Evidence against tetrapod-wide digit identities and for a limited frame shift in bird wings

Thomas A. Stewart[1,2]*, Cong Liang[3], Justin L. Cotney[4], James P. Noonan[5], Thomas J. Sanger[6], and Günter P. Wagner[1,3]*

[1]Department of Ecology and Evolutionary Biology, Yale University
[2]Minnesota Center for Philosophy of Science, University of Minnesota
[3]Systems Biology Institute, Yale University
[4]Department of Genetics and Genome Sciences, UConn Health
[5]Department of Genetics, Yale School of Medicine
[6]Department of Biology, Loyola University in Chicago
*To whom correspondence should be addressed

Summary: In crown group tetrapods, individual digits are homologized in relation to a pentadactyl ground plan. However, testing hypotheses of digit homology is challenging because it is unclear whether digits develop with distinct and conserved gene regulatory states. Here we show dramatic evolutionary dynamism in the gene expression profiles of digits, challenging the notion that five digit identities are conserved across amniotes. Transcriptomics of developing limbs shows diversity in the patterns of genetic differentiation of digits, although the anterior-most digit has a unique, conserved expression profile. Further, we identify a core set of transcription factors that are differentially expressed among the digits of amniote limbs; their spatial expression domains, however, vary between species. In light of these results, we reevaluate the frame shift hypothesis of avian wing evolution and conclude that only the anterior-most digit has shifted position, suggesting a 1,3,4 digit identity in the bird wing.

Limbs evolved from paired fins in the Late Devonian, and early tetrapods possessed more than five digits on the fore- and hindlimbs[1,2]. Later in the tetrapod stem, a pentadactyl pattern stabilized as the ground plan for the limb. Individual digits are homologized between species and between fore- and hindlimbs in reference to this pentadactyl ground plan[3]. However, it remains controversial whether such hypotheses of identity correspond to distinct developmental programs among the digits (developmental identities), or just the relative position of digits along the limb's anteroposterior axis (positional identities)[4-7].

The anterior-most digit (D1) (*e.g.,* human thumb) appears to have a distinct developmental identity as compared to the more posterior digits (D2-D5). D1 is marked by a

unique gene expression profile—low expression of *HoxD11* and *HoxD12* and high expression of *Zic3* relative to other digits[7-9]—and it appears able to develop independently of *Shh* signaling[9-11]. Additionally, an analysis of morphological variation in primates identified a high degree of variational independence of D1 relative to the more posterior digits[12]. Models of posterior digit identity have been proposed according to the relative exposure of limb bud mesenchymal cells to *Shh,* which emanates from the zone of polarizing activity prior to digit condensation[10,11]. However, broadly conserved marker genes for individual posterior digits have not been identified in the interdigital mesenchyme, a signaling center that patterns digits[13,14]. For instance, while the combinatorial expression of *Tbx2* and *Tbx3* is necessary to generate the phenotypes of D3 and D4 in chicken hindlimb[15], it is questionable whether these developmental identities are conserved in other species, like mouse, with limited morphological differentiation of the posterior digits.

Debates of digit homology are especially challenging to resolve when limbs have fewer than five digits. This problem has been most actively investigated in the tridactyl avian wing, because of an apparent conflict between paleontological and developmental data[16]. The fossil record of theropod dinosaurs shows a clear pattern of reduction of the posterior two digits in the lineage leading to birds, yet digits in the wing have been described as developing in the middle three positions of a pentadactyl developmental groundplan[17-21]. To explain this discrepancy, the frame shift hypothesis was proposed[16]. It posited that a homeotic shift occurred in the avian stem such that the developmental programs that were once expressed in D1, D2, and D3 are now executed in the digits that develop in positions D2, D3, and D4 respectively. Comparative analyses of gene expression have found support for this hypothesis: *in situ* hybridization and transcriptomics have revealed similarity between the anterior digit of the adult avian wing, which develops in position 2, and D1 of other limbs[7,22].

Studies that aim to test digit homology assume that developmental identities (1) were present in a common ancestor, (2) are conserved among the descendent lineages, and (3) are reflected in gene expression profiles. Here we present comparative transcriptomic data from five species that challenge these assumptions among amniotes by documenting a surprising diversity of digital gene expression profiles. Analyses further reveal a core set of transcription factor genes differentially expressed among digits and suggest a new model for the evolution of the bird wing.

Disparity in digit expression profiles

To characterize the gene expression profiles of digits in pentadactyl amniote limbs, we sequenced RNA of developing digits and their associated posterior interdigital mesenchyme from the forelimbs of mouse, green anole (*Anolis*), and American alligator (Fig. 1 a). In each species, hierarchical cluster analysis (HCA) and principal component analysis (PCA) of the transcriptomes shows a weak signal of sample clustering by digit (Extended Data Fig. 1). The strongest signals of digit-specific expression profiles are observed in D1 of mouse and D4 of alligator. Groupings of

77 the other digit samples are not well supported. We interpret this result to imply that any signal of
78 gene expression differentiation among digits is overwhelmed by noise when all genes are
79 considered, because most genes are likely irrelevant to the developmental identity of digits. If
80 such a signal exists, we predict that it will be reflected preferentially in the expression of
81 transcription factor and signaling genes. Therefore, we again performed HCA and PCA on the
82 samples of each species, this time using two gene lists: a curated set of known limb patterning
83 genes that are sensitive to *Shh* signaling (N=159)[23], and transcription factor genes (N=2138)[24].
84 In mouse and alligator, HCA and PCA of known limb patterning genes results in
85 clustering of samples by digit (Fig. 1 b, c). In mouse, D1 is strongly differentiated from the other
86 digits. In alligator, an anterior cluster, comprised of digits D1, D2, and D3, is differentiated from a
87 posterior cluster, comprised of D4 and D5. By contrast, analysis of known limb patterning genes
88 in *Anolis* shows weak clustering of samples by digits (Fig. 1 d). This suggests a level of
89 homogeneity among *Anolis* digits that is not observed in either mouse or alligator. Analysis of all
90 transcription factors for these species yields comparable results to what is recovered for limb
91 patterning genes, but with generally lower adjusted uncertainty values in HCAs (Extended Data
92 Fig. 2).
93 To further test the hypothesis that there is limited gene expression differentiation among
94 *Anolis* digits as compared to the other pentadactyl limbs, we took advantage of a result from
95 multiple testing theory[25]: If a differential expression analysis is conducted on two sample types
96 that are not genetically differentiated, then the resultant frequency distribution of *p* values will be
97 uniform within the [0, 1] interval. On the other hand, if there are truly differentially expressed
98 genes among the compared sample types, then the *p* value distribution is expected to be biased
99 towards p=0. We conducted differential expression analyses of adjacent digits of the forelimbs of
100 mouse, alligator, and *Anolis* using EdgeR[26,27] and inspected *p* value distributions (Extended Data
101 Fig. 3). In *Anolis*, all comparisons of adjacent digits result in *p* value distributions that are close to
102 uniform, suggesting that there is very weak, if any, genetic differentiation of adjacent fingers. We
103 note that this result is independent of any *p*-value significance threshold. By contrast, most of
104 adjacent pairwise digit comparisons for mouse and alligator show a strongly biased *p* value
105 distribution, the exception being D2 and D3 in mouse. This is consistent with the idea that, in
106 general, most digits in a limb are genetically differentiated, while in *Anolis* genetic differentiation
107 of digits is minimal or absent.
108 Given that these limbs differ in their broad patterns of gene expression differentiation of
109 digits, we next asked whether individual genes show divergent or constrained expression patterns
110 across the forelimb in the different species. Specifically, we compared adjacent digits, identified
111 differentially expressed transcription factor genes, and then assessed which differences are
112 shared among mouse, alligator, and *Anolis*. Of the 1066 transcription factor genes that are one-
113 to-one orthologs in these three species, only four genes are differentially expressed in a

114 conserved pattern among corresponding adjacent digits (Extended Data Fig. 4). There are three
115 genes that differentiate D1 from D2 (*Hoxd11, Hoxd12*, and *Sall1)*, and there is one that
116 differentiates D4 from D5 (*Tbx15*) in all three species. No transcription factors are differentially
117 expressed in all three species between the median digits (*i.e.*, differentiating D2 from D3, or D3
118 from D4).

119 If the homogeneity observed among *Anolis* forelimb digits is a derived condition, then this
120 could limit our ability to diagnose plesiomorphic developmental identities. Therefore, we also
121 considered the chicken hindlimb, which has digits D1-D4. We reanalyzed published
122 transcriptomic data for hindlimb digits[7], mapping reads to a new chicken genome (Galgal5.0).
123 HCA and PCA of digits of the chicken hindlimb show a unique pattern of similarity as compared to
124 pentadactyl limbs: an anterior cluster, comprised of D1 and D2, is differentiated from the posterior
125 cluster, comprised of D3 and D4 (Extended Data Fig. 5). As before, we tested for differential
126 expression in adjacent digits and identified one-to-one orthologous transcription factor genes that
127 are differentially expressed at the same position between mouse and alligator forelimb and
128 chicken hindlimb (Fig. 2 a, Extended Data Fig. 6). Of the 1049 transcription factor genes, ten
129 differentiate D1 and D2 (n=10), none distinguish D2 and D3, and one (*Tbx3*) differentiates D3
130 from D4 in all three species (Fig. 2 a).

131 Overall, data from these four species do not support the hypothesis that amniote digits
132 have conserved developmental identities. The exception appears to be D1, which likely had a
133 distinct developmental program in the most recent common ancestor of amniotes. We further
134 tested whether D1 has a conserved gene expression profile by sequencing RNA from developing
135 human fore- and hindlimb, which were partitioned into D1 and the posterior digital plate (D2-5). Of
136 the ten genes identified above as differentiating D1 and D2, six show conserved patterns of
137 expression change at this position: in all limbs sampled *Hand2, Hoxd11, Hoxd12, and Tfap2b* are
138 more highly expressed in D2 than D1, and *Alx1 and Pax9* are more lowly expressed in D2 than
139 D1 (Fig. 2 b).

140 T-box family genes are predicted to regulate the identities of posterior digits[15]. Our data
141 provide some support for the hypothesis that this function is conserved across amniotes (Fig. 2
142 c). *Tbx2*, which was previously shown to regulate posterior digit identity in the chicken hindlimb[14],
143 shows divergent patterns of expression in the posterior digits of other species. *Tbx3* differentiates
144 D3 from D4 in mouse, alligator, and chicken hindlimb, and the likelihood that it was recovered by
145 chance alone is 7.2×10^{-6} (binomial test). However, it is not differentially expressed at this position
146 in *Anolis* forelimb. *Tbx15* differentiates D4 from D5 among pentadactyl limbs, and the likelihood
147 that it was recovered by chance alone is 1.9×10^{-5} (binomial test).

