J Mol Evol (2008) 67:581–593
DOI 10.1007/s00239-008-9156-7

A Molecular Footprint of Limb Loss: Sequence Variation of the Autopodial Identity Gene *Hoxa-13*

Tiana Kohlsdorf · Michael P. Cummings · Vincent J. Lynch · Geffrey F. Stopper · Kazuhiko Takahashi · Günter P. Wagner

Received: 31 January 2008 / Accepted: 5 August 2008 / Published online: 15 October 2008

Abstract The homeobox gene *Hoxa-13* codes for a transcription factor involved in multiple functions, including body axis and hand/foot development in tetrapods. In this study we investigate whether the loss of one function (e.g., limb loss in snakes) left a molecular footprint in exon 1 of *Hoxa-13* that could be associated with the release of functional constraints caused by limb loss. Fragments of the *Hoxa-13* exon 1 were sequenced from 13 species and analyzed, with additional published sequences of the same region, using relative rates and likelihood-ratio tests. Five amino acid sites in exon 1 of *Hoxa-13* were detected as evolving under positive selection in the stem lineage of snakes. To further investigate whether there is an association between limb loss and sequence variation in *Hoxa-13*, we used the random forest method on an alignment that included shark, basal fish lineages, and "eutetrapods" such as mammals, turtle, alligator, and birds. The random forest method approaches the problem as one of classification, where we seek to predict the presence or absence of autopodium based on amino acid variation in *Hoxa-13* sequences. Different alignments tested were associated with similar error rates (18.42%). The random forest method suggested that phenotypic states (autopodium present and absent) can often be correctly predicted based on *Hoxa-13* sequences. Basal, nontetrapod gnathostomes that never had an autopodium were consistently classified as limbless together with the snakes, while eutetrapods without any history of limb loss in their phylogeny were also consistently classified as having a limb. Misclassifications affected mostly lizards, which, as a group, have a history of limb loss and limb re-evolution, and the urodele and caecilian in our sample. We conclude that a molecular footprint can be detected in *Hoxa-13* that is associated with the lack of an autopodium; groups with classification ambiguity (lizards) are characterized by a history of repeated limb loss and possible limb re-evolution.

Keywords *Hoxa-13* · Molecular evolution · Limb loss · Squamata

Electronic supplementary material The online version of this article (doi:10.1007/s00239-008-9156-7) contains supplementary material, which is available to authorized users.

T. Kohlsdorf · V. J. Lynch · G. F. Stopper · K. Takahashi · G. P. Wagner (✉)
Department of Ecology and Evolutionary Biology, Yale University, 165 Prospect Street, New Haven, CT 06520, USA
e-mail: gunter.wagner@yale.edu

Present Address:
T. Kohlsdorf
Departamento de Biologia, FFCLRP, University of São Paulo, Av. Bandeirantes 3900, Bairro Monte Alegre, Ribeirao Preto, SP 14040-901, Brazil

M. P. Cummings
Center for Bioinformatics and Computational Biology, University of Maryland, College Park, MD 20742-3360, USA

Present Address:
K. Takahashi
National Institute for Basic Biology, Okazaki National Research Institutes, 38 Nishigonaka, Myodaijicho, Okazaki, Aichi 444-8585, Japan

Introduction

Genes coding for a lost character might be maintained in the genome due to pleiotropic effects on other characters that are of selective value to the organism (Wright 1968; Lande 1978; Teotónio and Rose 2001; West-Eberhard

2003), and it is estimated that reactivation of long unexpressed genes (silenced for more than 10 million years) is only possible if gene function is maintained by other selective constraints (Marshall et al. 1994). Many developmental genes are involved in multiple functions (e.g., limb development and axial patterning [see Veraksa et al. 2000; Carroll 2005]) and are thus unlikely to be lost because of the loss of one of these characters (e.g., limbs). It is possible, however, that the gene has organ specific elements (e.g., associated with heart or kidney development), like a character specific enhancer or protein-protein interaction motif (Wagner and Lynch 2008). In the case of an interaction motif, the loss of one specific character (e.g., the limb) could lead to the release of some of the selective constraints on the protein coding sequence. In the present study we investigate whether the pattern of variation in the coding sequence of a developmental gene involved in limb development, *Hoxa-13*, is affected by limb loss.

Hox genes code for transcription factors involved in multiple developmental functions (Carroll 2005), including limb and fin development (Nelson et al. 1996; Hérault et al. 1999; Innis et al. 2002; Knosp et al. 2004; Metscher et al. 2005; Davis et al. 2007). The expression patterns of many Hox genes during limb development are very dynamic and present distinct phases of expression and complex domains for each gene (Duboule 1994; Davis et al. 1995; Favier et al. 1996; Fromental-Ramain et al. 1996; Nelson et al. 1996). For example, during vertebrate limb development *Hoxa-13* localizes to the autopodial paddle (Nelson et al. 1996; Hérault et al. 1999; Innis et al. 2002), and later to discrete domains within the interdigital regions of the autopodium, where its function is required for interdigital programmed cell death, digit outgrowth, and chondrogenesis (Fromental-Ramain et al. 1996; Stadler et al. 2001; Knosp et al. 2004). A knockout of *Hoxa-13* leads to a specific loss of digit 1 (Fromental-Ramain et al. 1996). Spontaneous mutants of *Hoxa-13* exhibit digit and limb malformation (e.g., Post and Innis 1999; Zákány and Duboule 1999; Innis et al. 2002; Knosp et al. 2004), and mutations in *Hoxa-13* in humans can lead to the 'hand-foot-genital syndrome' (Mortlock and Innis 1997; Mortlock et al. 2000; Innis et al. 2004). The *Hoxa-13* gene consists of two exons, and while exon 1 is experiencing evolutionary changes, exon 2 mostly contains the homeobox and does not change on the taxonomic level of species we investigated (Lynch et al. 2004; Wagner et al. 2005). Considering the role of *Hoxa-13* in autopodium development, an interesting question is, What are the changes observed in the sequence of this gene in lineages that do not exhibit autopodium development because limbs either were lost or never developed?

