SCIENCE CHINA
Life Sciences

•RESEARCH PAPER•

October 2023 Vol.66 No.10: 2399–2414
https://doi.org/10.1007/s11427-023-2362-5

Developmental regulation of conserved non-coding element evolution provides insights into limb loss in squamates

Zeng Wang[1,2], Changjun Peng[1,2], Wei Wu[1,2], Chaochao Yan[1], Yunyun Lv[1,3] & Jia-Tang Li[1,2,4*]

[1]*CAS Key Laboratory of Mountain Ecological Restoration and Bioresource Utilization & Ecological Restoration and Biodiversity Conservation Key Laboratory of Sichuan Province, Chengdu Institute of Biology, Chinese Academy of Sciences, Chengdu 610041, China;*
[2]*University of Chinese Academy of Sciences, Beijing 100049, China;*
[3]*College of Life Science, Neijiang Normal University, Neijiang 641100, China;*
[4]*Southeast Asia Biodiversity Research Institute, Chinese Academy of Sciences, Yezin Nay Pyi Taw 05282, Myanmar*

Received April 14, 2023; accepted May 9, 2023; published online May 22, 2023

Limb loss shows recurrent phenotypic evolution across squamate lineages. Here, based on three *de novo*-assembled genomes of limbless lizards from different lineages, we showed that divergence of conserved non-coding elements (CNEs) played an important role in limb development. These CNEs were associated with genes required for limb initiation and outgrowth, and with regulatory signals in the early stage of limb development. Importantly, we identified the extensive existence of insertions and deletions (InDels) in the CNEs, with the numbers ranging from 111 to 756. Most of these CNEs with InDels were lineage-specific in the limbless squamates. Nearby genes of these InDel CNEs were important to early limb formation, such as *Tbx4*, *Fgf10*, and *Gli3*. Based on functional experiments, we found that nucleotide mutations and InDels both affected the regulatory function of the CNEs. Our study provides molecular evidence underlying limb loss in squamate reptiles from a developmental perspective and sheds light on the importance of regulatory element InDels in phenotypic evolution.

squamates, limb loss, *cis*-regulatory elements, early limb development

Citation: Wang, Z., Peng, C., Wu, W., Yan, C., Lv, Y., and Li, J.T. (2023). Developmental regulation of conserved non-coding element evolution provides insights into limb loss in squamates. Sci China Life Sci 66, 2399–2414. https://doi.org/10.1007/s11427-023-2362-5

INTRODUCTION

The development of limbs was a crucial event to the origin and evolution of tetrapods, providing multiple functions and abilities such as grasping, digging, running, swimming, and flying (Bejder and Hall, 2002; Burgess, 2016; Mann et al., 2022; Parra-Olea and Wake, 2001). Although limb loss has occurred in many clades, it has evolved independently at least 26 times in Squamata (Mann et al., 2022; Royle et al., 2021; Wiens et al., 2006). Recurrent limb reduction in squamate reptiles has long been considered a classic example for studying phenotypic evolution, developmental biology, and comparative genomics (Camaiti et al., 2021; Cohn and Tickle, 1999; Kvon et al., 2016; Leal and Cohn, 2016; Leal and Cohn, 2018; Wiens et al., 2006). Limb reduction or loss generally results from failures or disruptions at different stages of limb development (Camaiti et al., 2021). In limbed tetrapods, limb bud development is controlled by two instructive signaling centers, i.e., the apical ectodermal ridge (AER), which forms from the distal epithelium shortly after limb bud emergence, and the zone of polarizing activity (ZPA), which determines the polarity of limb bud outgrowths (Niswander et al., 1994; Ohuchi et al., 1997; Zuniga and Zeller, 2020). Studies of representative species, such as basal snakes and limbless lizards (e.g., *Malayopython reticulatus* and *Anguis fragilis*), suggest that limb loss results from the truncation of limb bud growth (Cohn and Tickle, 1999;

*Corresponding author (email: lijt@cib.ac.cn)

Rahmanl, 1974; Raynaud, 1985; Raynaud, 1990) due to necrosis and subsequent regression of the AER at the embryonic stage (Rahmanl, 1974). This provides the opportunity to study the occurrence of genetic changes during development.

Genetically, the evolution of enhancers provides a new perspective on the modulation of embryonic limb initiation and outgrowth in squamates (Royle et al., 2021). Previous studies have suggested that point mutations in conserved non-coding elements (CNEs) may affect gene pathways and contribute to limb loss in squamates (Roscito et al., 2018; Roscito et al., 2022). In fact, short insertions and deletions (InDels) are comparable to nucleotide changes in genomes and may exert considerable phenotypic effects (Tóth-Petróczy and Tawfik, 2013). Snakes are known to have a specific 17-bp deletion in the distal ZPA regulatory sequence (ZRS) enhancer of Sonic hedgehog (*Shh*) (Kvon et al., 2016). Segment deletions within the ZRS enhancer can reduce and eliminate *Shh* expression, preventing limb and girdle growth in early embryos, as seen in pythons (Leal and Cohn, 2016). This provides a new perspective for investigating the influence of InDels on limb loss (Kvon et al., 2016). In vertebrates, most genetic circuits are conserved in the early stages of limb development (Royle et al., 2021; Schneider and Shubin, 2013). These processes require the combined action of transcription factors (TFs), morphogens, and signaling pathways, and depend on the temporal regulation of gene expression by *cis*-regulatory elements (CREs) (Lettice et al., 2017; Newton and Smith, 2021). Although genetic divergence in non-coding regions, especially within CREs, is pervasive across genomes, its influences on the evolution of limb loss in squamates are still limited.

Here, we sequenced three limbless lizard species and showed that the shared and divergent evolution of CNEs in limbless squamates played a preferential role in the early stages of limb development. We also identified prevalent InDels within the CNEs, which were predominantly lineage-specific among the limbless squamates. Functional experiments indicated that both mutations and InDels contributed to the regulatory activity of the CNEs. Overall, our study reveals the evolutionary role of CNEs underlying convergent limb loss in squamates and provides insight into the regulatory function of InDels in these CNEs during early limb development.

RESULTS

Genomic landscapes of three limbless lizards

To investigate the genetic basis of limb loss in squamates, we sequenced three limbless lizards from two families (Anguidae and Dibamidae) (Table S1 in Supporting Information). Anguidae is a highly representative clade in which limb loss has occurred (Wiens and Slingluff, 2001), including some well-known limbless lizards such as glass lizards. Dibamidae is the most ancient lineage of squamates, comprising some rare and little-known lizards without limbs (Quah et al., 2017). We generated 140 and 151 Gb of long-read sequences from *Pseudopus apodus* (European glass lizard, Oxford Nanopore Technology) and *Dopasia harti* (Hart's glass lizard, PacBio Circular Long Reads), respectively (Figure 1A; Table S2 in Supporting Information). The assemblies were polished with Illumina short reads (Table S2 in Supporting Information) and showed relatively high continuity (European glass lizard: contig N50, 93 Mb, Hart's glass lizard: contig N50, 12 Mb) compared with previously published limbless lizard genomes (Figure S1 and Table S3 in Supporting Information). We also yielded 393 Gb of Illumina short reads to assemble the draft genome of *Dibamus bourreti* (Bourret's blind skink) (Table S2 in Supporting Information). Overall, the assembled genomes ranged in size from 1.73 Gb (European glass lizard) to 2.86 Gb (Bourret's blind skink), comparable to their lizard counterparts (Figure S2 and Table S4 in Supporting Information). Based on benchmarking of universal single-copy orthologs (BUSCO), we retrieved 97.87% (European glass lizard) and 97.10% (Hart's glass lizard) complete BUSCO genes (Table S5 in Supporting Information).

The genomes contained 55.24% (European glass lizard) and 55.08% (Hart's glass lizard) repetitive sequences (Figure S3 and Table S6 in Supporting Information). After masking the repetitive sequences, genome-wide annotation was performed by combining homology-based prediction, *de novo* prediction, and transcriptome sequencing-based gene prediction. In total, 21,987 and 19,857 protein-coding genes were annotated for the European glass lizard and Hart's glass lizard, respectively (Table S7 in Supporting Information). Furthermore, 96.37% and 92.34% of genes were functionally annotated for the two species (Table S8 in Supporting Information).

Independent evolution of limb loss in squamate lineages

We performed multiple genome alignments using the European glass lizard as a reference. The orthologous segments of the genome alignments were used to construct a phylogenetic tree (Table S9 in Supporting Information). The reconstructed topology showed that limb loss evolved independently in several lineages (Figure 1B; Figure S4 in Supporting Information). Based on six fossil calibrations (Table S10 in Supporting Information), we estimated that several independent limb losses occurred at different times (Figure S5 in Supporting Information). Bourret's blind skink was derived from the ancestral lineage of Squamata nearly 198 million years ago (Ma), belonging to the sister group of all living squamates. The snake lineage diverged about

