Nature Communications

https://doi.org/10.1038/s41467-026-73821-7

Article in Press

Evolution of limb and digit identity genes since the tetrapod ancestor

Received: 28 November 2025

Accepted: 19 May 2026

Cite this article as: Kang, W., Fei, J.-F., Liang, C. *et al.* Evolution of limb and digit identity genes since the tetrapod ancestor. *Nat Commun* (2026). https://doi.org/10.1038/s41467-026-73821-7

Wen Kang, Ji-Feng Fei, Cong Liang, Günter Wagner & Qi Zhou

Evolution of limb and digit identity genes since the tetrapod ancestor

Wen Kang[1,2], Ji-Feng Fei[3,4], Cong Liang[5], Günter Wagner[6,7,8,9,*], Qi Zhou[1,2,10,11,*]

1. Center for Reproductive Medicine, Second Affiliated Hospital of Zhejiang University School of Medicine, Life Sciences Institute, Zhejiang University, Hangzhou 310058, Zhejiang, China

2. Center for Evolutionary & Organismal Biology, Zhejiang University, Hangzhou 310058, China

3. Department of Pathology, Guangdong Provincial People's Hospital (Guangdong Academy of Medical Sciences), Southern Medical University, Guangzhou, Guangdong 510080, China

4. School of Medicine, South China University of Technology, Guangzhou, Guangdong 510006, China

5. Center for Applied Mathematics, Tianjin University, Tianjin 300072, China

6. Department of Ecology and Evolutionary Biology, Yale University, New Haven, CT, USA

7. Yale Systems Biology Institute, Yale University, New Haven, CT, USA

8. Department of Evolutionary Biology, University of Vienna, Vienna, Austria

9. Hagler Institute for Advanced Study, Texas A&M University, College Station, TX, USA

10. Zhejiang Provincial Key Laboratory for Cancer Molecular Cell Biology, Life Sciences Institute, Zhejiang University, Hangzhou 310058, Zhejiang, China

11. State Key Laboratory of Transvascular Implantation Devices, Hangzhou, China

* corresponding authors: zhouqi1982@zju.edu.cn or gunter.wagner@yale.edu

Abstract

Natural selection has sculpted tetrapod limbs and digits into a tremendous morphological diversity, but the underlying developmental mechanism remains unclear. This leads to an enduring debate concerning the identity of the three digits of bird wing. Here we use comparative digit transcriptomics across six tetrapod species to examine gene expression signatures of digit evolution. Here we show that the avian wing digits are homologous to reptilian forelimb digits 1,3,4. We find that birds uniquely increase fore-hindlimb transcriptome divergence during development, relaxing constraints and facilitating wing and foot digit specialization. We identify many limb- or digit- identity associated genes that have a biased expression between fore- or hind-limb, or toward specific digits. There are more limb and digit biased genes among the avian lineages than in other tetrapod lineages. Our work reveals ancient regulatory networks that maintain limb and digit molecular identity through persistent differential expression. We also show that this conserved network does not preclude independent digit evolution, especially when developmental constraints are reduced.

Introduction

Crown tetrapods have evolved remarkable interspecific variations in limbs, particularly in the number and morphology of hand or foot digits from a pentadactyl ancestor dated about 350 million years ago[1]. This comprises one of the most prominent examples for adaptive evolution, by which an appendage is sculpted for different locomotive (e.g., running, flying and swimming) and prehensile functions. Hence the evolutionary and developmental mechanisms of limb and digit patterning are at the forefront of interdisciplinary studies, which in turn generate many enduring controversies among paleontologists, evolutionary-, and developmental biologists[2]. Most of these controversies reflect the prevalent difficulty or the inconsistency in establishing homology of traits across species (e.g., fish fin and teleost limb[3]) using different (e.g., morphological vs. gene expression) criteria[4]. One classic debate that has lasted for nearly 200 years[5, 6, 7, 8, 9, 10] concerns the identities of avian wing digits. Extant adult birds have between one and three wing digits, and between two and four foot digits. Traceable reduction of forelimb digit numbers from five to three in fossils from the more distant to their closer avian relatives provides a key evidence for the dinosaur origin of birds[11]. It also supports a parsimonious scenario that the avian ancestor might have lost the digit 4 (ring finger) and 5 (pinky), i.e., with an identity of 1,2,3 of the remaining digits during the limb-to-wing transition. However, embryological and transcriptome data of chicken embryonic digits suggest an identity of either 2,3,4[7, 12, 13] or 1,3,4[6, 11, 14] respectively. The outcome of the ongoing debate about interpretations of digit identity from different biological disciplines not only has broad implications for inferring homology of many other serial homologous body parts (e.g., vertebrae) but provide a testable paradigm to advance our understanding of the concept of homology.

Two mutually non-exclusive hypotheses have been proposed to reconcile the issue, and both acknowledge the most anterior digit (MAD) has a distinguished expression profile from others during its outgrowth. The difference is that the MAD hypothesis emphasizes the distinctive expression pattern of 5'-Hox genes at the MAD, i.e., activation of *Hoxa13*, and reduced expression of *Hoxd12*, regardless of its positional identity. It even questions the necessity of distinguishing between non-MAD digits or numbering them with different identities[15]. This seems to be supported by a recent work that found almost no transcriptional factor that is differentially expressed among non-MAD digit positions and that is shared among

all tetrapods[14]. While the 'frameshift' hypothesis acknowledges the existence of digit identity and does not focus on specific genes. It assumes the digit identity is not rigidly linked to a condensation position ('positional identity') specified in the limb bud but that in birds the wing digit 1 with a distinguished expression profile from others during its outgrowth ('outgrowth identity 1', D1) grows at the condensation position of digit 2 (positional identity 2, C2)[16]. Both hypotheses were recently challenged by a broad examination across 18 amniotes of three posterior *Hoxd* genes' expression in their digits[17] that showed many species-specific variations of expression in the most anterior digit; although some cases of variations could reflect species-specific degeneration (e.g., the emu wing) or adaptation (e.g., the mole limbs[18]). Another challenge specifically to the MAD hypothesis is that in the silkie mutant with polydactyl, the foot has a digit with the morphology and gene expression of digit D1 but is not the most anterior digit[19], showing that digit position does not necessarily determine the developmental digit identity.

The mechanism of such evolutionary lability of digit identity can be understood with respect to sequential phases of limb development[20, 21, 22]. Natural selection can target genes that function at any phase of development and convergently impact the eventual number and morphology of digits. For example, digit loss in artiodactyl mammals and jerboa are caused by mutations either tinkering with the expression of Sonic hedgehog (*Shh*) receptor PTCH1 in the early stage of limb bud[23], or those causing apoptosis of digit cells at a much later stage[24]. The classic morphogen model, proposed by Wolpert back to the 1960s, proposes that *Shh* from the Zone of Polarizing Activity at the posterior part of limb bud instructs the positional identity of digits along the anterior-posterior (AP) axis[25]. However, recent studies suggested the requirement for *Shh* for digit specification is very time-restricted and involves a downstream relay signal that is still not known[26]. On the other hand, studies in chicken hindlimb showed that each digit is characterized with a unique *SMAD1/5/8* activity specified by the BMP signaling at each respective posterior interdigit tissue[27]. In short, models for specifying the early digit positional identity remain controversial[28]. What generates the molecular fingerprint of the late outgrowth identity of each digit also remains unclear.

In this work, we collected and compared transcriptomes of individual developing digits in hindlimbs and forelimbs across six representative species spanning the tetrapod phylogeny (**Fig.**

1a), to elucidate the development and evolution mechanisms underlying the interspecific variation of digits. Previous works used the four chicken hindlimb digits as reference to infer the outgrowth identities of three forelimb digits[6, 14]. A better solution is to align the avian digits of respective limbs against those of their pentadactyl outgroup. Moreover, little attention has been paid to the molecular and functional disparity between the fore- and hind-limbs (FL-HL), despite it has been known for long that they are initiated by different (*Tbx5* or *Tbx4*) transcription factors[29]. Such disparity is expected to be greatly amplified in birds and humans, the only two groups of animals that independently evolved obligate bipedalism (compared to facultative bipedalism in some lizards and other primates)[30]. In this work, besides revisiting the two hypotheses about avian digit identity[15, 16], we propose and test a new hypothesis here that obligate bipedalism has contributed to overcoming ancient FL-HL developmental constraints and facilitated the functional specialization of avian digits. Besides the varied number of digits of both wings and feet, birds exhibit a great diversity of digit configurations in the foot. For example, toe orientation (e.g., anisodactyl of most passerines vs. zygodactyl of parrots) has been proposed to have played an important role during the avian speciation[31, 32, 33]. With the newly identified molecular fingerprint genes of each individual digit of each species, we reconstruct the stepwise integration of different genes into the regulatory network specifying the digit outgrowth identity during tetrapod evolution.

Results

Digit transcriptome and molecular fingerprint genes

We choose to study three birds (chicken, ostrich, emu), two reptiles (Siamese crocodile and Chinese softshell turtle), and one amphibian, the axolotl, taking both the phylogenetic breadth and digit diversity into account (**Fig. 1a**). To study development and molecular evolution of these species' digits, we collected a total of 112 transcriptomes of limb bud and individual developing digits. Samples were collected in replicates with high repeatability (Pearson's correlation, $P<0.001$, $R>0.9$, **Supplementary Fig. 2**), separately from forelimbs and hindlimbs spanning three developmental stages, i.e., Hamburger-Hamilton stage 21 (HH21, limb bud), HH28 and HH32 (digits) of three birds, and corresponding stages of other species (**Supplementary Data 1**) according to the limb morphology. All studied species but axolotl (**Supplementary Note**) do not exhibit pronounced heterochrony between FL-HL morphology in

development[34]. Each digit was dissected with its posterior interdigital tissue, i.e., the digit-specific outgrowth identity signaling molecules[35]. The digit samples are named as C1 to C5 from anterior to posterior reflecting their positional identity. The avian wing digits hence were labeled as C2,3,4[5] (**Fig. 1a**).

This large transcriptome dataset allows us to define the molecular fingerprint genes (MFG) as those genes that are preferentially upregulated or downregulated in a certain digit relative to others in either hindlimb or forelimb in at least one of two sampled stages, based on the distribution of expression breadth matrices of all genes across all digit samples (**Materials and Methods**). The MFGs defined here are used to be cross examined with the previously reported characteristic genes for each digit. More importantly, they enable later comparison of individual digit to understand the specific evolutionary changes of likely different functional categories across stage and species. Indeed, the differential expression patterns of these genes are consistent with their reported RNA *in situ* hybridization results among mouse or chicken digits (**Supplementary Data 2**). Particularly, identification of MFGs in most digits indicates existence of their molecular identity other than the MAD (see below). While the other global analyses, including the unsupervised clustering of these digit transcriptomes and the total numbers of differentially expressed genes (DEGs) between any of the two digits are used for establishing the homologous relationship and comparing the overall digit divergence level between different species, as performed in[14].

By including published data[14] of two other species, the green-anole lizard, and the mouse, in the subsequent analyses, this work comprises the largest limb/digit comparative transcriptome dataset known to date. Overall, birds have a significantly (Proportion test, $P < 0.05$) higher digit individuality, measured by normalised number of differentially expressed genes (DEGs) between any of the two digits (**Supplementary Fig. 3, Supplementary Data 3**), than other studied tetrapods. This provides evidence for greater developmental individuation and possibly functional specialization of avian digits. Among the studied non-avian reptiles, the anole lizard was shown to have an extremely low level of digit individuality[14]. This is probably a derived state, because we found in this work crocodile, turtle and mouse have a similarly higher level of digit individuality than the anole (**Supplementary Fig. 3**). Both turtle and crocodile are pentadactyl from embryo to adult except for the crocodile HD5 (**Fig. 1a**), and they are known to

have a much slower genome evolution rate than other vertebrates[36]. This makes them a more appropriate reference than other tetrapods to be compared with birds.