148

149 ***A core set of digit patterning genes***

150 Given our result that the gene expression profile of digits is evolutionarily dynamic, we
151 tested whether a conserved set of genes might pattern amniote autopods, albeit in different
152 spatial patterns. Specifically, we reanalyzed transcriptomic data of mouse, alligator and *Anolis*
153 forelimbs and chicken hindlimb, conducting ANOVA to test for genes that were differentially
154 expressed between any two digits in the limb, not just adjacent digits. This analysis recovers
155 genes that are differentially expressed between some digits in the limb, but it does not indicate
156 between which digits a gene is differentially expressed. The number of differentially expressed
157 transcription factor genes differs greatly among species: 356 in mouse, 377 in alligator, 34 in
158 *Anolis*, and 211 in chicken hindlimb (FDR < 0.05, Fig. 3 a). This is consistent with previous results
159 that showed *Anolis* forelimb to be more homogeneous than other sampled limbs. Therefore, we
160 focused on transcription factor genes that are one-to-one orthologous between mouse, alligator,
161 and chicken and identified a set of 49 genes that are differentially expressed in these three limbs
162 (Fig. 3 b). We call these *conserved differentially expressed genes* (CDEGs). The expected
163 number of overlapping genes among these sets by chance alone is 7.57, and the probability of
164 observing an overlap of 49 genes or more by chance is $<10^{-6}$ (binomial test). To assess whether
165 this gene set is biologically meaningful, we performed HCA and PCA on the samples of each
166 species using the 49 CDEGs. In *Anolis,* we considered the subset (n=42) that are one-to-one
167 orthologs across all four species. In combination, CDEGs can produce unique expression profiles
168 of each digit within a limb (Fig. 3 c) and show patterns similar to those generated by analyses of
169 known limb patterning genes (Fig. 1 b-d, Fig. 2 a).
170 Analysis of amniote limbs showed that targeted gene lists generated either
171 experimentally (*i.e.*, known limb patterning genes[23]), by gene ontology (*i.e.*, all transcription
172 factors[24]), or statistically (*i.e.*, 49 CDEGs, see above), can reveal distinct gene expression profiles
173 among digits of a limb, which are not observed in the full transcriptome. These digit specific
174 expression profiles across digits, however, are species specific. In light of these results, we
175 reevaluated the frame shift hypothesis of bird wing origin[16].
176
177 ***Reevaluating the frame shift hypothesis***
178 The frame shift hypothesis predicts that the three digits of the avian forelimb, which we
179 refer to here as D2, D3, and D4 according to their developmental position[17-21], will express the
180 developmental programs observed in the digits D1, D2, and D3 of other limbs[16]. This hypothesis
181 was tested previously by analyzing the transcriptomes of chicken fore- and hindlimb digits[7]. That
182 study found correspondence between forelimb D2 and hindlimb D1, consistent with the frame
183 shift hypothesis. However, correspondence of more posterior digits was not detected[7].
184 We re-analyzed published transcriptomic data of digits from the chicken forelimb[7] and
185 compared them to digits of the chicken hindlimb. Surprisingly, when the 49 CDEGs are
186 considered, gene expression profiles of forelimb digits D2, D3, and D4 correspond to hindlimb

digits D1, D3, and D4, respectively (Fig. 4 a). Analyses of transcription factor genes and known limb patterning genes show similar results (Extended Data Fig. 7). Similarity between the posterior two digits of the fore- and hindlimb can also be seen in the expression patterns of numerous individual genes known to be involved in digit patterning (Fig. 4 b).

To assess whether spatial gene expression profiles are conserved between the fore- and hindlimbs of a species, even when they differ in digit number, we performed *in situ* hybridization in alligator. We evaluated expression of *Tbx2*, *Tbx3*, and *Sall1*, three transcription factor genes identified as differentially expressed between alligator forelimb D3 and D4. *In situ* hybridization confirms their expression in the posterior interdigital mesenchyme (Fig. 4 c) and shows conserved positional expression patterns for *Tbx3* and *Sall1* between the forelimb and hindlimb. It is unclear whether the pattern also holds for *Tbx2*, where difference in expression level detected from RNA sequencing appear to reflect the proximodistal extent of gene expression.

We further tested the frame shift hypothesis by comparing the limbs of chicken to the pentadactyl forelimbs of other species. For each pentadactyl species, PCA were run using the CDEGs (49 in mouse and alligator, and 42 in *Anolis*), and chicken samples were projected into the reference PCA plane as supplementary observations. CDEGs were used because they can produce digit-specific expression profiles for mouse and alligator forelimb and chicken hindlimb, and because these patterns are reflective of more inclusive gene lists, described above. These projections show that the digits D2, D3, and D4 of the bird wing consistently fall into regions of the PCA plane corresponding to the digits D1, D3, and D4 of other limbs. Although it is difficult to differentiate D3 and D4 expression in all species, it is clear that D3 of the chicken forelimb does not correspond to the D2 other limbs (Fig 4 d).

It has been proposed on the basis of *Shh* expression that the digits in the avian wing are of positions D1, D2, and D3[28]. This model, like the frame-shift hypothesis, predicts wing digits have developmental identities corresponding to the digits D1, D2, D3 of other limbs. The results presented here, and also the morphological evidence for five digit condensations[17-21], suggest *Shh* expression is not a conserved marker of digit identity or position in limbs with highly reduced digit number.

Discussion

Serial homologs are repeated body parts, generated by a common developmental program. For example, in the case of digits, chondrogenic condensations are generated by a reaction-diffusion Turing-type mechanism[29,30]. Serial homologs can be developmentally identical (homomorph parts) or they can assume distinct developmental identities through the differential expression of regulatory genes (paramorph parts)[31]. The degree to which serial homologs are individuated can be difficult to assess from morphology alone, because the same developmental program can lead to different morphological outcomes depending on the developmental

environment[32,33]. However, detailed analyses of gene expression and regulation can identify developmentally individualized body parts.

In this study, we performed a comparative analysis of whole genome expression data to test the hypothesis that digits have conserved developmental identities. In interpreting our data, we acknowledge that gene expression does not demonstrate gene function. Nevertheless, a lack of differential gene expression between digits is evidence of a lack of developmental individuation, and a high level of differential gene expression is evidence for distinct gene regulatory states.

Our analyses show that patterns of regulatory gene expression in digits are evolutionarily dynamic (Fig. 5 a). This implies that the developmental identities of digits are evolving across amniotes and can be lineage-specific. The exception is the conserved developmental identity that characterizes D1 of mouse, alligator, *Anolis* and human forelimbs, and the hindlimbs of chicken and human (Fig. 2b). Our analysis also identified a core set of regulatory genes, which we call CDEGs, that are differentially expressed among digits, although species differ in which digits differentially express the genes. We propose that the CDEGs represent a “digit differentiation tool kit” deployed for the individuation of different sets of digits in different lineages, depending on the adaptive needs of the species. Twenty-eight the 49 CDEGs have known roles in patterning distal limb skeleton (Extended Data Table 1).

In *Anolis* most fingers, though they differ in number of phalanges, lack developmental individuality and, thus, appear to be homomorphic. Of the CDEGs, only 15 are differentially expressed across the *Anolis* forelimb. Thus, this homogeneity appears to be a derived condition among the taxa sampled, as it is unlikely that the other 34 CDEGs reflect homoplasy between mammals and archosaurs.

Finally, we reassessed the homology of fingers in the bird wing and obtain the novel result that the three digits in the bird wing reflect a combination of translocated digit identities and conserved identities. The anterior-most digit in the chicken wing, although it develops in position D2, is homologous to D1 of other limbs, consistent with previous studies[7,22]. The more posterior wing digits (D3 and D4) do not show evidence of translocation. These results suggest a new model for the origin of the avian wing digits (Fig. 5 b). The possibility of a 1-3-4 pattern of digit identity in the bird wing has been noted previously[34,35] on the basis of experimental studies[35]. Such models might help to reveal the mechanisms that generated the unique avian wing pattern described here.

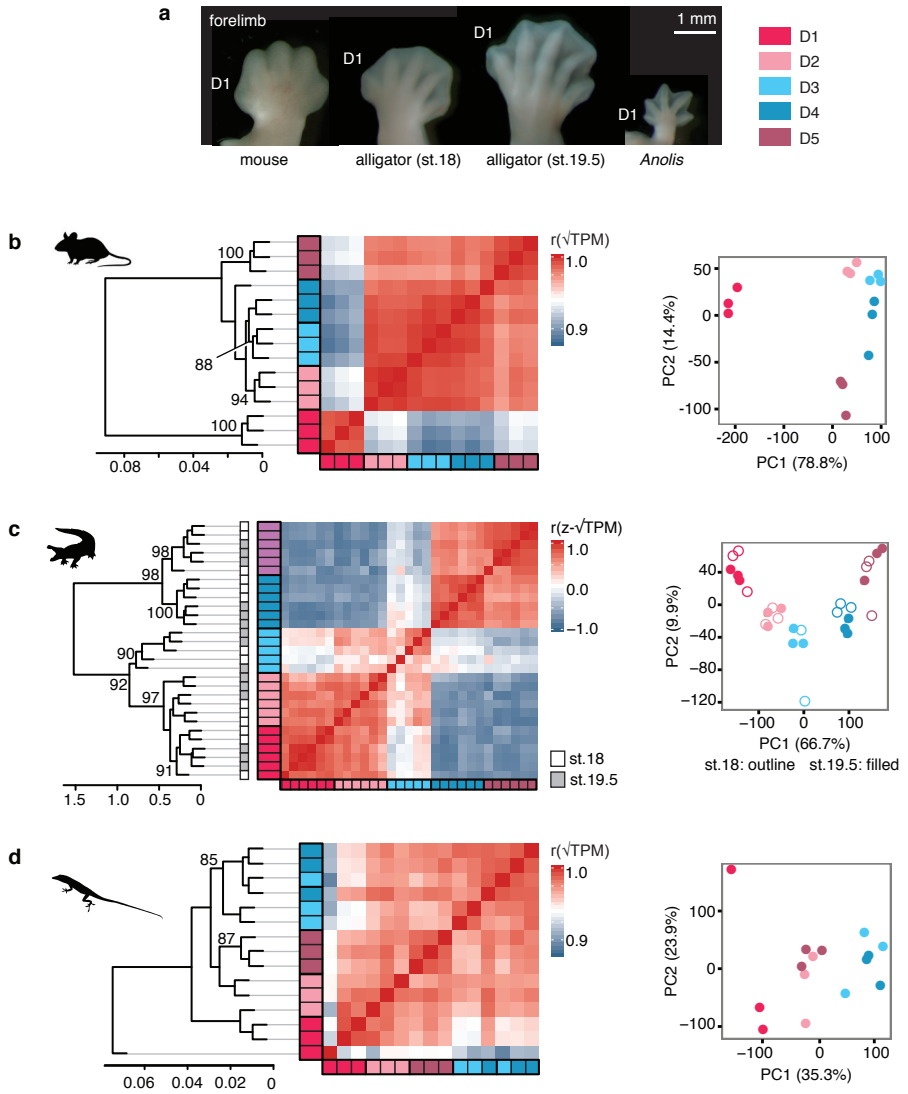


Figure 1 | Pentadactyl amniote limbs have disparate patterns of genetic differentiation of digits. (a) Photographs of right forelimbs at the stages sampled, dorsal perspective. Analyses of limb patterning genes show that in (b) mouse and (c) alligator, replicates of each digit form clusters, indicating that the digits have distinct gene expression profiles. By contrast, (c) *Anolis* digits do not show clear differentiation of gene expression profiles.