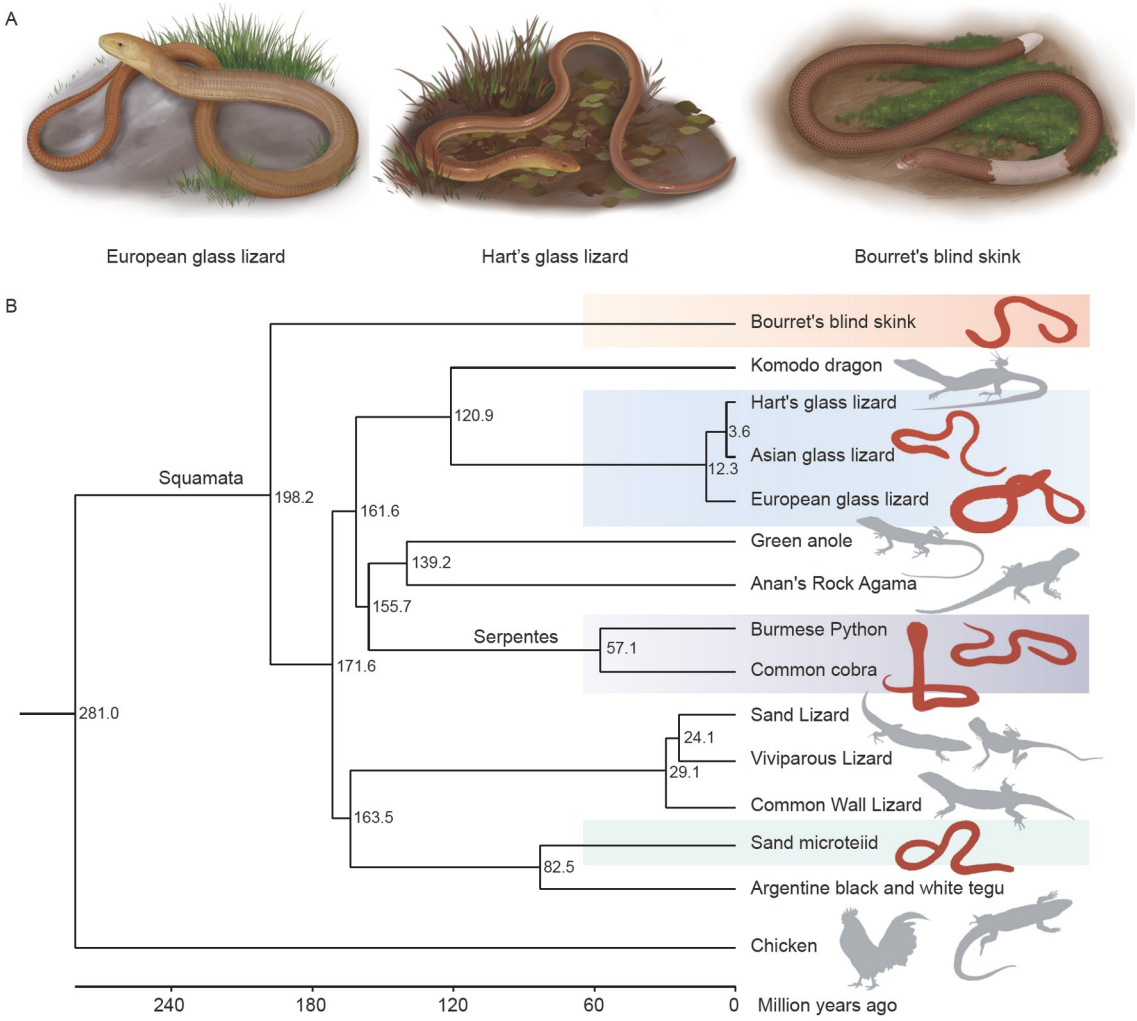

Figure 1 Genome assemblies of three limbless lizards and independent limb loss in Squamata. A, Diagrams of three limbless lizards newly sequenced in this study. From left to right, European glass lizard (*Pseudopus apodus*), Hart's glass Lizard (*Dopasia harti*), and Bourret's blind skink (*Dibamus bourreti*). B, Phylogenetic tree and divergence time of squamate reptiles. Limbless lineages are highlighted with different background colors. Limbless species are shown in red.

155 Ma, while the sand microteiid lineage diverged nearly 82 Ma. The most recent common ancestor of the three anguids (i.e., European glass lizard, Asian glass lizard, and Hart's glass lizard) diverged around 12 Ma, suggesting that limb loss in the anguid lineage was much younger than that in the other three lineages (Figure 1B; Figure S5 in Supporting Information). These four independent limbless lineages provided the foundation for subsequent genomic analyses.

Potential regulatory function of convergent evolution of dCNEs

We first detected convergent signals in snakes and glass lizards. A total of 796 divergent CNEs (dCNEs) were found in the pair of snakes and glass lizards. Integrating six ATAC-seq regulatory datasets (heart, limb, flank, liver, brain, and embryo) from the tegu lizard (Roscito et al., 2018), 81 dCNEs were identified as potentially functional. Among these dCNEs, approximately 53% overlapped with limb regulatory signals, showing a relatively higher percentage than signals from other tissues (Figure 2A). In addition to snakes, limb loss has evolved in multiple lizard lineages (Wiens et al., 2006). We identified 438 convergent dCNEs in the two limbless lizard lineages (glass lizards and Bourret's blind skink), 34 of which overlapped with regulatory peaks. These dCNEs also showed a relatively higher percentage (almost 50%) in limb regulatory signals (Figure 2A). Furthermore, the dCNEs in the two limbless lineage pairs were significantly enriched in limb regulation (Figure S6A in Supporting Information), suggesting a potential function in limb development. This differs from the previous study (Roscito et al., 2022), showing little shared limb-relevant dCNEs between limbless lizards. In addition, we used time-series ordered ATAC-seq datasets of mouse limb development stages from E9.5 to E14.5 to explore the involvement of the convergent dCNEs. Our analysis revealed that a significantly

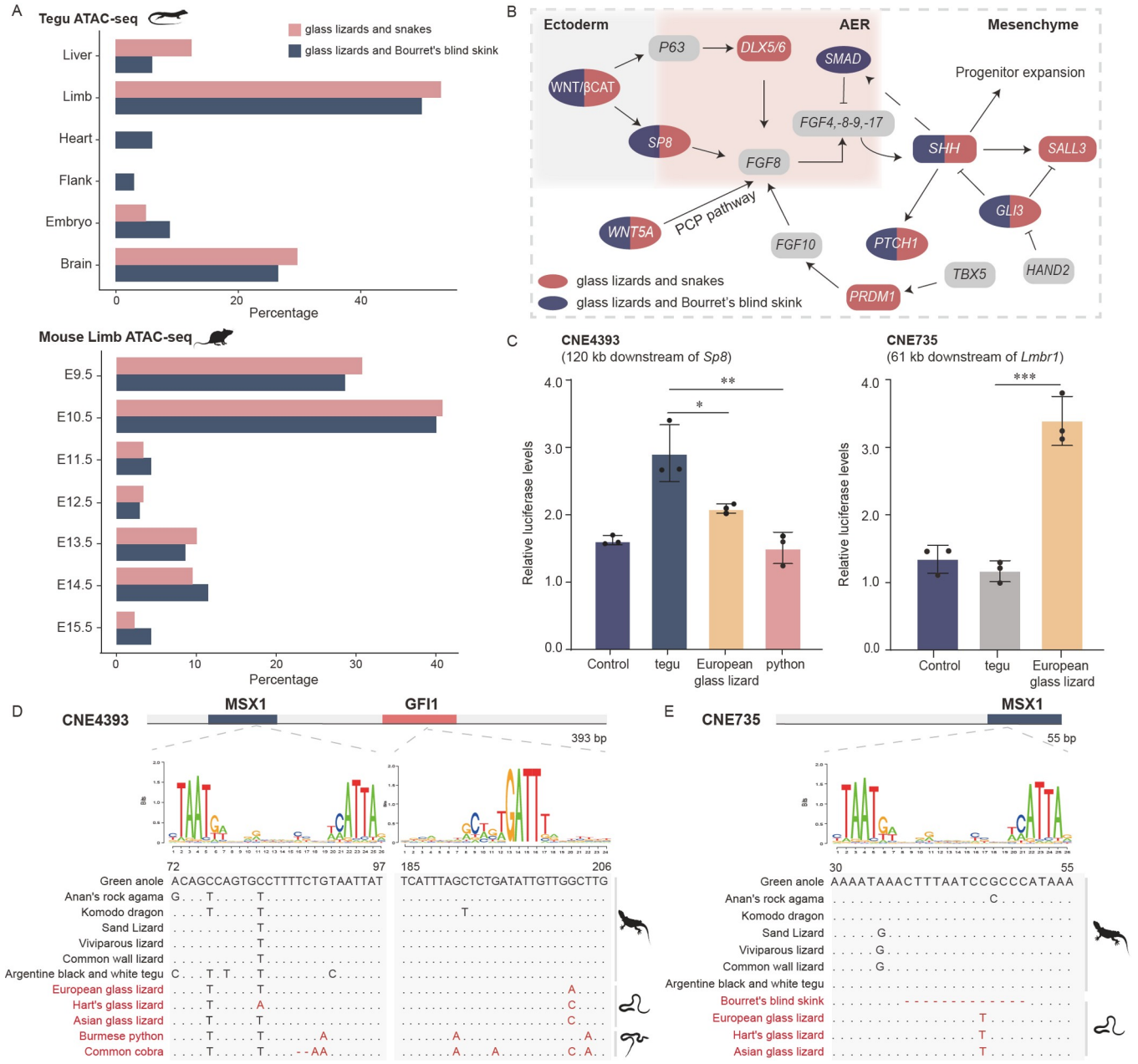

Figure 2 Convergent dCNEs associated with limb regulatory elements and limb bud formation. A, For convergent dCNEs that overlap with regulatory signals in limbless lineage pairs (glass lizards and Bourret's blind skink, glass lizards and snakes), histograms show percentage of dCNEs that overlap with peaks in different tegu lizard tissues (top) and time-ordered mouse embryo stages (bottom). B, Neighboring genes of convergent dCNEs in two limbless lineage pairs are involved in limb bud formation networks. C, Comparisons of the regulatory activity between limbed and limbless squamates in CNE4393 and CNE735 by dual-luciferase assays. Data are mean±standard error of three individual experiments. *P* value was calculated using Student's *t*-test (*, $P<0.05$; **, $P<0.01$; ***, $P<0.001$). D, Mutations in two TFBSs of three glass lizards and two snakes in CNE4393, a CNE overlapping with experimentally tested limb enhancer (Figure S7A in Supporting Information). Mutations and limbless species are highlighted in red. Dots indicate the same nucleotides as above. E, Mutations in TFBS of three glass lizards and Bourret's blind skink in CNE735, a CNE overlapping a limb ATAC-seq peak in tegu (Figure S7B in Supporting Information). Mutations and limbless species are highlighted in red. Dots indicate the same nucleotides as above. Deletions in Bourret's blind skink are shown by short lines.

higher percentage of dCNEs in both two pairs might have functional significance in the E9.5 and E10.5 stages (Figure 2A; Figure S6B in Supporting Information), during which forelimb and hindlimb buds are initiated (Zuniga, 2015).

Regarding the potential target genes for dCNEs, we associated CNE with the relevant genes that were located 300 kb upstream or downstream of it. We found that multiple genes were broadly implicated in the pathways governing limb initiation and outgrowth, including *Shh*, *Gli3*, *Ptch1*, and *Twist1* (Figure 2B; Tables S11 and S12 in Supporting Information). TWIST acts as a regulator of epithelial-to-mesenchymal transition in the limb bud formation regions, and

the TWIST1 TF is essential for mesoderm formation (Hirsch et al., 2018; Yang et al., 2004). SHH, encoded by *Shh*, is a major signaling center in the ZPA and is involved in the self-regulatory properties of the SHH/GREM1/AER-FGF signaling system, which is comprised of several interlinked feedback loops that are crucial during limb bud formation (Benazet et al., 2009; Zuniga and Zeller, 2020). Furthermore, *Ptch1* is a transcriptional indicator of SHH signaling. The previous study on python embryogenesis has shown that *Ptch1* is expressed in the hindlimb buds on the day of oviposition, but the expression domain is weaker than that expressed in limbed lizards (Leal and Cohn, 2016). These results suggest that the convergent dCNEs may putatively regulate gene expression during early limb development.

To determine whether divergent sequences affect regulatory function, we selected dCNEs that were located near genes involved in early limb development in each of the two limbless pairs and performed luciferase assays. Among the dCNEs, CNE4393 was detected in the snake and glass lizard pair and located 120 kb downstream of *Sp8* (Figure 2D; Figure S7A in Supporting Information). It is a TF expressed in the AER and involved in interactions between the ectoderm and mesoderm of the early limb bud (Haro et al., 2014). This CNE also overlapped with an enhancer and exhibited AER limb expression when tested *in vivo* (VanderMeer et al., 2014). In addition, CNE735 was identified 61 kb downstream of *Lmbr1* (Figure S7B in Supporting Information) and overlapped with a limb-specific ATAC-seq peak from the tegu embryonic stage, corresponding to the E10.5 stage in mouse development (Roscito et al., 2018). Luciferase assays of the two dCNEs showed significant differences in activity between the limbed and limbless squamate groups (Figure 2C). Notably, the limbless European glass lizard sequence of CNE735 showed significantly higher expression compared with the tegu with full limbs, suggesting that the divergence of the European glass lizard sequence may have released a potential repressive activity. TFs bind to specific *cis*-regulatory regions to regulate gene expression, and mutations in transcription factor binding sites (TFBS) act by disrupting or altering TF binding (Roscito et al., 2022). We identified 14 motifs in CNE4393 and 7 motifs in CNE735, respectively (Table S13 in Supporting Information). By screening the binding TFs in the motifs, we found that mutations in the TFBSs of the two CNEs might be relevant for limb development (Figure 2D and E; Table S13 in Supporting Information). These observations suggest that TFBS mutations may contribute to the regulatory functions of dCNEs, but more experiments are required in further studies.

