Appendage expression driven by the *Hoxd* Global Control Region is an ancient gnathostome feature

Igor Schneider[a], Ivy Aneas[b], Andrew R. Gehrke[a], Randall D. Dahn[c], Marcelo A. Nobrega[b], and Neil H. Shubin[a,1]

[a]Department of Organismal Biology and Anatomy, and [b]Department of Human Genetics, University of Chicago, Chicago, IL 60637; and [c]Mount Desert Island Biological Laboratory, Salisbury Cove, ME 04672

Contributed by Neil H. Shubin, June 21, 2011 (sent for review June 17, 2011)

The evolutionary transition of the fins of fish into tetrapod limbs involved genetic changes to developmental systems that resulted in novel skeletal patterns and functions. Approaches to understanding this issue have entailed the search for antecedents of limb structure in fossils, genes, and embryos. Comparative genetic analyses have produced ambiguous results: although studies of posterior *Hox* genes from homology group 13 (Hoxa-13 and *Hoxd*-13) reveal similarities in gene expression between the distal segments of fins and limbs, this functional homology has not been supported by genomic comparisons of the activity of their *cis*-regulatory elements, namely the *Hoxd* Global Control Region. Here, we show that *cis*-regulatory elements driving *Hoxd* gene expression in distal limbs are present in fish. Using an interspecies transgenesis approach, we find functional conservation between gnathostome *Hoxd* enhancers, demonstrating that orthologous sequences from tetrapods, zebrafish and skate can drive reporter gene expression in mouse limbs and zebrafish fins. Our results support the notion that some of the novelties associated with tetrapod limbs arose by modification of deeply conserved *cis*- and *trans*-acting mechanisms of *Hox* regulation in gnathostomes.

autopod | development | evolution

One of the classic issues of vertebrate evolution is the origin of the tetrapod limb. First seen in Devonian ecosystems, the limb skeletal pattern has been highly conserved during the wide adaptive diversification to the extremes of tetrapod morphology seen in running, hopping, digging, flying, and swimming forms. The limb skeleton consists of three segments: a proximal stylopod with a humerus or femur, the zeugopod with a radius/ulna and tibia/fibula, and an autopod with wrist/ankle and digits. Comparative morphology reveals that the stylopod and zeugopod have homologs in the fins of fish, but hypotheses of the origin of the autopod have been controversial (1). Developmental approaches to this question have relied on the observation that there are distinct sets of the activity of *Hox* genes in the limb: an early phase that patterns the stylopod and zeugopod, and a late phase that is implicated in the patterning of the autopod (Fig. 1*A*) (2). Although comparative studies support late-phase *Hoxd* expression as an ancestral gnathostome characteristic (3–6), studies of *Hoxd* regulation suggest other possibilities (7).

In tetrapod limbs, the early and late phases of *Hoxd* expression are partitioned into distinct regulatory domains: enhancers located downstream of the *Hoxd* cluster control early-phase expression (8–11), whereas late-phase *Hoxd* expression is controlled by enhancers upstream of the *Hoxd* cluster (Fig. 1 *B* and *C*) (7, 12). The late-phase regulatory apparatus is composed of a Global Control Region (GCR) that contains two regulatory elements, CsB and CsA, yet only CsB is required and sufficient to recapitulate all GCR regulatory potentials, including autopod, genital bud, and neural tube expression (12). A third regulatory element, termed CsC, maps to the intergenic interval between the genes *Lnp* and *Evx2* and can also direct distal limb expression. Together, the GCR (i.e., CsB) and CsC recapitulate most aspects of late-phase autopod *Hoxd* expression in mouse limbs (Fig. 1*B*) (12).

Genomic sequence comparisons reveal that CsB exhibits high conservation between mammals and teleost fish (Fig. 1*C*). This observation raised the intriguing possibility that fishes may be equipped with a late-phase *Hoxd* *cis*-regulatory element, providing support for the hypothesis of equivalence between late-phase *Hoxd* expression in fishes and tetrapods. However, despite its high conservation to tetrapod sequences, a pufferfish CsB ortholog was unable to elicit limb expression in a transgenic mouse assay, and it was proposed that this *Hoxd* *cis*-regulatory machinery controlling distal limb expression is unique to tetrapods (7, 12).

A key challenge involving comparisons between tetrapod and teleost enhancer activity arises from the fact that ray-finned fishes have undergone a lineage-specific whole-genome duplication event (13). As a result, enhancers that were duplicated in the teleost lineage may have been retained with diverged regulatory capacities, providing an additional level of complexity to analyses involving interspecies transgenesis. In fact, the pufferfish genome has retained two copies of the *Hoxd* gene cluster and may potentially possess two CsB enhancers (14). To circumvent this difficulty, we chose to study CsB function of a teleost species that has retained a single *Hoxd* cluster (zebrafish, *Brachydanio rerio*) and of a chondrichthyan species (little skate, *Leucoraja erinacea*), which diverged from bony fishes before the ray-finned fish whole-genome duplication (14).