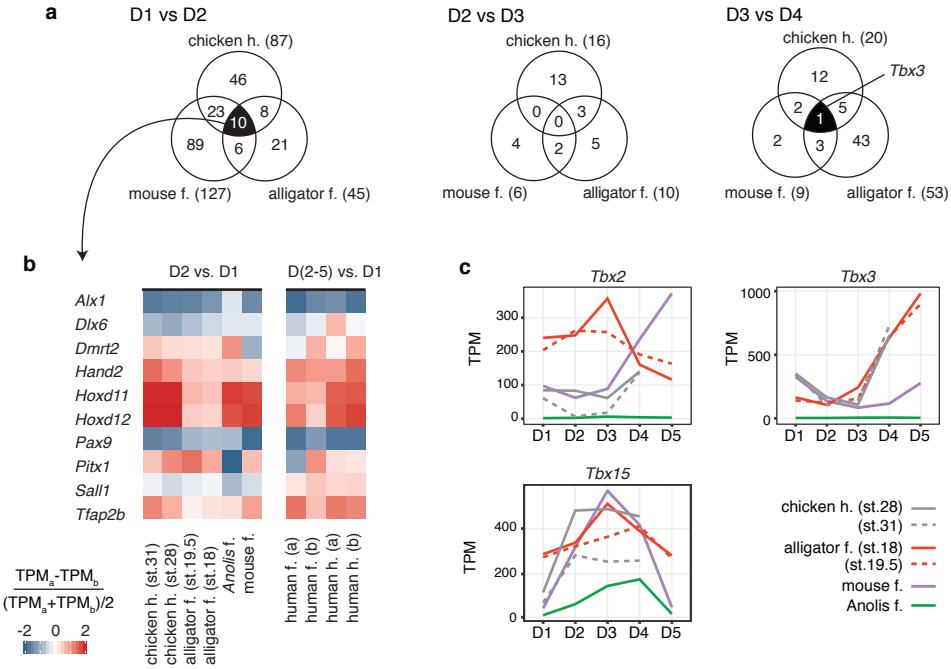


Figure 2 | D1 has a conserved, unique gene expression profile. (a) Venn diagrams of one-to-one orthologous transcription factors genes of mouse, alligator, and chicken that are differentially expressed between adjacent digits (FDR threshold of 0.05). (b) Heatmap showing relative expression of genes in D2 and D1. Human transcriptomic data provides additional support for the hypothesis that D1 has a conserved developmental identity across amniotes. (c) Expression levels of *T-box* family genes across the autopod. Data represented in transcripts per million (TPM) calculated from the gene list of one-to-one orthologous genes in mouse, alligator, *Anolis*, chicken, and human.

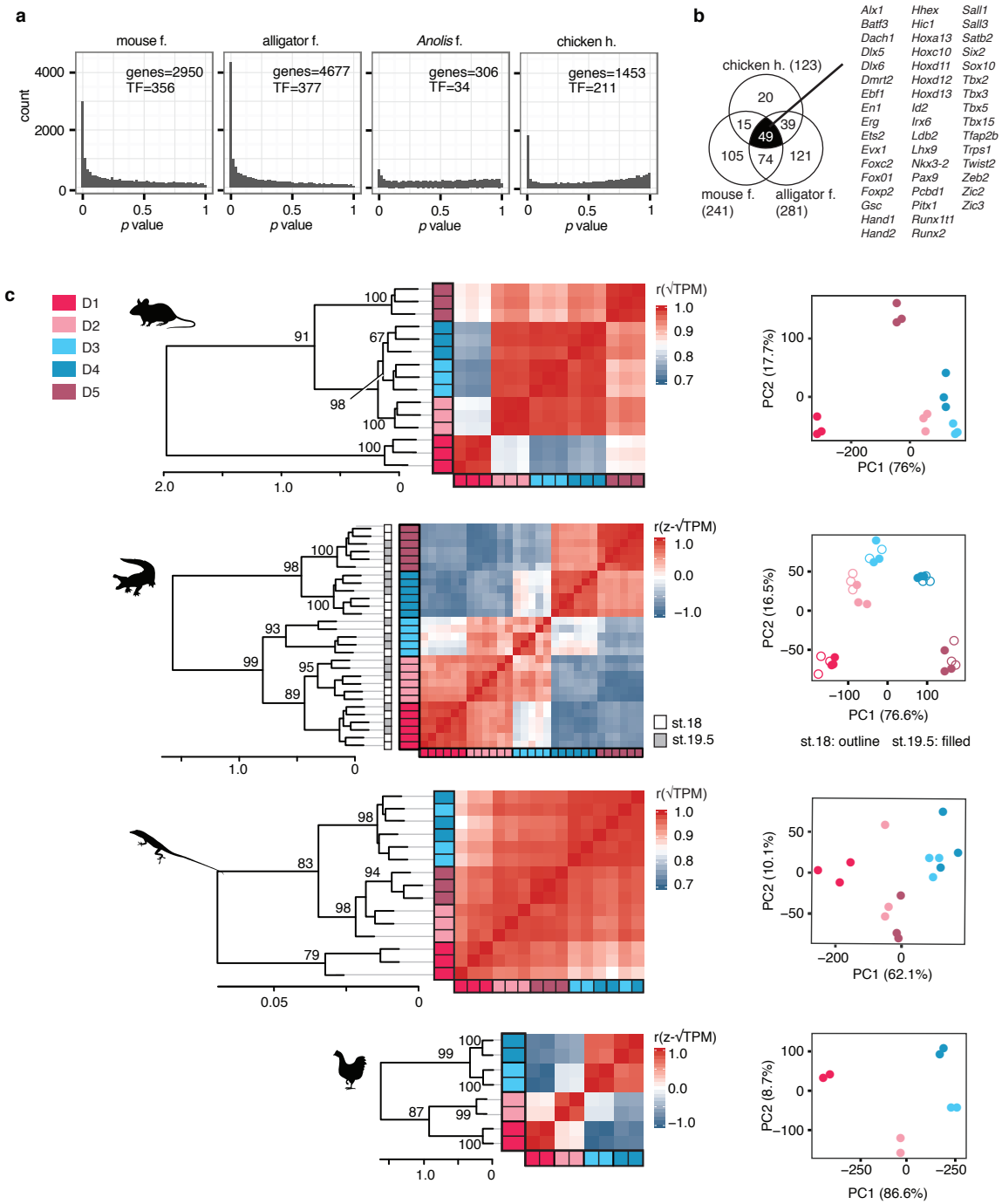


Figure 3 | A conserved core set of digit patterning genes in amniotes. (a) ANOVA comparing all combination of digits within limbs. The number of genes identified as differentially expressed at a FDR threshold of 0.05 are noted in each panel as 'genes.' The number of transcription factors among these are noted as 'TF.' (b) Venn diagram showing one-to-one orthologous transcription factor genes that are differentially expressed across mouse, alligator, and chicken limbs. (c) HCA and PCA show these genes can recover patterns of digit correspondence similar to analyses of limb patterning genes.

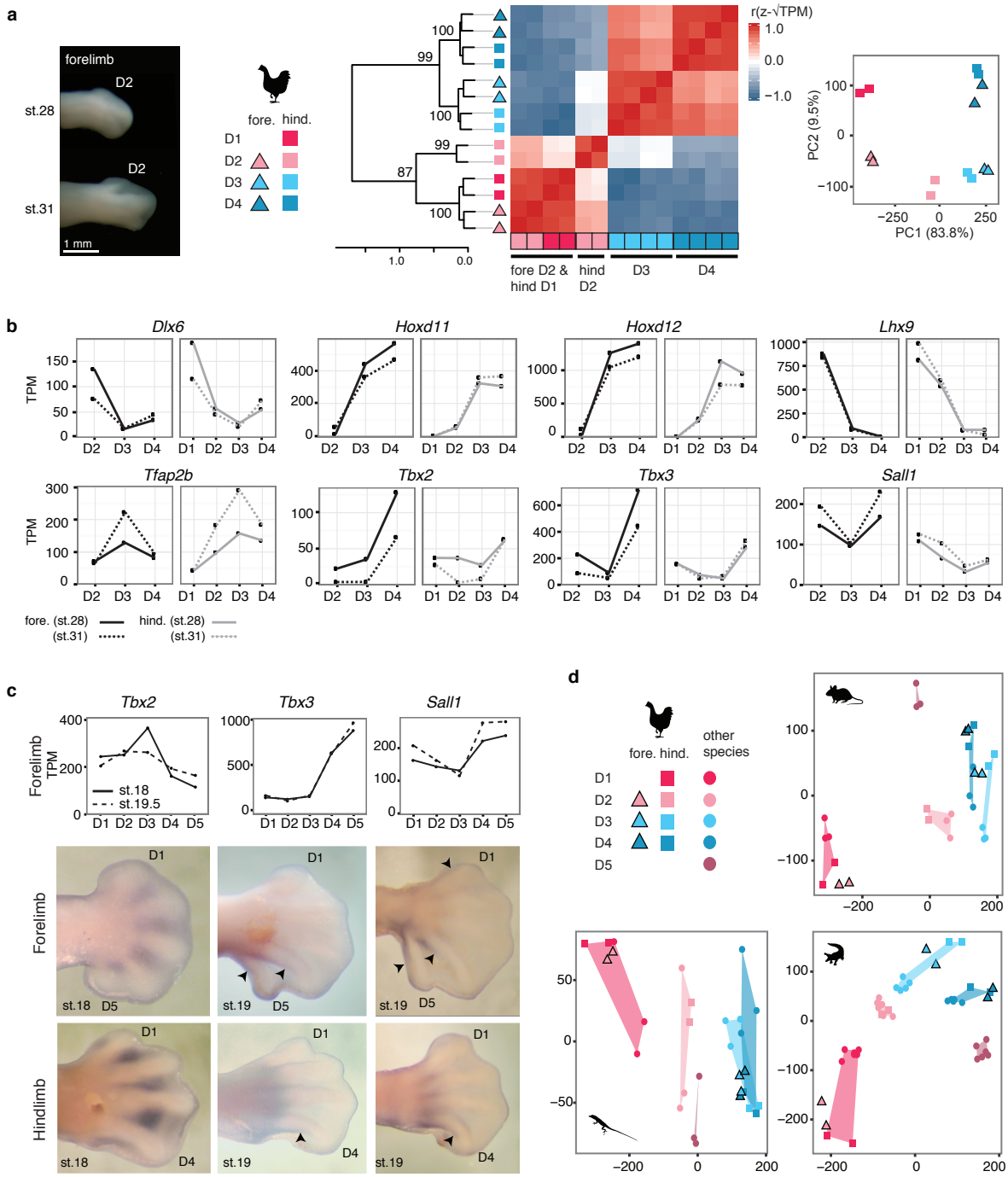


Figure 4 | Digits D2, D3, D4 of the chicken forelimb correspond to D1, D3, and D4 of other limbs. This pattern is observed by (a) HCA and PCA of CDEGs for digits of chicken wing and hindlimb. (b) Expression of individual transcription factor genes in chicken. (c) *In situ* hybridization of alligator embryos confirms that the position of differential expression of transcription factors can be conserved between fore- and hindlimbs that differ in digit number. Arrowheads indicate distal-most expression along a developing digit. (d) Projection of chicken digit data upon PCA of CDEGs for mouse, alligator, and *Anolis*.

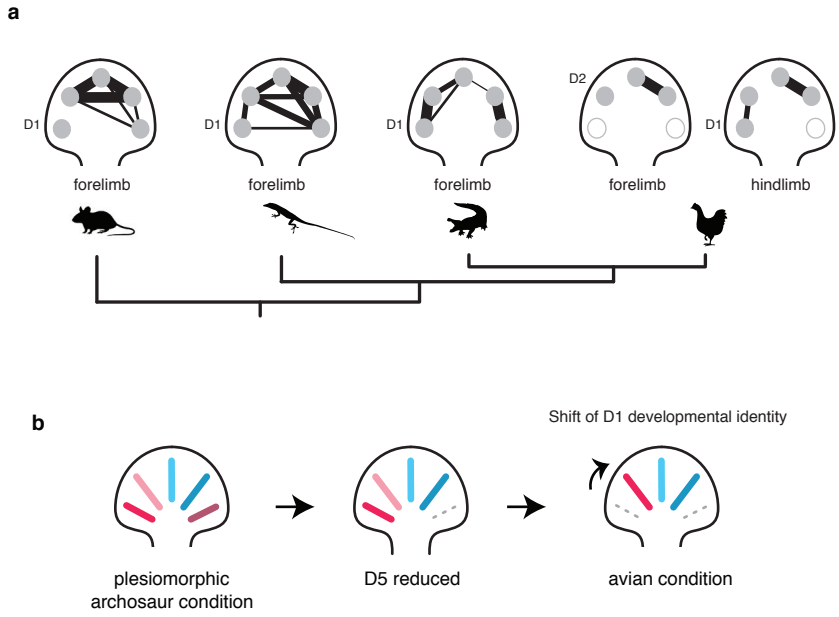


Figure 5 | The evolution of digit gene expression in amniotes is highly dynamic. (a) A phylogeny of the taxa sampled by this study and schematic graphs summarizing the relative similarity of digits within limbs, where connections and line thickness reflect degree of similarity in gene expression profiles. (b) Schematic of a limited frame shift model for evolutionary origin of the avian wing in which the developmental identity of D1 was translocated to position D2.