Contribution of InDels to CNE evolution in limbless squamates

To identify CNEs with InDels (InDel CNEs), we analyzed four datasets containing four independent limbless lineages, including the Bourret's blind skink, glass lizard, sand microteiid, and snake lineages (Table S14 in Supporting Information). We identified 545, 756, 584, and only 111 InDel CNEs in the Bourret's blind skink, sand microteiid, snake, and glass lizard lineages (Figure 3A). The fewer InDel CNEs in glass lizard may be explained by their short divergence time, resulting in the accumulation of fewer mutations than in the three "older" limbless lineages. Notably, the InDel CNEs among the four lineages exhibited little overlap, indicating that these CNEs were predominantly lineage-specific (Figure 3A). After screening potential regulatory functions, we identified 72, 123, 10, and 58 InDel CNEs in the Bourret's blind skink, sand microteiid, glass lizard, and snake lineages, which overlapped with regulatory signals. Overlap with limb regulatory signals was higher than that with signals from other tissues (32%, 37%, 60%, and 41% in the Bourret's blind skink, sand microteiid, glass lizard, and snake lineages) (Figure 3B). In addition, a higher percentage of InDel CNEs overlapped with regulatory peaks at stages E9.5 and E10.5 in the four lineages (Figure 3B), which is consistent with the dCNEs trend and indicates an abundant divergence of CNEs in the early stages of limb development in limbless squamates.

Neighboring genes of InDel CNEs in the four lineages were involved in limb initiation and limb bud formation (Figure 3C). For example, *Tbx4* and *Pitx1* are required for the initiation of hindlimb development (Chapman et al., 1996; Infante et al., 2013). We also detected *Fgf10*, which is expressed in the limb mesenchyme and forms an epithelial-mesenchymal feedback signal with *Fgf8* in the AER (Ohuchi et al., 1997). Interestingly, we found that limb genes near the CNEs in the four lineages, including *Shh* and *Gli3*, were involved in the SHH signaling pathway in the ZPA. Mutual antagonism of the HAND2 and GLI3 TFs polarizes the anterior-posterior axis of the nascent limb bud, resulting in the activation and posterior restriction of *Shh* expression in the mesenchyme (te Welscher et al., 2002; Zuniga, 2015). Thus, our results indicate that limb-related InDels in the four lineages affect pathways involved in early limb development, particularly limb field positioning and initiation, and subsequent outgrowth processes, in which SHH functions with AER-FGFs signals in a positive feedback loop (Laufer et al., 1994; Niswander et al., 1994).

Contribution of InDels to regulatory functions of CNEs

InDels in CNEs can strongly influence *cis*-regulatory activity underlying phenotypic evolution (Frankel et al., 2011; Jeong et al., 2008; McLean et al., 2011). As we identified several candidate regulatory regions with InDels in each limbless lineage, we next determined whether these InDels affected gene expression. We first focused on CNEs that overlapped

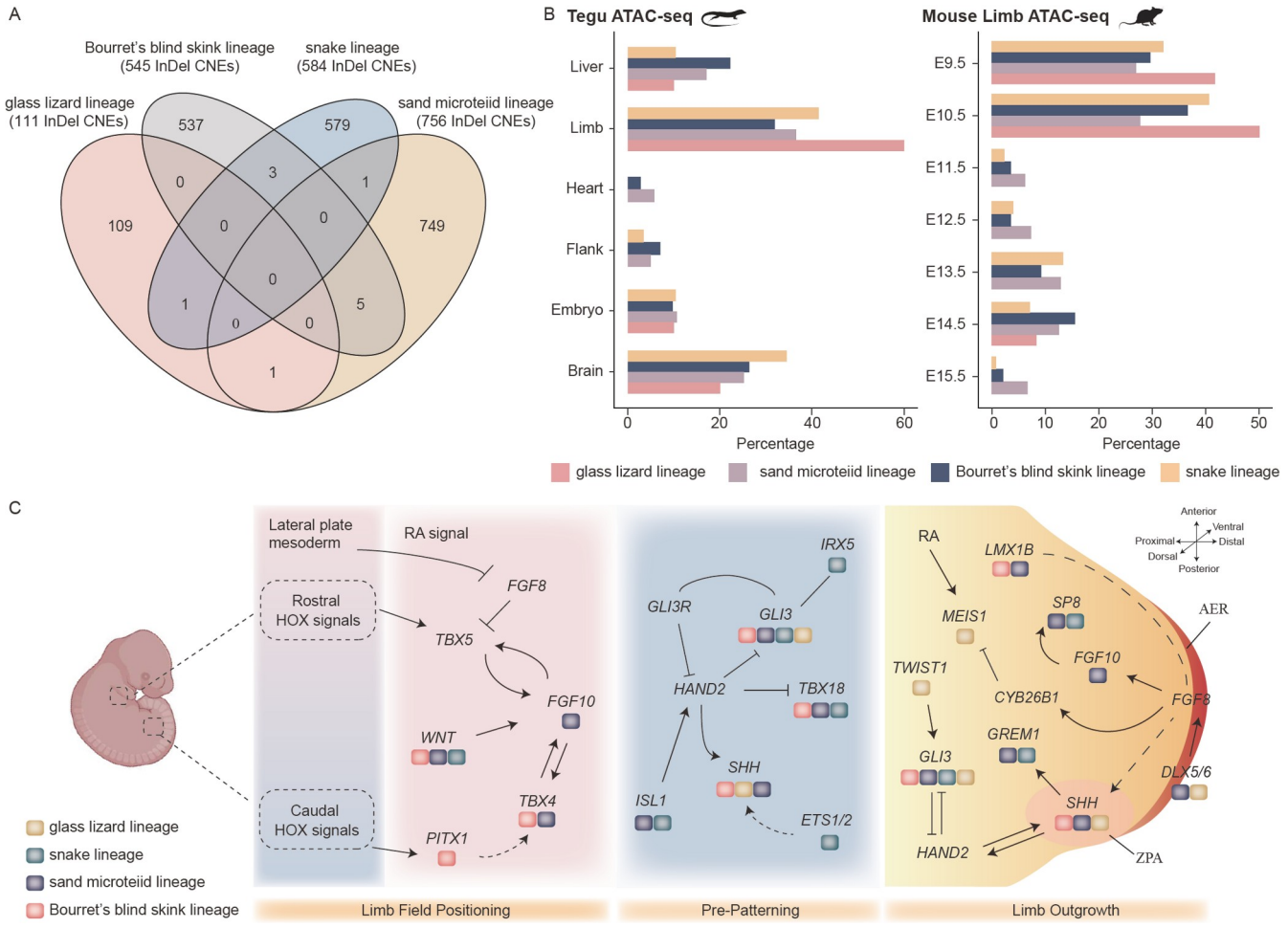

Figure 3 InDel CNEs are primarily limbless-lineage specific and participate in regulatory networks during early limb development. A, The overlapping of InDel CNEs in four limbless lineages. B, For InDel CNEs that overlap with regulatory signals in the four lineages, histograms show percentages of InDel CNEs that overlap with peaks in different tegu lizard tissues (left) and time-ordered mouse embryo stages (right). C, Genes associated with CNEs are broadly involved in limb initiation and outgrowth pathways. Two signaling centers, AER at the distal tip of the limb bud and polarizing region (ZPA) in the posterior limb bud mesenchyme, are indicated. The cartoon of the embryo is created with BioRender.com. Different colors correspond to four limbless lineages.

with experimentally validated enhancers using published *cis*-regulatory datasets. We found that glass lizard-specific insertions in CNE5611 were located in the exonic enhancer (eExon) of *Hdac9*, which regulates *Twist1* expression in limb buds (Hirsch et al., 2018) (Figure 4A). By screening neighboring genes, we also identified CNE3902 with a specific 52-bp insertion in Bourret's blind skink located near *Shh*, an important gene expressed in the ZPA during limb initiation. This CNE overlapped with a DNase signal peak in mouse limb development (Figure 4B). Deletions were also detected, including two CNEs (Figure 4C and D) with specific deletions in the Bourret's blind skink and snake lineages, respectively. These CNEs overlapped with a validated enhancer and were located in eExon17 of *Dync1l1*, which drives the expression of *Dlx5/6* in the limb bud mesenchyme (Birnbaum et al., 2012).

To verify our findings, we constructed groups of insertions and deletions using luciferase assays. Two candidate regulatory regions (CNE5611 and CNE3902) showed potential regulatory activities between limbed and limbless squamate groups (Figure 4E). The activities between Bourret's blind skink and wild-type tegu in CNE1639 showed no significant differences, possibly due to other mutations in the sequence apart from the 12-bp deletion. Overall, the constructed InDels in the wild-type tegu sequence of the three tested CNEs significantly influenced the regulatory function. Two CNEs (CNE5611 and CNE3902) showed increased expression when insertions were introduced to the wild-type tegu sequence, while CNE1639 showed a decreasing trend when the deletion was introduced. These results suggest that InDels may affect the regulatory activities of CNEs, resulting in increase or decrease in the expression of target genes.