One group that exhibits substantial variation in limb and foot morphology is the Squamata (lizards including snakes [Pough et al. 1998]), which contains several limbless lineages (e.g., Greer 1991; Bejder and Hall 2002; Caldwell 2003; Kohlsdorf and Wagner 2006; Wiens et al. 2006). In the snake clade, most species lack front and hind limbs entirely and all lack an autopodium. Despite the ancestral terrestrial fossil (*Najash rionegrina* from the Cretaceous [Apesteguía and Zaher 2006]) and three other fossils of marine snakes (Caldwell and Lee 1997; Rage and Escuillié 2000; Tchernov et al. 2000), all extant snake species either are entirely limbless or present only vestigial limb elements; therefore, the autopodium may have been absent in this group for about 100 million years (the oldest snakes are probably from the mid-Cretaceous [Rage 2001]). In the present study we investigate whether sequence changes in *Hoxa-13* are associated with the presence or absence of the autopodium. We first focus on Squamata, comparing limbed lizards with snakes, and we then compare these results with sequence variation in a broader sample of gnathostomes, using the random forest method (Breiman 2001a).

Materials and Methods

DNA Extraction and Amplification

Genomic DNA was extracted from frozen muscle tissues with the DNeasy Tissue Kit (Qiagen) according to the manufacturer's instructions. A detailed list of taxa included (with related museum numbers or GenBank accession numbers) is included in Supplementary Appendix I. Gene fragments of *Hoxa-13* (375 to 455 bp) were amplified from one individual of each species by polymerase chain reaction (PCR), using 10 to 100 ng of DNA, a 0.1 to 0.2 μM concentration of each primer, and PCR Master Mix (Reddymix, 2.5 mM $MgCl_2$; Abgene, Inc.) to obtain a final reaction volume of 50 μl. The thermocycling conditions were 30 cycles of 1 min denaturation at 94°C/1 min annealing at 52–56°C/1 min extension at 72°C, followed by 1 cycle of 8 min at 72°C for final extension. The sequences for all primers are presented below. PCR products were gel-purified using the QIAquick Gel Purification System (Qiagen). Genes were cloned using the pGEM-T vector system (Promega) and *E. coli*-competent cells (Promega). Sequencing was performed in both directions with the vector primers T7 and SP6 on an ABI 3100 Genetic Analyzer (Applied Biosystems, Foster City, CA). Four to ten clones of each gene were sequenced to control for errors during PCR amplification. Sequences for all loci were deposited in GenBank under accession numbers EF456654 to EF46667.

Hoxa-13 primers were as follows: FA13LZ, CTC CATCCCCGCTGGATCGA; RA13LZ, GGCATATCCA GGTAGCC; KTA13F1, CTYCATYCCCGCTGGATYGA;

KTA13R1, CTTTKACYCKYCTGTTYTGRAACC; KTA 13R3, CATGGCTGGTARCTTTCCA (From Fry et al. 2006); F54, ACCAACAGCYTGGARGAGATYAACAA; R2B, TGGTAGAAAGCAAACTCCTGG; and R2C, GCC CTGGTAGAARGCRAACTCCT.

Molecular Evolution Analyses

To investigate possible changes in *Hoxa-13* sequences that could be related to the lack of autopodium development, we applied analyses of molecular evolution that characterized two complementary scenarios. First, we analyzed molecular evolution of *Hoxa-13* in the scenario of limb loss in Squamata by comparing sequence length, GC content, and substitution rates between Serpentes and limbed lizards. Then we analyzed patterns of molecular evolution of *Hoxa-13* on a larger phylogenetic scale, using random forest analysis to identify possible changes in amino acid positions that might be associated with the presence or absence of autopodium in Gnathostomata.

The first approach (molecular evolution of *Hoxa-13* in lizards and snakes) required a phylogenetic hypothesis available for Squamata. Therefore, we performed phylogenetic analyses to determine a topology for Squamata, based on data from entire mitochondrial genomes downloaded from GenBank, as described in detail in Supplementary Appendix II. The relationships among major clades of Squamata that are relevant to the present study (Agamidae, Iguanidae, Anguimorpha, and Serpentes; see Fig. 1) were congruent between the phylogenetic analyses performed here, based on entire mitochondrial genomes, and other studies using molecular data (Vidal and Hedges 2005; Wiens et al. 2006; Fry et al. 2006) (see Supplementary Appendix II).

