a new graphical interface. *Bioinformatics*, **35**(20): 4168–4169.

Dong XR, Wan SM, Zhou JJ, et al. 2022. Functional differentiation of *BMP7* genes in zebrafish: *bmp7a* for dorsal-ventral pattern and *bmp7b* for melanin synthesis and eye development. *Frontiers in Cell and Developmental Biology*, **10**: 838721.

Fan JW, Li F, Lou LY, et al. 2025. Gonadal transcriptome profiling via RNA-seq during temperature-dependent sex determination and differentiation in the red-eared slider (*Trachemys scripta elegans*). *Comparative Biochemistry and Physiology Part D: Genomics and Proteomics*, **56**: 101621.

Feschotte C, Pritham EJ. 2007. DNA transposons and the evolution of eukaryotic genomes. *Annual Review of Genetics*, **41**: 331–368.

Flores MV, Tsang VWK, Hu WJ, et al. 2004. Duplicate zebrafish *runx2* orthologues are expressed in developing skeletal elements. *Gene Expression Patterns*, **4**(5): 573–581.

Georges A, Auguste A, Bessière L, et al. 2014. FOXL2: a central transcription factor of the ovary. *Journal of Molecular Endocrinology*, **52**(1): R17–R33.

Golling G, Li LH, Pepling M, et al. 1996. *Drosophila* homologs of the proto-oncogene product PEBP2/CBFβ regulate the DNA-binding properties of Runt. *Molecular and Cellular Biology*, **16**(3): 932–942.

Guerois R, Nielsen JE, Serrano L. 2002. Predicting changes in the stability of proteins and protein complexes: a study of more than 1000 mutations. *Journal of Molecular Biology*, **320**(2): 369–387.

Hecht J, Stricker S, Wiecha U, et al. 2008. Evolution of a core gene network for skeletogenesis in chordates. *Public Library of Science Genetics*, **4**(3): e1000025.

Heyn P, Kalinka AT, Tomancak P, et al. 2015. Introns and gene expression: cellular constraints, transcriptional regulation, and evolutionary consequences. *BioEssays*, **37**(2): 148–154.

Hovland HN, Mchaina EK, Høberg-Vetti H, et al. 2023. Functional analyses of rare germline missense *BRCA1* variants located within and outside protein domains with known functions. *Genes (Basel)*, **14**(2): 262.

Huang BY, Lou LY, Fan JW, et al. 2023. Evolutionary analyses and sequence feature of *Hox* genes in turtles. *Acta Hydrobiologica Sinica*, **47**(10): 1617–1627. (in Chinese)

Hui HB, Xiao L, Sun W, et al. 2021. *Sox9* is indispensable for testis differentiation in the red-eared slider turtle, a reptile with temperature-dependent sex determination. *Zoological Research*, 42(6): 721-725.

Inoue KI, Shiga T, Ito Y. 2008. Runx transcription factors in neuronal development. *Neural Development*, **3**: 20.

Ito Y, Miyazono K. 2003. RUNX transcription factors as key targets of TGF-β superfamily signaling. *Current Opinion in Genetics & Development*, **13**(1): 43–47.

Kalyaanamoorthy S, Minh BQ, Wong TKF, et al. 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature Methods*, **14**(6): 587–589.

Katoh K, Misawa K, Kuma KI, et al. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research*, **30**(14): 3059–3066.

Kim D, Langmead B, Salzberg SL. 2015. HISAT: a fast spliced aligner with low memory requirements. *Nature Methods*, **12**(4): 357–360.

Kimchi-Sarfaty C, Oh JM, Kim IW, et al. 2007. A "silent" polymorphism in the *MDR1* gene changes substrate specificity. *Science*, **315**(5811): 525–528.

Komori T. 2009. Regulation of osteoblast differentiation by Runx2. *In*: Choi Y. Osteoimmunology. New York: Springer, 43–49.

Kuraku S, Usuda R, Kuratani S. 2005. Comprehensive survey of carapacial ridge-specific genes in turtle implies co-option of some regulatory genes in carapace evolution. *Evolution & Development*, **7**(1): 3–17.

Lee KS, Kim HJ, Li QL, et al. 2000. Runx2 is a common target of transforming growth factor β1 and bone morphogenetic protein 2, and cooperation between Runx2 and Smad5 induces osteoblast-specific gene expression in the pluripotent mesenchymal precursor cell line C2C12. *Molecular and Cellular Biology*, **20**(23): 8783–8792.

Levanon D, Groner Y. 2004. Structure and regulated expression of mammalian *RUNX* genes. *Oncogene*, **23**(24): 4211–4219.

Li H, Coghlan A, Ruan J, et al. 2006. TreeFam: a curated database of phylogenetic trees of animal gene families. *Nucleic Acids Research*, **34**(Suppl_1): D572–D580.