Evidence for the frameshift and MAD hypotheses about the avian digit identity

To resolve the digit identity of birds, we aligned the transcriptome of individual digit of birds against those of turtle and crocodile by hierarchical cluster analysis (HCA) of limb-related genes (**Materials and Methods, Supplementary Fig. 4**). Under the frameshift hypothesis[14], we expected that the avian C2 wing digits to be clustered with the crocodile/turtle C1/D1 (in pentadactyl species, C1 is equal to D1 and so on) digits, or other mismatches in any digit between its positional and outgrowth identities. Interestingly, while we confirm frameshift of wing digits, we unexpectedly found evidence of frameshift also in foot digits: the C2 digits of all three bird species, except for the ostrich lacking the C2 foot digit, are clustered with the C1 forelimb and hindlimb digit of crocodile/turtle, supporting that in both avian wings and feet, a digit with a D1 outgrowth identity grows from a C2 position (**Fig. 1b**). PCA analyses between forelimb and hindlimb digit transcriptomes within each bird showed consistent clusterings of forelimb C2 with hindlimb C1[6, 14] (**Supplementary Fig. 5**). In avian feet, the clustering of both C1 and C2 digits with a non-avian reptilian D1 suggests a possible two-step model of how the frameshift could have occurred: the digit at the C2 position first substituted its outgrowth identity with D1, followed by the loss of digit at the C1 position. Most other wing digits C3 and C4 of three birds, except for the ostrich wing digit C4, do not show frameshifts and are clustered respectively with the reptilian D3 and D4. These results together support that of Stewart et al. 2019[14], and an avian wing digit identity of 1,3,4, as well as the hypothesis of Xu and Mackam[11]. We discovered here other new cases of frameshift, they include the ostrich foot C5 and axolotl foot C5 clustered with the reptilian D4, and the emu foot C3 and C4 digits clustered respectively with reptilian D4 and D5. Unlike MAD, few genes have been previously reported to have specific or biased expression patterns in other digits. These newly inferred cases of frameshift require future tests of the newly identified genes (e.g., *Hoxc12*, *Bmp3*) shared between digits of different positions between different species (**Supplementary Data 4**).

A second expectation from the frameshift hypothesis is that digits of one positional identity would be characterized with MFGs indicating another outgrowth identity. Indeed, MFGs of most frameshifted digits (e.g., avian wing digits C2) have the highest overlapping percentage

with the orthologous MFGs of another digit in the non-avian reptiles (e.g., FD1) (**Fig. 1c, Supplementary Fig. 6**), except for certain digits influenced by lineage-specific degeneration (e.g., ostrich foot C5, **Fig. 1a, Supplementary Note**). Specifically, the frameshift hypothesis predicts the C2 digits of avian wings to exhibit characteristic expression patterns of the classic 'MADness' genes[15, 19], i.e., absence of the posterior Hox genes *Hoxd10-12*. Indeed, these genes show a conserved downregulation in the D1 relative to the other digits across almost all studied tetrapods except for axolotl, in the mouse C1 but in the avian wing C2 digits (**Fig. 1d, Supplementary Fig. 7**). Another gene, *Hand2*, shows a similarly conserved low expression pattern in the D1. It is an upstream positive regulator of *Shh* [37], and seems to have persisted its polarized expression toward the posterior end from limb bud to digit stages of both forelimbs and hindlimbs, forming an expression gradient across digits of most studied species except for ostrich (**Fig. 1d, Supplementary Fig. 7**). After taking species specific variation into account, we identified many other genes that have a conserved pattern of upregulation in the D1 relative to other digits, in either early or late digit stage, as new candidates of D1 identity genes. They include *Alx1/4*, *Dlx5/6*, *Lhx9*, *Msx1*, *Pax9*, *Zic3*, etc.. Contrary to the specific downregulation of *Hand2* the anterior digit, all of them were reported to have an anterior specific or biased expression in the limb bud[38, 39, 40, 41] (see below). Mouse mutants of many of these genes show specifically in the thumb or anterior digits phenotypes of polydactyly or digit loss (e.g., *Alx4*[42], *Msx1*[43], *Pax9*[44], *Dlx5/6*[45]). As for the overlapped MFGs of D3 and D4 of birds, despite the limited overlapping size between birds and non-avian reptiles (**Supplementary Fig.6**), we provide the full list of those shared MFGs (**Supplementary Data 7**, **Supplementary Fig.8**). Among these, several candidates are noteworthy: for example, *Smoc1*, which is shared in the D3, has been reported to associate with syndactyly affecting the middle digits in mice[46], and *Shh*, identified in the D4 comparison, has well-established roles in posterior digit identity[26]. These examples of genes offer preliminary molecular support for the proposed digit homology. One caveat is that there are only 2 to 80 orthologous MFGs that can be compared between birds and reptiles, a much lower number than the 4741 genes that are involved in the above clustering and PCA analyses. Therefore, we only consider the pattern of MFG here as side evidence for the frameshift hypothesis, similarly for the MAD hypothesis below.

In addition to the posterior *Hoxd* genes, conserved expression of other D1 identity genes (**Fig. 1d, Supplementary Fig. 7**) seems to support the MAD hypothesis in most studied species except for axolotl (**Supplementary Note**). Among amniotes, if we exclude species with a degenerating D1 (**Fig. 1a**) given its unusual expression pattern (see below), the MAD hypothesis is only supported in chicken, crocodile and mouse. These species show the highest numbers of MFGs of D1 compared to those of other digits (**Supplementary Fig. 3**), suggesting a 'stronger' identity of D1. However, other digits clearly have their respective MFGs that are highly conserved across tetrapods (see below). In addition, hierarchical cluster analysis of digit transcriptome within species shows a mixed picture among species, with D1 being clustered separately from all other digits only in forelimbs of two birds and mouse but not in other limb samples (**Supplementary Fig. 9**). Overall, the MAD hypothesis that the characteristic expression pattern of posterior Hoxd and their interacting genes in the D1 shaping its unique morphology, is supported in some but not all examined tetrapods in our results.

Evolution of transcriptome disparity between forelimbs and hindlimbs in birds

With the established identity and homology of individual avian digits vs. those of reptiles, we are able to compare each corresponding digit between FL-HL, within and between species. In the limb bud stage, the transcriptomes of all three bird species evolve significantly (z-test, $P < 0.05$) faster than the crocodile, with turtle as an outgroup (**Fig. 2a, Supplementary Fig. 10a**). In the digit stages, all birds have a significantly (Proportion test, $P < 0.05$) higher total numbers of FL-HL DEGs (Wald test, adjusted $P < 0.05$, Fig.9Fold-change>2) between all corresponding digits than the two non-avian reptiles (**Fig. 2a**). This is particularly pronounced in the D1: compared to two other reptiles, all three birds show a significant (Fisher's exact test, adjusted $P < 0.01$) decrease of shared MFGs between FL-HL, but a significant increase of wing or feet specific MFGs (**Supplementary Fig. 11**). Overall, there is a 3 to 9 fold increase in the total number of non-redundant MFGs of D1,3,4 of forelimbs and hindlimbs in birds (D2/5 are not counted because they are absent in most birds and not comparable) than in other tetrapods. These results indicate a higher diversification of digits within and between bird species. This can be attributed to evolution of obligate bipedalism, and relaxation of developmental constraints on forelimbs

and hindlimbs in theropod ancestor of birds[47]. We next ask specifically what genes are involved in the divergence between FL-HL and how they dynamically change during development.

During limb development of all three birds, the number of FL-HL DEGs continuously increases from limb bud to later between each pair of corresponding digits in birds, but remains at a low and constant level in the two non-avian reptiles (**Fig. 2b, Supplementary Fig. 10b**), a constant level after an increase in axolotl (**Supplementary Note**). This provides direct evidence for elevated FL-HL functional divergence in both development and evolution of birds compared to other tetrapods. The increased DEG number of birds suggests that expression changes in a few upstream genes in limb buds, mostly Hox genes (**Fig. 2a, Supplementary Data 5**), consistent to the established role of Hox genes as master regulators of limb patterning[48, 49], and amplified the functional specialization in the later developed digits. These Hox genes that only recently became a FL-HL DEG in the ancestor of Archosauria or birds, show changes of expression specifically in the forelimb (e.g., *Hoxa11* and *Hoxd8*) or hindlimb (e.g., *Hoxd10*) or in both (*Hoxc12*) relative to their orthologs in outgroup species (**Fig. 2c, Supplementary Fig. 12-13**). The forelimb bud biased genes mainly include the *Hox* paralogous group genes 4 and 5 (*HoxPG4* and *5*) that express along the lateral plate mesenchyme and regulate the forelimb position[48, 50, 51] and several posterior *Hoxd8-10* genes (**Supplementary Fig. 13**). Most of these Hox genes were shown to be predominantly expressed in the proximal parts of chicken and mice limb buds[48, 52, 53, 54], presumably dictating the upper arm/leg differences between birds and mammals, except for *Hoxb4*, *d10* and *c12* (**Supplementary Data 6**).

The phylogenetic distribution of all FL-HL limb bud DEGs reveals a layered architecture of limb identity specification, with a conserved core network shared across tetrapods progressively supplemented by lineage-specific additions toward the avian ancestor. This ancestral regulatory network is inferred from the conserved patterns of FL/HL bud differential expression across all studied tetrapods or aminotes (**Fig. 2c**). It includes the reported forelimb initiation transcription factor (TF) TBX5, hindlimb initiation TFs TBX4, PITX1, ISL1[55, 56, 57], and posterior *Hox* genes *Hoxc9-11* and *Hoxb9*. Consistently, these four posterior Hox genes have been reported to have a hindlimb bud specific expression pattern in chicken[54], and some act as *Tbx5* suppressors[58]. Our results suggest that they have acquired the hindlimb patterning role already in the tetrapod ancestor. In addition, *Tbx5/4*, *Pitx1*, *Hoxb9* and *Hoxc10* were reported to

have differential expressions in either pectoral or pelvic fin bud in zebrafish and little skate[59, 60, 61, 62, 63, 64] suggesting their functions of specifying different appendage identity can be dated even to the origin of paired appendages in stem gnathostomes.

Interestingly many genes have preserved their FL-HL DE patterns from the limb bud until the digit stages. The tetrapod-wide FL-HL bud DEGs *Tbx4/5*, *HoxC10/11* (**Fig. 2c**) are differentially expressed between FL-HL in almost all tetrapod digits at both examined stages, and Archosauria- and avian-wide FL-HL bud DEGs *Hoxd10* and *Hoxd8* exhibit DE in fewer digits, stages and species (**Fig. 2d, Supplementary Fig. 14-15**). These genes could coordinate with the other 31 identified Archosauria- and avian-wide FL-HL digit DEGs (**Supplementary Data 5**) and contribute to the much more pronounced morphological differences between FL-HL digits of birds than other tetrapods. Remarkably, among the 33 digit DEGs, 11 were reported to have functions related to the AP patterning or feather formation (**Fig. 2d, Supplementary Data 6**). A prominent avian wing digit bias gene of AP patterning is *Tbx3*, which was reported to specify D3 and D4 in chicken hindlimb[65] and particularly, also have much higher and prolonged expression in the bat wing digits relative to mouse digits[66]. Other feather development genes include the feather primordium inhibitor *Bmp4* differentially expressed in the hindlimb digits and the forelimb digit biased stimulators *Fgf20* and *Wnt7a*, as well as other feather (*Edar*, *Gata3*, *Satb1*[67, 68, 69], or specifically flight feather (*Sim1*[70]) patterning genes. It remains to be examined whether and how these feather patterning genes are regulated by the avian-wide increase of *Hoxd8* (in both limb bud and digits), and decrease of *Bmp4* expression in the wing (**Fig. 2c-d**), given that *Hoxd8* was also reported to co-express with *Shh* in the feather buds[71]. In conclusion, we identify across the limb development course candidate genes which have been specifically altered in birds for their expressions and likely underlie the elevated functional disparity between FL-HL.

Evolution of molecular fingerprint genes across tetrapods

There are clearly more expression diversifications between digits of forelimb or/and hindlimb in birds than in other studied tetrapods (**Supplementary Fig. 3, Fig. 2a-b**). We next ask how obligate bipedalism can potentially contribute to functional specialization of avian digits evolved for flight or running. We propose that the transition to obligate bipedalism in the dinosaur

ancestor, i.e., the uncoupling of forelimbs from functional and developmental constraints of weight-bearing locomotion[72], might lead to relaxation of purifying selection or adaptive evolution and transcriptional changes in both forelimbs and hindlimbs. One possible scenario that we can test in this work is that birds might have evolved specific pleiotropic genes that are differentially expressed between FL-HL and also among the digits. This requires us to first map all MFGs onto the tetrapod phylogeny and identify the avian specific MFGs, based on their presence/absence of orthologous MFGs in other species.