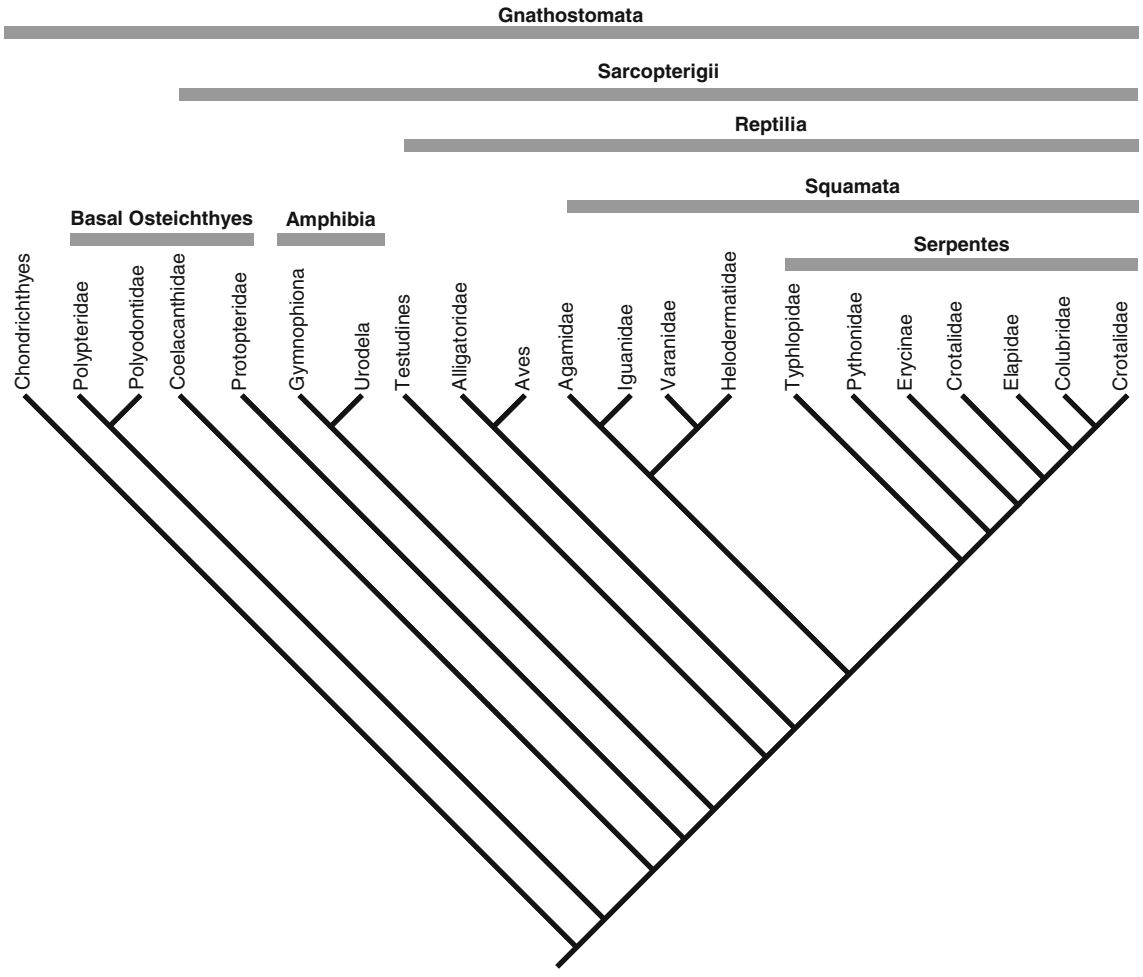

Fig. 1 Phylogenetic hypothesis used to perform tests of molecular evolution of *Hoxa-13*, *c-mos*, *RAG-1*, and *BDNF* in two comparisons: (1) Squamata—Serpentes versus limbed squamates (Helodermatidae, Varanidae, Iguanidae, and Agamidae); (2) Gnathostomata. Phylogenetic relationships among Squamata taxa were confirmed by the molecular phylogenetic analyses described in Supplementary Appendix II

Molecular Evolution of *Hoxa-13* in Squamata: Serpentes Versus Limbed Lizards

Hoxa-13 sequences obtained for species of Squamata were aligned using the Clustal-V algorithm (Higgins et al. 1992), implemented by the software BioEdit sequence alignment editor. The alignment was manually improved based on sequences translated into amino acids. Codon bias was assessed by information on GC content calculated for each individual sequence in the software BioEdit Sequence Alignment Editor and analyzed among groups in the computer program RRTree (Robinson et al. 1998; Rechavi and Huchon 2000).

We analyzed the molecular evolution of *Hoxa-13* sequences in Squamata using two complementary approaches. First, we analyzed the pattern of insertions and deletions in the gene sequences, and how those related to Serpentes in comparison with limbed species of squamates. We then analyzed the patterns of amino acid substitutions in these two clades based on the topology shown in Fig. 1. In this second approach, we deleted insertions and deletions from the alignment due to computational demands. To see whether the pattern found in *Hoxa-13* is gene specific we applied the same analyses to three other nuclear genes: *c-mos*, *RAG-1*, and *brain-derived neurotrophic factor* (*BDNF*). Sequences for these three genes were retrieved from GenBank for several species, as described in Supplementary Appendix I.

In the first approach, focusing on patterns of insertions and deletions, total sequence length was inferred in BioEdit and compared among groups (limbed [Anguimorphs + Iguanidae + Agamidae] versus Serpentes) in SPSS 12.0 for PC using ANOVA. Changes in sequence length were visualized in a phylogenetic context using the program Mesquite 1.05 (parsimony reconstructions [Maddison and Maddison 2004]). Related species share a common evolutionary history, and therefore phylogenetic analyses (based on independent contrasts [Felsenstein 1985]) were also applied using the MS-DOS computer program PDTREE (Garland et al. 1992, 1993, 1999; Garland and Ives 2000), with arbitrary branch lengths following Pagel (1992). Phylogenetic ANOVAs were performed by simulating character evolution (1000 simulations) under the Gradual Brownian Motion Process in the MS-DOS computer program PDSIMUL and analyzing the simulated data in PDANOVA (Garland et al. 1993).