Loss of limbs accompanies molecular changes in skeletal development

To assess the significance of limb loss in squamates, we examined the selective pressures exerted on limbless squa-

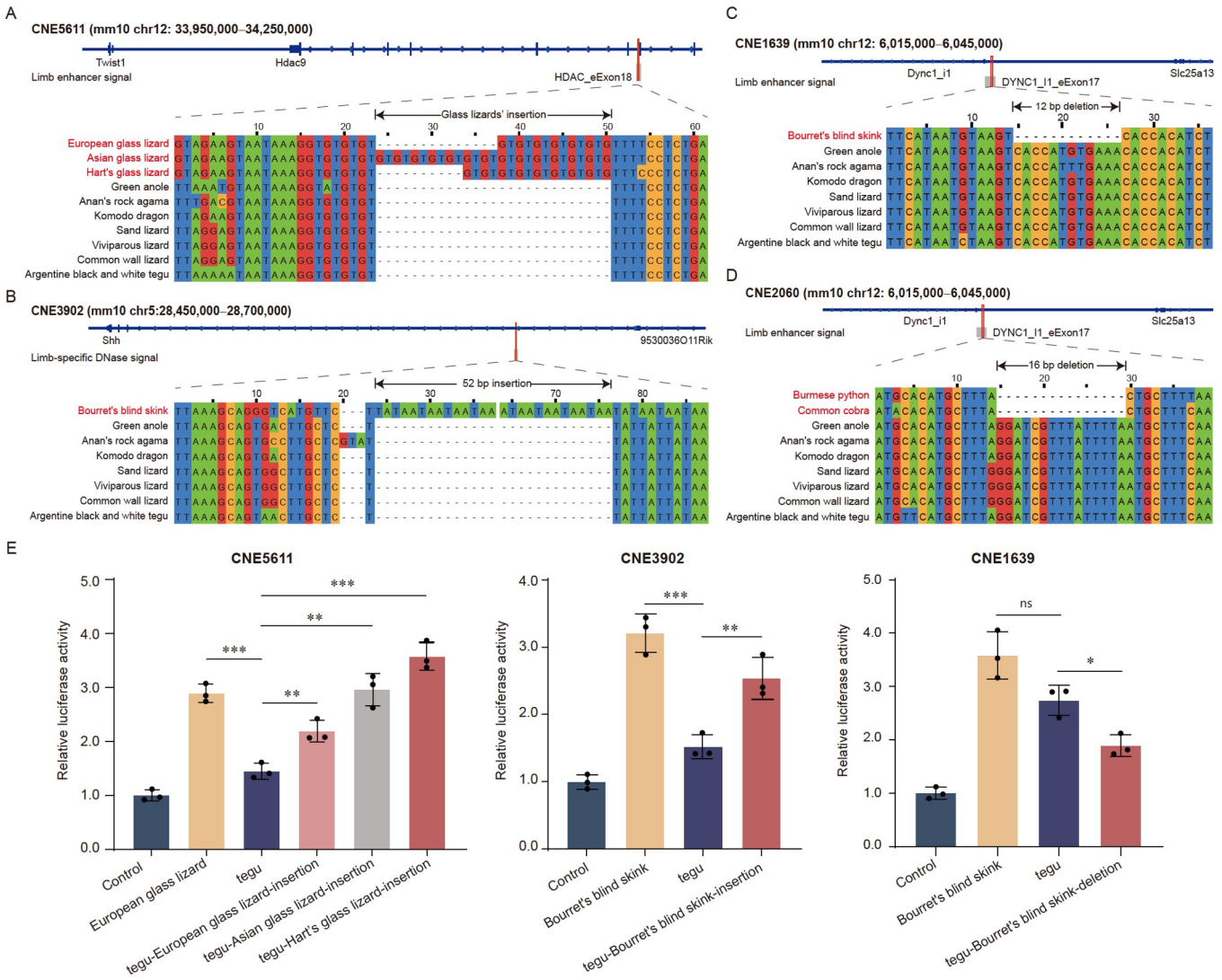

Figure 4 Influence of InDels on regulatory activity detected by dual-luciferase assays. A, Glass lizard-specific and variable insertions in CNE5611 overlapped with eExon of *Hdac9*, an experimentally validated enhancer of *Twist1*. B, A 52-bp insertion of Bourret's blind skink in CNE3902, located near the important limb gene *Shh*. Corresponding mouse limb-specific DNase signal is shown. C, A 12-bp deletion of Bourret's blind skink in CNE1639 overlapped with eExon17 of *Dync1I1*, an experimentally validated enhancer of *Dlx5/6*. D, A 16-bp deletion of snakes in CNE2060, located in eExon17 of *Dync1I1*. E, Dual-luciferase assays detected function of InDels in CNEs. Data are mean±standard error of three individual experiments. *P* value was calculated using Student's *t*-test (ns, not significant; *, $P<0.05$; **, $P<0.01$; ***, $P<0.001$).

mates, given that limb loss evolved as an adaptive trait for fossorial or cluttered habitats (Gans, 1975; Wiens et al., 2006). We used European glass lizard and Hart's glass lizard to compare genetic changes between limbless and limbed lizards, as their ancestral traits have been reconstructed as limbless (Wiens and Slingluff, 2001). In total, 17,193 orthologous gene families were identified and used to screen for selection with branch-site and branch models. We detected 1,762 positively selected genes (PSGs) and 1,314 rapidly evolving genes (REGs) in the most recent common ancestor of the two glass lizards (*P* value<0.05) (Tables S15 and S16 in Supporting Information). Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis indicated that these candidates were significantly associated with ossification (adjusted *P* value=3.40×10^{-2}) and skeletal system development (adjusted *P* value=3.87×10^{-2}) (Figure S8, Tables S17 and S18 in Supporting Information). Bone is composed of osteocytes that originate from mesenchymal stem cells, grow into osteoprogenitor cells, and differentiate into osteoblasts. This developmental process requires multiple signaling pathways, including the Transforming growth factor-beta/Bone morphogenic protein (TGF-β/BMP), Fibroblast Growth Factor-FGF Receptor (FGF-FGFR), and WNT signaling pathways. We identified a number of PSGs and REGs involved in these essential pathways (Figure 5), suggesting that glass lizards underwent considerable genetic transformations during skeletal development. Thus, limb loss may be accompanied

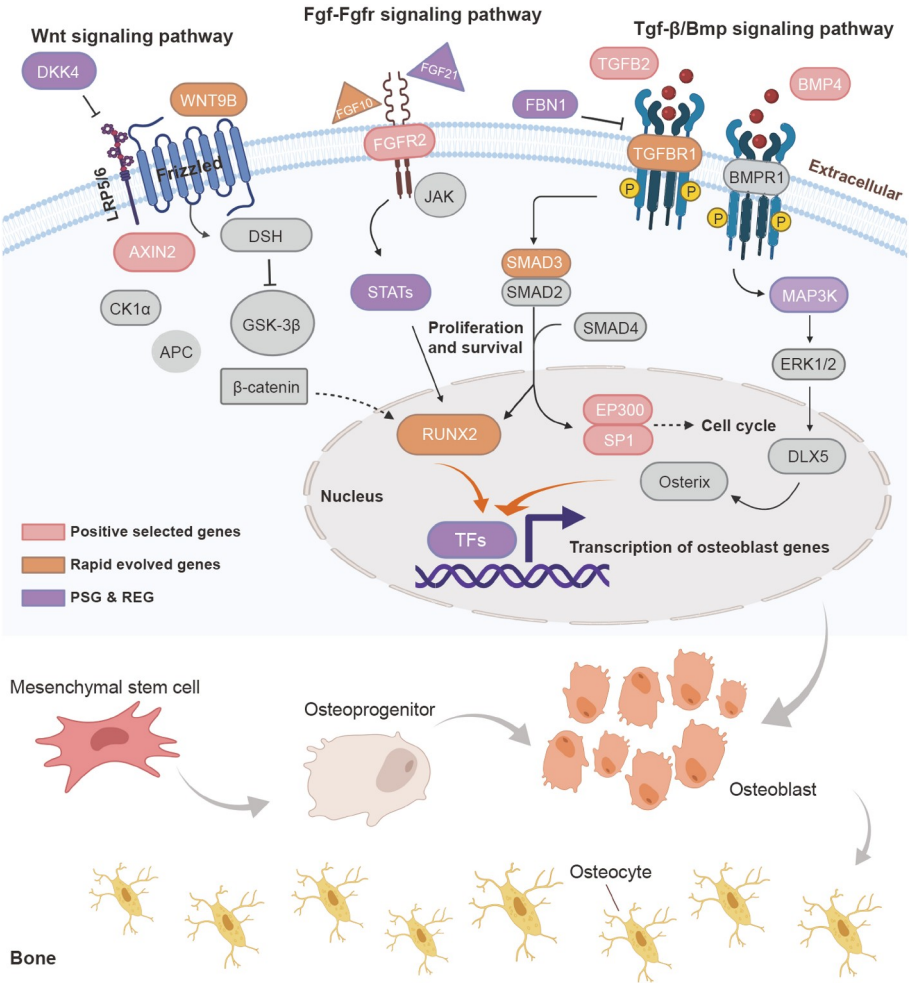

Figure 5 Genes under positive selection are involved in key skeletal development pathways. Created with BioRender.com.

by changes in other skeletal systems to adapt to different environments.

We also performed amino acid replacement analysis (Xu et al., 2017) to capture convergent signals in protein-coding regions. In total, 293 genes with convergent amino acid substitutions were identified in snakes and glass lizards (Table S19 in Supporting Information). Although GO and KEGG enrichment analyses showed little correlation with skeletal development (Tables S20 and S21 in Supporting Information), among the identified convergent genes, *Chsy1* is essential for bone development, with *Chsy1*$^{-/-}$ mice displaying chondrodysplasia and reduced bone density (Wilson et al., 2012). We also found that *Ofd1*, which is required for correct ciliogenesis during skeletal patterning (Bimonte et al., 2011), contained convergent substitutions. Interestingly, these skeletal development-related genes are involved in the morphogenesis of the limb autopod. For example, *Chsy1*$^{-/-}$ mice have profound defects in distal digit patterning (Wilson et al., 2012), and mutant mice in which *Ofd1* is inactivated in the mesenchyme display severe polydactyly with loss of antero-posterior digit patterning and shortened long bones (Bimonte et al., 2011). This raises the possibility that the molecular changes underlying body axis elongation in limbless squamates may be correlated with limb development. However, further research is needed to investigate the underlying genetic links between body elongation and limb loss.

DISCUSSION

Vertebrates share widely conserved genetic cascades involved in early limb development (Royle et al., 2021). In the current study, we investigated the genomic changes underlying convergent limb loss in squamate reptiles. Most reports of limb loss suggest that changes in gene expression and regulation play an important role in the early stages of limb development (Farlie et al., 2017; Smith et al., 2016; Thewissen et al., 2006; Young et al., 2019). Our analyses indicate that CNEs with mutations or InDels in limbless squamates contribute to regulatory function and are involved in gene pathways during limb positioning, initiation, and bud

formation. In basal snakes, the signals that drive hindlimb development are initiated, but subsequently arrested and regressed due to apoptosis in the AER, which plays a crucial role in the feedback loop with SHH in the ZPA (Cohn and Tickle, 1999; Leal and Cohn, 2016; Petit et al., 2017). Similarly, previous studies in limbless lizards have shown that the main causative factor inhibiting sustained limb bud outgrowth is the onset of AER necrosis and its subsequent decline (Rahmanl, 1974; Raynaud, 1985). Based on genomic evidence, we found that many genes associated with CNEs in the limbless squamates interacted with genes expressed in the AER and ZPA. Interestingly, research on the developmental mechanism underpinning hindlimb loss in dolphins suggests that the AER formation occurs in cetacean hindlimb buds, but it is poorly maintained and the ZPA formation fails (Thewissen et al., 2006). This implied limb loss in distant vertebrate lineages may follow similar evolutionary pathways, affecting gene networks of the two signaling centers (AER and ZPA). These studies suggest that changes in genetic pathways affecting AER and ZPA interactions may be common determinants of limb loss in vertebrates.