Results

Previously, a 4.7-kb fragment of the human CsB was shown sufficient to promote reporter expression in neuronal tissues, external genitalia, and distal limbs of transgenic mice (12). Here we have focused on the 2.6-kb sequence of human CsB that shares blocks of sequence conservation with other tetrapods. Using this shorter human CsB as a reference, we aligned orthologous CsB sequences from mouse, chicken, frog, zebrafish, and skate (Fig. 1*C*). The zebrafish and skate CsBs are 4.3-kb and 2.5-kb long, respectively, and share three domains of sequence conservation with tetrapod CsBs. The skate CsB is more similar to its tetrapod orthologs than is the zebrafish, in line with previous reports revealing conservation between noncoding sequences of chondrichthyans and tetrapods (15, 16).

To explore fish CsB regulatory activity, we examined enhancer function in vivo in a zebrafish-based reporter assay (17). The zebrafish CsB element was PCR-amplified from genomic DNA

Author contributions: I.S. and N.H.S. designed research; I.S., A.R.G., and R.D.D. performed research; I.A. and M.A.N. contributed new reagents/analytic tools; I.S., I.A., M.A.N., and N.H.S. analyzed data; and I.S. and N.H.S. wrote the paper.

The authors declare no conflict of interest.

Freely available online through the PNAS open access option.

Data deposition: The sequence reported in this paper has been deposited in the GenBank database (accession no. JN160846).

[1]To whom correspondence should be addressed. E-mail: nshubin@uchicago.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1109993108/-/DCSupplemental.