Initial exploratory analyses were performed in HyPhy (version HyPhy Kernel for Windows [Pond et al. 2005]) and suggested differences in evolutionary rates between snakes and limbed lizards (using alligator as outgroup), which were visualized using the program MacClade (version 4.08 for MacOS X [Maddison and Maddison 1992]). Two complementary tests based on phylogenetic topologies were then applied, in order to verify the hypothesis that Serpentes and limbed squamates differ in their rates of *Hoxa-13* evolution. The same tests were performed for the three nuclear neutral genes (*c-mos*, *RAG-1*, and *BDNF*).

Substitution rates between sequences were compared using relative rate tests based on phylogenetic trees (topology following Fig. 1), implemented by the computer program RRTree (Robinson et al. 1998; Rechavi and Huchon 2000). In these tests, alligator was defined as the outgroup and the Serpentes clade was compared with a clade of limbed squamates (Iguanidae and Agamidae + Varanidae and Helodermatidae). Rates of both synonymous and nonsynonymous substitutions per (synonymous or nonsynonymous) site were analyzed.

Based on methods proposed by Yang (1997; implemented using the computer program PAML), two-branch-site maximum likelihood (ML) tests of positive selection were applied together, namely, Test I and Test II (described in detail by Zhang et al. 2005; Arbiza et al. 2006). The phylogeny used is shown in Figs. 1 and 3, with an unrooted topology. This approach allows detection of lineage-specific events while distinguishing true cases of positive selection from likely cases of relaxation of selective constraints, by first comparing the likelihood of a 'nearly neutral model' (MIa) with that of an alternative model testing for positive selection in the branch of interest (MA), and then comparing the likelihood of the model testing for positive selection (MA) with that of an alternative model where ω of only the foreground lineage is fixed to 1 (MB). If Test I (MIa $\times$ MA) is significant but the likelihoods compared in Test II (MA $\times$ MB) are not statistically different, then the sites detected in Test I to be under positive selection are actually experiencing relaxation after release of selective constraints (described in detail by Zhang et al. [2005] and Arbiza et al. [2006] and in the PAML manual). RRTree and PAML do not perform well if too many gaps are present, so the alignments were trimmed when necessary in order to minimize the number of gaps.

Predicting the Presence or Absence of Autopodium: Random Forest Analysis

To determine whether the increased rates of *Hoxa-13* evolution observed in Squamata (see Results) could be explained by limb loss or indicative of snake-specific changes, we analyzed the sequence variation of this gene in a broad sample of gnathostoms (Fig. 1). Additional sequences for exon 1 of *Hoxa-13* were obtained for five fish ([basal] Osteichthyes + Chondrichthyes), three amphibians (a frog, a caecilian, and a salamander), one bird, one turtle, six additional limbless lizards (*Ophisaurus apodus*, *Isopachys gyldenstolpei*, *Rhineura floridana, Trogonophis*

wiegmanii, *Amphisbaena cubana*, and *Anniella pulchra*; see Supplementary Appendix I for details), and six additional limbed lizards (*Sceloporus undulatus*, *Eumeces inexpectatus, Ameiva auberi, Cordylus giganteus, Xantusia riversiana,* and *Takydromus sexlineatus*; see Supplementary Appendix I for details). The complete alignments, comprising 38 taxa, were obtained from amino acid sequences of *Hoxa-13* using different algorithms (ClustalV [Higgins et al. 1992]; ClustalX [Thompson et al. 1997]).

The random forest method approaches the problem as one of classification, where we seek to classify (or predict) the presence or absence of autopodium (the response variable) based on amino acids at positions in the *Hoxa-13* amino sequences, the predictor variables. Analyses were done using the statistical/machine learning method known as random forest (Breiman 2001a) with the randomForest package (Breiman et al. 2007) for the R statistical programming system (Ihaka and Gentleman 1996). Breiman (2001a, b) has demonstrated that relatively high levels of classification accuracy can be achieved through an ensemble of tree-based statistical models (Breiman et al. 1984; Clark and Pregibon 1993), where each tree in the ensemble includes some randomization in observations and predictor variables considered. Final predictions are obtained by aggregating (voting) over the ensemble. The mechanism by which aggregating models reduces prediction error for unstable predictors, such as trees, is well understood in terms of variance reduction resulting from averaging (Breiman 1996; Hastie et al. 2001). Further variance reduction can be achieved by decreasing the correlation between the quantities being averaged. Random forests seek to decrease correlation by incorporation of additional randomness by considering only a random subset of the predictor variables for determining the optimal split of a given node of a (constituent) tree rather than all predictors. The method has exceptional prediction accuracy and eliminates overfitting because ensemble members are generated independently (Breiman 2001a, b).

Our application of the random forest method follows previous successful application of tree-based statistical models (Segal et al. 2001) and random forests to genetic data (Cummings and Myers 2004; Cummings and Segal 2004). Here, our random forests comprised 1×10^5 individual tree-based statistical models. Variable importance was measured in terms of the increase in group purity when partitioning data based on the permutation accuracy importance procedure (Breiman et al. 2007). The procedure is broadly similar to other uses of permutation tests. A dataset is analyzed and the prediction accuracy is quantified. A variable is permuted in the dataset, the other variables remaining unpermuted, and the analysis is repeated. The difference in the prediction accuracy between the original (unpermuted) and the variable-permuted datasets is a measure of importance of the variable. If a predictor variable is strongly associated with response (i.e., contributes substantially to the prediction), a marked decrease in prediction accuracy will result from the permutation. Among the several advantages of permutation accuracy importance is that it provides an assessment of the importance of the variable in interaction, positive or negative, with other variables.