CREs can precisely control developmental processes by orchestrating spatio-temporal gene expression (Malkmus et al., 2021; Wittkopp and Kalay, 2011). A long-standing question is how *cis*-regulatory sequences have evolved to influence the developmental regulatory networks underlying phenotypic evolution. Previous studies have demonstrated that large segment deletions and insertions (i.e., structural variations) contribute to phenotypic evolution in a variety of taxa (Alonge et al., 2020; Guan et al., 2021; He et al., 2019; Li et al., 2021; Wang et al., 2019; Yang et al., 2019), but the large evolutionary distances make it difficult to detect structural variations that may contribute to limblessness in squamates. In our study, we identified InDels, including segmental insertions and deletions, and investigated their impacts on regulatory function. We found that InDels were ubiquitous in limbless squamate reptiles and showed broad lineage specificity. In addition to point mutations, InDels were also sufficient to alter the *cis*-regulatory activity of CNEs. Indeed, *in vitro* assay showed that deletion in Bourret's blind skink reduced regulatory activity, consistent with the function of CNE deletions in snakes, three-spined sticklebacks, and fruit flies (Chan et al., 2010; Frankel et al., 2011; Leal and Cohn, 2016). These studies indicate that loss of CNE function tends to be associated with segmental deletions. However, we identified a variable microsatellite expansion insertion in the CNE of three glass lizards, which showed different enhancer activity capacities. Similarly, microsatellite expansion has been identified in the candidate *cis*-regulatory region of artiodactyls (Lopez-Rios et al., 2014). This discrepancy may be because shared mutations in CNEs are maintained in the ancestral state, while each species evolves individually during the threshold period.

Combined with time-series ordered ATAC-seq in mice, we identified many CNEs that overlapped with limb regulatory signals at stages E9.5 and E10.5, suggesting that modifications in these CNEs could potentially influence early developmental processes. However, most CNEs tested as putative enhancers were not found at either of the embryonic stages, and only CNE5611 was identified with insertions at E10.5. One possible explanation may be the distant evolutionary relationship between squamates and mice. Notably, we found a relatively high percentage of CNEs with InDels in the four limbless lineages at stages E13.5 and E14.5. Research on developmental regulatory genes suggests that enhancers would be redundant in acting as buffers against numerous mutations (Malkmus et al., 2021; Osterwalder et al., 2018). Therefore, the maintenance of these CNEs may be due to pleiotropy (Infante et al., 2018), implying that they may play additional roles in embryogenesis and development other than limb morphogenesis. For example, many enhancers play dual roles in regulating both limb and genital appendage development (Infante et al., 2015). Nevertheless, the lack of a *cis*-regulatory landscape of essential developmental stages in limbless squamates hinders our understanding of specific developmental signals, which remains to be addressed in future investigations.

Limb loss is often associated with multiple morphological changes, with axial elongation being the most prominent feature in vertebrates (Caldwell, 2003; Collar et al., 2016; Law et al., 2019; Urošević et al., 2016). Body lengthening can be achieved by lengthening the trunk, tail, or both (Brandley et al., 2008; Ward and Brainerd, 2007). Here, we detected abundant protein-coding genes enriched in ossification and skeletal development under positive selection pressure. This may support the prevailing hypothesis that body elongation accompanied by limb loss is an adaptation to fossorial or cluttered habitats (Gans, 1975; Mehta et al., 2010; Sharpe et al., 2015). Previous research has shown that heterochronic shifts in ossification are correlated with limb reduction in scincid lizards (Hugi et al., 2012). Our study provided molecular evidence that genes in several key pathways associated with ossification were positively selected during skeletal development, suggesting a potential link between limb loss, locomotor patterns (Bergmann et al., 2020), and environmental adaptation in limbless squamates.

Although genomic changes in coding genes and noncoding elements have been investigated, exploring the genetic basis for limb loss in squamate reptiles is challenging due to their long evolutionary history, considerable diversification, and global distribution (Burbrink et al., 2020). The order Squamata also comprises approximately 11,820 extant lizards and snakes (Uetz et al., 2022), with limb loss having evolved in multiple lineages (Anguidae, Cordylidae, Dibamidae, Gerrhosauridae, Gymnophthalmidae, Pygopodidae, Scincidae, and serpents) (Brandley et al., 2008). The limited

number of squamate genomes, especially those of limbless squamates, remains a major limitation of the study. This could be addressed by global collaboration and sequencing of more samples, which will further contribute to our understanding of evolutionary genomics.

MATERIALS AND METHODS

Sampling and genome sequencing

We newly sequenced and *de novo* assembled the genomes of three lizards. Liver samples from *Pseudopus apodus* (European glass lizard) and *Dopasia harti* (Hart's glass lizard) were dissected and immediately frozen in liquid nitrogen for subsequent genome sequencing (Table S1 in Supporting Information). Muscle samples were obtained from *Dibamus bourreti* (Bourret's blind skink) and preserved in 95% alcohol until used. All specimens were deposited at the Herpetological Museum, Chengdu Institute of Biology, Chinese Academy of Sciences, Chengdu, China. All experimental procedures, including animal collection and sampling, were performed with the approval of the Animal Experiment Ethics Committee, Chengdu Institute of Biology, Chinese Academy of Sciences.

DNA was extracted using a QIAGEN® Genomic Kit (Qiagen, USA) and the concentration was determined using the Qubit® 4.0 Fluorometer (Invitrogen, USA) according to the manufacturer's standard instructions. Libraries were further constructed for qualified samples (Table S2 in Supporting Information). For the European glass lizard, a 700 ng DNA library was constructed and sequenced on a Nanopore PromethION sequencing instrument (Oxford Nanopore Technologies, UK) at the Genome Center of GrandOmics (Wuhan, China). For the Hart's glass lizard, genomic DNA was sheared to a length of 20 kb using a g-TUBE device (Covaris, USA), then further purified and concentrated using AmpureXP beads (Agencourt, USA). A Single-Molecular Real-Time (SMRTbell) library was constructed using the PacBio 20-kb template preparation protocol on the Sequel II platform. A whole-genome shotgun sequencing strategy was used to generate short reads. Paired-end libraries were sequenced on the MGISEQ 2000 platform for the European glass lizard and the Illumina NovaSeq 6000 platform for Hart's glass lizard. For the Bourret's blind skink, paired-end libraries with insert lengths ranging from 300 to 500 bp were constructed, then sequenced on the Illumina Hiseq 2500 platform.

Transcriptome sequencing

Multiple tissues were dissected from the European glass lizard and Hart's glass lizard, respectively (Table S22 in Supporting Information). For each lizard, tissues were weighed in equal mass and mixed for subsequent transcriptome sequencing. Total RNA was extracted from pooled tissues using a TRIzol kit (Life Technologies, USA) following the manufacturer's standard protocols. PolyA mRNAs were isolated and disrupted into short segments using oligonucleotide (dT) magnetic beads. cDNA was then synthesized using random primers and reverse transcriptase. Finally, after adapter-ligation, end-repair, and polymerase chain reaction (PCR) amplification, cDNA libraries were constructed and sequenced using the Illumina Hiseq 2000 and NovaSeq 6000 strategies (Table S22 in Supporting Information).

Genome assembly and evaluation

For the European glass lizard, NextDenovo v1.0 (https://github.com/Nextomics/NextDenovo) was used for self-correction with the parameters "read_cuoff=1k, seed_cutoff=25k". Reads were then assembled using Wtdbg2 (Ruan and Li, 2020) and polished three times using NextPolish v1.0.1 (Hu et al., 2020) with default parameters. After quality control, the short reads were mapped to the assembled genome using BWA software and polished four times using NextPolish v1.0.1 to correct errors. For the Hart's glass lizard, we used a hybrid method following our previous study (Zhang et al., 2022). Illumina reads were first assembled into initial contigs using Platanus v1.2.4 (Kajitani et al., 2014) with the parameters "-k 31 -t 8 -d 0.3 -m 200". The contigs were then aligned to PacBio long reads using DBG2OLC (Ye et al., 2016) to generate consensus contigs. Base errors were polished with NextPolish v1.0.1. Finally, BUSCO v5.1.3 (Simão et al., 2015) was used to assess the integrity of the genome assemblies with vertebrate genome models and lineage data.

De novo genome assembly of Bourret's blind skink

To obtain high-quality data, we filtered raw reads as follows. First, duplicate reads were removed. Reads with adapter contamination were filtered and low-quality regions were trimmed using Trimmomatic (Bolger et al., 2014). Reads containing more than 5% N bases were discarded. After filtering, all clean reads were *de novo* assembled using DISCOVAR (https://www.broadinstitute.org/software/discovar/blog) with default parameters.

Genome annotation

Repeat annotation

Transposable elements (TEs) were annotated using similarity search and *de novo* prediction with RepeatModeler v1.0.11 and RepeatMasker v1.331 against the RepBase transposable element library (20170127) (Jurka, 1998). Tandem repeats were predicted using Tandem Repeats Finder (TRF) (Benson, 1999), and long terminal repeats (LTRs) were identified

using LTR_FINDER (Xu and Wang, 2007) with default parameters. Finally, the RepBase and *de novo* annotations were merged.

Gene prediction

Genome annotation was performed by combining transcriptome alignment, *de novo* prediction, and homology-based strategies. For transcriptome alignment, clean data were aligned to the corresponding reference genome using STAR v2.6.1d (Dobin et al., 2013). The transcript locus in the assemblies was obtained with StringTie v1.3.4 (Pertea et al., 2015), and the longest transcripts were selected as the representative of each gene. Untranslated regions (UTR) and alternative splicing regions were predicted by PASA v2.3.3 (Haas et al., 2003). For *de novo* prediction, PASA v2.3.3 was used to generate a training set and Augustus v3.3.1 (Stanke et al., 2006) was used with default parameters for ab initio gene prediction with the training set. For the homology-based strategy, protein sequences of five species (Komodo dragon, Argentine black and white tegu, common wall lizard, Asian glass lizard, and green anole) were downloaded from the Swiss-Prot database, and gene structures were predicted using GeMaMo v1.6.1 (Keilwagen et al., 2016). Results from the three methods were further integrated with EvidenceModeler v1.1.1 (Haas et al., 2008). Final gene predictions were obtained by removing transposons and genes containing stop codons with TransposonPSI (Urasaki et al., 2017).

Gene functional annotation

To assess the function of these gene sets, BLASTP v2.7.1 (Altschul et al., 1990) was conducted against the NR (non-redundant protein sequences in NCBI) (v20190401), SwissProt (v 20200709), and KEGG (v20200712) databases with an E-value cut-off of 1e-5. The results were processed to retrieve associated GO terms from the idmapping database.

Multiple genome alignment and phylogenetic construction

We performed whole-genome alignments between the European glass lizard and 14 other genome assemblies (Table S9 in Supporting Information) using LASTZ (https://www.bx.psu.edu/~rsharris/lastz/, parameters: K=2400, L=3000, Y=3000, H=2000, and the HoxD55 scoring matrix). The axtChain program (http://www.soe.ucsc.edu/~kent) was used to build co-linear alignment chains, which were further converted to alignment nets using chainNet (http://www.soe.ucsc.edu/~kent). Low-scoring alignments nets with a span in the reference and query genome of at least 4 kb and a score of 20,000 were removed. Filtered pairwise alignment nets were used as input to construct multiple alignments with Multiz (http://bio.cse.psu.edu/). Four-fold degenerated (4d) sites were extracted from the multiple alignments using the msa_view program in the PHAST package (Hubisz et al., 2011), then concatenated with a Perl script to further construct a phylogenetic tree using IQ-TREE v1.6.8 (Nguyen et al., 2015) (parameters: -alrt 2000 -bb 2000 -nt AUTO -bnni). To estimate divergence times among species, we downloaded six fossil records (Table S10 in Supporting Information) from TimeTree (Kumar et al., 2022) for calibration. The phylogenetic tree was then entered into r8s v1.71 (Sanderson, 2003) for calculation using the Langley-Fitch (LF) method. The calculated calibrated tree was finally inferred using the same fossil calibrations with MCMCTREE program in PAML v4.9i (Yang, 2007) based on the Bayesian relaxed molecular clock approach.