We find none of the avian-specific MFGs are generated by gene duplications (**Supplementary Fig. 16**) : they are instead pre-existing genes that have acquired avian-specific expression biases toward certain digit through regulatory changes. And acquisition of avian MFGs seems to have occurred probably after the species divergence, because the number of species specific MFGs is 2 to 8 fold higher than that of the inferred avian ancestor, and different birds show different distributions of MFG numbers toward forelimb or hindlimb (**Fig. 3a**). And this elevated number of MFGs in birds is not a technical artifact resulting from their reduced digit number (three vs. five), both a comparison restricted to homologous digits (D1, D3, and D4) and a permutation test (randomly subsampling three digits from pentadactyl species, n = 1,000 simulations) consistently reproduced significantly higher MFG counts in avian lineages, with all simulated distributions falling below the observed avian values (empirical $P < 0.001$; **Supplementary Fig. 17**). Furthermore, the MFG genes identified across all three analyses showed high concordance, with around 80% of the original MFG set consistently recovered in the other two analyses, confirming that the three approaches identify largely the same set of MFGs. This suggests that functional specialization of FL-HL digits continued independently after obligate bipedalism originated in the avian ancestor, mainly through regulatory changes of shared orthologous genes specific to the respective avian lineage. For example, the emu-specific degenerating wing D4 has a significant (Fisher's exact test, adjusted $P < 0.05$) higher proportion of MFGs relative to the avian or Archosauria ancestor (**Supplementary Fig. 18**). Among the internal nodes of tetrapod phylogeny, MFGs of different tetrapod nodes share the enriched (one-sided Hypergeometric test, adjusted $P < 0.05$) GO terms involved in the development of appendage, skeleton, cartilage and muscles. And MFGs of the avian ancestor are enriched for GOs of BMP signaling regulation (**Fig. 3c, Supplementary Data 9**). Notable BMP-interacting

avian-specific MFGs include the BMP antagonist *Grem1* that has shifted its expression from the posterior end in vertebrate ancestor[73] to the anterior digit in birds, and the key AER component *Fgf8* associated with digit elongation[74] that becomes biasedly expressed in the avian anterior wing digit but not in other tetrapods. In total, we identified 20% (108 out of 564) of total MFGs of the avian ancestor and each bird species is also differentially expressed between FL-HL buds or digits, indicating that a substantial fraction of digit identity genes in birds have acquired expression biases associated with limb type — consistent with a pleiotropic coupling between digit patterning and forelimb-hindlimb differentiation. However, few of them have been studied for their functions during limb bud or digit patterning, except for *Sim1*[70], and the emu-specific wing-biased *Nkx2.5* (**Supplementary Fig. 19**). *Sim1* has been found to be wing-biased (**Fig. 2d**), and also a D4 MFG shared by all birds. While the cardiac muscle transcription factor *Nkx2.5* was reported to be expressed in the wing buds of emu, but not in those of other birds, contributing to the overall wing muscle reduction of emu[75].

A previous study comparing adjacent digits identified few conserved DEGs shared across four amniotes, except for the D1/D2 comparison, and concluded lack of conserved developmental identities of amniote digits, with the exception of D1[14]. In contrast, we here defined the MFGs by comparing the focal vs. all the rest digits of either forelimb or hindlimb and also allow for lineage specific changes when mapping all MFGs onto each phylogenetic node (e.g., *Hand2* in **Fig. 1d, Materials and Methods**). This allows us to reveal that tetrapods in fact share some of MFGs that respectively define the conserved outgrowth identity of each individual digit (**Fig. 3a, Supplementary Data 7-8**). In addition, the scope of conservation of some MFGs likely can be extended beyond tetrapods: when we examined all of the genes reported to have AP biased expression in the fin buds of more basal types of jawed vertebrates (lungfish, zebrafish, catshark and little skate) [41, 73, 76, 77, 78], all of them show consistent expression biased toward the anterior or posterior tetrapod digits (**Supplementary Fig. 20**). For example, the genes reported to have anterior expression in the fin buds of catshark *Pax9*, *Zic3*, and *Alx4*[41] have a higher expression in the anterior digits of all studied tetrapods in this work, and vice versa for the posterior genes (**Fig. 3b, Fig. 1d, Supplementary Fig. 21**). This suggests the existence of an ancient jawed vertebrate regulatory network that has been co-opted to specify and maintain the AP polarity throughout development from limb bud to digits in extant tetrapods, in parallel to the

ancient network that distinguishes the anterior and posterior paired appendages (**Fig. 2d**). This AP specification network is more precisely confirmed by our analyses of single-cell RNA-seq data of mouse and emu limb bud[79, 80], for which we defined the anterior and posterior cell populations with the reported anterior (*Alx4* etc.) and posterior (*Shh* etc.) gene markers[81] (**Fig. 3d**). Several tetrapod-wide anterior MFGs preferentially expressed in the D1 (e.g., *Msx1*), and posterior MFGs of D5 (*Shox2*, *Hand2* etc.), as well as some amniote-wide D1 identity genes (e.g., *Hoxd11/12*) identified in this work have a biased expression in the anterior or posterior cells of mouse and emu limb buds (**Fig. 3e, Supplementary Fig. 22-23**).

In conclusion, we inferred an ancestral gene network dated to at least the jawed vertebrate ancestor, whose AP specification function probably continues to play a role during development of the limb bud and individual digits. We also identified substantial numbers of genes that have specifically evolved novel expression patterns, possibly novel functions among birds between FL-HL and also in certain digits.

Discussion

Serial homologs like tetrapod limbs, digits and insect wings comprise classic models for elucidating the mechanisms of developmental pattern formation[22, 82, 83]. On the other hand, their variations within and between species have been inspiring more interest than controversy[84, 85]. Because they also provide a paradigm for revealing the molecular mechanisms underlying various forms of evolutionary changes including gain, loss or modification of some of the repetitive body units without changing the cell fate[4], on top of their often shared and highly conserved underlying genetic and developmental modules[3]. A very challenging prerequisite to be established before reconstructing any of these evolutionary changes is the resolution of the homologous relationship and character identity, if any, of each of the serial homologous units. Here we focus on such a paradigm, namely that of tetrapod digits, and dispel speculations about the lack of different digit identities[14, 15] by reporting the MFGs (**Fig. 3, Supplementary Data 7**). We further provide comparative transcriptomic evidence insights into how the limb and digit identities are maintained by the conserved gene regulatory networks dated to the ancestor of jawed vertebrates and tetrapods (**Fig. 4**). Axolotl is considered in the network conservation against its divergent developmental program, e.g., pre-axial dominance and heterochrony[34]. As an obligate paedomorph, the axolotl retains larval features throughout adult life due to the

absence of thyroid hormone-driven metamorphosis[86], and these life history differences may influence digit transcriptomic profiles and thus complicate cross-species comparisons involving this taxon. Indeed, despite being an outlier in some of the aspects, our analysis reveals that core gene regulatory networks remain deeply conserved across all tetrapod, allowing us to robustly distinguish deep homology from lineage-specific adaptations.

By aligning individual avian digit transcriptomes against their pentadactyl outgroup reptiles, we first disentangle the long-term controversy about the avian wing digit identity and reveal that the three avian wing digits are respectively clustered with the reptile digit 1,3,4, confirming the results of [14], and identify new cases of frameshifts in the hindlimbs (**Fig. 1b-c**). These results together suggest that, the positional and outgrowth identities of a certain character can vary during the course of development, confounding the homology inference of the character. It is also parallel to the experiment, where ectopic expression of *Tbx4* in the wing field can induce a leg-like structure, and vice versa for that of *Tbx5*[56, 57].

Besides possible mutations or changes of expression in *Hox* genes, the underlying mechanism of generating such mismatches of avian digit identities can be also understood by the switch between separate signaling centers that are active and self-organizing at different phases of limb development[20, 21, 22], e.g. from limb bud to the digit outgrowth stages. The positional identity instructed early by ZPA/AER in limb buds can be disconnected from the developmental or outgrowth identity mainly by the influence of the posterior interdigital tissue, resulting in the frameshift of a certain digit. Such disconnection has long been implicated by the pattern that digits with the same positional identity can have drastically different morphologies between different species. For example, a dolphin digit 2 of fin has 9 phalanges compared to 3 phalanges of mouse and human forelimb digit 2[21], and in fact in the ancestral mammalian hand. Another implication of such disconnection is that after the positional identity is specified, for example for the HD5 across archosaurs[87], only vestigial (for crocodile and ostrich, **Fig. 1a**) or no digit (for other birds in this study) grows at this position. Interestingly, the avian frameshift cases observed in this work, regardless of wings or claws, are only found between digit identity 1 and 2, and between 4 and 5. Such non-random combinations might be caused by the fact that digits with positional identities of 4 and 5 arise from a common origin of Shh-expressing ZPA cells[88, 89], possibly making them more interchangeable with each other than other digits. For digit 1 we find

some evidence supporting the MAD hypothesis[15], and it has the highest number of MFGs, besides the reported Hox genes, compared to other digits in all studied species (**Fig. 3b**). This implicates its conserved stronger developmental identity than any other digits, thus possibly a stronger selective constraint to retain the MAD during evolution. By contrast, there are no conserved MFGs shared by any of the two studied species (except axolotl) in digit 2 (**Fig. 3a, Supplementary Fig. 16**), making it possibly more dispensable during development and evolution, contributing to its more frequent frameshifts with the nearby digit 1. This conclusion is strongly supported by developmental evidence but raises questions of how such a replacement is or is not reflected in the paleontological evidence. We encourage a re-examination of the fossil record in the light of these developmental findings.

With a different definition of characteristic genes or MFG of each digit from a previous study[14], we identified distinct sets of MFGs for respective digits at different tetrapod nodes, providing clear evidence that not only the MAD, but other digits (except for the digit 2) also have highly conserved outgrowth identities (**Fig. 3**). It is important to note that the MFGs defined in this work are those biasedly, but not necessarily uniquely expressed within a certain digit. This is a less stringent criteria to identify the characteristic genes associated with certain digit, which accounts for our different results, i.e., there are identity-associated genes, from the previous study[14]. We further show that such digit identities are specified or maintained by some genes (e.g., *Alx4* in D1, *Hoxd10*, *Hand2* in D5) that show differential AP expression in both fish fin and tetrapod limb buds, and have preserved such bias until the tetrapod digit stage. While some genes, e.g., *Hoxa13* and *Gli3* only have AP biased expression in the limb bud[89, 90, 91, 92], but have no expression bias found at the digit stage. Another set of highly conserved genes across the development and evolution processes involve those that specify the FL-HL identity[93]. Many genes that show a tetrapod-wide bias toward either fore- or hind-limb bud have preserved such bias until the digit stage (**Fig. 2c-d**). Additionally, some of these genes, e.g., *Tbx4/5*[59], *Hoxc10*, *Hoxb9*[94] have biased expression in the pectoral/pelvic fin buds of teleosts. All these results support that the extant limb and digit identity gene regulatory networks can be dated to the stem lineage of jawed vertebrates (**Fig. 4**), with additions of limb- (**Fig. 2a**) or digit- (**Fig. 3a**) identity genes into the networks during subsequent species divergence. A major gap in our current knowledge of the limb and digit patterning process is how cells can maintain their positional

identities during the developmental switch from one signaling center (ZPA) to another (PID/PFR). These persistently biased genes toward the anterior or posterior end of the autopod identified in this work likely play an important role in the cell memory and should be the focus of future experimental tests.

On top of such highly conserved limb and digit identity regulatory networks, in birds, we have found much larger developmental and molecular differences than other quadrupedal tetrapods between forelimbs and hindlimbs. And increased molecular and morphological diversity of digits probably have occurred after the avian species radiation. Although we have identified many genes, particularly Hox genes that have only recently acquired an expression bias toward FL-HL limb bud in the ancestor of birds or Archeosaur (**Fig. 2a**), we are not able to draw causal conclusions between the evolution of obligate bipedalism vs. increased avian digit diversity, due to the knowledge gap mentioned above. And whether and how relaxation of functional constraints resulting from obligate bipedalism contributes to the diversification of avian digits remains an intriguing topic. Our observation that ~20% of avian MFGs show biases in both limb type and digit identity points to a tight evolutionary coupling between digit patterning and forelimb-hindlimb (FL-HL) differentiation in birds. This supports the idea that such diversification may involve the direct evolution of pleiotropic, avian-specific genes that are differentially expressed both between FL-HL and among digits. While we currently lack reported experimental evidence linking upstream avian-specific FL-HL biased genes to downstream avian-specific MFGs, we cannot exclude the possibility that this occurs through the evolution of a new regulatory hierarchy between them. Further investigation is needed to elucidate these regulatory pathways. Ultimately, elucidating whether obligate bipedalism drove avian digit diversification through regulatory rewiring or modular evolution remains an intriguing frontier, and it is to be investigated at the wake of functional genomics at the single-cell level, from a hierarchical perspective of homology [84].