Results

Sequences from exon 1 of *Hoxa*-13 were PCR amplified, cloned, and sequenced for seven snake species and one outgroup (alligator) and, also, for two species of limbless squamates, one species of basal Osteichthyes, four species of limbed squamates, two species of amphibians, and one turtle species (see Supplementary Appendix I for details). Published sequences of *Hoxa-13* were downloaded from GenBank for three additional snake species, for three species of iguanids and two species of anguimorphs (as representatives of the 'limbed squamates' group), and also for four species of limbless squamates, one species of shark, three species of basal Osteichthyes, one frog species, two additional limbed squamates (used in the random forest analysis), and one bird species (details given in Supplementary Appendix I). For comparison, sequences of three nuclear genes, *c-mos*, *RAG-1*, and *brain-derived neurotrophic factor* (*BDNF*), were also obtained from GenBank: *RAG-1*—alligator, three species of iguanids, two species of anguimorphs, and six snake species; *c-mos*—alligator, three species of iguanids, two anguimorphs, and nine snake species; and *BDNF*—alligator, four species of iguanids, and six snake species (see Supplementary Appendix I for details). Phylogenetic analyses performed to determine the topology used for Squamata were based on data for entire mitochondrial genomes downloaded from GenBank (described in Supplementary Appendix II). Most of the analyses of *Hoxa-13* evolution were based on two main comparisons: (1) Serpentes versus 'limbed squamates' (the clade including Agamidae, Iguanidae, Varanidae, and Helodermatidae; Fig. 1) and (2) 'limbed' versus 'limbless' Ganathostomata (fish, amphibians, bird, turtle, and limbed and limbless squamates, the latter including snakes; see Fig. 1). Below we present our results in the form of (1) contrasts between five limbed squamates and Serpentes, in a phylogenetic context, and (2) classification analysis using random forest analysis of *Hoxa-13* sequences across Gnathostomata.

Serpentes Versus Limbed Squamates

Sequences of *Hoxa-13* were significantly longer in Serpentes than in limbed lizards ($p = 0.04$, $F_{14,15} = 12.490$), a

pattern that resulted from insertions of repeated elements, particularly the amino acids alanine and histidine (Fig. 2). Because related species share a common evolutionary history, we also performed a phylogenetic ANOVA in the MS-DOS computer program PDTREE (Garland et al. 1992, 1993, 1999; Garland and Ives 2000). This approach suggested that the magnitude of the differences was mostly explained by relatedness (which is not surprising given that the trend toward sequence elongation is associated with the Serpentes clade), as the phylogenetic ANOVA was not significant (F of the real data between the 19th and the 20th percentile of F values from simulated data; $p = 0.19$). This pattern seems to be specific to *Hoxa-13* because no length variation was observed in the other three nuclear genes used as control (*c-mos*, *RAG-1*, and *BDNF*). Serpentes exhibited a weak trend toward lower GC content in all genes analyzed (*Hoxa-13*, snakes = 66.5% GC, limbed squamates = 68.7% GC; *c-mos*, snakes = 44.4% GC, limbed squamates = 45.9% GC; *RAG-1*, snakes = 40.8% GC, limbed squamates = 42.1% GC; and *BDNF*, snakes = 50.7% GC, limbed squamates = 50.9% GC).

The second approach used to analyze molecular evolution of *Hoxa-13* was based on patterns of amino acid substitutions in the two clades compared, using two complementary programs that take into account a phylogenetic topology: RRTree (Rechavi and Huchon 2000) and PAML (Yang 1997). As described next, both parsimony and ML methods revealed increased rates of nonsynonymous substitutions in *Hoxa-13* associated with the Serpentes clade (Table 1). Distance methods implemented by RRTree (Robinson et al. 1998; Rechavi and Huchon 2000) showed that snakes exhibit a rate of nonsynonymous substitutions that is significantly higher than that of limbed squamates (Ka = 0.078, vs. 0.052; $p = 0.019$), as shown in Fig. 3. Rates of synonymous substitutions (Ks) did not differ between Serpentes (Ks = 0.594) and limbed (Ks = 0.440) squamates ($p = 0.083$; see Table 1).

To determine whether *Hoxa-13* experienced directional selection or relaxed selection in the stem lineage of snakes, we performed two complementary analyses based on likelihood-ratio tests (LRTs), implemented by PAML (Yang 1997). The first analysis, called Test I (described by Zhang et al. 2005; Arbiza et al. 2006), compared the likelihood of a 'nearly neutral model' (MIa), where rates of all branches were fixed to neutral, with that of an alternative model (MA) testing for positive selection in the branch leading to the Serpentes clade (indicated by the arrow in Fig. 3). This test detected five sites evolving under positive selection in the stem lineage of Serpentes (positions 54, 70, 79, 129, and 180 of the alignment presented in Fig. 2 and Table 1; $p < 0.05$). LRTs investigating positive selection in internal and tip branches of the Serpentes clade were not significant (Table 1). To distinguish between positive selection and relaxed selection after release of a constraint (i.e., selection in *Hoxa-13* after limb loss), we compared the likelihood of the model testing for positive selection in Test I (MA) with that of an alternative model (MB) where ω of the foreground lineage was fixed to 1 (Test II, as described by Zhang et al. [2005] and Arbiza et al. [2006]). The LRT of Test II was strongly significant (Table 1; $p < 0.001$) for the node leading to the Serpentes clade (Fig. 3), suggesting that the five sites identified by Test I (Fig. 2) experienced positive selection, instead of relaxed selection. It is important to note, in the alignment presented in Fig. 2, that these five sites change in Serpentes and are maintained within the clade, showing positive and then stabilizing selection in these amino acid positions. Test II did not show selection in *Hoxa-13* in other internal or tip branches in the Serpentes clade. Therefore we conclude that *Hoxa-13* has experienced positive selection in the stem lineage of Serpentes, but there is no evidence for relaxed selection on the gene in this clade.