Li QL, Ito K, Sakakura C, et al. 2002. Causal relationship between the loss of *RUNX3* expression and gastric cancer. *Cell*, **109**(1): 113–124.

Lyson TR, Bever GS, Bhullar BAS, et al. 2010. Transitional fossils and the origin of turtles. *Biology Letters*, **6**(6): 830–833.

Minh BQ, Schmidt HA, Chernomor O, et al. 2020. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Molecular Biology and Evolution*, **37**(5): 1530–1534.

Nagashima H, Kuraku S, Uchida K, et al. 2007. On the carapacial ridge in turtle embryos: its developmental origin, function and the chelonian body plan. *Development*, **134**(12): 2219–2226.

Nah GSS, Tay BH, Brenner S, et al. 2014. Characterization of the *Runx* gene family in a jawless vertebrate, the Japanese lamprey (*Lethenteron japonicum*). *Public Library of Science One*, **9**(11): e113445.

Naillat F, Yan WY, Karjalainen R, et al. 2015. Identification of the genes regulated by *Wnt-4*, a critical signal for commitment of the ovary. *Experimental Cell Research*, **332**(2): 163–178.

Newton AH, Pask AJ. 2020. Evolution and expansion of the RUNX2 QA repeat corresponds with the emergence of vertebrate complexity. *Communications Biology*, **3**(1): 771.

Nicol B, Grimm SA, Chalmel F, et al. 2019. RUNX1 maintains the identity of the fetal ovary through an interplay with FOXL2. *Nature Communications*, **10**(1): 5116.

Nie CH, Wan SM, Chen YL, et al. 2022. Single-cell transcriptomes and *runx2b*$^{-/-}$ mutants reveal the genetic signatures of intermuscular bone formation in zebrafish. *National Science Review*, **9**(11): nwac152.

Ojima F, Saito Y, Tsuchiya Y, et al. 2019. Runx3 regulates folliculogenesis and steroidogenesis in granulosa cells of immature mice. *Cell and Tissue Research*, **375**(3): 743–754.

Okuda T, Nishimura M, Nakao M, et al. 2001. RUNX1/AML1: a central player in hematopoiesis. *International Journal of Hematology*, **74**(3): 252–257.

Otto F, Lübbert M, Stock M. 2003. Upstream and downstream targets of RUNX proteins. *Journal of Cellular Biochemistry*, **89**(1): 9–18.

Otto F, Thornell AP, Crompton T, et al. 1997. *Cbfa1*, a candidate gene for cleidocranial dysplasia syndrome, is essential for osteoblast differentiation and bone development. *Cell*, **89**(5): 765–771.

Pagani F, Raponi M, Baralle FE. 2005. Synonymous mutations in *CFTR* exon 12 affect splicing and are not neutral in evolution. *Proceedings of the National Academy of Sciences of the United States of America*, **102**(18): 6368–6372.

Paysan-Lafosse T, Blum M, Chuguransky S, et al. 2023. InterPro in 2022. *Nucleic Acids Research*, **51**(D1): D418–D427.

Pelassa I, Corà D, Cesano F, et al. 2014. Association of polyalanine and polyglutamine coiled coils mediates expansion disease-related protein aggregation and dysfunction. *Human Molecular Genetics*, **23**(13): 3402–3420.

Pertea M, Kim D, Pertea GM, et al. 2016. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nature Protocols*, **11**(9): 1650–1667.

Pertea M, Pertea GM, Antonescu CM, et al. 2015. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nature Biotechnology*, **33**(3): 290–295.

Phimphilai M, Zhao ZR, Boules H, et al. 2006. *BMP* signaling is required for *RUNX2*‐dependent induction of the osteoblast phenotype. *Journal of Bone and Mineral Research*, **21**(4): 637–646.

Rennert J, Coffman JA, Mushegian AR, et al. 2003. The evolution of Runx genes I. A comparative study of sequences from phylogenetically diverse model organisms. *BMC Evolutionary Biology*, **3**: 4.

Revell LJ. 2024. Phytools 2.0: an updated R ecosystem for phylogenetic comparative methods (and other things). *PeerJ*, **12**: e16505.

Revell LJ, Schliep K, Valderrama E, et al. 2018. Graphs in phylogenetic comparative analysis: Anscombe's quartet revisited. *Methods in Ecology and Evolution*, **9**(10): 2145–2154.

Rice R, Riccio P, Gilbert SF, et al. 2015. Emerging from the rib: resolving the turtle controversies. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution*, **324**(3): 208–220.

Rose A, Meier I. 2004. Scaffolds, levers, rods and springs: diverse cellular functions of long coiled-coil proteins. *Cellular and Molecular Life Sciences CMLS*, **61**(6): 1996–2009.

Sato H, Adachi N, Kondo S, et al. 2023. Turtle skull development unveils a molecular basis for amniote cranial diversity. *Science Advances*, **9**(46): eadi6765.