Downloaded at Palestinian Territory, occupied on December 10, 2021

EVOLUTION

www.manaraa.com

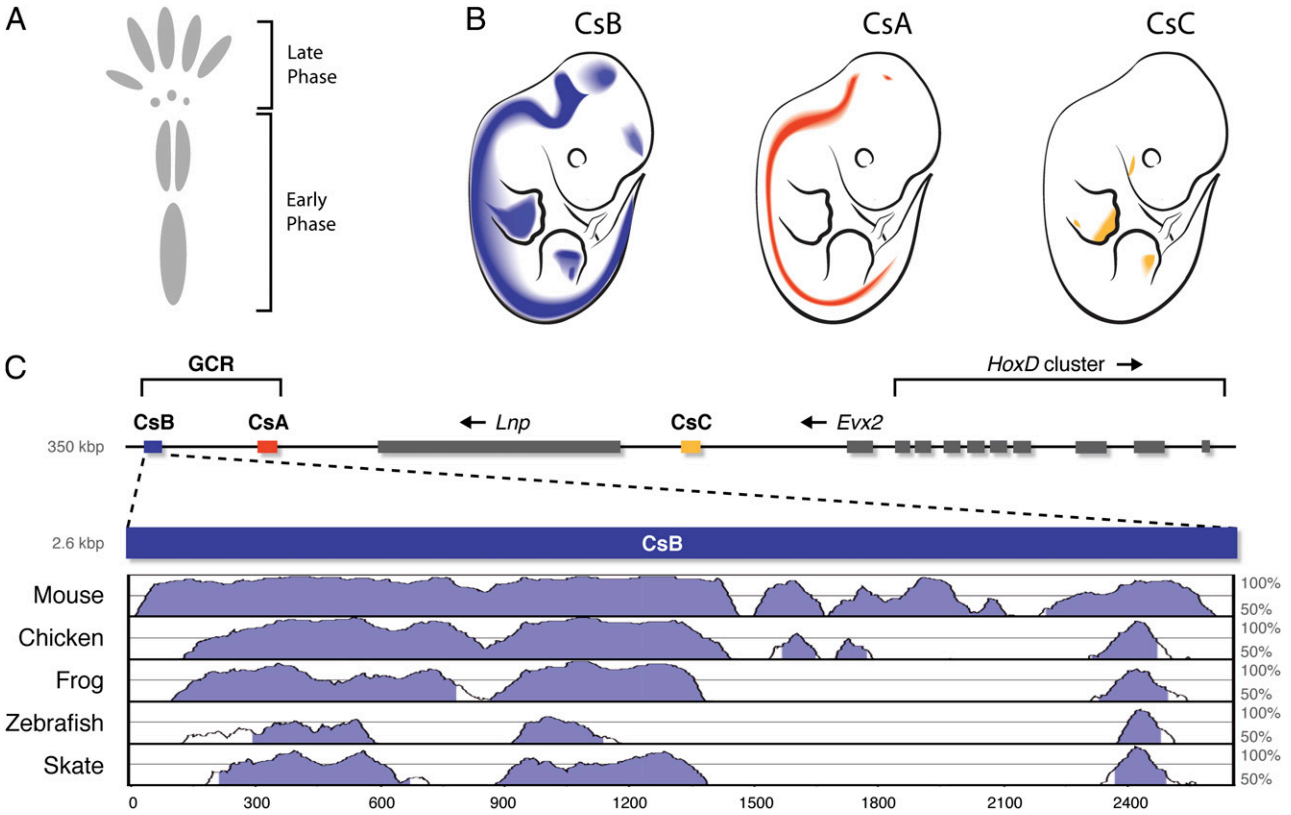

Fig. 1. Regulatory potential of HoxD enhancers. (*A*) Schematic of early- and late-phase *Hoxd* expression in a tetrapod limb. (*B*) Domains of expression denote regulatory potentials of CsB (blue), CsA (red), and CsC (yellow), adapted from transgenic embryos reported (12). (*C*) (*Upper*) Organization of the genomic locus containing CsB (blue), CsA (red), and CsC (yellow) and the genes *Lnp*, *Evx2*, and the HoxD cluster (gray). (*Lower*) Sequence comparison of human, mouse, chicken, frog, zebrafish, and skate CsB. Alignment (mVISTA program, homology threshold 70%) shows regions of homology between tetrapod, zebrafish, and skate CsB sequences.

and cloned into an eGFP reporter cassette, driven by a minimal c-fos promoter (18, 19), containing two *tol2* transposon sites for rapid and efficient integration of the transgene (17). Enhancer activity was evaluated in G0 fish embryos, as previous reports have demonstrated that the patterns observed in mosaic transgenic founders are reproduced in the germ-line transmission to G1 (19).

As early as 10 h postfertilization (hpf), eGFP signal was readily detected along the midline, on the presumptive neural plate of zebrafish CsB-injected embryos, and at 26 to 30 hpf, virtually all eGFP-positive embryos displayed neural tube and brain expression (Fig. S1 and Table S1). At 50 to 52 hpf, eGFP signal diminished in neuronal cell populations and reporter activity was first detected in the fin, in about half of the neural tube eGFP-positive embryos (Fig. 2 *A* and *B*, Fig. S2, and Table S1). The eGFP-positive cells were mostly found in the extreme distal regions, just proximal to the fin fold (Fig. 2 *C* and *D*).

We next tested whether a tetrapod CsB could also elicit distal fin expression in zebrafish. To this end, we cloned a 2.6-kb orthologous mouse CsB sequence into the same destination vector used in our zebrafish CsB analysis. Similar to the zebrafish CsB, eGFP signal was detected along the midline of mouse CsB-injected zebrafish embryos (Fig. S1). At 26 to 30 hpf, in addition to neural tube and brain expression, the mouse CsB construct also elicited strong eGFP signal in the notochord (Fig. S2 and Table S1). At 50 to 52 hpf, reporter activity ceased in neuronal and notochord cells and fin signal was detected in 43% of the embryos (Fig. 2 *E* and *F*, Fig. S2, and Table S1). Fin expression in mouse CsB-injected embryos was distal, yet more proximal and more intense than that elicited by zebrafish CsB (Fig. 2 *G* and *H*). Finally, injection of the *tol2*-eGFP vector alone ($n = 40$) could not elicit fin eGFP expression during the developmental stages assessed. Therefore, we concluded that zebrafish and mouse CsBs possess comparable regulatory potentials—perhaps in the form of conserved *trans*-acting factors—for neuronal tissues and, importantly, a previously unrecognized appendage expression potential of the zebrafish CsB ortholog.

Given our results showing eGFP signal in zebrafish fins, we examined a phylogenetic sample of CsB regulatory capacity in mouse transgenic assays. To achieve this examination, chicken, zebrafish, or skate CsB elements were inserted into a reporter cassette containing a β-galactosidase (*lacZ*) gene driven by a β-globin minimal promoter (β-globin-lacZ) and injected in fertilized mouse pronuclei, generating multiple independent transgenic embryos for each construct injected. Embryos were assayed for lacZ activity using X-gal at embryonic day (E) 12.5. As expected, chicken CsB transgenic mice displayed X-gal staining in distal limbs (two of four transgenics) (Fig. 3 *A* and *B*, and Fig. S3). X-gal signal was also scored in the brain (two of four), neural tube (two of four), face, heart, and external genitalia (one of four) (Fig. S3). In the autopod, X-gal signal extended to the base of the digits (Fig. 3*B*). Thus, the chicken CsB, used here as a tetrapod control element, could recapitulate the expression patterns previously reported for avian and murine CsBs (12).