To test whether the pattern of sequence evolution observed is a characteristic of snake genomes or if it is specific to the developmental gene *Hoxa-13*, we performed the same analyses for three other nuclear genes: *c-mos*, *RAG-1*, and *BDNF*. In the nuclear gene *c-mos*, the relative rate method (RRTree) did not detect significant differences in Ks (rate of synonymous substitutions) or Ka (rate of nonsynonymous substitutions) between limbed squamates and snakes, but LRTs detected relaxed selection in tip branches of the Serpentes clade (Table 1). In contrast, significant differences in the rate of nonsynonymous substitutions (Ka) between limbed squamates and Serpentes were detected in *RAG-1* by RRTree, but those did not reflect significant differences in ω values (nonsynonymous substitutions ratio/synonymous substitutions ratio, given by Ka/Ks), and LRTs showed that only tip branches in the Serpentes clade are experiencing relaxed (instead of positive) selection (Table 1). Therefore, Serpentes exhibited an overall increase in substitution rates in the gene *RAG-1*, which is not driven by directional selection in the stem lineage. Finally, neither parsimony nor likelihood tests provided evidence for significant differences in rates of nonsynonymous substitutions (Ka) between Serpentes and limbed squamates in the gene *BDNF*, but rates of neutral mutations (Ks) were higher in Serpentes (Table 1). Therefore, we conclude that the pattern of positive selection in the limb developmental gene *Hoxa-13* associated with the stem lineage of Serpentes is specific to *Hoxa-13*.

Random Forests and Limb Absence Among Gnathostomata

To determine whether the changes that occurred in *Hoxa-13* in the origin of Serpentes were at least in part related to

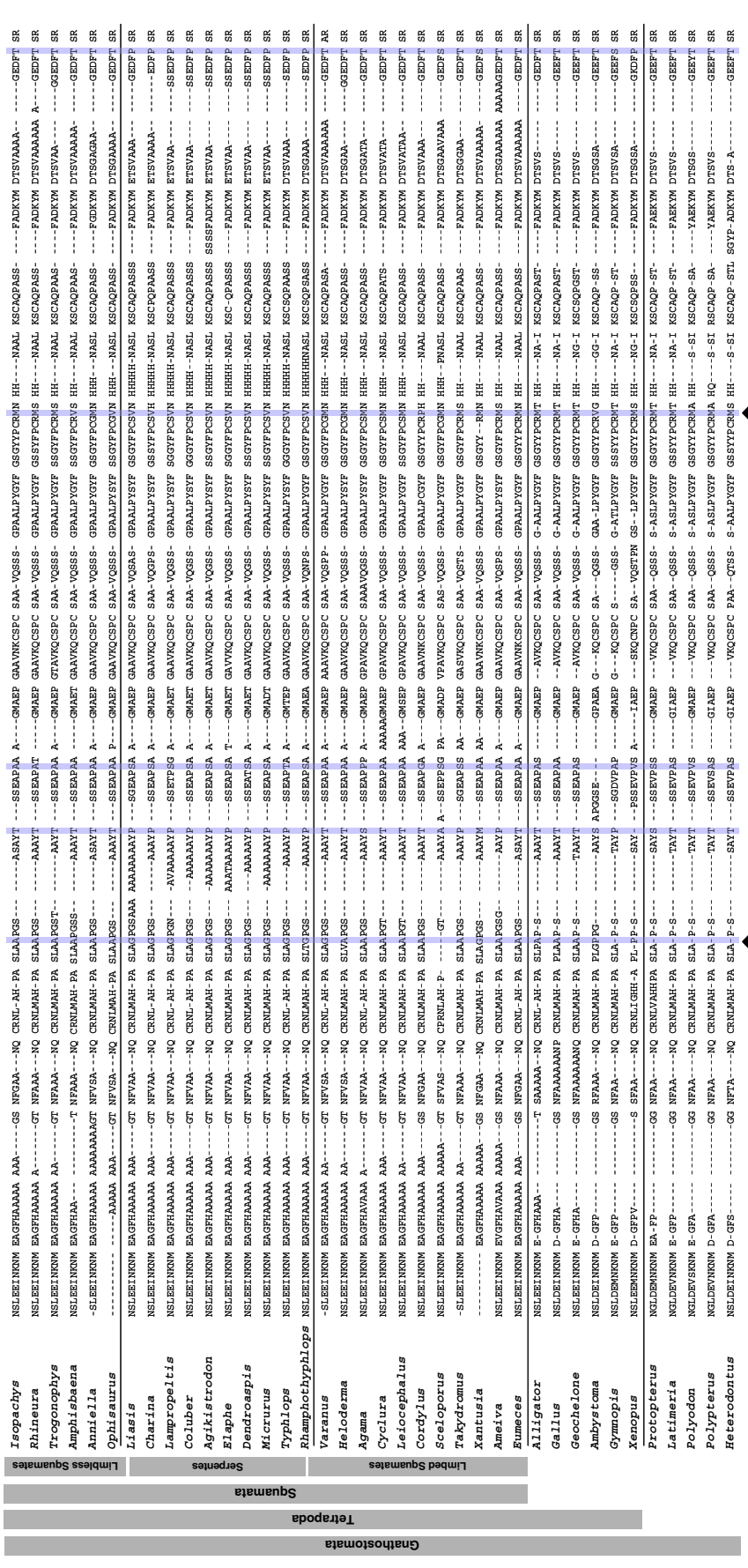

Fig. 2 Alignment of *Hoxa-13* for Gnathostomata. Gray columns correspond to the five sites detected to be under positive selection in comparisons between Serpentes and limbed squamates. Arrows indicate positions that were also confirmed to be among the most important for the classification of taxa according to 'presence' and 'absence' of autopodium in the random forest analysis