257 ***Methods***

258 Limbs of each species were sampled after digital condensations have formed and after inter-
259 digital webbing has begun to reduce. RNA was extracted from digits and their associated
260 posterior inter-digital webbing, a known source of digit patterning signals[13,14], following the
261 dissection strategy shown in Figure 1a of Wang *et* al.[7]. A summary of the taxonomic and tissue
262 sampling strategy is presented in Extended Data Fig. 8. Investigators were not blinded to the
263 group allocation during the experiment or when assessing outcomes.

264

265 ***Alligator mississipiensis*.** Fertilized eggs were collected from six nests of wild individuals at the
266 Rockefeller Wildlife Refuge in Grand Chenier, Louisiana (USA) in July 2015 by Dr. Ruth Elsey
267 and colleagues. Eggs were marked with pencil to indicate the side that was facing upwards in the
268 nest, so that embryos would not be injured by rotation of during transfer. Eggs were transported
269 to Yale University in mesh wire boxes containing original nesting material, and they were
270 incubated in a temperature-controlled room at 32°C. Eggs were placed on a plastic rack,
271 surrounded with original nesting material. Racks were suspended four inches above the bottom of
272 a 10 gallon aquarium. The base of the aquarium was filled with three inches of water, which was
273 heated to 90°F with a submerged aquarium heater. The top of the tank was covered with
274 plexiglass perforated with 1cm diameter holes to allow for airflow. Humidity within the tank was
275 maintained at 90%.

276 Embryos were collected at Ferguson[36] st.18 and 19.5. The left and right limbs of ten
277 individuals were dissected at each stage. For each stage, individuals sampled were of a single
278 nest and, therefore, at least half-siblings[37]. Embryos were extracted under sterile, RNAse-free
279 conditions. Individual digits and the associated posterior interdigital webbing were dissected with
280 fine scissors and forceps and placed immediately in room temperature RNAlater (Sigma-Aldrich).
281 Digits were pooled into a single vial (n=20 digits) and divided into four samples of five randomly
282 selected digits. RNA was extracted from each sample with TRIzol (Thermo Fisher Scientific)
283 following described methods[38]. For digit D3 stage 18, two of the extractions yielded too little RNA
284 for sequencing approach described below; therefore there are only two replicates of this sample
285 type. RNA quality was assessed using an Agilent Technologies 2100 Bioanalyzer, and samples
286 with RIN scores above 8.5 were submitted for sequencing at the Yale Genome Sequencing
287 Center. Sample size (three replicates per sample type) was selected for downstream differential
288 expression analyses, according to References 25 and 26. An in-house protocol was used to
289 generate strand-specific polyadenylated RNA libraries. Samples were multiplexed on an Illumina
290 Hiseq 2000. Each sample was sequenced to a depth of approximately 50 million reads (single-
291 stranded, 75 base pair length).

292 Reads were mapped to the American alligator genome assembly (allMis0.2) with genome
293 assembly described by Green *et* al.[39]. Sequenced reads were mapped to the genome using

294 Tophat2 v2.0.6 on Yale University*'s* Ruddle computing cluster. In Tophat2, reads were first
295 mapped to the transcriptome, and the remaining reads were then mapped to the genome.
296 Mapped reads were assigned to genes with *HTSeq v0.5.3p*[40]*,* which was implemented with
297 Python v2.7.2. In HTSeq, we required that reads be mapped to a specific strand, and to account
298 for reads that mapped to more than one feature, we ran with the setting "intersection-nonempty."
299
300 ***Mus musculus.*** Mice embryos (E13.5) were collected from a pregnant female of the strain
301 C57BL/6J (Jackson Laboratories) in accordance with Yale IACUC #2015-11-483. The female was
302 pregnant with nine embryos. Digits from the left and right forelimbs of each individual were
303 dissected as described for alligator and pooled. From these 18 digits, RNA was extracted for
304 three batches of five digits each. RNA extraction and sequencing methods are the same as
305 described above for alligator, with the exception of sequencing depth (30 million reads were
306 obtained for each mouse sample). Sequenced reads were mapped to the mouse genome
307 assembly GRCm38 with Ensembl annotation v85 and the same Bowtie2 and HTSeq settings as
308 described for alligator.
309
310 ***Anolis carolinensis***. Animals were bred according to published protocols[41] and in accordance
311 with Loyola University's IACUC protocol #1992. Fertilized eggs were collected and transferred to
312 petri dishes containing vermiculate moistened by equal mass water. Embryos were shipped to
313 Yale University and incubated in a Digital Sportsman Incubator (No. 1502) at 26°C. Tissues were
314 extracted and dissected according to methods described for alligator. Stage 10[42] embryos were
315 sampled, and RNA was extracted using Qiagen RNeasy Micro Kit. RNA quality was assessed
316 using with a BioAnalyzer. Samples with RIN scores above 9.0 were submitted for sequencing at
317 the Yale Genome Sequencing Center. The RNAseq library was prepared with Clontech's Ultra
318 Low V4 kit (cat# 634890). Each sample was sequenced to a depth of approximately 30 million
319 reads (single-stranded, 75 base pair length). Sequenced reads were mapped to the *Anolis*
320 genome assembly (AnoCar2.0, GCA_000090745.1) with Ensembl annotation v85.
321
322 ***Gallus gallus*.** Published transcriptomes of the digits of the fore- and hindlimbs of chicken[7], were
323 mapped to the newest chicken genome version (GalGal5.0) with Ensembl annotation v83
324 following analytic methods described for alligator.
325
326 ***Homo sapiens.*** Three individuals of Carnegie stage 18[43] were donated to Yale University's
327 Medical School. The fore- and hindlimbs were sampled, and the anterior-most digit and its
328 posterior interdigital webbing was dissected from the posterior digital plate. Dissections were
329 performed and RNA was extracted and sequenced as previously described[43]. This stage most
330 closely corresponds to E12.5 of mouse[44]. Sequenced reads were mapped to the human genome

assembly GRCh37 with Ensembl annotation v82 using the same Bowtie2 and HTSeq settings as described for alligator.

Hierarchical clustering analysis. To estimate relative mRNA abundance, we calculated transcripts per million (TPM)[45] for the genes of a given gene list (*i.e.*, full trancriptome, transcription factors, limb patterning genes, CDEGs). The TPM measurement standardizes for sequencing depth and transcript length. If multiple transcripts are described for a gene, then the median transcript length was used to calculate TPM. TPM measures were normalized by a square root transformation, and hierarchical clustering was performed on the normalized TPM data with the R package "pvclust"[46]. Clusters were generated from the correlation-based dissimilarity matrix using the average-linked method. Adjusted uncertainty values were calculated from 1000 bootstrapping analysis and are reported for some nodes. If analyses involved comparisons between developmental stages or limbs, a bulk correction was performed with a mean transformation[7], and in these instances Pearson's correlation coefficients rang from -1:1 (e.g., Fig. 1 c).

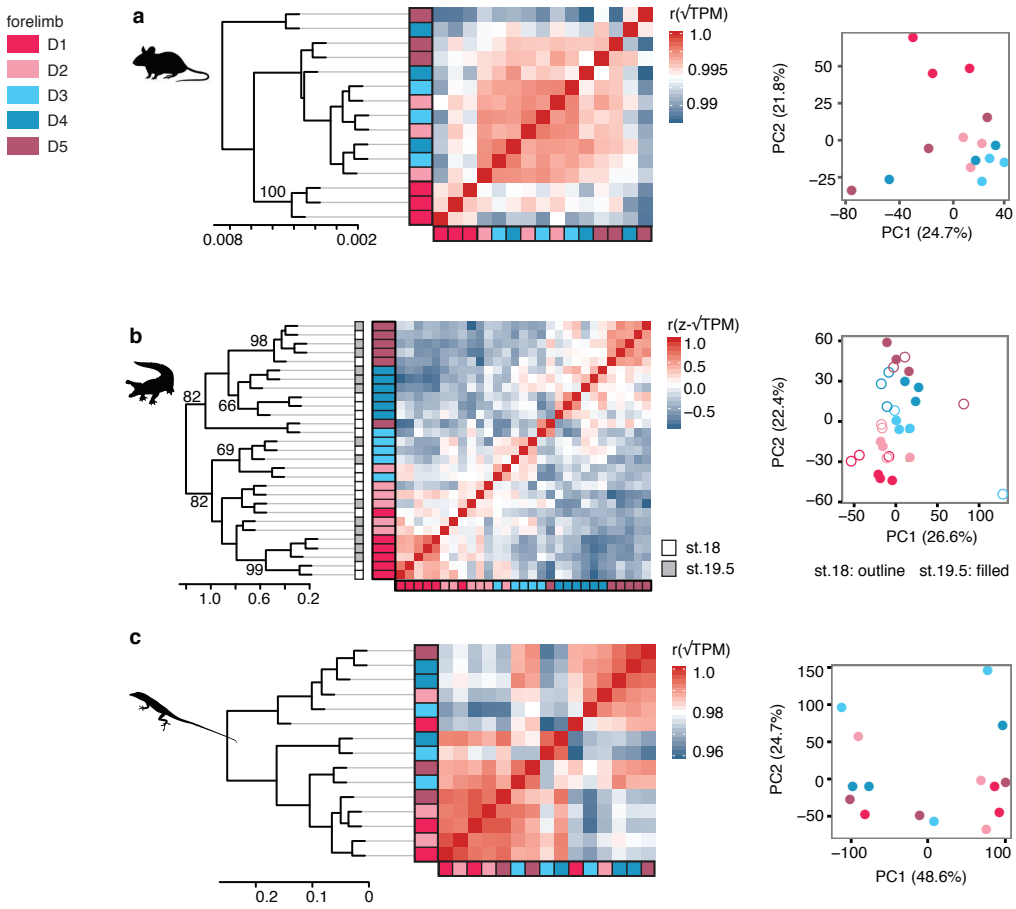
Principal component analyses. PCA were performed using the "prcomp" function in R for various gene lists using normalized TPMs as measures of relative mRNA abundance.

Differential expression testing of adjacent digits. EdgeR (Release 3.1)[26,27] was used to test for differential expression of adjacent digits (*e.g.*, D1 vs. D2) of mouse, alligator, *Anolis* and chicken. We used function glmFit and glmLRT in EdgeR, which implemented a generalized regression model for differential expression test. Specifically, in alligator and chicken, hierarchical analyses revealed stable grouping of samples by digit number across stages. Therefore, subsequent analyses consider these data from two stages simultaneously.