Methods

Data collection

All of the animal embryos were sampled after the approval from the Laboratory Animal Welfare and Ethics Committee of Zhejiang University. For the three birds, fertilized chicken (White Leghorns), emu and ostrich eggs were obtained from farms nearby Hangzhou, Lishui and Xi'An

and the eggs were incubated at 37.5, 36.5, 37 ℃ respectively with a humidity around 65-75%. We collected the samples at Hamburger and Hamilton stage 21, stage 28 and stage 32 embryos[95, 96]. The hindlimb shape was used as a morphological criterion to stage embryos, with the Hamburger-Hamilton stage for chickens and also for emu and ostrich embryos. Stage 21 was characterized by slightly asymmetrical leg buds, while stage 28 featured a second digit and third toe longer than the others, creating a pointed contour in the digital and toe plates. By stage 32, all digits and toes have significantly lengthened, with thin, concave-webbed structures forming between them. Corresponding stages of other vertebrate species were chosen according to the similarity of their digit morphology with chicken (**Supplementary Data 1**). We collected the Chinese softshell turtle eggs from a farm nearby Hangzhou, and incubated them at 30°C. We collected limb bud and digit samples on 10, 18 and 28 days according to the staging work of[97]. We collected the Siamese crocodile eggs from a farm in Beihai, and incubated them at 33°C. The samples were collected on 12, 24, and 31-33 days according to staging work of[98]. Axolotl larvae were bred in freshly dechlorinated tap water at 18–20 °C with a 12 -hour light-dark cycle. The axolotls were deeply anesthetized using 0.03% ethyl-p-aminobenzoate, and the limbs were collected separately according to their limb morphology that are morphologically consistent with other species. The forelimb buds were collected on 6-8 dph (days post hatching), and hindlimb buds and early stage forelimb digital primordia were collected on 18-21 dph due to the developmental heterochrony[98]. Late stage forelimb and early hindlimb digital primordia were collected on 28-32 dph. Late hindlimb digits were collected at around 84 dph[98]. For all samples, we collected each digit including its posterior interdigital mesoderm (**Supplementary Fig. 24**) from forelimb and hindlimb respectively. Left and right limb bud or digit samples were collected as replicates from the same individual embryo for 6 species, and placed immediately in the RNAlater (Sigma-Aldrich) solution. RNA was extracted from each sample with TRIzol® (Thermo Fisher Scientific)[99]. Then paired-end libraries were constructed using NEBNext® UltraTM RNA Library Prep Kit for Illumina® (NEB, USA) and 2Gb paired-end reads of 150bp long were produced for each library. We also collected additional forelimb and hindlimb samples from six species and fixed them in ethanol overnight for Alcian staining. The samples were then transferred to acetone overnight. Alcian blue and alizarin red staining were performed for 1–2

days. Embryos were cleared with 1% KOH, then placed in 50% glycerol/KOH solution until transparent, and finally transferred to 100% glycerol for long-term storage[100].

Sequence Alignment, gene counting, and normalization

We used the genomes of chicken genome (GRCg6a,GCA_000002315.5), emu (ZJU2, GCA_016128335.2), ostrich (ASM69896v1, GCA_000698965.1), Australian saltwater crocodile (CroPor_comp1, GCA_001723895.1, which is the closest-related species with available genome to Siamese crocodile), Chinese softshell turtle (PelSin_1.0, GCA_000230535.1), and axolotl (AmexG_v6.0-DD) as reference for aligning the RNA-seq reads with HISAT2 (2.1.0)[101]. The alignment rates between the RNA-seq reads vs. the reference genomes range between 70 to 90 % (**Supplementary Data 10**). Raw gene counts were counted by featureCounts(v2.0.1) and normalized by transcripts per million (TPM) method for estimating the gene expression level[102]. Other published transcriptomes of the digits of the forelimbs of *Anolis*, and mouse[14], were mapped to the *Anolis* genome (AnoCar2.0 (GCA_000090745.1)) and mouse genome (GRCm38.p6 (GCA_000001635.8)) following the same pipeline mentioned above. 4,741 limb-related genes were compiled using Mouse Genome Informatics (MGI)[103] and Chicken Embryo Gene Expression Database (Geisha)[38]. Specifically, we extracted the genes that have been characterized with limb or digit expression or mutant phenotypes in limbs or digits or annotated with limb or digit development GO terms in MGI. We also included the genes that have been reported to have expression evidence (by whole mount RNA *in situ* hybridization) in chicken limb.

Molecular fingerprint genes (MFGs) and Degree of Digit Individuality (DDI)

We define MFGs as those that are upregulated or downregulated relative to the rest digits in at least one of the sampled developmental stages. The upregulation or downregulation index was measured by $\frac{\sum_{i=1}^{n} 1-\widehat{x_i}}{n-1}$; $\widehat{x_i} = \frac{x_i}{max_{1\le i\le n(x_i)}}$ (1) or $\frac{\sum_{i=1}^{n} 1-\widehat{x_i}}{1-n}$; $\widehat{x_i} = \frac{1}{min_{1\le i\le n(x_i)}}$ (2) according to[104], where the x is the TPM of the respective digit, and n is the tissue number. A value higher than 0.5 was used as the cutoff for defining MFGs according to the distribution of the index values of each species (**Supplementary Fig. 25)**. For each clade, clade-specific MFGs were identified as genes exhibiting a conserved pattern of upregulation or downregulation in at least half of the

species, including the most ancestral ones (e.g., four out of the seven tetrapod species studied, including axolotl). We used *OrthoFinder* (version 2.2.6)[105] with the *-f* command to perform a comparison of the complete protein sequences from three bird species (chicken, emu and ostrich), the Chinese softshell turtle, and the Australian saltwater crocodile. Then we identified avian-specific gene duplications, and we compared them with the identified avian-specific MFGs to determine whether any of the avian-specific gene duplications were also our defined MFGs. We define DDI by pairwise *ANOVA* test to find the genes that were differentially expressed between any two digits in the limb as in[14]. glmFit and glmLRT functions of the *EdgeR* (Release 3.1) package[106] were used to test for differential expression of adjacent digits of chicken, emu, ostrich, crocodile, softshell turtle, *Anolis*, mouse, and axolotl. We filtered the genes with fold change of more than 2 as candidates. Numbers of each differential expressed candidate gene were counted in each species, and then we divided the numbers by the digit number in each limb of different species as the DDI value.

Gene expression analyses

To exclude the reported confounding factors from the lowly-expressed or housekeeping genes when performing the HCA[14], we extracted a non-redundant set of 4741 limb patterning orthologous gene groups, which either have a mutant with defective limb/digit phenotypes (e.g., *Hoxd13*, *Pitx1*)[107, 108] or a limb/digit specific expression pattern in mouse or chicken (e.g., *Hoxc10,Tbx5*, **Supplementary Fig. 4**)[38, 109]. Indeed, using chicken hindlimb digits as a test, clusters of individual digits can be better separated by transcriptomes of limb-related genes than those of all genes (**Supplementary Fig. 26**). Hierarchical clustering and 1000 bootstrapping analyses were performed on the square root transformed TPM values with the R package *pvclust*[110]. We generated the clusters of different samples from the correlation-based dissimilarity matrix using the average-linked method[14]. We used *TreeExp*[111] to estimate the branch lengths of gene expression with TPM values of limb related genes. Relative rate tests of Z-scores were calculated based on sampling variance of the expression distance. We defined the DEGs between fore- and hind-limb buds with the *deseq* function of *DESeq2*[112]. Genes with log2FoldChange more than 1 and padj less than 0.05 were defined as DE genes. We inferred the branch-specific forelimb and hindlimb expression changes by *PhyDGET* [113]. We first normalized and

transformed the RNA-seq read counts of forelimb and hindlimb digits of each species into log2(cpm) (log2 transform of counts per million), then tested the expression difference across seven species using *BayesTrait* (v3.0.2)[114] under a null model that hypothesizes constant rate changes and alternative models. Marginal log-likelihoods were compared to calculate the Bayes factors, which improved alternative model fit. And the genes can be defined with branch-specific changes of expression levels in the forelimb or hindlimb relative to their orthologs in outgroups.

Analysis of the scRNA-seq data

Published mouse limb bud E10.5 and emu limb HH25 data[79, 80] were downloaded and aligned to the mouse (GRCm38.p6 (GCA_000001635.8)) and emu (ZJU2, GCA_016128335.2) genome respectively. We used *CellRanger*[115] with default parameters and then the matrices were processed to *Seurat*[116]. Counts were scaled by *LogNormalize*, and they were clustered using *FindNeighbors* and *FindClusters*. We used *FindAllMarkers* to identify specific markers, and then the cell clusters were annotated according to reported limb markers[79]. Mesenchymal clusters were retained and re-clustered. The markers were re-identified only in mesenchymal clusters, and then the overlapping genes with MFGs are found.

GO analysis

The GO enrichments were produced by Metascape[117], and the orthologous gene to mouse of MFGs are given to this online tool. All genes in the genome were used as the enrichment background.

Statistics & Reproducibility

No statistical method was used to predetermine sample size. No data were excluded from the analyses; The experiments were not randomized; The Investigators were not blinded to allocation during experiments and outcome assessment. Two biological replicates were collected per limb bud/ digit per developmental stage per species.

The following statistical tests were used in this study:

Pearson's correlation was used to assess the repeatability of biological replicates ($P < 0.001$, $R > 0.9$).

Proportion test was used to compare the normalised number of DEGs between digits across species, and to compare the total number of FL-HL DEGs between birds and non-avian reptiles ($P < 0.05$).

Wald test (adjusted $P < 0.05$, fold-change > 2) was used to identify differentially expressed genes (DEGs) between forelimb and hindlimb digits.

Fisher's exact test (adjusted $P < 0.01$ or $P < 0.05$) was used to compare the proportion of shared MFGs between FL-HL, and to assess the proportion of MFGs in degenerating digits relative to ancestral nodes.

z-test ($P < 0.05$) was used to compare transcriptome evolutionary rates between birds and crocodile.

One-sided Hypergeometric test (adjusted $P < 0.05$) was used for GO enrichment analysis of MFGs across tetrapod phylogenetic nodes.

Permutation test ($n = 1{,}000$ simulations, empirical $P < 0.001$) was used to confirm that elevated MFG counts in avian lineages are not a technical artifact of reduced digit number.

Data availability

The RNA-seq data generated in this work for axolotl, turtle, crocodile, chicken, emu and ostrich are available in the NCBI Sequence Read Archive (SRA) database: PRJNA1258050 [https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1258050]. The published RNA-seq data of mouse and Anolis digits used in this study are available in the Gene Expression Omnibus (GEO) repository: GSE108337 [https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE108337]. The published scRNA-seq data of mouse limb used in this study are available in the ENCODE portal: ENCSR713GIS [https://www.encodeproject.org/experiments/ENCSR713GIS/]

Code availability

Scripts for analysing data and plot figures are available on Github at https://github.com/KangWwen/Evolution-of-limb-and-digit-identity/

Reference

1. Clack JA. An early tetrapod from 'Romer's Gap'. *Nature* **418**, 72-76 (2002).
2. Shubin N, Tabin C, Carroll S. Fossils, genes and the evolution of animal limbs. *Nature* **388**, 639-648 (1997).
3. Shubin N, Tabin C, Carroll S. Deep homology and the origins of evolutionary novelty. *Nature* **457**, 818-823 (2009).
4. Wagner GP. *Homology, Genes, and Evolutionary Innovation*. Princeton University Press (2018).
5. Larsson HCE, Wagner GP. Pentadactyl ground state of the avian wing. *J Exp Zool* **294**, 146-151 (2002).
6. Wang Z, Young RL, Xue H, Wagner GP. Transcriptomic analysis of avian digits reveals conserved and derived digit identities in birds. *Nature* **477**, 583-586 (2011).
7. Burke AC, Feduccia A. Developmental Patterns and the Identification of Homologies in the Avian Hand. *Science* **278**, 666-668 (1997).
8. Feduccia A, Nowicki J. The hand of birds revealed by early ostrich embryos. *Naturwissenschaften* **89**, 391-393 (2002).
9. Kundrát M, Seichert V, Russell AP, Smetana K, Jr. Pentadactyl pattern of the avian wing autopodium and pyramid reduction hypothesis. *J Exp Zool* **294**, 152-159 (2002).
10. Welten MCM, Verbeek FJ, Meijer AH, Richardson MK. Gene expression and digit homology in the chicken embryo wing. *Evol Dev* **7**, 18-28 (2005).
11. Xu X, Mackem S. Tracing the evolution of avian wing digits. *Curr Biol* **23**, R538-544 (2013).
12. Hinchliffe JR, Hecht MK. Homology of the Bird Wing Skeleton. *Evolutionary Biology*, 21-39 (1984).
13. Müller GB, Alberch P. Ontogeny of the limb skeleton in Alligator mississippiensis: developmental invariance and change in the evolution of archosaur limbs. *J Morphol* **203**, 151-164 (1990).
14. Stewart TA, Liang C, Cotney JL, Noonan JP, Sanger TJ, Wagner GP. Evidence against tetrapod-wide digit identities and for a limited frame shift in bird wings. *Nat Commun* **10**, 3244 (2019).
15. Woltering JM, Duboule D. The origin of digits: expression patterns versus regulatory mechanisms. *Dev Cell* **18**, 526-532 (2010).
16. Wagner GP, Gauthier JA. 1,2,3 = 2,3,4: a solution to the problem of the homology of the digits in the avian hand. *Proc Natl Acad Sci U S A* **96**, 5111-5116 (1999).
17. de Bakker MAG*, et al.* Selection on Phalanx Development in the Evolution of the Bird Wing. *Mol Biol Evol* **38**, 4222-4237 (2021).
18. Bickelmann C, van der Vos W, de Bakker MAG, Jiménez R, Maas S, Sánchez-Villagra MR. Hox gene expression in the specialized limbs of the Iberian mole (Talpa occidentalis). *Evol Dev* **19**, 3-8 (2017).
19. Vargas AO, Fallon JF. Birds have dinosaur wings: The molecular evidence. *J Exp Zool B Mol Dev Evol* **304**, 86-90 (2005).
20. Suzuki T. How is digit identity determined during limb development? *Dev Growth Differ* **55**, 130-138 (2013).
21. Huang B-L, Mackem S. Rethinking positional information and digit identity: The role of late interdigit signaling. *Dev Dyn* **251**, 1414-1422 (2022).
22. Young JJ, Tabin CJ. Saunders's framework for understanding limb development as a platform for investigating limb evolution. *Dev Biol* **429**, 401-408 (2017).
23. Lopez-Rios J*, et al.* Attenuated sensing of SHH by Ptch1 underlies evolution of bovine limbs. *Nature* **511**, 46-51 (2014).