Table 1 Results from the RRTree test comparing substitution rates in *Hoxa-13* and three other nuclear genes between snakes and a clade of limbed lizards (top); likelihood ratio test from branch-site models applied in PAML for *Hoxa-13* and three other nuclear genes (bottom)

Variable	*Hoxa-13*			*c-mos*			*RAG-1*			*BDNF*		
	Limbed	Snakes	*p*	Limbed	Snakes	*p*	Limbed	Snakes	*p*	Limbed	Snakes	*p*
RRTree												
$Ka_{(nonsyn\ sub/nonsyn\ site)}$	0.052	0.078	0.019*	0.218	0.228	0.722	0.034	0.053	0.018*	0.044	0.043	0.879
$Ks_{(syn\ sub/syn\ site)}$	0.440	0.594	0.083	1.171	1.043	0.595	0.877	1.060	0.274	0.473	0.608	0.041*
Ka/Ks	0.118	0.131		0.186	0.218		0.039	0.051		0.093	0.071	
Test	*Hoxa-13*			*c-mos*			*RAG-1*			*BDNF*		
	TI	TII	Selection	TI	TII	Selection	TI	TII	Selection	TI	TII	Selection
PAML												
Branch to snakes clade	6.508*	60.92**	Positive	0.000	0.000	–	0.000	0.000	–	0.000	0.000	–
Internal branches	0.038	0.036	–	0.000	0.000	–	2.923	0.001	–	0.000	0.000	–
Tip branches	0.759	0.001	–	6.550*	5.197	Relaxed	6.867*	0.081	Relaxed	0.000	0.000	–

Note: Nonsyn, nonsynonymous; sub, substitions; syn, synonymous; TI, Test I; TII, Test II. TI (df = 2), critical $\chi^2_{0.05} = 5.991$. TII: (df = 1), critical $\chi^2_{0.05} = 3.841$. Statistically significant differences: * $p < 0.05$; ** $p < 0.001$

limb loss, we analyzed the sequence variation of this gene in a broad sample of gnathostoms and aligned the snake sequences of *Hoxa-13* with all sequences available for fish, amphibians, and reptiles, including six additional limbless lizards and six additional limbed lizards (Fig. 2). We have applied the random forest approach using two different alignments: Clustal-V (Higgins et al. 1992) and Clustal-X (Thompson et al. 1997). The alternative alignments looked equally reasonable and generated similar results (both with an error estimate of 18.42%), and much of the variation among alignments was associated with regions with variable numbers of identical amino acids (tandem repeats). The random forest method is an ensemble (forest) of decision trees and, like many statistical analyses, has no specific minimum sample sizes. This approach proved to be suitable for our dataset, in the sense that phenotypic states (autopodium present and absent) can often be correctly predicted based on *Hoxa-13* sequences.

The analyses based on the Clustal-V alignment detected six most important positions in the *Hoxa-13* sequence to predict the presence or absence of the autopodium (positions 30, 54, 82, 129, 134, and 150 in the alignment). Of the 22 taxa with 'absent autopodium,' only 2 were misclassified as 'present' (the lizard *Isopachys* and the caecilian *Gymnopis*; class error = 0.0909), while of the 16 limbed species, 5 were misclassified as 'absent autopodium' based on the *Hoxa-13* sequences (the lizards *Heloderma, Agama, Varanus*, and *Tadydromus* and the salamander *Ambystoma*; class error = 0.3125). The analyses based on a different alignment, using Clustal-X, gave the same overall estimate of error (18.42%), but different class errors: of the 22 taxa with 'absent autopodium,' 17 were correctly classified as 'absent' and 5 were misclassified as 'present' (the lizards *Trogonophis, Anniella, Isopachys,* and *Rhineura* and the caecilian *Gymnopis*; class error = 0.2273), while of the 16 species that exhibit autopodium development, 14 were correctly classified as 'present' and 2 were misclassified as 'absent' (the lizard *Varanus* and the salamander *Ambystoma*; classification error = 0.1250).

In general, there was broad consistency in the results based on the set of reasonable alignments examined. All the nontetrapod taxa (shark, lungfish, coelacanth, paddle fish, and bichir) were correctly classified as autopodless, and most of the nonsquamate tetrapods (chick, alligator, turtle, *Xenopus*) were correctly classified as limb bearing. Some specific taxa (i.e., the amphibians *Ambystoma* and *Gymnopis*, the limbed lizard *Varanus*, and the limbless lizard *Isopachys*) were misclassified more often than others, and classification ambiguity among the alignments was greatest in the limbless squamates. Two of the five positions identified to be under positive selective pressures in the stem lineage of snakes were also detected among the six most important positions for the classification of 'presence' and 'absence' of autopodium in Gnathostomata using the same alignment (Clustal-V; positions 54 and 129 in the alignment shown in Fig. 2).