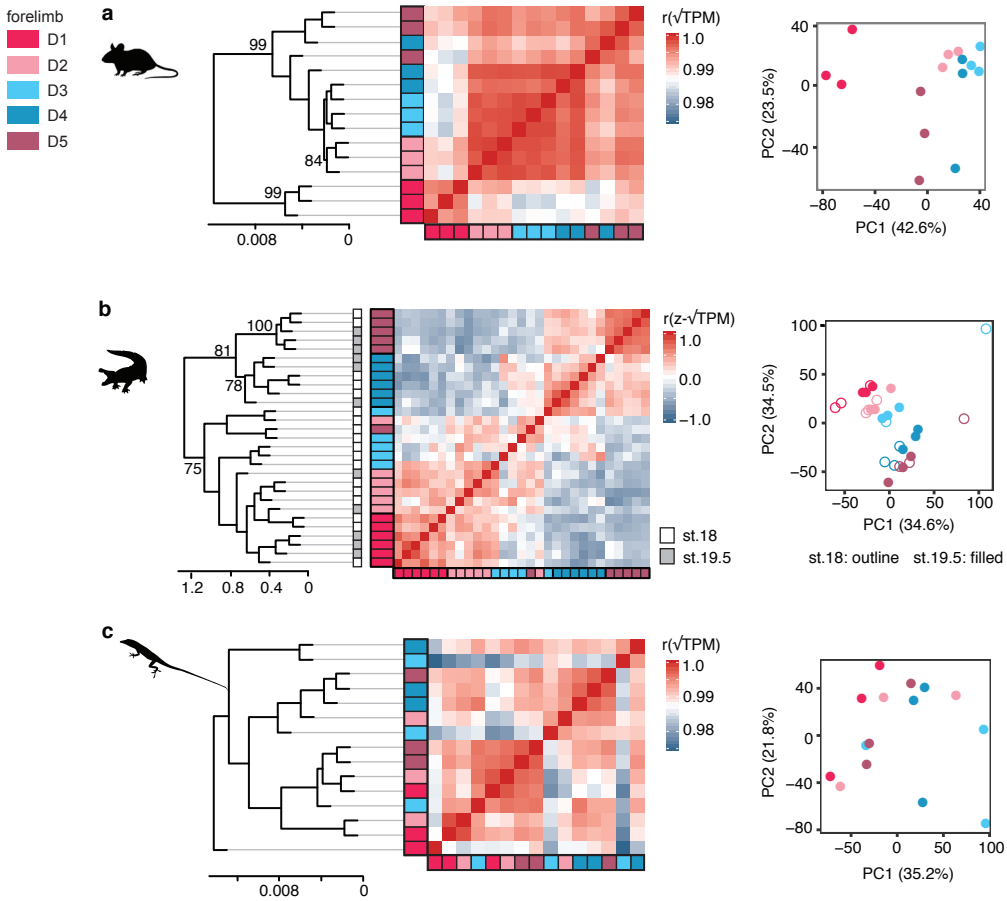
In *Anolis*, pairwise testing of samples revealed nonconventional *p* value distribution, with a decrease near zero and sometimes a bump near 0.5. Because correlation between replicates was lower than what was observed in either mouse or alligator (Extended Data Fig. 9 a), and because PCA of the full transcriptomes revealed two major clusters of data that did not correspond to biological phenomena (Extended Data Fig. 1 c), we corrected for the artifact of the non-biological clusters by including the first principle component in the regression model in EdgeR. Analyses were also run without this PC1 correction. Results presented in the manuscript are robust to both analytic approaches, although PC1 correction results in discovery of slightly more differentially expressed genes for a given false discovery rate. (*e.g.*, compare Extended Data Fig. 4 and 9 b). We tested whether lower correlation among *Anolis* replicates was driving the pattern of homogeneity observed in other analyses (*e.g.*, Fig. 1 d, Extended Data Fig. 3) by repeating all differential expression analyses considering only the two most highly correlated

368 samples of each digit for mouse, alligator, and *Anolis*. These results are consistent with analyses
369 of all samples (e.g., compare Extended Data Fig. 4 and Extended Data Fig. 9 c), indicating the
370 unique *Anolis* pattern does not reflect data quality. Further, comparison of mean dispersion
371 values among data sets (Extended Data Fig. 10 d) suggests that the unique *Anolis* pattern is not
372 a consequence of sensitivity in differential expression analyses, because chicken shows a
373 dispersion value similar to *Anolis*, but patterns of differential expression similar to mouse and
374 alligator.
375
376 **Differential expression testing between all digits.** EdgeR was also used to test for genes that
377 differed between any combination of digits within a limb. This was done by specifying multiple
378 coefficients to function glmLRT. As with the pairwise tests, we considered both stages of alligator
379 and chicken simultaneously and included the first principle component in the regression model for
380 *Anolis*.
381
382 **Transcription factors.** To identify transcription factor genes, we utilized a published atlas of
383 human and mouse transcription factors[24]. The published Entrez gene IDs were matched with
384 human Ensembl gene IDs in Ensembl assembly v85 using BioMart (N=2138). To recover the
385 species-specific lists of transcription factors, two approaches were taken. In mouse and *Anolis*,
386 orthologous genes were identified using BioMart's orthology predictions for Ensemble assembly
387 v85. In alligator and chicken, because these species were analyzed using different assembly
388 builds, orthology was determined by matching gene symbols to those of human from Ensembl
389 assembly v85. By this approach, we recovered transcription factors for each species as follows:
390 1839 in mouse, 1464 in *Anolis*, 1595 in Alligator, and 1214 in chicken. To identify human
391 orthologs of select genes identified by differential expression analyses, we first used gene
392 symbols and then confirmed that the ensemble IDs were consistent across Ensemble
393 assemblies.
394 To identify one-to-one orthologous transcription factor genes in mouse, alligator, and
395 *Anolis*, we used BioMart to generate a list of one-to-one orthologous genes between mouse and
396 *Anolis* in Ensembl assembly v85 that correspond to the published transcription factor Entrez
397 IDs[31]. This gene list was then matched to alligator and chicken by gene symbol to recover
398 transcription factor genes that are one-to-one orthologs in multiple species.
399
400 **Limb patterning genes.** A Ph.D. dissertation by Carkett[23] identified genes that are sensitive to
401 *Shh* signaling by experimental pertubations and *in silico* analyses. From these studies, a
402 summary list of genes that pattern the autopod was produced (pg. 172). This gene list includes
403 transcription factor and signaling genes. These gene symbols were matched with each species to

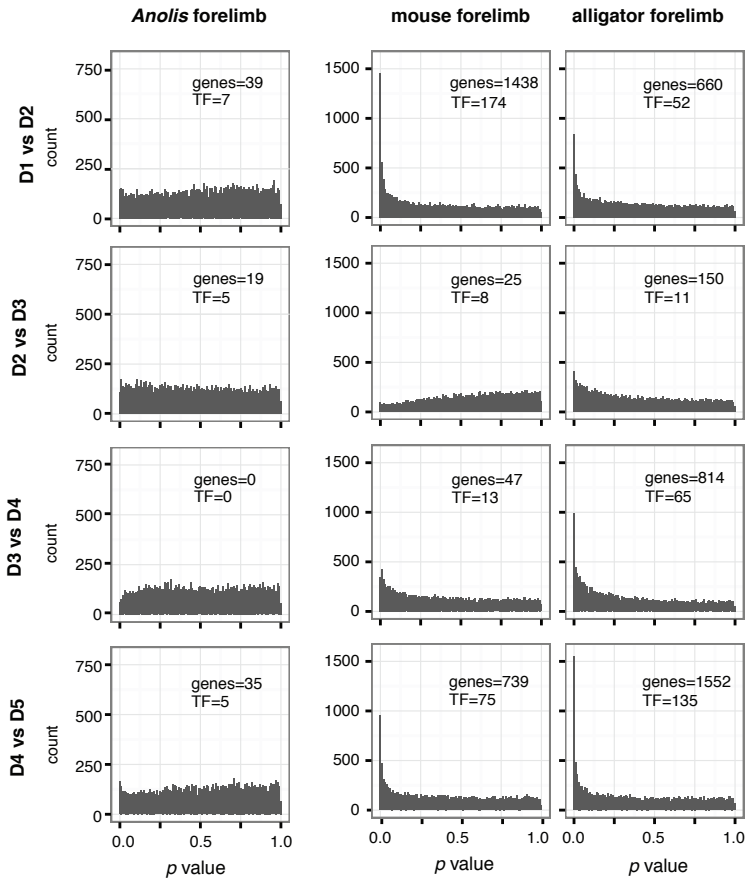
404 identify the subset of genes present in the genome assemblies that we considered for mouse
405 (n=152), alligator (n=152), Anolis (n=159), and chicken (n=133).
406
407 ***In situ* hybridization.** RNA from a stage 18 alligator limb was extracted, as described above for
408 sequencing, and cDNA was generated with the High Capacity cDNA Reverse Transcription kit
409 (Applied Biosystems). Primers were designed with Primer3[47] to amplify fragments of *Tbx2*
410 (forward, GACCTTGGGCCTTCTCCTAC; reverse, GGGAGTTGTTTGGGGTTTT), *Tbx3* (forward,
411 ACCAGGGGTGGATGAACATA; reverse, GCCCTAAAGCAGAGACATGC), and *Sall1* (forward,
412 CTCACAGCTCAACAACCCAC; reverse, AAACCACCAGCCTCTACCTC). PCR products were
413 purified with the QIAquick Gel Extraction kit (Qiagen) and cloned with the Topo TA cloning kit
414 (Invitrogen) into the pCR 4-TOPO vector. Vectors were transformed into DH5α-T1 competent
415 cells. Sense and antisense probes were prepared by linearizing plasmid with the restriction
416 endonucleases *Not1* or *Pme1* and then transcribing the linearized product with T7 or T3
417 polymerase, respectively. Probes are labeled with digoxigenin (Sigma-Aldrich) and hybridized
418 with the alligator embryos at 68C. Methods for *in situ* hybridization followed GEISHA Project
419 miRNA Detection Protocol Version 1.1 (http://geisha.arizona.edu/).
420
421 **Illustrations.** The authors modified existing images of the mouse forelimb skeleton[48], alligator
422 forelimb skeleton[49], chicken forelimb skeleton[50], and chicken and alligator silhouettes[50]. *Anolis*
423 silhouette by Sarah Werning.



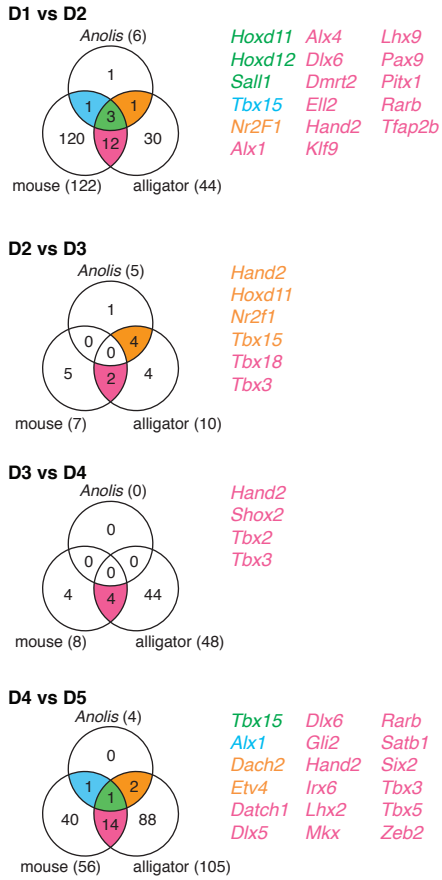
Extended Data Figure 1 | Analyses of full transcriptomes do not show clustering of samples by digit. (a) In mouse, only D1 forms a cluster in HCA, indicating that similarity can be diagnosed when the expression of all genes is considered; more-posterior digits do not form clusters of replicates. (b) In alligator, a cluster with low bootstrap support is observed for D4 samples. (c) In *Anolis,* samples do not reveal stable clustering of digit replicates.