24. Cooper KL*, et al.* Patterning and post-patterning modes of evolutionary digit loss in mammals. *Nature* **511**, 41-45 (2014).
25. Wolpert L. Positional information and the spatial pattern of cellular differentiation. *J Theor Biol* **25**, 1-47 (1969).
26. Zhu J, Patel R, Trofka A, Harfe BD, Mackem S. Sonic hedgehog is not a limb morphogen but acts as a trigger to specify all digits in mice. *Dev Cell* **57**, 2048-2062.e2044 (2022).
27. Suzuki T, Hasso SM, Fallon JF. Unique SMAD1/5/8 activity at the phalanx-forming region determines digit identity. *Proc Natl Acad Sci U S A* **105**, 4185-4190 (2008).
28. Zhu J, Mackem S. John Saunders' ZPA, Sonic hedgehog and digit identity - How does it really all work? *Dev Biol* **429**, 391-400 (2017).
29. Logan M. Finger or toe: the molecular basis of limb identity. *Development* **130**, 6401-6410 (2003).
30. Persons WS, Currie PJ. The functional origin of dinosaur bipedalism: Cumulative evidence from bipedally inclined reptiles and disinclined mammals. *J Theor Biol* **420**, 1-7 (2017).
31. Raikow RJ. Locomotor system. *Form and function in birds*, (1985).
32. Botelho JF, Smith-Paredes D, Nunez-Leon D, Soto-Acuna S, Vargas AO. The developmental origin of zygodactyl feet and its possible loss in the evolution of Passeriformes. *Proc Biol Sci* **281**, 20140765 (2014).
33. Francisco Botelho J, Smith-Paredes D, Soto-Acuña S, Mpodozis J, Palma V, Vargas AO. Skeletal plasticity in response to embryonic muscular activity underlies the development and evolution of the perching digit of birds. *Sci Rep* **5**, 9840 (2015).
34. Bininda-Emonds ORP*, et al.* Forelimb-hindlimb developmental timing changes across tetrapod phylogeny. *BMC Evol Biol* **7**, 182 (2007).
35. Suzuki T, Fallon JF. The dynamic spatial and temporal relationships between the phalanx-forming region and the interdigits determine digit identity in the chick limb autopod. *Dev Dyn* **250**, 1318-1329 (2021).
36. Green RE*, et al.* Three crocodilian genomes reveal ancestral patterns of evolution among archosaurs. *Science* **346**, 1254449 (2014).
37. Galli A*, et al.* Distinct roles of Hand2 in initiating polarity and posterior Shh expression during the onset of mouse limb bud development. *PLoS Genet* **6**, e1000901 (2010).
38. Antin PB, Yatskievych TA, Davey S, Darnell DK. GEISHA: an evolving gene expression resource for the chicken embryo. *Nucleic Acids Res* **42**, D933-937 (2014).
39. Kuijper S, Feitsma H, Sheth R, Korving J, Reijnen M, Meijlink F. Function and regulation of Alx4 in limb development: Complex genetic interactions with Gli3 and Shh. *Dev Biol* **285**, 533-544 (2005).
40. Vieux-Rochas M*, et al.* BMP-mediated functional cooperation between Dlx5;Dlx6 and Msx1;Msx2 during mammalian limb development. *PLoS One* **8**, e51700 (2013).
41. Onimaru K, Kuraku S, Takagi W, Hyodo S, Sharpe J, Tanaka M. A shift in anterior-posterior positional information underlies the fin-to-limb evolution. *Elife* **4**, (2015).
42. Qu S*, et al.* Polydactyly and ectopic ZPA formation in Alx-4 mutant mice. *Development* **124**, 3999-4008 (1997).
43. Bensoussan-Trigano V, Lallemand Y, Saint Cloment C, Robert B. Msx1 and Msx2 in limb mesenchyme modulate digit number and identity. *Dev Dyn* **240**, 1190-1202 (2011).
44. Peters H, Neubüser A, Kratochwil K, Balling R. Pax9-deficient mice lack pharyngeal pouch derivatives and teeth and exhibit craniofacial and limb abnormalities. *Genes Dev* **12**, 2735-2747 (1998).

45. Robledo RF, Rajan L, Li X, Lufkin T. The Dlx5 and Dlx6 homeobox genes are essential for craniofacial, axial, and appendicular skeletal development. *Genes Dev* **16**, 1089-1101 (2002).
46. Okada I*, et al.* SMOC1 is essential for ocular and limb development in humans and mice. *Am J Hum Genet* **88**, 30-41 (2011).
47. Hutchinson JR, Allen V. The evolutionary continuum of limb function from early theropods to birds. *Naturwissenschaften* **96**, 423-448 (2009).
48. Yakushiji-Kaminatsui N, Lopez-Delisle L, Bolt CC, Andrey G, Beccari L, Duboule D. Similarities and differences in the regulation of HoxD genes during chick and mouse limb development. *PLoS Biol* **16**, e3000004 (2018).
49. Zakany J, Duboule D. The role of Hox genes during vertebrate limb development. *Curr Opin Genet Dev* **17**, 359-366 (2007).
50. Moreau C*, et al.* Timed Collinear Activation of Hox Genes during Gastrulation Controls the Avian Forelimb Position. *Curr Biol* **29**, 35-50.e34 (2019).
51. Rancourt DE, Tsuzuki T, Capecchi MR. Genetic interaction between hoxb-5 and hoxb-6 is revealed by nonallelic noncomplementation. *Genes Dev* **9**, 108-122 (1995).
52. Tarchini B, Duboule D. Control of Hoxd genes' collinearity during early limb development. *Dev Cell* **10**, 93-103 (2006).
53. Peichel CL, Prabhakaran B, Vogt TF. The mouse Ulnaless mutation deregulates posterior HoxD gene expression and alters appendicular patterning. *Development* **124**, 3481-3492 (1997).
54. Nelson CE*, et al.* Analysis of Hox gene expression in the chick limb bud. *Development* **122**, 1449-1466 (1996).
55. Duboc V*, et al.* Tbx4 function during hindlimb development reveals a mechanism that explains the origins of proximal limb defects. *Development* **148**, (2021).
56. Rodriguez-Esteban C, Tsukui T, Yonei S, Magallon J, Tamura K, Izpisua Belmonte JC. The T-box genes Tbx4 and Tbx5 regulate limb outgrowth and identity. *Nature* **398**, 814-818 (1999).
57. Takeuchi JK*, et al.* Tbx5 and Tbx4 genes determine the wing/leg identity of limb buds. *Nature* **398**, 810-814 (1999).
58. Nishimoto S, Logan MPO. Subdivision of the lateral plate mesoderm and specification of the forelimb and hindlimb forming domains. *Semin Cell Dev Biol* **49**, 102-108 (2016).
59. Tamura K, Yonei-Tamura S, Izpisúa Belmonte JC. Differential expression of Tbx4 and Tbx5 in Zebrafish fin buds. *Mech Dev* **87**, 181-184 (1999).
60. Ruvinsky I, Oates AC, Silver LM, Ho RK. The evolution of paired appendages in vertebrates: T-box genes in the zebrafish. *Dev Genes Evol* **210**, (2000).
61. Cole NJ, Tanaka M, Prescott A, Tickle C. Expression of limb initiation genes and clues to the morphological diversification of threespine stickleback. *Curr Biol* **13**, (2003).
62. Cohn MJ, Patel K, Krumlauf R, Wilkinson DG, Clarke JD, Tickle C. Hox9 genes and vertebrate limb specification. *Nature* **387**, 97-101 (1997).
63. Choe A, Phun HQ, Tieu DD, Hu YH, Carpenter EM. Expression patterns of Hox10 paralogous genes during lumbar spinal cord development. *Gene Expr Patterns* **6**, (2006).
64. Murata Y*, et al.* Allometric growth of the trunk leads to the rostral shift of the pelvic fin in teleost fishes. *Dev Biol* **347**, (2010).
65. Suzuki T, Takeuchi J, Koshiba-Takeuchi K, Ogura T. Tbx Genes Specify Posterior Digit Identity through Shh and BMP Signaling. *Dev Cell* **6**, 43-53 (2004).
66. Wang Z*, et al.* Unique expression patterns of multiple key genes associated with the evolution of mammalian flight. *Proc Biol Sci* **281**, 20133133 (2014).

67. Ho WKW*, et al.* Feather arrays are patterned by interacting signalling and cell density waves. *PLoS Biol* **17**, e3000132 (2019).
68. Fessing MY*, et al.* p63 regulates Satb1 to control tissue-specific chromatin remodeling during development of the epidermis. *J Cell Biol* **194**, 825-839 (2011).
69. Musser JM*, et al.* Subdivision of ancestral scale genetic program underlies origin of feathers and avian scutate scales. *bioRxiv*, (2018).
70. Seki R*, et al.* Functional roles of Aves class-specific cis-regulatory elements on macroevolution of bird-specific features. *Nat Commun* **8**, 14229 (2017).
71. Hintermann A*, et al.* Developmental and evolutionary comparative analysis of a regulatory landscape in mouse and chicken. *Development* **149**, (2022).
72. Gatesy SM, Dial KP. LOCOMOTOR MODULES AND THE EVOLUTION OF AVIAN FLIGHT. *Evolution* **50**, 331-340 (1996).
73. Malkmus J*, et al.* Spatial regulation by multiple Gremlin1 enhancers provides digit development with cis-regulatory robustness and evolutionary plasticity. *Nat Commun* **12**, 5557 (2021).
74. Sanz-Ezquerro JJ, Tickle C. Fgf signaling controls the number of phalanges and tip formation in developing digits. *Curr Biol* **13**, 1830-1836 (2003).
75. Farlie PG*, et al.* Co-option of the cardiac transcription factor Nkx2.5 during development of the emu wing. *Nat Commun* **8**, 132 (2017).
76. Cass AN, Elias A, Fudala ML, Knick BD, Davis MC. Conserved Mechanisms, Novel Anatomies: The Developmental Basis of Fin Evolution and the Origin of Limbs. *Diversity* **13**, 384 (2021).
77. Ribeiro I*, et al.* Tbx2 and Tbx3 Regulate the Dynamics of Cell Proliferation during Heart Remodeling. *PLoS One* **2**, e398 (2007).
78. Woltering JM, Irisarri I, Ericsson R, Joss JMP, Sordino P, Meyer A. Sarcopterygian fin ontogeny elucidates the origin of hands with digits. *Science Advances* **6**, eabc3510 (2020).
79. He P*, et al.* The changing mouse embryo transcriptome at whole tissue and single-cell resolution. *Nature* **583**, 760-767 (2020).
80. Tsuboi E*, et al.* Immobilization secondary to cell death of muscle precursors with a dual transcriptional signature contributes to the emu wing skeletal pattern. *Nat Commun* **15**, 1-14 (2024).
81. Palacio V*, et al.* Single-cell profiling of penta- and tetradactyl mouse limb buds identifies mesenchymal progenitors controlling digit numbers and identities. *bioRxiv*, (2024).
82. Hu Y, Moczek AP. Wing serial homologues and the diversification of insect outgrowths: insights from the pupae of scarab beetles. *Proc Biol Sci* **288**, 20202828 (2021).
83. Hiscock TW, Tschopp P, Tabin CJ. On the Formation of Digits and Joints during Limb Development. *Dev Cell* **41**, 459-465 (2017).
84. DiFrisco J, Love AC, Wagner GP. The hierarchical basis of serial homology and evolutionary novelty. *J Morphol* **284**, e21531 (2023).
85. Fusco G. Serial Homology. *Biological Theory* **17**, 114-119 (2022).
86. Crowner A, Khatri S, Blichmann D, Voss SR. Rediscovering the Axolotl as a Model for Thyroid Hormone Dependent Development. *Front Endocrinol (Lausanne)* **10**, 237 (2019).
87. de Bakker MAG*, et al.* Digit loss in archosaur evolution and the interplay between selection and constraints. *Nature* **500**, 445-448 (2013).
88. Harfe BD, Scherz PJ, Nissim S, Tian H, McMahon AP, Tabin CJ. Evidence for an expansion-based temporal Shh gradient in specifying vertebrate digit identities. *Cell* **118**, 517-528 (2004).
89. Towers M, Signolet J, Sherman A, Sang H, Tickle C. Insights into bird wing evolution and digit specification from polarizing region fate maps. *Nat Commun* **2**, 426 (2011).