Remarkably, in the mice carrying the zebrafish CsB, three of eight embryos displayed staining in the limbs, which was detected in the zeugopod (one of three) and at the wrist and base of the digits in two of the three embryos (Fig. 3 *C* and *D* and Fig. S3). X-gal signal, whether in the wrist or digits, was restricted to the posterior and extended anteriorly, reminiscent of tetrapod *Hoxd*

Downloaded at Palestinian Territory, occupied on December 10, 2021
www.manaraa.com

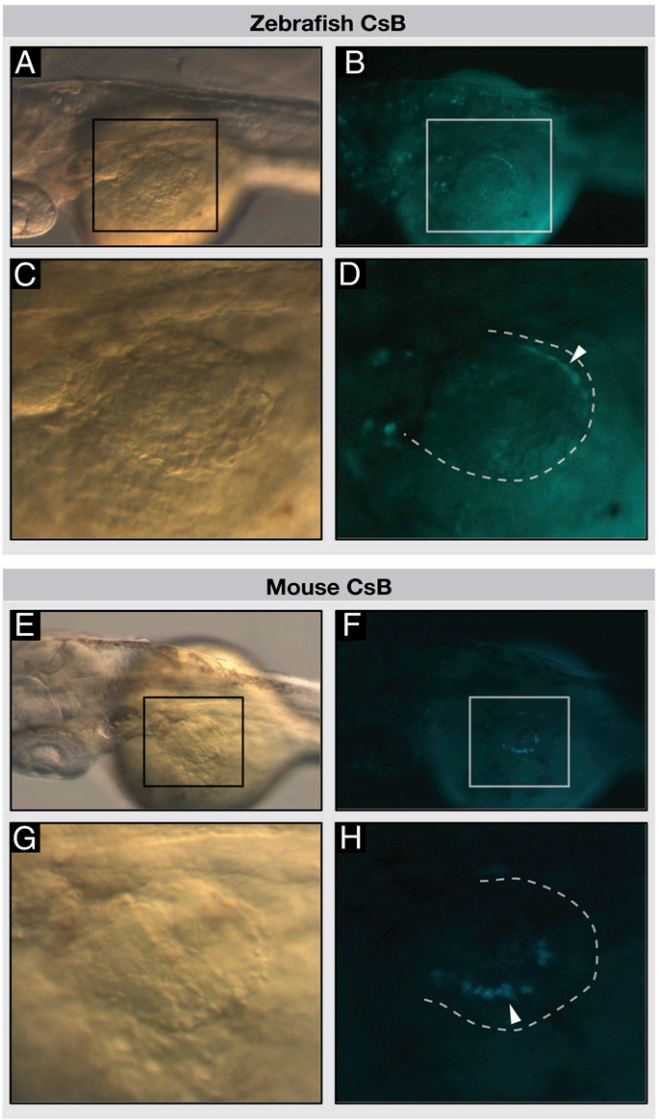

Fig. 2. Vertebrate CsB enhancers drive neural tube and distal fin GFP expression in transgenic zebrafish. Transgenic zebrafish embryos injected with zebrafish CsB (*A*–*D*) or mouse CsB (*E*–*H*), at 52 hpf, anterior to the left. Dorsal view of zebrafish embryos; black and white boxes indicate position of the fin in bright field (*A* and *E*) and fluorescence (*B* and *F*), respectively. (*C* and *G*) Bright-field images show dorsal view of fins. (*D* and *H*) Dotted lines indicate position of fins, white arrowheads indicate eGFP signal in fins.

expression (Fig. 3*D*). Neuronal signal was also scored and, unlike chicken CsB, it was mostly limited to the brain (seven of eight) (Fig. S3). In addition, X-gal staining was detected in the face (three of eight), external genitalia (two of eight), and heart (one of eight) (Fig. S3).

The CsB element of the skate, which has a more ancient common ancestor with tetrapods than does the zebrafish, also produced *lacZ*-positive transgenic mice. All 20 transgenic embryos obtained displayed X-gal signal in the brain (20 of 20) that extended along the neural tube (13 of 20) (Fig. 3*E* and Fig. S3). Significantly, limb signal was present in 7 of 20 embryos and, as seen in zebrafish CsB transgenesis, expression was detected in the zeugopod and autopod (Fig. 3 *E* and *F* and Fig. S3). Autopod signal was concentrated in the wrist and extended to the base of the digits (three of seven) (Fig. 3*F*). Staining was also detected in the face (6 of 20) and external genitalia (1 of 20) (Fig. S3).