Schymkowitz J, Borg J, Stricher F, et al. 2005. The FoldX web server: an online force field. *Nucleic Acids Research*, **33**(suppl_2): W382–W388.

Sears KE, Goswami A, Flynn JJ, et al. 2007. The correlated evolution of *Runx2* tandem repeats, transcriptional activity, and facial length in Carnivora. *Evolution & Development*, **9**(6): 555–565.

Session AM, Uno Y, Kwon T, et al. 2016. Genome evolution in the allotetraploid frog *Xenopus laevis*. *Nature*, **538**(7625): 336–343.

Shaffer HB, McCartney-Melstad E, Near TJ, et al. 2017. Phylogenomic analyses of 539 highly informative loci dates a fully resolved time tree for the major clades of living turtles (Testudines). *Molecular Phylogenetics and Evolution*, **115**: 7–15.

Soneson C, Love MI, Robinson MD. 2016. Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. *F1000Research*, **4**: 1521.

Stricker S, Fundele R, Vortkamp A, et al. 2002. Role of *Runx* genes in chondrocyte differentiation. *Developmental Biology*, **245**(1): 95–108.

Thirunavukkarasu K, Mahajan M, McLarren KW, et al. 1998. Two domains unique to osteoblast-specific transcription factor Osf2/Cbfa1 contribute to its transactivation function and its inability to heterodimerize with Cbfβ. *Molecular and Cellular Biology*, **18**(7): 4197–4208.

Thomson RC, Spinks PQ, Shaffer HB. 2021. A global phylogeny of turtles reveals a burst of climate-associated diversification on continental margins. *Proceedings of the National Academy of Sciences of the United States of America*, **118**(7): e2012215118.

Toriseva M, Björkgren I, Junnila A, et al. 2024. RUNX transcription factors are essential in maintaining epididymal epithelial differentiation. *Cellular and Molecular Life Sciences*, **81**(1): 183.

Trukhina AV, Lukina NA, Wackerow-Kouzova ND, et al. 2013. The variety of vertebrate mechanisms of sex determination. *BioMed Research International*, **2013**: 587460.

Tung Ho LS, Ané C. 2014. A linear-time algorithm for Gaussian and non-Gaussian trait evolution models. *Systematic Biology*, **63**(3): 397–408.

Wang J, Xu YH, Zhu HW, et al. 2025. ChelonianTraits: a comprehensive trait database of global turtles and tortoises. *Scientific Data*, **12**(1): 840.

Waterhouse AM, Procter JB, Martin DMA, et al. 2009. Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics*, **25**(9): 1189–1191.

Westendorf JJ, Hiebert SW. 1999. Mammalian runt‐domain proteins and their roles in hematopoiesis, osteogenesis, and leukemia. *Journal of Cellular Biochemistry*, **75**(S32): 51–58.

Wheeler JC, Shigesada K, Gergen JP, et al. 2000. Mechanisms of transcriptional regulation by Runt domain proteins. *Seminars in Cell & Developmental Biology*, **11**(5): 369–375.

Wu PF, Wang XF, Ge CT, et al. 2024. pSTAT3 activation of *Foxl2* initiates the female pathway underlying temperature-dependent sex determination. *Proceedings of the National Academy of Sciences of the United States of America*, **121**(37): e2401752121.

Yang JY, Xia YY, Li SH, et al. 2022. Gene regulation during carapacial ridge development of *Mauremys reevesii*: the development of carapacial ridge, ribs and scutes. *Genes*, **13**(9): 1676.

Yang ZH. 2007. PAML 4: phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution*, **24**(8): 1586–1591.

Yano-Sakamoto K, Kitai Y, Toriu N, et al. 2024. Expression pattern of Runt-related transcription factor (RUNX) family members and the role of RUNX1 during kidney development. *Biochemical and Biophysical Research Communications*, **722**: 150155.

Yoshida CA, Yamamoto H, Fujita T, et al. 2004. Runx2 and Runx3 are essential for chondrocyte maturation, and Runx2 regulates limb growth through induction of *Indian hedgehog*. *Genes & Development*, **18**(8): 952–963.

Yu XJ, Zheng HK, Wang J, et al. 2006. Detecting lineage-specific adaptive evolution of brain-expressed genes in human using rhesus macaque as outgroup. *Genomics*, **88**(6): 745–751.

Zangerl R. 1969. The turtle shell. *The Biology of the Reptilia*, **1**: 311–319.

Zhang J, Yu P, Zhao Y, et al. 2021. Global analysis of transcriptome and translatome revealed that coordinated WNT and FGF regulate the carapacial ridge development of Chinese soft-shell turtle. *International Journal of Molecular Sciences*, **22**(22): 12441.

Zhong YY, Li LY, He YT, et al. 2020. Activation of steroidogenesis, anti-apoptotic activity, and proliferation in porcine granulosa cells by *RUNX1* is negatively regulated by H3K27me3 transcriptional repression. *Genes*, **11**(5): 495.