90. Fromental-Ramain C, Warot X, Messadecq N, LeMeur M, Dollé P, Chambon P. Hoxa-13 and Hoxd-13 play a crucial role in the patterning of the limb autopod. *Development* **122**, 2997-3011 (1996).
91. Bastida MF*, et al.* The formation of the thumb requires direct modulation of transcription by Hoxa13. *Proc Natl Acad Sci U S A* **117**, 1090-1096 (2020).
92. Wang B, Fallon JF, Beachy PA. Hedgehog-regulated processing of Gli3 produces an anterior/posterior repressor gradient in the developing vertebrate limb. *Cell* **100**, 423-434 (2000).
93. Boer EF, Van Hollebeke HF, Park S, Infante CR, Menke DB, Shapiro MD. Pigeon foot feathering reveals conserved limb identity networks. *Dev Biol* **454**, 128-144 (2019).
94. Don EK, Currie PD, Cole NJ. The evolutionary history of the development of the pelvic fin/hindlimb. *J Anat* **222**, 114-133 (2013).
95. Brand Z, Cloete SWP, Malecki IA, Brown CR. Ostrich (Struthio camelus) embryonic development from 7 to 42 days of incubation. *Br Poult Sci* **58**, 139-143 (2017).
96. Hamburger V, Hamilton HL. A series of normal stages in the development of the chick embryo. *Developmental Dynamics* **195**, 231-272 (1992).
97. Tokita M, Kuratani S. Normal Embryonic Stages of the Chinese Softshelled Turtle Pelodiscus sinensis (Trionychidae). *jzoo* **18**, 705-715 (2001).
98. Nye HLD, Cameron JA, Chernoff EAG, Stocum DL. Extending the table of stages of normal development of the axolotl: limb development. *Dev Dyn* **226**, 555-560 (2003).
99. Chomczynski P, Mackey K. Short technical reports. Modification of the TRI reagent procedure for isolation of RNA from polysaccharide- and proteoglycan-rich sources. *Biotechniques* **19**, 942-945 (1995).
100. Ovchinnikov D. Alcian blue/alizarin red staining of cartilage and bone in mouse. *Cold Spring Harb Protoc* **2009**, db.prot5170 (2009).
101. Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat Biotechnol* **37**, 907-915 (2019).
102. Bedre R. Gene expression units explained: RPM, RPKM, FPKM, TPM, DESeq, TMM, SCnorm, GeTMM, and ComBat-Seq. (2017).
103. Eppig JT. Mouse Genome Informatics (MGI) Resource: Genetic, Genomic, and Biological Knowledgebase for the Laboratory Mouse. *ILAR J* **58**, 17-41 (2017).
104. Kryuchkova-Mostacci N, Robinson-Rechavi M. A benchmark of gene expression tissue-specificity metrics. *Brief Bioinform* **18**, 205-214 (2017).
105. Emms DM, Kelly S. OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biology* **16**, 1-14 (2015).
106. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139-140 (2010).
107. Johnson KR, Sweet HO, Donahue LR, Ward-Bailey P, Bronson RT, Davisson MT. A new spontaneous mouse mutation of Hoxd13 with a polyalanine expansion and phenotype similar to human synpolydactyly. *Hum Mol Genet* **7**, 1033-1038 (1998).
108. Lanctôt C, Moreau A, Chamberland M, Tremblay ML, Drouin J. Hindlimb patterning and mandible development require the Ptx1 gene. *Development* **126**, 1805-1810 (1999).
109. Baldarelli RM*, et al.* The mouse Gene Expression Database (GXD): 2021 update. *Nucleic Acids Res* **49**, D924-D931 (2021).
110. Suzuki R, Shimodaira H. Pvclust: an R package for assessing the uncertainty in hierarchical clustering. *Bioinformatics* **22**, 1540-1542 (2006).

111. Yang J. Tutorial: Use TreeExp for Phylogenetic Transcriptome Analysis. (2019).
112. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550 (2014).
113. Pease JB*, et al.* Layered evolution of gene expression in "superfast" muscles for courtship. *Proceedings of the National Academy of Sciences* **119**, e2119671119 (2022).
114. Pagel M, Meade A, Barker D. Bayesian estimation of ancestral character states on phylogenies. *Syst Biol* **53**, (2004).
115. Zheng GXY*, et al.* Massively parallel digital transcriptional profiling of single cells. *Nature Communications* **8**, 14049 (2017).
116. Hao Y*, et al.* Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573-3587.e3529 (2021).
117. Zhou Y*, et al.* Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun* **10**, 1523 (2019).
118. Purushothaman S, Elewa A, Seifert AW. Fgf-signaling is compartmentalized within the mesenchyme and controls proliferation during salamander limb development. *Elife* **8**, (2019).
119. Nacu E, Gromberg E, Oliveira CR, Drechsel D, Tanaka EM. FGF8 and SHH substitute for anterior-posterior tissue interactions to induce limb regeneration. *Nature* **533**, 407-410 (2016).
120. Furukawa S, Yamamoto S, Kashimoto R, Morishita Y, Satoh A. Variable Shh and Fgf8 positioning in regenerating axolotl limb guarantees consistent limb morphogenesis in different limb sizes. In: *bioRxiv*) (2022).
121. Chatterjee S, Kraus P, Lufkin T. A symphony of inner ear developmental control genes. *BMC Genet* **11**, 68 (2010).
122. Tripathi PP, Bozzi Y. The role of dopaminergic and serotonergic systems in neurodevelopmental disorders: a focus on epilepsy and seizure susceptibility. *Bioimpacts* **5**, 97-102 (2015).
123. Alfaro AC, Roberts B, Kwong L, Bijlsma MF, Roelink H. Ptch2 mediates the Shh response in Ptch1-/- cells. *Development* **141**, 3331-3339 (2014).
124. Nagai H, Mak S-S, Weng W, Nakaya Y, Ladher R, Sheng G. Embryonic development of the emu, Dromaius novaehollandiae. *Dev Dyn* **240**, 162-175 (2011).
125. Bai S*, et al.* Comparison of embryonic development, from HH21 to HH40, between ostrich (Struthio camelus) and chicken (Gallus gallus). *Dev Dyn* **252**, 668-681 (2023).
126. Galli A*, et al.* Distinct roles of Hand2 in initiating polarity and posterior Shh expression during the onset of mouse limb bud development.
127. Pearse RV, 2nd, Vogan KJ, Tabin CJ. Ptc1 and Ptc2 transcripts provide distinct readouts of Hedgehog signaling activity during chick embryogenesis. *Dev Biol* **239**, 15-29 (2001).
128. Damle S, Davidson EH. Precise cis-regulatory control of spatial and temporal expression of the alx-1 gene in the skeletogenic lineage of s. purpuratus. *Dev Biol* **357**, 505-517 (2011).
129. Takahashi M*, et al.* The role of Alx-4 in the establishment of anteroposterior polarity during vertebrate limb development. *Development* **125**, 4417-4425 (1998).
130. Kraus P, Lufkin T. Dlx homeobox gene control of mammalian limb and craniofacial development. *Am J Med Genet A* **140**, 1366-1374 (2006).
131. Wang C-KL*, et al.* Function of BMPs in the apical ectoderm of the developing mouse limb. *Dev Biol* **269**, 109-122 (2004).
132. LeClair EE, Bonfiglio L Fau - Tuan RS, Tuan RS. Expression of the paired-box genes Pax-1 and Pax-9 in limb skeleton development.
133. Quinn ME, Haaning A, Ware SM. Preaxial polydactyly caused by Gli3 haploinsufficiency is rescued by Zic3 loss of function in mice. *Hum Mol Genet* **21**, 1888-1896 (2012).

134. Krause A*, et al.* Tbx5 and Tbx4 transcription factors interact with a new chicken PDZ-LIM protein in limb and heart development. *Dev Biol* **273**, 106-120 (2004).
135. Gibson-Brown JJ*, et al.* Evidence of a role for T-box genes in the evolution of limb morphogenesis and the specification of forelimb/hindlimb identity. *Mech Dev* **56**, 93-101 (1996).
136. Szeto DP*, et al.* Role of the Bicoid-related homeodomain factor Pitx1 in specifying hindlimb morphogenesis and pituitary development. *Genes Dev* **13**, 484-494 (1999).
137. Butterfield NC, Qian C, Logan MPO. Pitx1 determines characteristic hindlimb morphologies in cartilage micromass culture. *PLoS One* **12**, e0180453 (2017).
138. Golzio C*, et al.* ISL1 directly regulates FGF10 transcription during human cardiac outflow formation. *PLoS One* **7**, e30677 (2012).
139. Itou J*, et al.* Islet1 regulates establishment of the posterior hindlimb field upstream of the Hand2-Shh morphoregulatory gene network in mouse embryos. *Development* **139**, 1620-1629 (2012).
140. Chen F, Capecchi MR. Targeted mutations in hoxa-9 and hoxb-9 reveal synergistic interactions. *Dev Biol* **181**, 186-196 (1997).
141. Xu B, Wellik DM. Axial Hox9 activity establishes the posterior field in the developing forelimb. *Proc Natl Acad Sci U S A* **108**, 4888-4891 (2011).
142. Hostikka SL, Gong J, Carpenter EM. Axial and appendicular skeletal transformations, ligament alterations, and motor neuron loss in Hoxc10 mutants. *Int J Biol Sci* **5**, 397-410 (2009).
143. Antin PB, Yatskievych TA, Davey S, Darnell DK. GEISHA: an evolving gene expression resource for the chicken embryo. *Nucleic Acids Res* **42**, D933-D937 (2014).
144. Pineault KM, Song JY, Kozloff KM, Lucas D, Wellik DM. Hox11 expressing regional skeletal stem cells are progenitors for osteoblasts, chondrocytes and adipocytes throughout life. *Nat Commun* **10**, 3168 (2019).
145. Fernandez-Guerrero M*, et al.* Mammalian-specific ectodermal enhancers control the expression of Hoxc genes in developing nails and hair follicles. *Proc Natl Acad Sci U S A* **117**, 30509-30519 (2020).
146. Huang D, Chen SW, Gudas LJ. Analysis of two distinct retinoic acid response elements in the homeobox gene Hoxb1 in transgenic mice. *Dev Dyn* **223**, 353-370 (2002).
147. Bel-Vialar S, Itasaki N, Krumlauf R. Initiating Hox gene expression: in the early chick neural tube differential sensitivity to FGF and RA signaling subdivides the HoxB genes in two distinct groups. *Development* **129**, 5103-5115 (2002).
148. Minguillon C, Nishimoto S, Wood S, Vendrell E, Gibson-Brown JJ, Logan MPO. Hox genes regulate the onset of Tbx5 expression in the forelimb. *Development* **139**, 3180-3188 (2012).
149. Kim Y-S, Lewandoski M, Perantoni AO, Kurebayashi S, Nakanishi G, Jetten AM. Identification of Glis1, a novel Gli-related, Kruppel-like zinc finger protein containing transactivation and repressor functions. *J Biol Chem* **277**, 30901-30913 (2002).

Acknowledgments

We thank Jingwen Yang for providing advice and guidance on the application of *TreeExp*.

Funding

Qi Zhou is supported by the National Key Research and Development Program of China (2023YFA1800500, 2024YFA1802500), National Natural Science Foundation of China

(32170415) and Fundamental Research Funds for the Central Universities (226-2024-00055) from Zhejiang University.

Author Contributions Statement: W.K. and Q.Z. conceived this project, with assistance from G.W. and C.L.; W.K. provided overall project leadership, collected and dissected and obtained RNAs from animal samples; J.-F.F. contributed to the collection of axolotl samples; W.K. performed all the bioinformatics analyses and designed the probes and performed in situ hybridization experiments; W.K., Q.Z., G.W., J.-F.F., and C.L. wrote and revised the manuscript; and all authors edited the manuscript and approved the final draft.

Competing Interests Statement: The authors declare to have no competing interests.