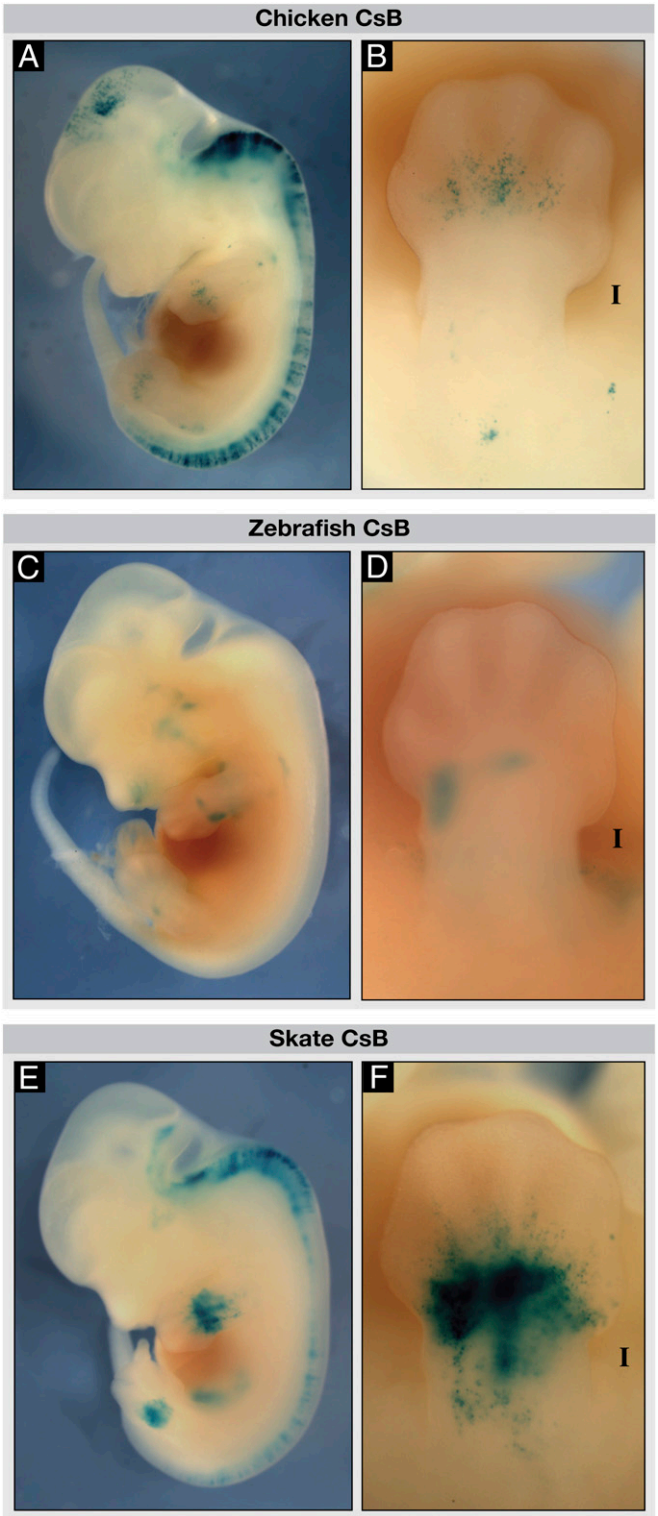

Fig. 3. Skate and zebrafish CsB enhancers promote distal limb expression in transgenic mice. Transgenic mouse embryos at E12.5 stained for X-gal. (*A* and *B*) Chicken CsB embryo showing X-gal signal in the brain and neural tube; limb signal extends across the base of the digits. (*C* and *D*) Zebrafish CsB embryo shows X-gal staining in face; limb signal extends anteriorly at the base of digits 3, 4, and 5. (*E* and *F*) Skate CsB embryo displaying X-gal staining along the neural tube; limb signal is present in distal zeugopod, as well as the wrist and at the base of the digits. Position of digit I is noted in *B*, *D*, and *F*.

Downloaded at Palestinian Territory, occupied on December 10, 2021

www.manaraa.com

Overall, our transgenic mouse assay reveals the presence of a *cis*-regulatory capacity in fish CsB to drive reporter gene expression in proximal regions of the autopod, namely the wrist and proximal digits.