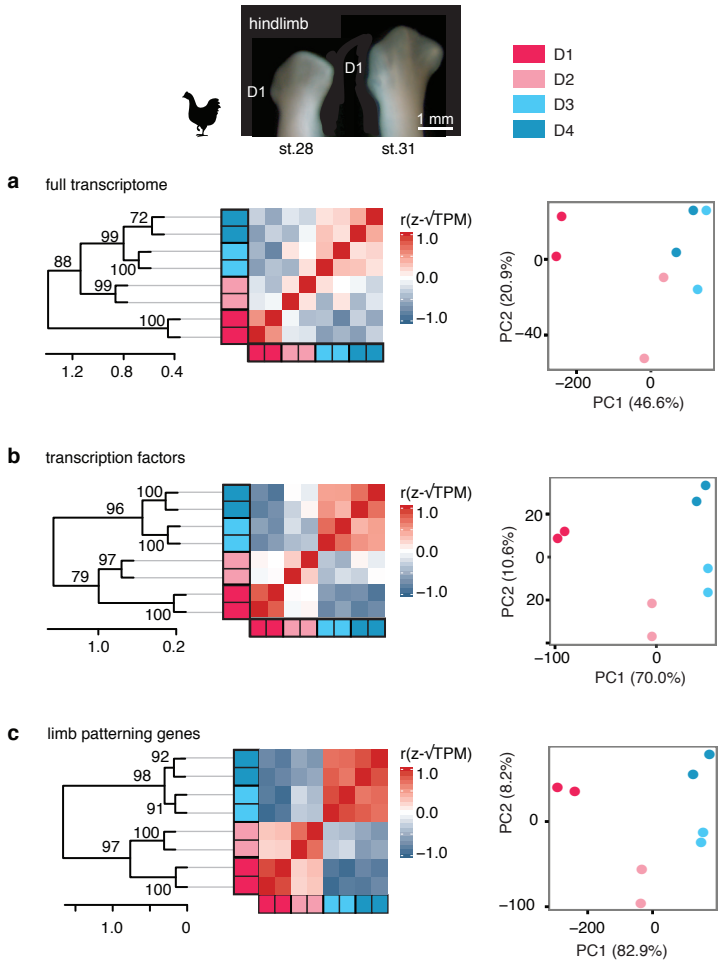
Extended Data Figure 2 | Analyses of all transcription factors show patterns of sample clustering consistent with analyses of limb patterning genes. (a) In mouse, D1 is markedly distinct from the posterior digits. (b) In alligator, two major clusters of digits are observed: (D1, D2, D3)(D4, D5). (c) In *Anolis*, digits do not show gene expression differentiation when all transcription factors are considered. Broadly, adjusted uncertainty values are lower for a given cluster (*e.g.*, alligator D4) than when limb patterning genes are analyzed.



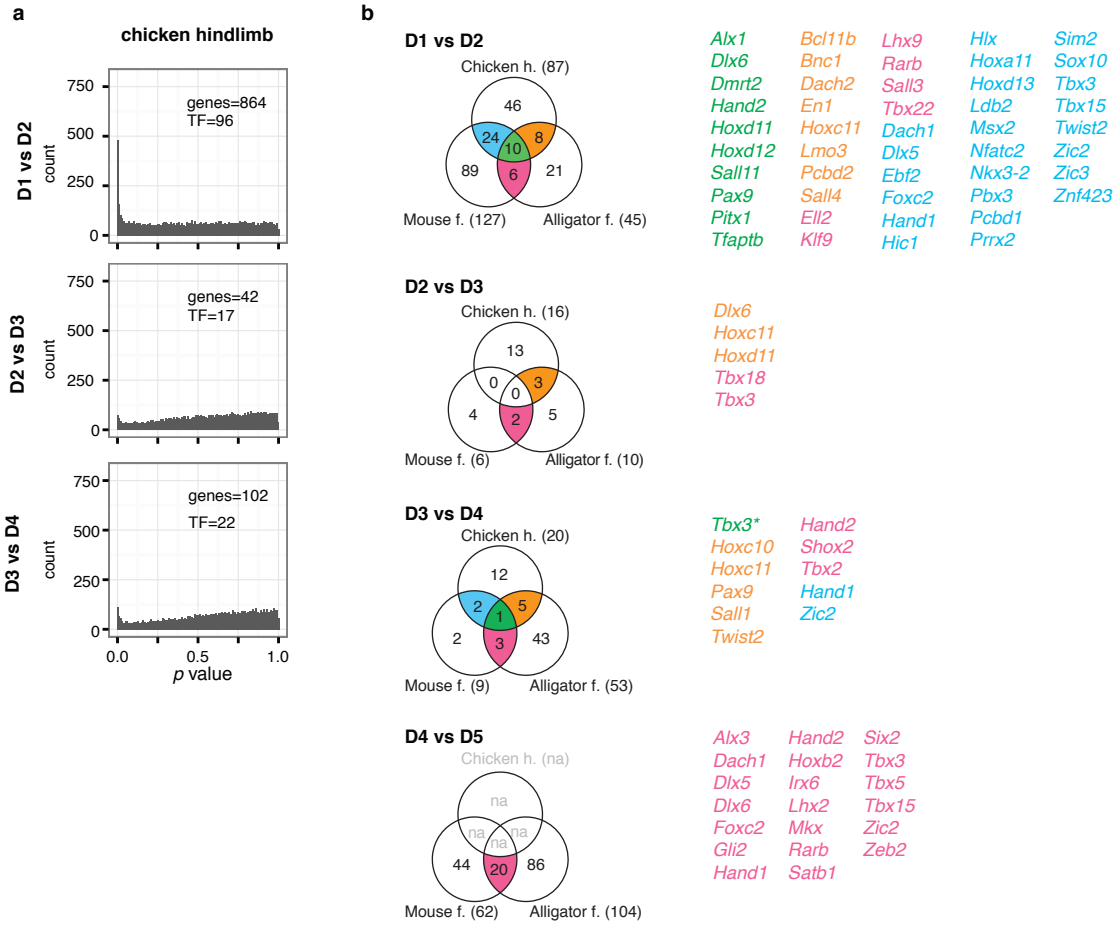
Extended Data Figure 3 | Differential expression analyses suggest homogeneity among *Anolis* digits. Histograms show the distribution of *p* values from differential expression analyses. The number of genes that are identified as differentially expressed at a FDR threshold of 0.05 are noted in each panel as “genes” and the number of transcription factors among these are noted in each panel as “TF.”



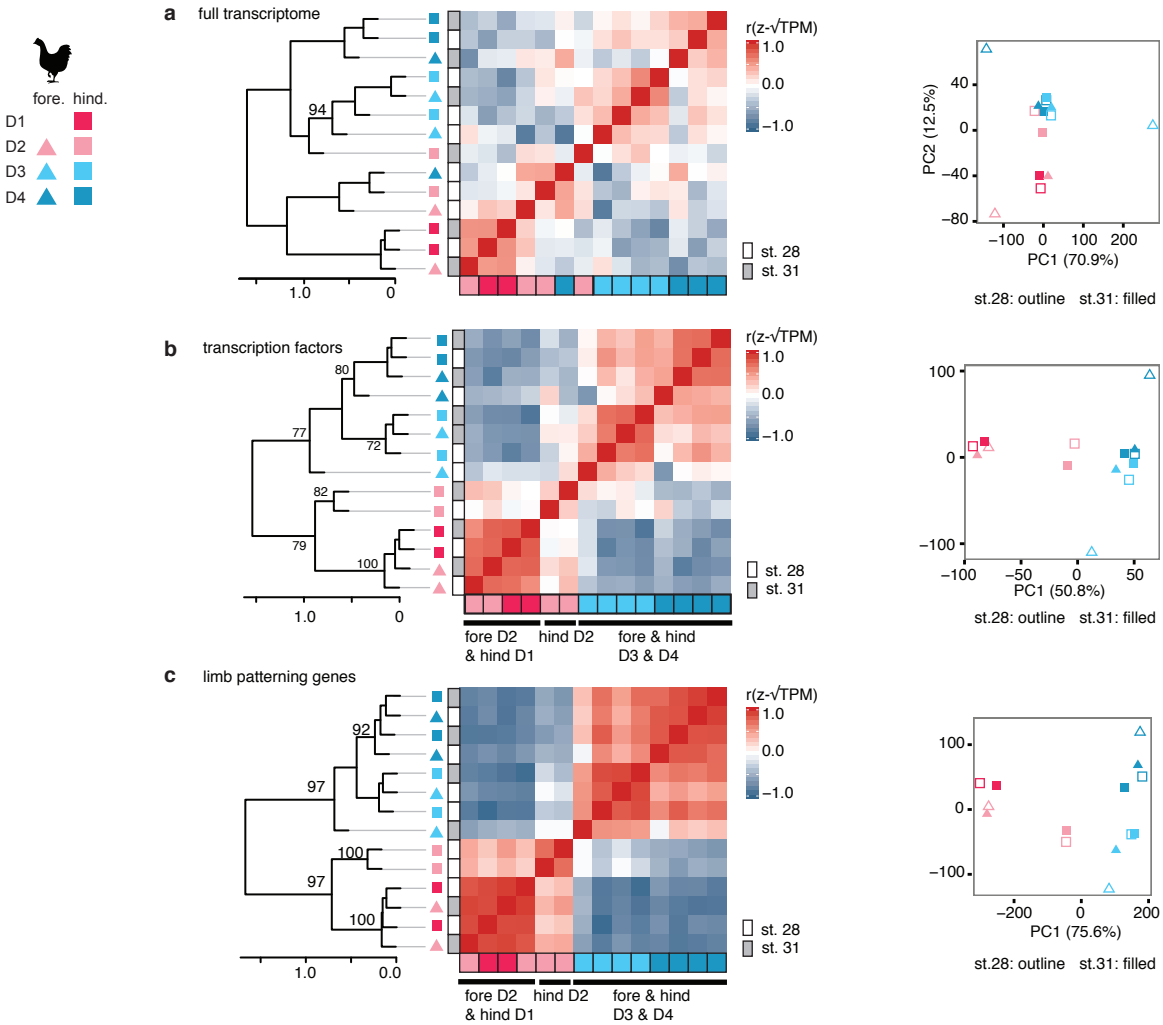
Extended Data Figure 4 | Few genes are differentially expressed at the same position between pentadactyl limbs. Venn diagrams of one-to-one orthologous transcription factors genes for mouse, alligator, and *Anolis* that were identified as differentially expressed between adjacent digits with a FDR threshold of 0.05.