Figure Legends/Captions (for main text figures)

Fig. 1. Development of tetrapod digits and molecular evidence for the frameshift and MAD hypotheses. a We studied six tetrapod species (chicken, ostrich, emu, Siameses crocodile, Chinese soft-shell turtle, and axolotl) for their limb bud and individual developing digit in forelimbs and hindlimbs at three corresponding embryonic stages. Cartilage staining with Alcian blue of each collected digit sample is shown here. **b** Hierarchical clustering of transcriptomes of individual digits of three bird species and axolotl with those of turtle and crocodile. We label the clustering patterns that indicate a frameshift by boxes, and show nodes with bootstrap values of 80 or lower, after 1000 bootstraps. **c** Percentage of overlapping MFGs between the frameshifted digits of birds/axolotl and the other digits of turtle or crocodile to the overall MFGs in each digit. **d** Normalized expression levels of MADness orthologous genes in mouse, crocodile, chicken, ostrich and axolotl, each of which shows downregulation (the upper panel) or upregulation (the lower panel) relative to other digits. All animal icons were created by the authors.

Fig 2. Increased gene expression disparity between forelimb and hindlimb in birds. a We show the scaled rates of gene expression evolution in limb buds of birds and crocodiles, taking the turtle as an outgroup on the phylogenetic branches. Forelimb bud biased genes (red) and

hindlimb bud biased genes (blue) are shown on the respective branch. Radar plots show for each species the number of forelimb and hindlimb biased genes for each digit in the early (dashed lines) and late (solid lines) stages. **b** We show the numbers of nonredundant FL-HL DEGs (two-sided Wald test, fold-change > 2, adjusted $P < 0.05$) during development for the six studied tetrapod species, and limb bud stage of mouse. **c** Heatmap of Log2 fold-change differences of gene expression level between forelimb- or hindlimb bud, with forelimb biased gene names shown in red, hindlimb biased gene names shown in blue. Colored arrow labels indicate genes with significant branch-specific expression changes relative to outgroup orthologs (Wilcoxon test, adjusted $P < 0.05$): red labels indicate significant changes in forelimb expression in the focal lineage (upward arrow: higher expression; downward arrow: lower expression); blue labels indicate significant changes in hindlimb expression (upward arrow: higher expression; downward arrow: lower expression). **d** Heatmap of Log2 fold-change differences of gene expression levels between each corresponding digit (D1, D3, and D4) of forelimbs and hindlimbs of each species in early and late stage digits. DEGs shared between the limb bud stage and digit stages are highlighted in bold. Forelimb-biased gene names are shown in red, hindlimb biased gene names are shown in blue, with genes that are specifically upregulated relative to outgroups marked with an upward arrow, and those downregulated with a downward arrow. Both outgroups in **b** and **c** are defined hierarchically: Archosauria-level comparisons use non-archosaurian reptiles as the outgroup; Aves-level comparisons use non-avian archosaurs as the outgroup.

Fig. 3. Molecular fingerprint genes across tetrapods. a The phylogenetic distribution of MFGs across tetrapods. The tree panel shows the scaled color branches to the summed number of MFGs from D1 to D4. D5 is not counted here given its degeneration in archosaurs. Limb-related MFGs for each digit are shown on the corresponding branches. **b** Reported anterior (upper panel) of posterior (lower panel) biased genes in the fin bud show the same bias among the tetrapod digits. Density plots show the normalised gene expression levels (TPM) of orthologs of these anterior/posterior fin bud genes[44, 73, 76, 77, 78] in all of the tetrapod anterior (D1,2) digits and posterior (D4,5) digits. **c** Gene Ontology enrichment of MFGs assigned to each phylogenetic branch shown in a. The size of the circles is scaled to the number of genes enriched in each GO

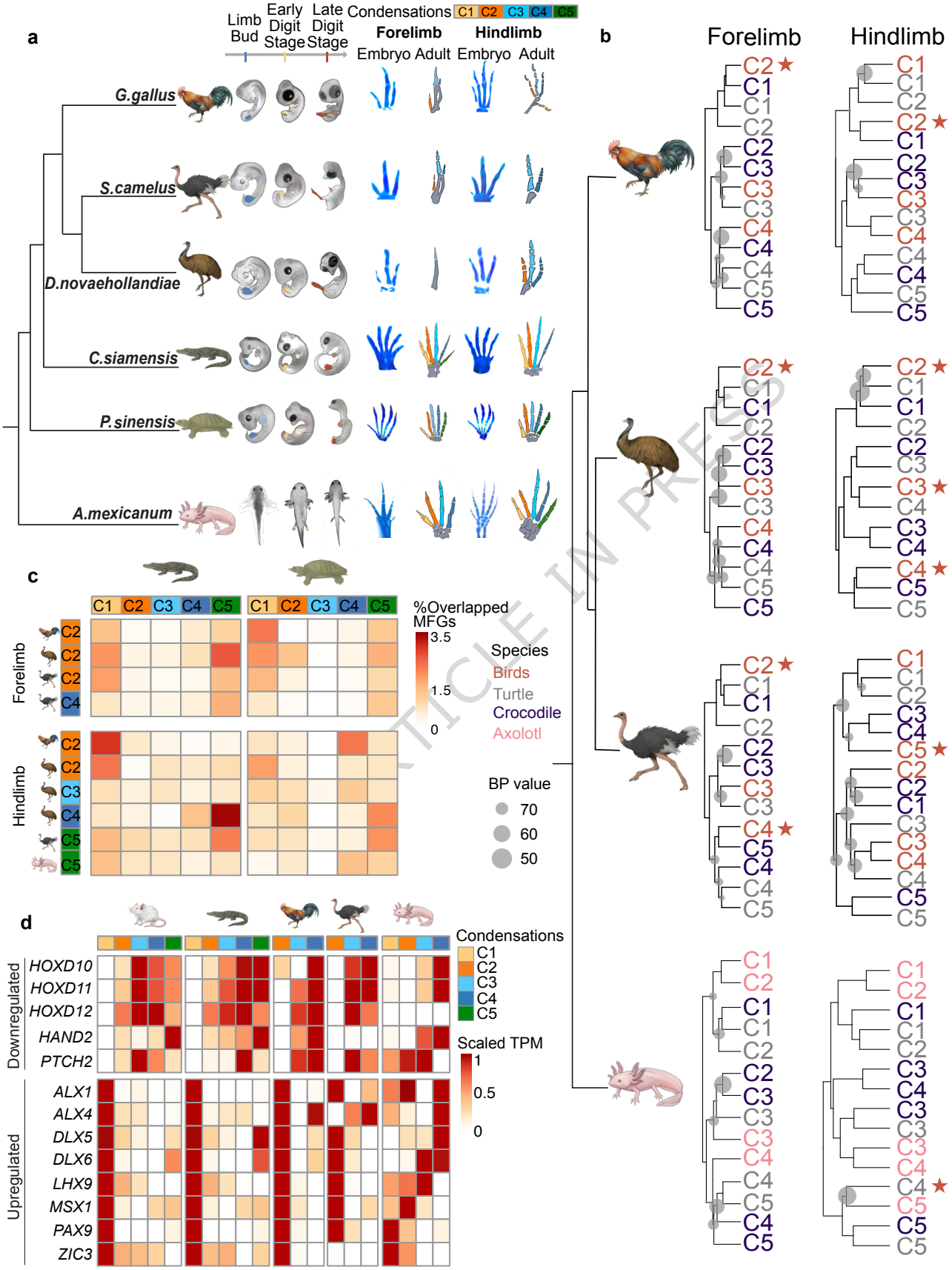
term, while the color indicates the negative log10 p-value. **d** UMAP visualization of E10.5 mouse forelimb bud scRNA-seq data, colored by the average normalized expression of anterior and posterior marker genes. **e** UMAP visualization of E10.5 mouse forelimb bud scRNA-seq data of average expression for several tetrapod-wide anterior or posterior MFGs.

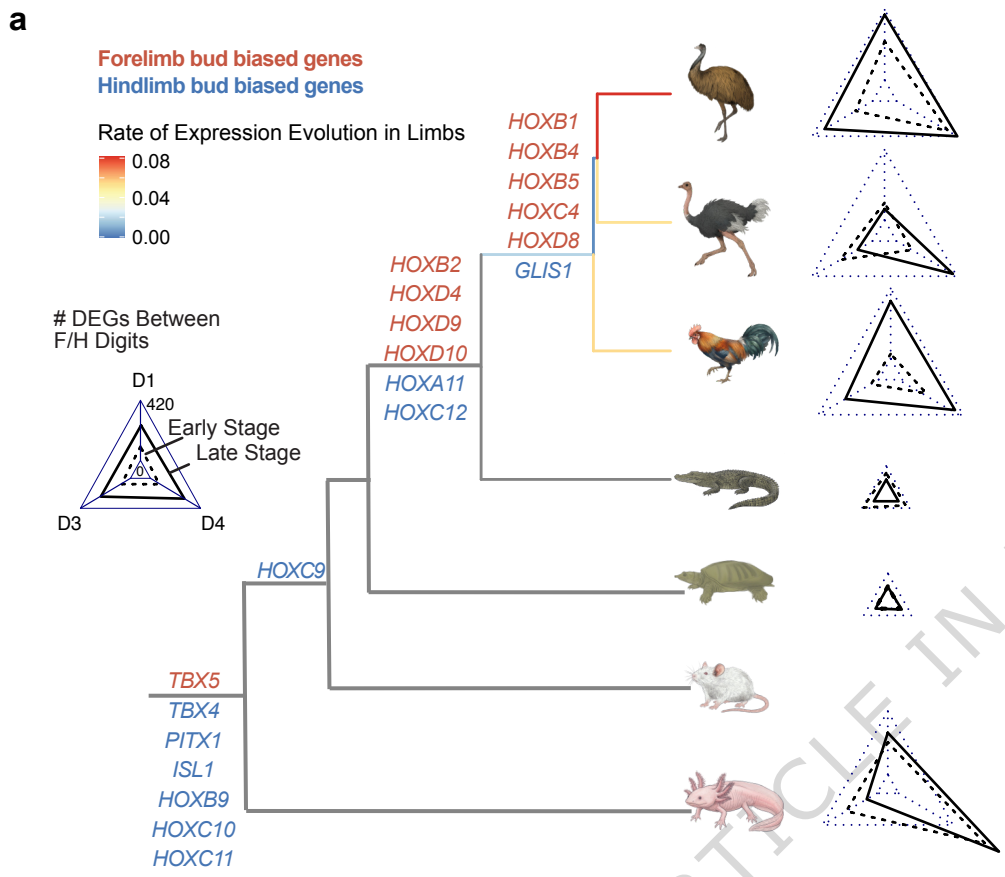
Fig. 4. Evolution of appendage and digit identity gene networks. We identify components of these two inferred ancient networks and show them at the ancestor branch of jawed vertebrates, based on their conserved expression in the forelimb/hindlimb or in the anterior/posterior side of limb bud or digit shown in Fig. 2 and 3. Some genes might be added into the network during the subsequent divergence of tetrapods (e.g., *Isl1*, *Hoxc11*, *Msx1*), or that of Aves (e.g., those related feather development).

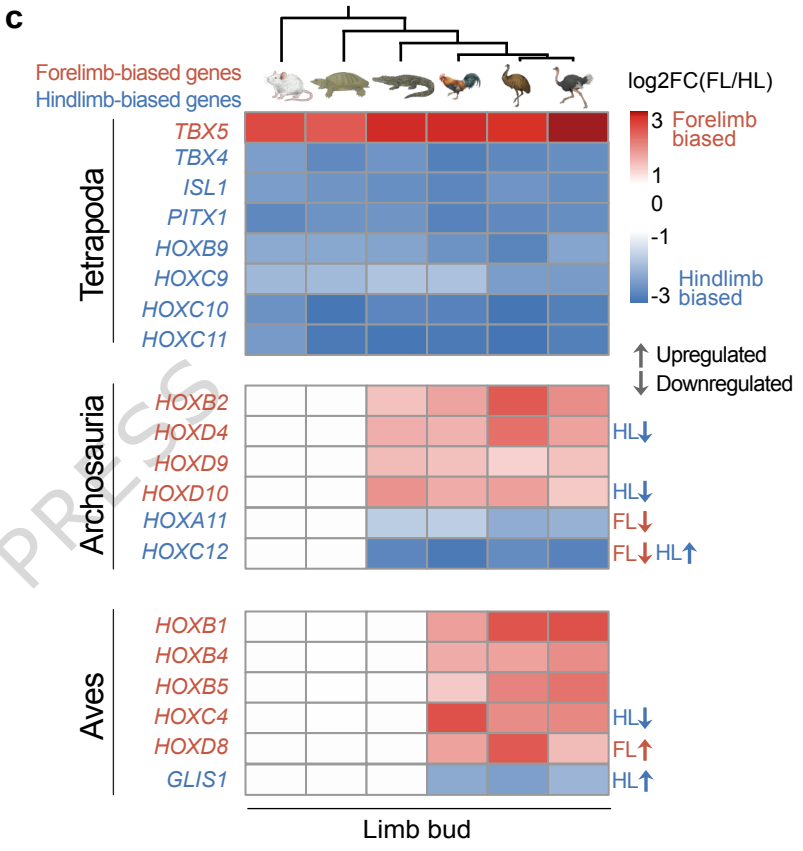
Editorial Summary:

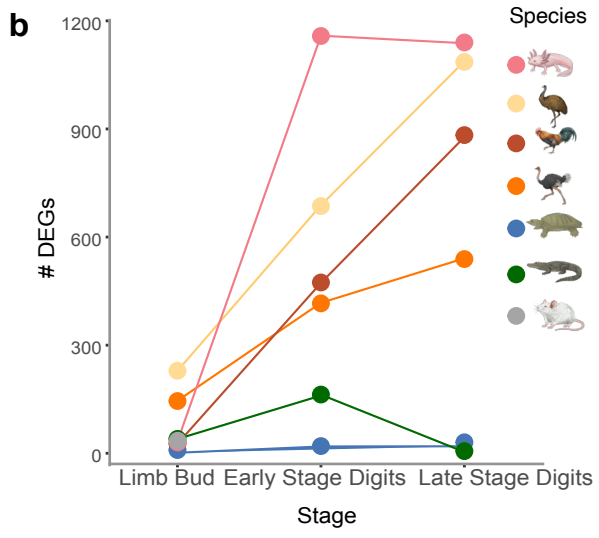
Comparative transcriptomics of digits from six tetrapod species reveals that bird wing digits are homologous to reptilian digits 1, 3, 4, and that bipedalism relaxed developmental constraints, enabling diversification of avian forelimb digits.