Evolution of conserved non-coding elements

Identification of CNEs

Whole-genome multiple alignments were obtained from the LASTZ pairwise alignments and 4d third codon positions were extracted for phylogenetic tree construction using IQ-TREE v1.6.8 (Nguyen et al., 2015), which served as an input into phyloFit (PhastCons v1.3 package, parameters: -EM -precision HIGH -subst-mod REV) to estimate neutral branch length. PhastCons (Siepel et al., 2005) (parameters: expectedlength=45, target coverage=0.3, rho=0.3) was then performed to identify the evolutionarily conserved elements. Regions overlapping with coding positions were filtered with a custom Perl script and Bedtools v2.25.0 (Quinlan and Hall, 2010). Finally, CNEs of at least 30 bp were retained for subsequent analysis.

Detection of convergent dCNEs

We implemented the "Forward Genomics" branch method (https://github.com/hillerlab/ForwardGenomics) (Hiller et al., 2012) (parameter: -thresholdConserved=0) to detect significantly dCNEs in the limbless pairs compared with other limbed species. We utilized Phylogeny-aware PRANK v150803 alignment (parameters: -keep -showtree -showanc -prunetree -seed=10) to reconstruct the most likely ancestral sequence of each CNE. The local percent identity of sequences between each species and the reconstructed ancestral state was computed using Phytools (Revell, 2012) and a custom Perl script in the R packages. Finally, dCNEs with a P value less than 0.005 were retained for further analysis.

Detection of InDels in CNEs

We defined two types of InDels in CNEs, including segmental insertions and deletions. Specifically, we defined four datasets containing a single limbless clade (i.e., Bourret's blind skink, glass lizard, sand microteiid, and snake lineages), using the tegu genome as the reference (Table S14 in Supporting Information). Considering that total CNE length

was relatively short (shortest CNE was 30 bp long), we defined an InDel as a continuous deleted or inserted sequence in the foreground species that was at least 10 bp long and present in all background species.

Enrichment of CNEs

We used the regulatory domain concept of GREAT (McLean et al., 2010) to associate dCNEs and CNEs with InDels with their potential target genes and to perform further functional enrichment. We defined a basal domain as 5 kb upstream and 1 kb downstream of the transcription start site. A distal domain was also defined as extending from the basal domain to the basal domain of the next gene or a maximum of 300 kb in either direction. Potential genes related to limb development were selected from datasets collected in previous studies (Roscito et al., 2018; Roscito et al., 2022).

To further identify the potential regulatory functions of CNEs, we retrieved the VISTA enhancer (Visel et al., 2007) and tegu lizard ATAC-seq datasets (Roscito et al., 2018). Time-ordered ATAC-seq data from different limb development stages in mice were obtained from a previous study (Davis et al., 2018). All peaks were mapped to the tegu genome or mouse genome (mm10) using the halLiftover program implemented in Progressive Cactus (Armstrong et al., 2020). We screened overlapping CNEs by crossing the position of CNEs with peaks, and finally calculated the proportion of CNEs that overlapped with peaks from different tissues or stages.

Identification of the TFBSs

The human motif library was downloaded from the CisBP database (Weirauch et al., 2014) and used as a reference to perform PoSSuMsearch v2.0 in PoSSuM (Beckstette et al., 2006) to identify motifs in the European glass lizard. All information, including coordinates and motifs, was finally integrated.

Dual-luciferase assays

CNE fragments (Table S23 in Supporting Information) were chemically synthesized and cloned into the pGL4.23 [luc2/minP] vector (Promega, USA). NIH-3T3 cells were cultured in a Dulbecco's Modified Eagle Medium (DMEM) containing 10% fetal bovine serum at 37°C, 5% CO_2, and 100% humidity. Before transfection, cells were washed with Trypsin and seeded into 24-well plates. After detachment, fresh complete medium solution was added for incubation at 37°C. CNE constructs were transfected 24 h after seeding, using Lipofectamine 2000 as the transfection reagent according to the manufacturer's instructions. Briefly, 2 μL of transfection reagent was mixed in 25 μL of basic medium. After sitting for 5 to 20 min at room temperature, 200 ng of the constructed pGL4.23 vectors was co-transfected with 10 ng of pGL4.73 and then co-transfected into NIH-3T3 cells after an incubation period of 20 min at room temperature. Cellular lysates were collected 48 h after transfection using passive lysis buffer. Luciferase activity was measured using the Dual-Luciferase Reporter Assay System (Promega) on a multi-mode microplate reader. The assay for each CNE was repeated three times. Luciferase activity was calculated as the ratio of firefly luciferase activity to sea kidney luciferase activity in each well. Significance was assessed by Student's *t*-test.

Identification of PSGs and REGs

We first identified orthologs in two glass lizards (European glass lizard and Hart's glass lizard) and nine limbed lizards (viviparous lizard, sand lizard, Anan's rock agama, green anole, European green lizard, Schlegel's Japanese gecko, Komodo dragon, Argentine black and white tegu, and common wall lizard) using the reciprocal best hit methodology by following Chen et al. (2019). Briefly, the longest transcript was selected to represent each gene and then translated into proteins. All proteins were aligned using BLASTP v2.7.1 (Altschul et al., 1990) with an *E*-value cutoff of 1×10^{-5} using the European glass lizard as a reference. A total of 17,193 orthologous gene families were finally retrieved (Figure S9 in Supporting Information). Single-copy genes were generated using OrthoFinder v2.2.7 (Emms and Kelly, 2015) and further concatenated with coding sequences (CDS) to construct the phylogenetic tree using IQ-TREE v1.6.8. The codeml program in PAML 4.9i was used to evaluate lineage-specific evolutionary rates based on the topology. The branch-site model was used to detect PSGs, with the most recent common ancestor (MRCA) of two glass lizards set as the foreground. The alternative and null models were conducted to estimate ω (ratio of nonsynonymous to synonymous substitutions), and the χ^2 test was used to calculate the variance between the two models. In addition, potential positively selected sites were assessed using Bayes Empirical Bayes (BEB) analysis. Genes with a threshold *P* value<0.05 and sites with a BEB value>0.95 were considered as PSGs. For identification of REGs, a branch model was performed with the same foreground as for PSGs. The significance of differences between ω values was calculated using the Likelihood Ratio Test (LRT). Genes with an LRT *P* value<0.05 and higher foreground ω were treated as REGs. All PSGs and REGs were combined and subsequently subjected to KEGG and GO enrichment analysis using an in-house Perl script.

Convergent evolution of protein-coding genes

Orthologs were identified using OrthoFinder v2.2.7 in 13

species, including species used in the previous selection analysis and two snakes (common cobra and Burmese python). After alignment and filtering of unaligned sequences using PRANK v.150803 (Löytynoja, 2014) and trimAL v1.4 (Capella-Gutiérrez et al., 2009), 3,279 single-copy orthologous genes were obtained for downstream analysis. All coding sequences were concatenated using an in-house Perl script for phylogenetic relationship analysis using IQ-TREE v1.6.8 (Nguyen et al., 2015). The MRCA of the two snakes and the MRCA of the two glass lizards were defined as convergent branches (foregrounds), since both ancestors are limbless (Mann et al., 2022; Wiens and Slingluff, 2001). The amino acid status of each node was simulated based on the acid frequency and tree topology under the JTT+gamma amino acid substitution model in PAML 4.9i to infer ancestral status. The highest posterior probabilities were considered as the ancestral status and accuracy was estimated by comparing the background to the simulated status. Non-convergent branches (background) were required to share the same ancestral character. Derived sites shared within foregrounds but were different from backgrounds, and the MRCA of all species was finally identified as convergent sites.

Availability of data

Raw sequencing data and genome assemblies were deposited in the China National GeneBank Nucleotide Sequence Archive under accession number CNP0003553. All other data are available in the article and/or supporting information.

Compliance and ethics *The author(s) declare that they have no conflict of interest.*

Acknowledgements *This work was supported by the Strategic Priority Research Program of Chinese Academy of Sciences (XDB31000000), the National Natural Science Foundation of China (32220103004, 32000296), the Second Tibetan Plateau Scientific Expedition and Research Program (2019QZKK0501), the International Partnership Program of Chinese Academy of Sciences (151751KYSB20190024), and the Sichuan Science and Technology Program (2021JDJQ0002). The authors would like to thank Ke Jiang for providing the specimens of Hart's glass lizard and Bourret's blind skink, and Jin-Long Ren for providing the photographs of the lizards and sample processing.*

References

Alonge, M., Wang, X., Benoit, M., Soyk, S., Pereira, L., Zhang, L., Suresh, H., Ramakrishnan, S., Maumus, F., Ciren, D., et al. (2020). Major impacts of widespread structural variation on gene expression and crop improvement in tomato. Cell 182, 145–161.e23.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. J Mol Biol 215, 403–410.

Armstrong, J., Hickey, G., Diekhans, M., Fiddes, I.T., Novak, A.M., Deran, A., Fang, Q., Xie, D., Feng, S., Stiller, J., et al. (2020). Progressive Cactus is a multiple-genome aligner for the thousand-genome era. Nature 587, 246–251.

Beckstette, M., Homann, R., Giegerich, R., and Kurtz, S. (2006). Fast index based algorithms and software for matching position specific scoring matrices. BMC Bioinf 7, 389.

Bejder, L., and Hall, B.K. (2002). Limbs in whales and limblessness in other vertebrates: mechanisms of evolutionary and developmental transformation and loss. Evol Dev 4, 445–458.

Benazet, J.D., Bischofberger, M., Tiecke, E., Goncalves, A., Martin, J.F., Zuniga, A., Naef, F., and Zeller, R. (2009). A self-regulatory system of interlinked signaling feedback loops controls mouse limb patterning. Science 323, 1050–1053.

Benson, G. (1999). Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res 27, 573–580.

Bergmann, P.J., Mann, S.D.W., Morinaga, G., Freitas, E.S., and Siler, C.D. (2020). Convergent evolution of elongate forms in craniates and of locomotion in elongate squamate reptiles. Integr Comp Biol 60, 190–201.

Bimonte, S., De Angelis, A., Quagliata, L., Giusti, F., Tammaro, R., Dallai, R., Ascenzi, M.G., Diez-Roux, G., and Franco, B. (2011). Ofd1 is required in limb bud patterning and endochondral bone development. Dev Biol 349, 179–191.

Birnbaum, R.Y., Clowney, E.J., Agamy, O., Kim, M.J., Zhao, J., Yamanaka, T., Pappalardo, Z., Clarke, S.L., Wenger, A.M., Nguyen, L., et al. (2012). Coding exons function as tissue-specific enhancers of nearby genes. Genome Res 22, 1059–1068.

Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114–2120.

Brandley, M.C., Huelsenbeck, J.P., and Wiens, J.J. (2008). Rates and patterns in the evolution of snake-like body form in squamate reptiles: evidence for repeated re-evolution of lost digits and long-term persistence of intermediate body forms. Evolution 62, 2042–2064.