Discussion

Fossil discoveries have revealed that the fins of lobe-finned fish possess homologs of carpal and tarsal bones (i.e., the ulnare, fibulare, and intermedium) set in a skeleton that retains the characteristic dermal rays of fish fins. Furthermore, the endoskeleton of the fin of osteolepiforms, elpistostegalians, and rhizodontids, such as *Panderichthys* (20), *Tiktaalik* (21), and *Sauripterus* (22), extends well distally to the boundary of the zeugopod, often with an extensive array of endochondral elements. These morphological conditions have led to the speculation that the autopod segment may be primitive to tetrapod limbs, despite the apparent lack of morphological homologs of digits in fins. Although studies of gene expression have lent support to this view, attempts to uncover an ancestral *cis*-regulatory code specifying the autopod program in fish fins have led to controversy. A significant argument against a fish autopod genetic program is the apparent lack of functional homology in *cis*-regulatory sequences involved in the expression of genes responsible for autopod development in tetrapods.

Here we have tested whether the *Hoxd* CsB enhancer, a key regulatory component of the *Hoxd* autopodal expression in tetrapods, has a functional counterpart in fish. We demonstrate that tetrapod and fish CsBs possess distal appendage enhancer activities in reciprocal transgenic experiments in zebrafish and mouse. The notably discrete domain of reporter gene expression driven by the zebrafish CsB in both transgenic fins and limbs is suggestive of a limited regulatory capacity for appendage expression. This observation correlates with reports that the *Hoxd13*, *Evx2*, and *Lnp* genes, known to be influenced by the GCR in mouse, fail to display similar expression patterns in zebrafish distal fins (3). In more basal actinopterygians, such as the paddlefish (*Polyodon spathula*) and in a chondrichthyan (*Scyliorhinus canicula*), the expression patterns of *Lnp* and *Evx2* are not known, yet *Hoxd13* expression bears greater resemblance to the late-phase expression seen in tetrapod limbs (4, 5). Likewise, the skate CsB, which shows high homology to its tetrapod ortholog and elicits broad expression in distal limbs of transgenic mice, may be a better model than the teleost CsB for comparative genetic analyses of tetrapod *Hoxd* regulation.