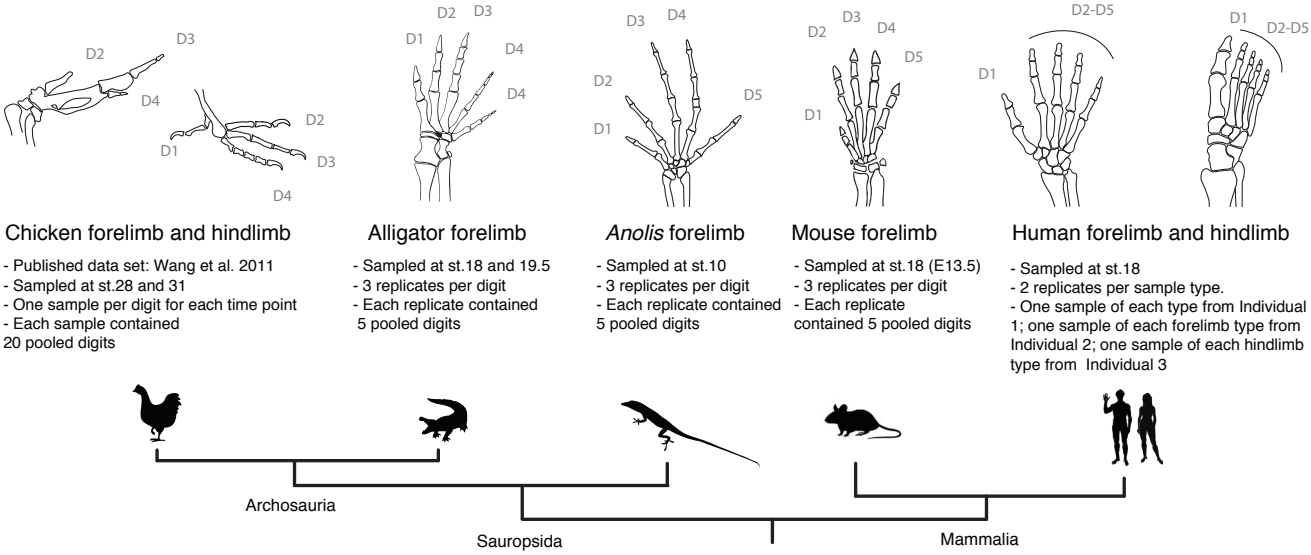
Extended Data Figure 5 | Transcriptomic analysis of the chicken hindlimb. PCA, heatmap of Pearson's correlations and HCA of (a) the full transcriptome and (b) transcription factor genes, (c) and known limb patterning genes. Each digit is represented by two data points that correspond to one sample from stage 28 and another from stage 31. Clustering analyses show stable gene expression profiles for individual digits over this developmental window, even after phalangeal number has been established.



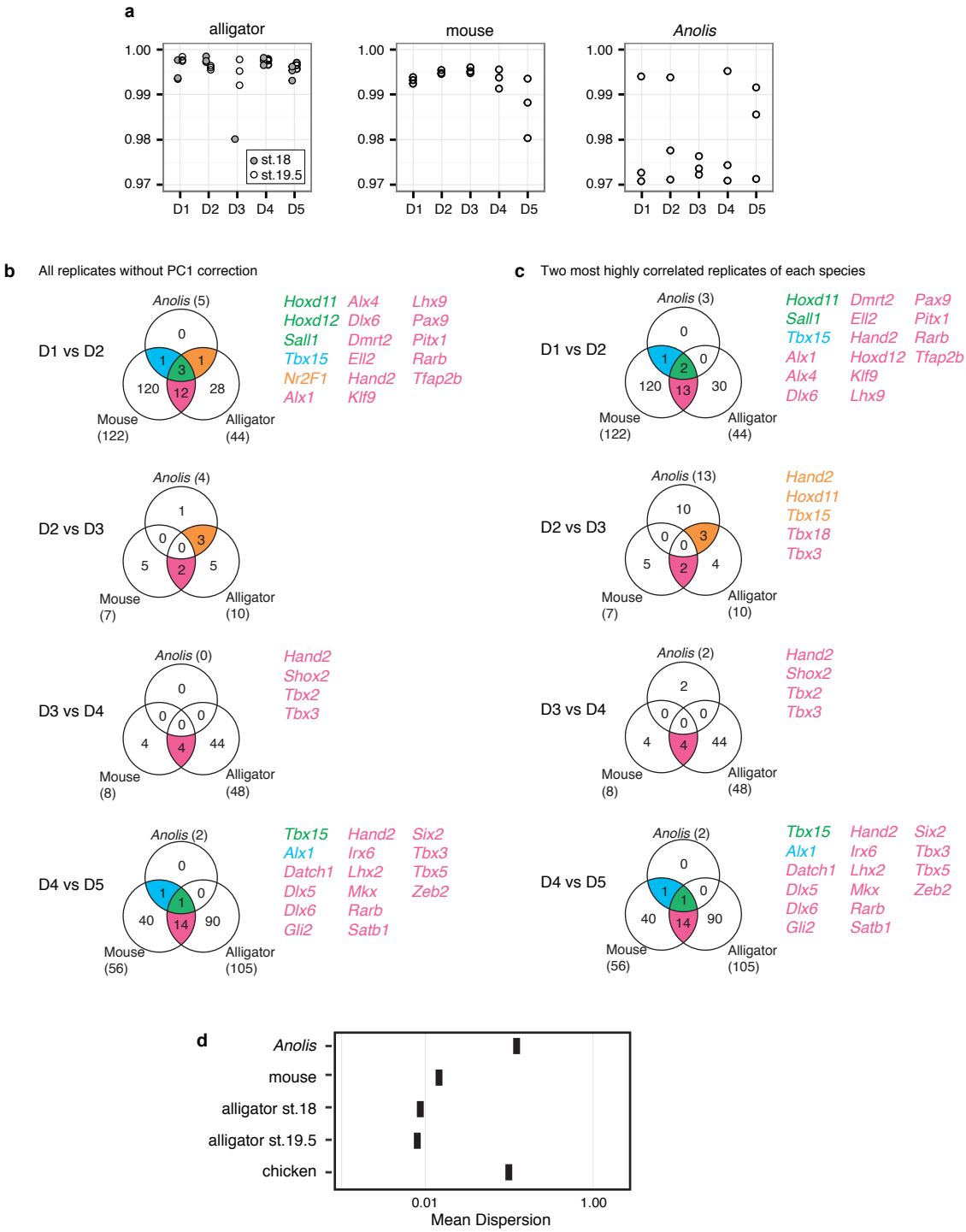
Extended Data Figure 6 | Differential expression analyses of the chicken hindlimb. (a) *p* value distrubitions for pairwise tests of differential expression. The number of genes that are identified as differentially expressed at a FDR threshold of 0.05 are noted in each panel as “genes” and of these the number of transcription factors are noted in each panel as “TF.” (b) Venn diagrams showing the genes that are one-to-one orthologous transcription factors and differentially expressed in each species to a FDR threshold of 0.05.



Extended Data Figure 7 | Correspondence of chicken forelimb and hindlimb digits. (a) PCA and HCA of full transcriptomes does not reveal correspondence between digits of the fore- and hindlimb. However, analyses of (b) of all transcription factors and (c) limb patterning genes show that the digits in the avian wing correspond to hindlimb digits D1, D3, and D4.



Extended Data Figure 8 | Summary of the data sampled. Illustrations are of the adult skeletons from the dorsal perspective, anterior is left.



Extended Data Figure 9 | Analyses to assess the robustness of *Anolis* data. (a) Pearson's correlation values of sample replicates. (b) Venn diagrams of differential expression analyses of adjacent digits where *Anolis* data was not treated with PC1 correction. (c) Venn diagrams of differential expression analyses of only the two most highly correlated samples of each digit for all limbs. (d) Mean dispersion values of different data sets, calculated by edgeR[25,26].

Extended Data Table 1 | Human and mouse phenotypes have been described for 28 of the 49 CDEGs

gene	human phenptype or syndrome	mouse phenotypes
Alx1	Camptodactyly [2]	polydatyly [4]
Dlx5	Split hand/foot malformation [1,3]	ectrodactyly, monodactyly, syndactyly [4]
Dlx6		brachydactyly, ectrodactyly, syndactyly [4]
En1		25 phenotypes including adactyly, ectopic digits, polydactyly, syndactyly, truncation of digits, fused phalanges [4]
Ets2	Chitayat syndrome [3]	
Hand1		hypoplastic limb buds [4]
Hand2		19 phenotypes including oligodactyly, polydactyly, abnormal pollex morphology [4]
Hic1	Hand-foot-genital syndrome [1]	abnormal fore- and hindlimb morphology [4]
Hoxa13	Guttmacher syndrome [1,2], Hand-foot-genital syndrome [1,3], Postaxial hand polydactyly [2]	21 phenotypes including brachydactyly, clinodactyly, syndactyly [4]
Hoxd11		7 phenotypes including abnormal phalanx, abnormal and fused carpals [4]
Hoxd12		16 phenotypes including brachyphalangia, brachydactyly, clinodactyly, oligodactyly [4]
Hoxd13	Brachydactyly-syndactyly syndrome [1], VACTERL association, Brachydactyly type E [2]	29 phenotypes including brachydactyly, ectrodactyly, polydactyly, polysyndactyly [4]
Nkx3-2	Spondylo-megaepiphyseal-metaphyseal dysplasia [3]	
Pax9		polydactyly, polysyndactyly [4]
Pitx1	Clubfoot and Lower limb malformations [1], Liebenberg syndrome [3]	11 phenotypes including brachydactyly, oligodactyly, and clubfoot [4]
Runx2	Brachydactyly [1], MDMHB [3], cleidocranial dysplasia (which includes abnormal thumbs and brachydactyly) [4]	10 phenotypes including abnormal phalanx morphology [4]
Sall1	Townes-Brocks Syndrome [1,2,3] , Lenz microphthalmia syndrome [2]	10 phenotypes including oligodactyly, preaxial polydactyly, syndactyly and triphalangia [4]
Satb2	chromosome 2q32-q33 deletion syndrome (which includes clinodactyly archnodactyly, and Talpes equinovarus [4]	
Sox10	kallmann-syndrome [3], Klein-Waardenburg syndrome [3], PCWH syndrome (which includes Pes cavus) [4]	
Tbx2	17q23.1q23.2 microdeletion syndrome [2]	Polydactyly, postaxial polydactyly [4]
Tbx3	Ulnar-mammary syndrome [1], Limb-mammary syndrome [2], post-axial polydactyly [2]	16 phenotypes including oligodactyly [4]

Tbx5	Holt-Oram syndrome [1,2]	10 phenotypes including abnormal phalanx morphology [4]
Tbx15	Cousin syndrome (which includes brachydactyly) [1,2]	9 phenotypes including abnormal phalanx morphology [4]
Tfap2b	Char syndrome [1,2]	Polydactyly, postaxial polydactyly [4]
Trps1	Trichorhinophalangeal syndrome type 1 [1,2], type 2 [2], and type 3 [2]	
Twist2	Ablepharon macrostomia syndrome [2]	
Zic2		6 phenotypes mostly restricted to carpals [4]
Zic3	VACTERL association [2], Radial abnormalities [2], aplasia/Hypoplasia of the radius [2]	

References

[1] Stenson *et al*. The Human Gene Mutation Database (HGMD®): 2003 Update. Hum. Mutat. **21**, 577-581 (2003)

[2] NIH Genetic and Rare Disease Information Center. http://rarediseases.info.nih.gov.

[3] National Library of Medicine (US). Genetics Home Reference [Internet]. https://ghr.nlm.nih.gov (2013).

[4] Blake JA *et al*. Mouse Genome Database (MGD)-2017: community knowledge resource for the laboratory mouse. Nucl. Acids Res. 4, D723-D729 (2017).