Burbrink, F.T., Grazziotin, F.G., Pyron, R.A., Cundall, D., Donnellan, S., Irish, F., Keogh, J.S., Kraus, F., Murphy, R.W., Noonan, B., et al. (2020). Interrogating genomic-scale data for Squamata (lizards, snakes, and amphisbaenians) shows no support for key traditional morphological relationships. Syst Biol 69, 502–520.

Burgess, D.J. (2016). Sonic snakes and regulation of limb formation. Nat Rev Genet 17, 715.

Caldwell, M.W. (2003). "Without a leg to stand on": on the evolution and development of axial elongation and limblessness in tetrapods. Can J Earth Sci 40, 573–588.

Camaiti, M., Evans, A.R., Hipsley, C.A., and Chapple, D.G. (2021). A farewell to arms and legs: a review of limb reduction in squamates. Biol Rev 96, 1035–1050.

Capella-Gutiérrez, S., Silla-Martínez, J.M., and Gabaldón, T. (2009). TrimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics 25, 1972–1973.

Chan, Y.F., Marks, M.E., Jones, F.C., Villarreal, G. Jr, Shapiro, M.D., Brady, S.D., Southwick, A.M., Absher, D.M., Grimwood, J., Schmutz, J., et al. (2010). Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a *Pitx1* enhancer. Science 327, 302–305.

Chapman, D.L., Garvey, N., Hancock, S., Alexiou, M., Agulnik, S.I., Gibson-Brown, J.J., Cebra-Thomas, J., Bollag, R.J., Silver, L.M., and Papaioannou, V.E. (1996). Expression of the T-box family genes, *Tbx1–Tbx5*, during early mouse development. Dev Dyn 206, 379–390.

Chen, L., Qiu, Q., Jiang, Y., Wang, K., Lin, Z., Li, Z., Bibi, F., Yang, Y., Wang, J., Nie, W., et al. (2019). Large-scale ruminant genome sequencing provides insights into their evolution and distinct traits. Science 364, 1152.

Cohn, M.J., and Tickle, C. (1999). Developmental basis of limblessness and axial patterning in snakes. Nature 399, 474–479.

Collar, D.C., Quintero, M., Buttler, B., Ward, A.B., and Mehta, R.S. (2016). Body shape transformation along a shared axis of anatomical evolution in labyrinth fishes (Anabantoidei). Evolution 70, 555–567.

Davis, C.A., Hitz, B.C., Sloan, C.A., Chan, E.T., Davidson, J.M., Gabdank, I., Hilton, J.A., Jain, K., Baymuradov, U.K., Narayanan, A.K., et al. (2018). The Encyclopedia of DNA elements (ENCODE): data portal update. Nucleic Acids Res 46, D794–D801.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15–21.
Emms, D.M., and Kelly, S. (2015). OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. Genome Biol 16, 157.
Farlie, P.G., Davidson, N.M., Baker, N.L., Raabus, M., Roeszler, K.N., Hirst, C., Major, A., Mariette, M.M., Lambert, D.M., Oshlack, A., et al. (2017). Co-option of the cardiac transcription factor *Nkx2.5* during development of the emu wing. Nat Commun 8, 132.
Frankel, N., Erezyilmaz, D.F., McGregor, A.P., Wang, S., Payre, F., and Stern, D.L. (2011). Morphological evolution caused by many subtle-effect substitutions in regulatory DNA. Nature 474, 598–603.
Gans, C. (1975). Tetrapod limblessness: evolution and functional corollaries. Am Zool 15, 455–467.
Guan, J., Xu, Y., Yu, Y., Fu, J., Ren, F., Guo, J., Zhao, J., Jiang, Q., Wei, J., and Xie, H. (2021). Genome structure variation analyses of peach reveal population dynamics and a 1.67 Mb causal inversion for fruit shape. Genome Biol 22, 13.
Haas, B.J., Delcher, A.L., Mount, S.M., Wortman, J.R., Smith, R.K. Jr, Hannick, L.I., Maiti, R., Ronning, C.M., Rusch, D.B., Town, C.D., et al. (2003). Improving the *Arabidopsis* genome annotation using maximal transcript alignment assemblies. Nucleic Acids Res 31, 5654–5666.
Haas, B.J., Salzberg, S.L., Zhu, W., Pertea, M., Allen, J.E., Orvis, J., White, O., Buell, C.R., and Wortman, J.R. (2008). Automated eukaryotic gene structure annotation using EVidenceModeler and the program to assemble spliced alignments. Genome Biol 9, R7.
Haro, E., Delgado, I., Junco, M., Yamada, Y., Mansouri, A., Oberg, K.C., and Ros, M.A. (2014). Sp6 and Sp8 transcription factors control AER formation and dorsal-ventral patterning in limb development. PLoS Genet 10, e1004468.
He, Y., Luo, X., Zhou, B., Hu, T., Meng, X., Audano, P.A., Kronenberg, Z. N., Eichler, E.E., Jin, J., Guo, Y., et al. (2019). Long-read assembly of the Chinese rhesus macaque genome and identification of ape-specific structural variants. Nat Commun 10, 4233.
Hiller, M., Schaar, B.T., Indjeian, V.B., Kingsley, D.M., Hagey, L.R., and Bejerano, G. (2012). A "Forward Genomics" approach links genotype to phenotype using Independent phenotypic losses among related species. Cell Rep 2, 817–823.
Hirsch, N., Eshel, R., Bar Yaacov, R., Shahar, T., Shmulevich, F., Dahan, I., Levaot, N., Kaplan, T., Lupianez, D.G., and Birnbaum, R.Y. (2018). Unraveling the transcriptional regulation of TWIST1 in limb development. PLoS Genet 14, e1007738.
Hu, J., Fan, J., Sun, Z., and Liu, S. (2020). NextPolish: a fast and efficient genome polishing tool for long-read assembly. Bioinformatics 36, 2253–2255.
Hubisz, M.J., Pollard, K.S., and Siepel, A. (2011). PHAST and RPHAST: phylogenetic analysis with space/time models. Briefings Bioinf 12, 41–51.
Hugi, J., Hutchinson, M.N., Koyabu, D., and Sánchez-Villagra, M.R. (2012). Heterochronic shifts in the ossification sequences of surface- and subsurface-dwelling skinks are correlated with the degree of limb reduction. Zoology 115, 188–198.
Infante, C.R., Mihala, A.G., Park, S., Wang, J.S., Johnson, K.K., Lauderdale, J.D., and Menke, D.B. (2015). Shared enhancer activity in the limbs and phallus and functional divergence of a limb-genital *cis*-regulatory element in snakes. Dev Cell 35, 107–119.
Infante, C.R., Park, S., Mihala, A.G., Kingsley, D.M., and Menke, D.B. (2013). *Pitx1* broadly associates with limb enhancers and is enriched on hindlimb *cis*-regulatory elements. Dev Biol 374, 234–244.
Infante, C.R., Rasys, A.M., and Menke, D.B. (2018). Appendages and gene regulatory networks: lessons from the limbless. Genesis 56, e23078.
Jeong, S., Rebeiz, M., Andolfatto, P., Werner, T., True, J., and Carroll, S.B. (2008). The evolution of gene regulation underlies a morphological difference between two *Drosophila* sister species. Cell 132, 783–793.
Jurka, J. (1998). Repeats in genomic DNA: mining and meaning. Curr Opin Struct Biol 8, 333–337.
Kajitani, R., Toshimoto, K., Noguchi, H., Toyoda, A., Ogura, Y., Okuno, M., Yabana, M., Harada, M., Nagayasu, E., Maruyama, H., et al. (2014). Efficient *de novo* assembly of highly heterozygous genomes from whole-genome shotgun short reads. Genome Res 24, 1384–1395.
Keilwagen, J., Wenk, M., Erickson, J.L., Schattat, M.H., Grau, J., and Hartung, F. (2016). Using intron position conservation for homology-based gene prediction. Nucleic Acids Res 44, e89.
Kumar, S., Suleski, M., Craig, J.M., Kasprowicz, A.E., Sanderford, M., Li, M., Stecher, G., and Hedges, S.B. (2022). TimeTree 5: an expanded resource for species divergence times. Mol Biol Evol 39, msac174.
Kvon, E.Z., Kamneva, O.K., Melo, U.S., Barozzi, I., Osterwalder, M., Mannion, B.J., Tissières, V., Pickle, C.S., Plajzer-Frick, I., Lee, E.A., et al. (2016). Progressive loss of function in a limb enhancer during snake evolution. Cell 167, 633–642.e11.
Laufer, E., Nelson, C.E., Johnson, R.L., Morgan, B.A., and Tabin, C. (1994). Sonic hedgehog and Fgf-4 act through a signaling cascade and feedback loop to integrate growth and patterning of the developing limb bud. Cell 79, 993–1003.
Law, C.J., Slater, G.J., and Mehta, R.S. (2019). Shared extremes by ectotherms and endotherms: Body elongation in mustelids is associated with small size and reduced limbs. Evolution 73, 735–749.
Leal, F., and Cohn, M.J. (2016). Loss and re-emergence of legs in snakes by modular evolution of sonic hedgehog and *HOXD* enhancers. Curr Biol 26, 2966–2973.
Leal, F., and Cohn, M.J. (2018). Developmental, genetic, and genomic insights into the evolutionary loss of limbs in snakes. Genesis 56, e23077.
Lettice, L.A., Devenney, P., De Angelis, C., and Hill, R.E. (2017). The conserved sonic hedgehog limb enhancer consists of discrete functional elements that regulate precise spatial expression. Cell Rep 20, 1396–1408.
Li, R., Gong, M., Zhang, X., Wang, F., Liu, Z., Zhang, L., Xu, M., Zhang, Y., Dai, X., Zhang, Z., et al. (2021). The first sheep graph-based pan-genome reveals the spectrum of structural variations and their effects on tail phenotypes. bioRxiv, doi: 10.1101/2021.12.22.472709.
Lopez-Rios, J., Duchesne, A., Speziale, D., Andrey, G., Peterson, K.A., Germann, P., Ünal, E., Liu, J., Floriot, S., Barbey, S., et al. (2014). Attenuated sensing of SHH by *Ptch1* underlies evolution of bovine limbs. Nature 511, 46–51.
Löytynoja, A. (2014). Phylogeny-aware alignment with PRANK. In: Russell, D., ed. Multiple Sequence Alignment Methods. Methods in Molecular Biology. Totowa: Humana Press. 155–170.
Malkmus, J., Ramos Martins, L., Jhanwar, S., Kircher, B., Palacio, V., Sheth, R., Leal, F., Duchesne, A., Lopez-Rios, J., Peterson, K.A., et al. (2021). Spatial regulation by multiple *Gremlin1* enhancers provides digit development with *cis*-regulatory robustness and evolutionary plasticity. Nat Commun 12, 5557.
Mann, A., Pardo, J.D., and Maddin, H.C. (2022). Snake-like limb loss in a Carboniferous amniote. Nat Ecol Evol 6, 614–621.
McLean, C.Y., Bristor, D., Hiller, M., Clarke, S.L., Schaar, B.T., Lowe, C. B., Wenger, A.M., and Bejerano, G. (2010). GREAT improves functional interpretation of *cis*-regulatory regions. Nat Biotechnol 28, 495–501.
McLean, C.Y., Reno, P.L., Pollen, A.A., Bassan, A.I., Capellini, T.D., Guenther, C., Indjeian, V.B., Lim, X., Menke, D.B., Schaar, B.T., et al. (2011). Human-specific loss of regulatory DNA and the evolution of human-specific traits. Nature 471, 216–219.
Mehta, R.S., Ward, A.B., Alfaro, M.E., and Wainwright, P.C. (2010). Elongation of the body in eels. Integr Comp Biol 50, 1091–1105.
Newton, A.H., and Smith, C.A. (2021). Regulation of vertebrate forelimb development and wing reduction in the flightless emu. Dev Dyn 250, 1248–1263.
Nguyen, L.T., Schmidt, H.A., von Haeseler, A., and Minh, B.Q. (2015). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol 32, 268–274.
Niswander, L., Jeffrey, S., Martin, G.R., and Tickle, C. (1994). A positive feedback loop coordinates growth and patterning in the vertebrate limb.