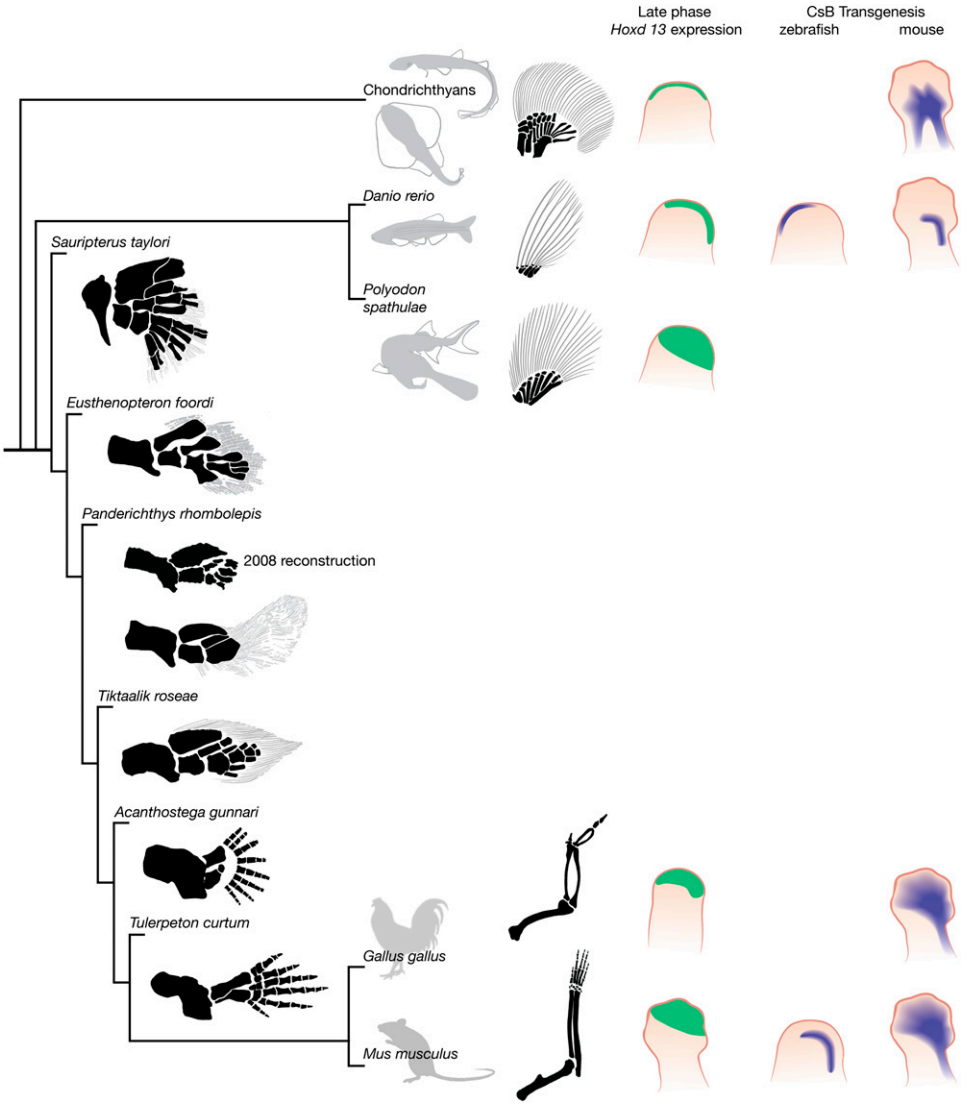

Fig. 4. Evolution of limb and fin morphology and *Hoxd* regulation. Cladogram shows extant and extinct taxa, their corresponding fin/limb morphologies, late-phase *Hoxd* expression pattern, and CsB regulatory potential in transgenic assays using zebrafish and mouse.

Downloaded at Palestinian Territory, occupied on December 10, 2021
www.manaraa.com

Taking these data together, this study reveals great conservation of both *cis*- and *trans*-acting factors involved in vertebrate appendage development (summarized in Fig. 4).

The expression driven by the skate and zebrafish CsBs was mostly limited to the wrist and base of the digits, in contrast to the mouse CsB that includes these domains but extends more distally in the autopod (12). This finding implies that expansion of the *Hoxd* expression domain is associated with the enhanced distal endochondral skeleton of tetrapods. Additionally, these data support the hypothesis that the distal fin of teleosts, composing the boundary between the proximal and distal radials, compares most favorably with the mesopod and proximal autopod of limbed vertebrates. Our data support the hypothesis that the origin of the tetrapod limb involved an expansion of *Hoxd* expression (5), perhaps by acquisition of additional *cis*-regulatory elements, such as the CsC, an element with no obvious orthologous sequence in fish (12). Taken in light of studies of fossils, comparative morphology, and gene expression, the demonstration of conserved and ancient regulatory potentials reveals the deep homology of the autopod segment and now allows us to target derived regions of the genome that are involved in the origin of digits themselves.

Methods

Plasmid Generation. The pBS-βgLacZ.gw vector was constructed by cloning a Gateway frame A cassette (Invitrogen) and human minimal β-globin promoter upstream of the LacZ/SV40polyA reporter gene into a PBS KSII vector. Genomic DNA fragments ranging from 2.5 to 4.3 Kbp, corresponding to mouse, chicken, zebrafish, and skate CsB were isolated using the Expand High Fidelity PCR System (Invitrogen). The DNA oligonucleotide sequence can be found in *SI Methods*. Fragments were cloned into an entry vector (pENTR/D-TOPO) and transferred to the destination vectors using the LR recombination reaction (Invitrogen). All plasmids were verified by restriction analysis and direct sequencing.

Zebrafish Transgenesis. Zebrafish embryos were collected from natural spawning. Staging was according to standard conditions (23). Genomic DNA fragments corresponding to mouse and zebrafish CsB sequences were transferred from pENTR/D-TOPO to the transposon pXIG-cFos-EGFP vector (24). Transposase RNA was transcribed in vitro using the mMessage mMachine Sp6 kit (Ambion). Injection solutions were made with 175 ng/μL transposase RNA and 125 ng/μL circular plasmid in water. DNA was injected into the yolk of wild-type embryos at the one- to two-cell stage and embryos were analyzed and imaged using a Leica M205FA Microscope.

Mouse Transgenesis. Vertebrate CsB orthologs previously cloned were transferred to pBS-βgLacZ.gw by LR recombination reaction. Plasmid DNA was purified using the Wizard Plus SV miniprep Kit (Promega), and 50 μg of each plasmid was digested with SalI to excise the vector backbone. Following a gel-purification step using the QIAquick Gel Extraction Kit (Qiagen), the DNA to be injected was further purified using a standard ethanol precipitation and diluted to a concentration of 2 ng/μL. Purified, linearized plasmid DNA was then used for pronuclear injections of CD1 mouse embryos in accordance with standard protocols approved by the University of Chicago. Mouse embryos were harvested, stained and fixed (*SI Methods*).

ACKNOWLEDGMENTS. We thank Ilya Ruvinsky for discussions; Nadia Fröbisch, Natalia Taft, Victoria Prince, Robert Ho, Gunter Wagner, and Cliff Tabin for insightful comments on the manuscript, Kalliopi Monoyios for illustration work and graphics assistance, and Joe Davis for technical support; and Linda Degenstein for assistance in generating transgenic animals. This work was supported by a fellowship from the American Heart Association (to I.A.) and by the Division of Biological Sciences at The University of Chicago (N.H.S.).