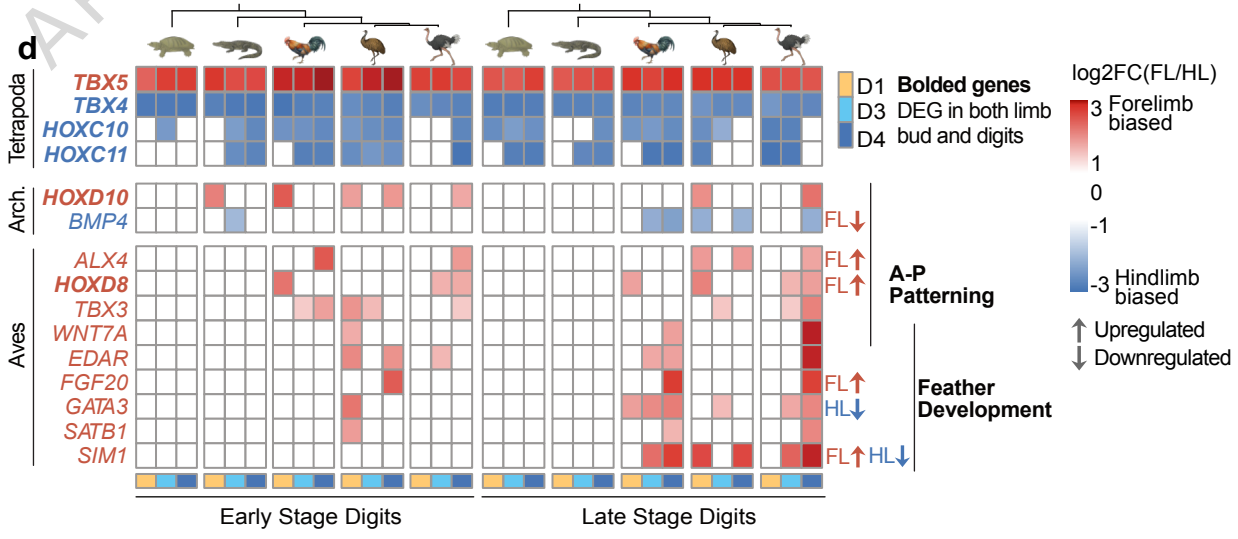
Peer review information: *Nature Communications* thanks the anonymous reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

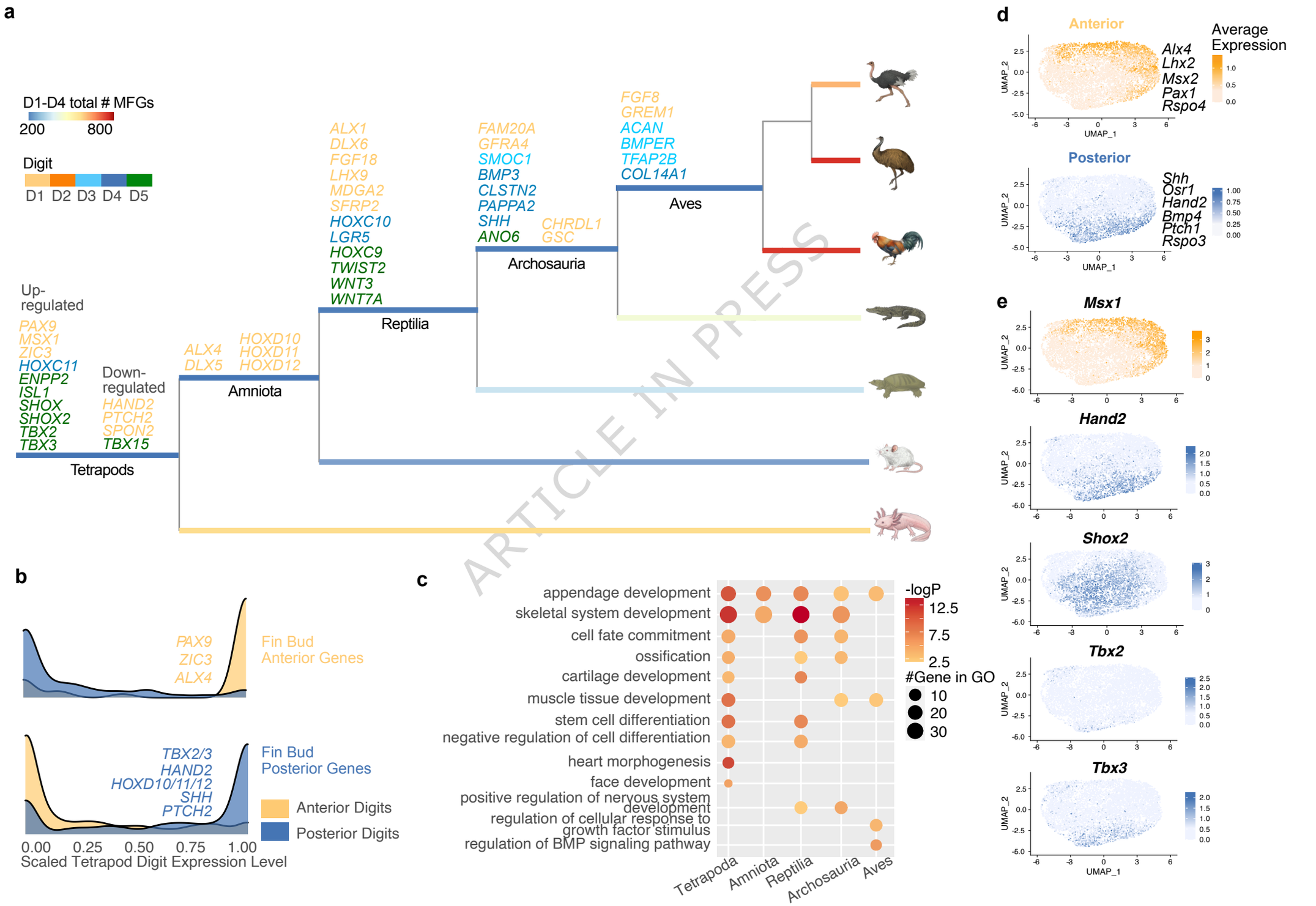

a
Limb Bud
Early Digit Stage
Late Digit Stage
Condensations
C1 C2 C3 C4 C5
Forelimb
Hindlimb
Embryo Adult Embryo Adult
G.gallus
S.camelus
D.novaehollandiae
C.siamensis
P.sinensis
A.mexicanum
b
Forelimb
Hindlimb
c
%Overlapped MFGs
3.5
1.5
0
Species
Birds
Turtle
Crocodile
Axolotl
BP value
70
60
50
Forelimb
Hindlimb
d
Condensations
C1
C2
C3
C4
C5
Scaled TPM
1
0.5
0
Downregulated
HOXD10
HOXD11
HOXD12
HAND2
PTCH2
Upregulated
ALX1
ALX4
DLX5
DLX6
LHX9
MSX1
PAX9
ZIC3

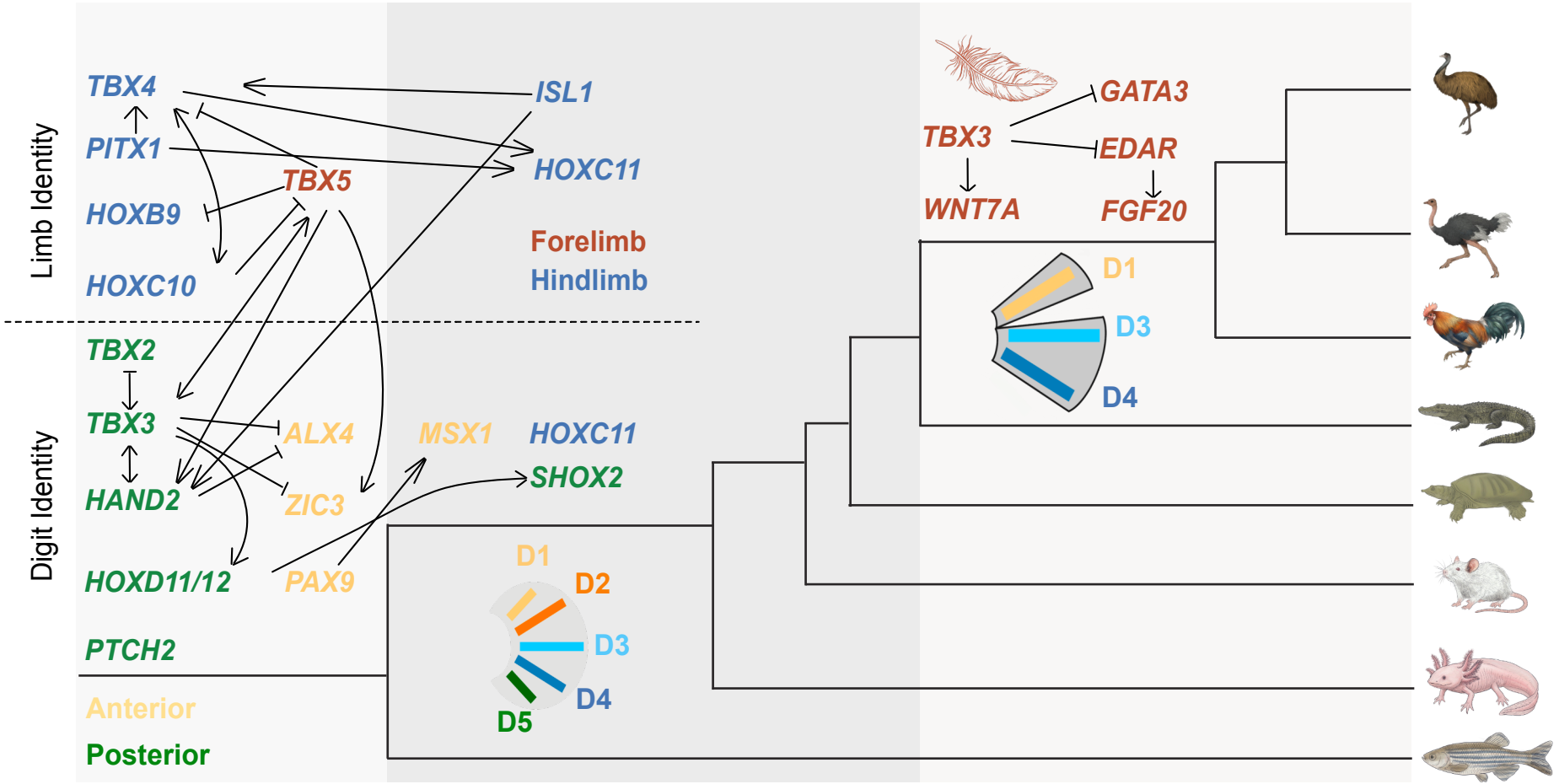











Limb Identity
Digit Identity
TBX4
PITX1
HOXB9
HOXC10
TBX5
ISL1
HOXC11
Forelimb
Hindlimb
TBX2
TBX3
HAND2
ALX4
ZIC3
MSX1
HOXC11
SHOX2
HOXD11/12
PAX9
PTCH2
Anterior
Posterior
D1
D2
D3
D4
D5
GATA3
TBX3
EDAR
WNT7A
FGF20
D1
D3
D4