Nature 371, 609–612.
Ohuchi, H., Nakagawa, T., Yamamoto, A., Araga, A., Ohata, T., Ishimaru, Y., Yoshioka, H., Kuwana, T., Nohno, T., Yamasaki, M., et al. (1997). The mesenchymal factor, FGF10, initiates and maintains the outgrowth of the chick limb bud through interaction with FGF8, an apical ectodermal factor. Development 124, 2235–2244.
Osterwalder, M., Barozzi, I., Tissières, V., Fukuda-Yuzawa, Y., Mannion, B.J., Afzal, S.Y., Lee, E.A., Zhu, Y., Plajzer-Frick, I., Pickle, C.S., et al. (2018). Enhancer redundancy provides phenotypic robustness in mammalian development. Nature 554, 239–243.
Parra-Olea, G., and Wake, D.B. (2001). Extreme morphological and ecological homoplasy in tropical salamanders. Proc Natl Acad Sci USA 98, 7888–7891.
Pertea, M., Pertea, G.M., Antonescu, C.M., Chang, T.C., Mendell, J.T., and Salzberg, S.L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nat Biotechnol 33, 290–295.
Petit, F., Sears, K.E., and Ahituv, N. (2017). Limb development: a paradigm of gene regulation. Nat Rev Genet 18, 245–258.
Quah, E.S.H., Sah, S.A.M., Grismer, L.L., and Grassby-Lewis, R. (2017). A new species of *Dibamus* Duméril & Bibron 1839 (Squamata: Dibamidae) from a hill station in Peninsular Malaysia. Raffles Bull Zool 65, 681–690.
Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26, 841–842.
Rahmanl, T.Z. (1974). Morphogenesis of the rudimentary hind-limb of the glass snake (*Ophisaurus apodus* Pallas). J Embryol Exp Morphol 32, 431–443.
Raynaud, A. (1985). Development of limbs and embryonic limb reduction. Biol Reptilia 15, 59–148.
Raynaud, A. (1990). Developmental mechanism involved in the embryonic reduction of limbs in reptiles. Int J Dev Biol 34, 233–243.
Revell, L.J. (2012). phytools: an R package for phylogenetic comparative biology (and other things). Methods Ecol Evol 3, 217–223.
Roscito, J.G., Sameith, K., Parra, G., Langer, B.E., Petzold, A., Moebius, C., Bickle, M., Rodrigues, M.T., and Hiller, M. (2018). Phenotype loss is associated with widespread divergence of the gene regulatory landscape in evolution. Nat Commun 9, 4737.
Roscito, J.G., Sameith, K., Kirilenko, B.M., Hecker, N., Winkler, S., Dahl, A., Rodrigues, M.T., and Hiller, M. (2022). Convergent and lineage-specific genomic differences in limb regulatory elements in limbless reptile lineages. Cell Rep 38, 110280.
Royle, S.R., Tabin, C.J., and Young, J.J. (2021). Limb positioning and initiation: an evolutionary context of pattern and formation. Dev Dyn 250, 1264–1279.
Ruan, J., and Li, H. (2020). Fast and accurate long-read assembly with wtdbg2. Nat Methods 17, 155–158.
Sanderson, M.J. (2003). R8s: inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. Bioinformatics 19, 301–302.
Schneider, I., and Shubin, N.H. (2013). The origin of the tetrapod limb: from expeditions to enhancers. Trends Genet 29, 419–426.
Sharpe, S.S., Koehler, S.A., Kuckuk, R.M., Serrano, M., Vela, P.A., Mendelson Iii, J., and Goldman, D.I. (2015). Locomotor benefits of being a slender and slick sand-swimmer. J Exp Biol 218, 1111.
Siepel, A., Bejerano, G., Pedersen, J.S., Hinrichs, A.S., Hou, M., Rosenbloom, K., Clawson, H., Spieth, J., Hillier, L.D.W., Richards, S., et al. (2005). Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. Genome Res 15, 1034–1050.
Simão, F.A., Waterhouse, R.M., Ioannidis, P., Kriventseva, E.V., and Zdobnov, E.M. (2015). BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics 31, 3210–3212.
Smith, C.A., Farlie, P.G., Davidson, N.M., Roeszler, K.N., Hirst, C., Oshlack, A., and Lambert, D.M. (2016). Limb patterning genes and heterochronic development of the emu wing bud. Evodevo 7, 26.
Stanke, M., Schöffmann, O., Morgenstern, B., and Waack, S. (2006). Gene prediction in eukaryotes with a generalized hidden Markov model that uses hints from external sources. BMC Bioinf 7, 62.
te Welscher, P., Zuniga, A., Kuijper, S., Drenth, T., Goedemans, H.J., Meijlink, F., and Zeller, R. (2002). Progression of vertebrate limb development through SHH-mediated counteraction of GLI3. Science 298, 827–830.
Thewissen, J.G.M., Cohn, M.J., Stevens, L.S., Bajpai, S., Heyning, J., and Horton W.~E., J. (2006). Developmental basis for hind-limb loss in dolphins and origin of the cetacean bodyplan. Proc Natl Acad Sci USA 103, 8414–8418.
Tóth-Petróczy, Á., and Tawfik, D.S. (2013). Protein insertions and deletions enabled by neutral roaming in sequence space. Mol Biol Evol 30, 761–771.
Uetz, P., Freed, P, Aguilar, R., Reyes, F., and Hošek, J. (2022). The Reptile Database. Available from URL: http://www.reptile-database.org.
Urasaki, N., Takagi, H., Natsume, S., Uemura, A., Taniai, N., Miyagi, N., Fukushima, M., Suzuki, S., Tarora, K., Tamaki, M., et al. (2017). Draft genome sequence of bitter gourd (*Momordica charantia*), a vegetable and medicinal plant in tropical and subtropical regions. DNA Res 24, 51–58.
Urošević, A., Slijepčević, M.D., Arntzen, J.W., and Ivanović, A. (2016). Vertebral shape and body elongation in Triturus newts. Zoology 119, 439–446.
VanderMeer, J.E., Smith, R.P., Jones, S.L., and Ahituv, N. (2014). Genome-wide identification of signaling center enhancers in the developing limb. Development 141, 4194–4198.
Visel, A., Minovitsky, S., Dubchak, I., and Pennacchio, L.A. (2007). VISTA Enhancer Browser—a database of tissue-specific human enhancers. Nucleic Acids Res 35, D88–D92.
Wang, G.D., Shao, X.J., Bai, B., Wang, J., Wang, X., Cao, X., Liu, Y.H., Wang, X., Yin, T.T., Zhang, S.J., et al. (2019). Structural variation during dog domestication: insights from gray wolf and dhole genomes. Natl Sci Rev 6, 110–122.
Ward, A.B., and Brainerd, E.L. (2007). Evolution of axial patterning in elongate fishes. Biol J Linnean Soc 90, 97–116.
Weirauch, M.T., Yang, A., Albu, M., Cote, A.G., Montenegro-Montero, A., Drewe, P., Najafabadi, H.S., Lambert, S.A., Mann, I., Cook, K., et al. (2014). Determination and inference of eukaryotic transcription factor sequence specificity. Cell 158, 1431–1443.
Wiens, J.J., and Slingluff, J.L. (2001). How lizards turn into snakes: a phylogenetic analysis of body-form evolution in anguid lizards. Evolution 55, 2303–2318.
Wiens, J.J., Brandley, M.C., and Reeder, T.W. (2006). Why does a trait evolve multiple times within a clade? Repeated evolution of snakeline body form in squamate reptiles. Evolution 60, 123–141.
Wilson, D.G., Phamluong, K., Lin, W.Y., Barck, K., Carano, R.A.D., Diehl, L., Peterson, A.S., Martin, F., and Solloway, M.J. (2012). Chondroitin sulfate synthase 1 (*Chsy1*) is required for bone development and digit patterning. Dev Biol 363, 413–425.
Wittkopp, P.J., and Kalay, G. (2011). *Cis*-regulatory elements: molecular mechanisms and evolutionary processes underlying divergence. Nat Rev Genet 13, 59–69.
Xu, S., He, Z., Guo, Z., Zhang, Z., Wyckoff, G.J., Greenberg, A., Wu, C.I., and Shi, S. (2017). Genome-wide convergence during evolution of mangroves from woody plants. Mol Biol Evol 34, 1008–1015.
Xu, Z., and Wang, H. (2007). LTR_FINDER: an efficient tool for the prediction of full-length LTR retrotransposons. Nucleic Acids Res 35, W265–W268.
Yang, J., Mani, S.A., Donaher, J.L., Ramaswamy, S., Itzykson, R.A., Come, C., Savagner, P., Gitelman, I., Richardson, A., and Weinberg, R.A. (2004). Twist, a master regulator of morphogenesis, plays an essential role in tumor metastasis. Cell 117, 927–939.
Yang, N., Liu, J., Gao, Q., Gui, S., Chen, L., Yang, L., Huang, J., Deng, T., Luo, J., He, L., et al. (2019). Genome assembly of a tropical maize inbred line provides insights into structural variation and crop improvement. Nat Genet 51, 1052–1059.
Yang, Z. (2007). PAML 4: phylogenetic analysis by maximum likelihood. Mol Biol Evol 24, 1586–1591.

Ye, C., Hill, C.M., Wu, S., Ruan, J., and Ma, Z.S. (2016). DBG2OLC: efficient assembly of large genomes using long erroneous reads of the third generation sequencing technologies. Sci Rep 6, 31900.
Young, J.J., Grayson, P., Edwards, S.V., and Tabin, C.J. (2019). Attenuated Fgf signaling underlies the forelimb heterochrony in the emu *Dromaius novaehollandiae*. Curr Biol 29, 3681–3691.e5.
Zhang, Z.Y., Lv, Y., Wu, W., Yan, C., Tang, C.Y., Peng, C., and Li, J.T. (2022). The structural and functional divergence of a neglected three-finger toxin subfamily in lethal elapids. Cell Rep 40, 111079.
Zuniga, A. (2015). Next generation limb development and evolution: old questions, new perspectives. Development 142, 3810–3820.
Zuniga, A., and Zeller, R. (2020). Dynamic and self-regulatory interactions among gene regulatory networks control vertebrate limb bud morphogenesis. Curr Top Dev Biol 139, 61–88.

SUPPORTING INFORMATION

The supporting information is available online at https://doi.org/10.1007/s11427-023-2362-5. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.