1. Shubin N, Tabin C, Carroll S (2009) Deep homology and the origins of evolutionary novelty. *Nature* 457:818–823.
2. Deschamps J (2004) Developmental biology. *Hox* genes in the limb: A play in two acts. *Science* 304:1610–1611.
3. Ahn D, Ho RK (2008) Tri-phasic expression of posterior *Hox* genes during development of pectoral fins in zebrafish: Implications for the evolution of vertebrate paired appendages. *Dev Biol* 322:220–233.
4. Davis MC, Dahn RD, Shubin NH (2007) An autopodial-like pattern of *Hox* expression in the fins of a basal actinopterygian fish. *Nature* 447:473–476.
5. Freitas R, Zhang G, Cohn MJ (2007) Biphasic *Hoxd* gene expression in shark paired fins reveals an ancient origin of the distal limb domain. *PLoS ONE* 2:e754.
6. Johanson Z, et al. (2007) Fish fingers: Digit homologues in sarcopterygian fish fins. *J Exp Zoolog B Mol Dev Evol* 308:757–768.
7. Spitz F, Gonzalez F, Duboule D (2003) A global control region defines a chromosomal regulatory landscape containing the HoxD cluster. *Cell* 113:405–417.
8. Zákány J, Kmita M, Duboule D (2004) A dual role for *Hox* genes in limb anterior-posterior asymmetry. *Science* 304:1669–1672.
9. Spitz F, Herkenne C, Morris MA, Duboule D (2005) Inversion-induced disruption of the *Hoxd* cluster leads to the partition of regulatory landscapes. *Nat Genet* 37:889–893.
10. Deschamps J (2007) Ancestral and recently recruited global control of the *Hox* genes in development. *Curr Opin Genet Dev* 17:422–427.
11. Tarchini B, Huynh TH, Cox GA, Duboule D (2005) HoxD cluster scanning deletions identify multiple defects leading to paralysis in the mouse mutant Ironside. *Genes Dev* 19:2862–2876.
12. Gonzalez F, Duboule D, Spitz F (2007) Transgenic analysis of *Hoxd* gene regulation during digit development. *Dev Biol* 306:847–859.
13. Jaillon O, et al. (2004) Genome duplication in the teleost fish *Tetraodon nigroviridis* reveals the early vertebrate proto-karyotype. *Nature* 431:946–957.
14. Amores A, et al. (2004) Developmental roles of pufferfish *Hox* clusters and genome evolution in ray-fin fish. *Genome Res* 14(1):1–10.
15. Venkatesh B, et al. (2007) Survey sequencing and comparative analysis of the elephant shark (*Callorhinchus milii*) genome. *PLoS Biol* 5(4):e101.
16. Ravi V, et al. (2009) Elephant shark (*Callorhinchus milii*) provides insights into the evolution of *Hox* gene clusters in gnathostomes. *Proc Natl Acad Sci USA* 106: 16327–16332.
17. Kawakami K, et al. (2004) A transposon-mediated gene trap approach identifies developmentally regulated genes in zebrafish. *Dev Cell* 7(1):133–144.
18. Fisher S, Grice EA, Vinton RM, Bessling SL, McCallion AS (2006) Conservation of RET regulatory function from human to zebrafish without sequence similarity. *Science* 312:276–279.
19. Fisher S, et al. (2006) Evaluating the biological relevance of putative enhancers using Tol2 transposon-mediated transgenesis in zebrafish. *Nat Protoc* 1:1297–1305.
20. Boisvert CA, Mark-Kurik E, Ahlberg PE (2008) The pectoral fin of Panderichthys and the origin of digits. *Nature* 456:636–638.
21. Shubin NH, Daeschler EB, Jenkins FA, Jr. (2006) The pectoral fin of *Tiktaalik roseae* and the origin of the tetrapod limb. *Nature* 440:764–771.
22. Davis MC, Shubin N, Daeschler EB (2004) A new specimen of *Sauripterus taylori* (Sarcopterygii, Osteichthyes) from the Famennian Catskill Formation of North America. *J Vertebr Paleontol* 24(1):26–40.
23. Kimmel CB, Ballard WW, Kimmel SR, Ullmann B, Schilling TF (1995) Stages of embryonic development of the zebrafish. *Dev Dyn* 203:253–310.
24. Sakabe NJ, Nobrega MA (2010) Genome-wide maps of transcription regulatory elements. *Wiley Interdiscip Rev Syst Biol Med* 2:422–437.

Downloaded at Palestinian Territory, occupied on December 10, 2021

www.manaraa.com