424 **References**

425 1 Coates, M. & Clack, J. Polydactyly in the earliest known tetrapod limbs. *Nature* **347**, 66-
426 69 (1990).
427 2 Lebedev, O. A. & Coates, M. I. The postcranial skeleton of the Devonian tetrapod
428 *Tulerpeton curtum* Lebedev. *Zoological Journal of the Linnean Society* **114**, 307-348,
429 doi:10.1111/j.1096-3642.1995.tb00119.x (1995).
430 3 Owen, R. *On the nature of limbs*. 132 (John Van Voorst, 1849).
431 4 Coates, M. Ancestors and homology. *Acta Biotheoretica* **41**, 411-424,
432 doi:10.1007/BF00709374 (1993).
433 5 Tabin, C. J. Why we have (only) five fingers per hand: hox genes and the evolution of
434 paired limbs. *Development* **116**, 289-296 (1992).
435 6 Zhu, J. & Mackem, S. John Saunders' ZPA, Sonic hedgehog and digit identity – How
436 does it really all work? *Developmental Biology*,
437 doi:http://dx.doi.org/10.1016/j.ydbio.2017.02.001 (2017).
438 7 Wang, Z., Young, R. L., Xue, H. & Wagner, G. P. Transcriptomic analysis of avian digits
439 reveals conserved and derived digit identities in birds. *Nature* **477**,
440 doi:10.1038/nature10391 (2011).
441 8 Montavon, T., Garrec, J.-F. L., Kerzberg, M. & Duboule, D. Modelling HOX genes
442 regulation in digits: reverse collinearity and the molecular origin of thumbness. *Genes*
443 *Dev* **22**, doi:10.1101/gad.1631708 (2008).
444 9 Chiang, C. *et al.* Manifestation of the Limb Prepattern: Limb Development in the Absence
445 of Sonic Hedgehog Function. *Developmental Biology* **236**, 421-435,
446 doi:http://dx.doi.org/10.1006/dbio.2001.0346 (2001).
447 10 Harfe, B. D. *et al.* Evidence for an expansion-based temporal Shh gradient in specifying
448 vertebrate digit identities. *Cell* **118**, doi:10.1016/j.cell.2004.07.024 (2004).
449 11 Scherz, P. J., McGlinn, E., Nissim, S. & Tabin, C. J. Extended exposure to sonic
450 hedgehog is required for patterning the posterior digits of the vertebrate limb. *Dev Biol*
451 **308**, doi:10.1016/j.ydbio.2007.05.030 (2007).
452 12 Reno, P. L. *et al.* Patterns of correlation and covariation of anthropoid distal forelimb
453 segments correspond to Hoxd expression territories. *Journal of Experimental Zoology*
454 *Part B: Molecular and Developmental Evolution* **310B**, 240-258, doi:10.1002/jez.b.21207
455 (2008).
456 13 Dahn, R. D. & Fallon, J. F. Interdigital regulation of digit identity and homeotic
457 transformation by modulated BMP signaling. *Science* **289**,
458 doi:10.1126/science.289.5478.438 (2000).

14 Suzuki, T., Hasso, S. M. & Fallon, J. F. Unique SMAD1/5/8 activity at the phalanx – forming region determines digit identity. *Proc Natl Acad Sci U S A* **105**, doi:10.1073/pnas.0707899105 (2008).

15 Suzuki, T., Takeuchi, J., Koshiba-Takeuchi, K. & Ogura, T. Tbx Genes Specify Posterior Digit Identity through Shh and BMP Signaling. *Developmental Cell* **6**, 43-53, doi:http://dx.doi.org/10.1016/S1534-5807(03)00401-5 (2004).

16 Wagner, G. P. & Gauthier, J. A. 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 USA* **96**, doi:10.1073/pnas.96.9.5111 (1999).

17 Larsson, H. C. E. & Wagner, G. P. Pentadactyl ground state of the avian wing. *Journal of Experimental Zoology* **294**, 146-151, doi:10.1002/jez.10153 (2002).

18 Feduccia, A. & Nowicki, J. The hand of birds revealed by early ostrich embryos. *Naturweissenschaften* **89**, 391-393 (2002).

19 KundráT, M., Seichert, V., Russell, A. P. & Smetana, K. Pentadactyl pattern of the avian wing autopodium and pyramid reduction hypothesis. *Journal of Experimental Zoology* **294**, 152-159, doi:10.1002/jez.10140 (2002).

20 Burke, A. C. & Feduccia, A. Developmental Patterns and the Identification of Homologies in the Avian Hand. *Science* **278**, 666-668, doi:10.1126/science.278.5338.666 (1997).

21 Welten, M. C. M., Verbeek, F. J., Meijer, A. H. & Richardson, M. K. Gene expression and digit homology in the chicken embryo wing. *Evolution & Development* **7**, 18-28, doi:10.1111/j.1525-142X.2005.05003.x (2005).

22 Vargas, A. & Fallon, J. F. Birds have dinosaur wings: the molecular evidence. *J Exp Zoolog B Mol Dev Evol* **304**, doi:10.1002/jez.b.21023 (2005).

23 Carkett, M. D. *Comparative analysis of Sonic Hedgehog signalling and the response to Sonic Hedegehog signalling in vertebrate forelimbs and hindlimbs* PhD thesis, University College London, (2015).

24 Ravasi, T. *et al.* An Atlas of Combinatorial Transcriptional Regulation in Mouse and Man. *Cell* **140**, 744-752, doi:http://dx.doi.org/10.1016/j.cell.2010.01.044 (2010).

25 Benjamini, Y. & Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)* **57**, 289-300, doi:citeulike-article-id:1042553 doi: 10.2307/2346101 (1995).

26 Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics (Oxford, England)* **26**, doi:10.1093/bioinformatics/btp616 (2010).

494 27 McCarthy, J. D., Chen, Y. & Smyth, G. K. Differential expression analysis of multifactor
495 RNA-Seq experiments with respect to biological variation. *Nucleic Acids Research* **40**, 9
496 (2012).
497 28 Towers, M., Signolet, J., Sherman, A., Sang, H. & Tickle, C. Insights into bird wing
498 evolution and digit specification from polarizing region fate maps. *Nat Commun* **2**,
499 doi:10.1038/ncomms1437 (2011).
500 29 Newman, S. A. & Frisch, H. L. Dynamics of skeletal pattern formation in developing chick
501 limb. *Science* **205**, 662-668 (1979).
502 30 Raspopovic, J., Marcon, L., Russo, L. & Sharpe, J. Digit patterning is controlled by a
503 Bmp-Sox9-Wnt Turing network modulated by morphogen gradients. *Science* **345**,
504 doi:10.1126/science.1252960 (2014).
505 31 Wagner, G. P. *Homology, Genes and Evolutionary Innovation*. (University Press, 2014).
506 32 Bateson, W. *Problems of Genetics*. (Oxford University Press, 1913).
507 33 Stewart, T. A. The origin of a new fin skeleton through tinkering. *Biol. Letts* **11**,
508 doi:10.1098/rsbl.2015.0415 (2015).
509 34 Xu, X. & Mackem, S. Tracing the evolution of Avian Wing Digits. *Curr Biol* **23**,
510 doi:10.1016/j.cub.2013.04.071 (2013).
511 35 Zhu, J. *et al.* Uncoupling Sonic Hedgehog Control of Pattern and Expansion of the
512 Developing Limb Bud. *Developmental Cell* **14**, 624-632,
513 doi:http://dx.doi.org/10.1016/j.devcel.2008.01.008 (2008).
514 36 Ferguson, M. W. J. in *Biology of the Reptilia* Vol. 14 (ed Carl Gans) Ch. 5, 329-481
515 (1985).
516 37 Davis, L. M., Glenn, T. C., Elsey, R. M., Dessauer, H. C. & Sawyer, R. H. Multiple
517 paternity and mating patterns in the American alligator, Alligator mississippiensis.
518 *Molecular Ecology* **10**, 1011-1024, doi:10.1046/j.1365-294X.2001.01241.x (2001).
519 38 Chomczynski, P. & Mackey, K. Short Technical Report - Modification of the TRIZOL
520 reagent procedure for isolation of RNA from Polysaccharaide and proteoglycan rich
521 sources. *Biotechniques* **19**, 942-945 (1995).
522 39 Green, R. E. *et al.* Three crocodilian genomes reveal ancestral patterns of evolution
523 among archosaurs. *Science* **346**, doi:10.1126/science.1254449 (2014).
524 40 Anders, S., Pyl, P. T. & Huber, W. HTSeq - A Python framework to work with high-
525 throughput sequencing data. *Bioinformatics*, doi:doi:10.1093/bioinformatics/btu638
526 (2014).
527 41 Sanger, T. J., Hime, P. M., Johnson, M. A., Diani, J. & Lsos, J. B. Laboratory protocols
528 for husbandry and embryo collection of *Anolis* lizards. *Herpetological Review* **39**, 58-63
529 (2008).

42 Sanger, T., Losos, J. & Gibson-Brown, J. A developmental staging series fo the lizard genus *AnloiAnolis*: a reptilian model system for the integration of evolution, development, and ecology. *J. Morph.* **269**, 129 - 137 (2008).

43 O'Rahilly, R. Early human devleopment and the chief sources of information on staged human embryos. *Europ. J. Obstet. Gynec. Reprod. Biol.* **9**, 273-280 (1979).

44 Cotney, J. *et al.* The Evolution of Lineage-Specific Regulatory Activities in the Human Embryonic Limb. *Cell* **154**, 185-196, doi:http://dx.doi.org/10.1016/j.cell.2013.05.056 (2013).

45 Wagner, G. P., Kin, K. & Lynch, V. J. Measurement of mRNA abundance using RNA-seq data: RPKM measure is inconsistent among samples. *Theory in Biosciences* **131**, 281-285, doi:10.1007/s12064-012-0162-3 (2012).

46 Suzuki, R. & Shimodaira, H. Pvclust: an R package for assessing the uncertainty in hierarchical clustering. *Bioinformatics* **22**, 1540-1542, doi:10.1093/bioinformatics/btl117 (2006).

47 Untergasser, A. *et al.* Primer3 - new capabilities and interfaces. . *Nucleic Acids Research* **40**, e115 (2012).

48 Krebs, O. *et al.* Replicated anterior zeugopod (raz): a polydactylous mouse mutant with lowered Shh signaling in the limb bud. *Development* **130**, 6037-6047, doi:10.1242/dev.00861 (2003).

49 Muller, G. B. & Alberch, P. Ontogeny of the Limb Skeleton in Alligator mississipiensis: Developmental Invariance and Change in the Evolution of Archosaur Limbs *Journal of Morphology* **203**, 151-164 (1990).

50 de Bakker, M. A. G. *et al.* Digit loss in archosaur evolution and the interplay between selection and constraints. *Nature* **500**, 445-448, doi:10.1038/nature12336 http://www.nature.com/nature/journal/v500/n7463/abs/nature12336.html - supplementary-information (2013).

561 **Data availability**

562 The RNA-sequencing data reported in this study will be made available on Gene Expression
563 Omnibus (GEO) repository (*https://www.ncbi.nlm.nih.gov/geo/*). Sequencing data for human limb
564 samples is available through the database of Genotypes and Phenotypes (dbGaP) under study
565 accession number phs001226.v1.p1. Code used for analyses is available upon request.

566

567 **Acknowledgements**

568 We thank the Wagner lab members and SE Newman for discussions on data interpretation. The
569 human embryonic material was provided by the Joint MRC (G0700089)/Wellcome Trust
570 (GR082557) Human Developmental Biology Resource. This work was supported by the John
571 Templeton Foundation (Integrating Generic and Genetic Explanations of Biological Phenomena;
572 ID 46919), by National Institutes of Health grant GM094780 (to J.P.N.), and by National Science
573 Foundation grant 1353691.

574

575 **Author Contributions**

576 T.A.S. and G.P.W. contributed project design and paper writing. T.A.S., T.J.S., J.L.C., and J.P.N.
577 contributed data for analysis. T.A.S., C.L., and G.P.W. contributed analyses of data.
578 Correspondence and requests for materials should be addressed to tomstewart@uchicago.edu
579 and gunter.wagner@yale.edu

580

581 **Additional Information**

582 Extended Data accompanies this paper.
583 Competing Financial Interests: The authors declare no competing financial interests.