Discussion

The present study investigates patterns of molecular evolution in a transcription factor gene, *Hoxa-13*, in contexts where one of the characters with which the gene is involved (development of the autopodium, i.e., hand/foot) is absent (by limb loss, in the case of Serpentes and limbless lizards, or was never present, where fish are also

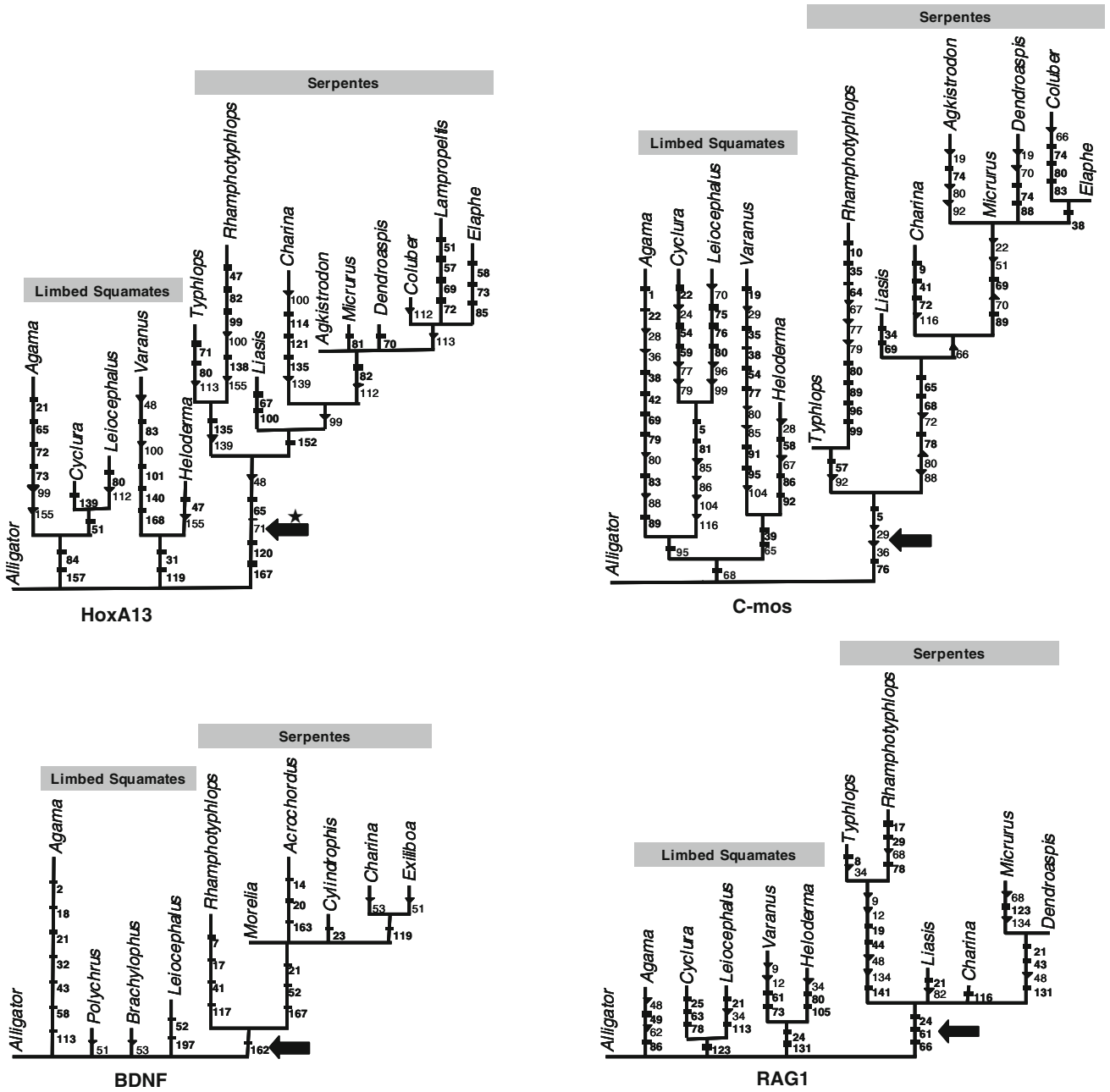

Fig. 3 Topologies of squamates (based on Fig. 1) showing the number of amino acid changes for *Hoxa-13*, *c-mos*, *BDNF*, and *RAG-1*, calculated in MacClade. Black arrows indicate the branch corresponding to the stem lineage that originates the Serpentes clade, and the star (present only in the topology of *Hoxa-13*) indicates significant changes in rates of evolution detected by RRTree and PAML

considered). In a scenario where one constraint (i.e., autopodium development) is released from a transcription factor (i.e., *Hoxa-13*) involved in multiple functions, one could expect relaxed selection in regions of the sequence that are associated with the function that was lost (Bull and Charnov 1985; Graur and Li 2000). This should be detectable as increased rates of amino acid sequence evolution in specific regions of the molecule in the lineage that lost limbs during evolution (i.e., Serpentes). Alternatively, after the release of the constraint associated with limb development these sites could also be free for acquiring new functions or for responding to formerly latent directional selection and, thus, may be under directional selection after limb loss.

The analyses based on relative rates and LRTs performed in the present study detected five amino acid sites in *Hoxa-13* evolving under positive selection among the snakes. Nonsynonymous substitutions occurred in the

origin of Serpentes in *Hoxa-13* positions 54, 70, 79, 129, and 180, and the amino acid replacements were maintained along the evolution of all Serpentes lineages, suggesting stabilizing selection after the substitution. These changes occurred concomitantly with the loss of limbs (and autopodium) in the snake clade, and two of them (positions 54 and 129) also seem to be related to the presence or absence of autopodium along the evolution of Gnathostomata. It is possible that the positive selection in the stem lineage of snakes not only is related to limb loss, but also is associated with the evolution of new functions during the development of regions where *Hoxa-13* is expressed in snakes. For example, *Hoxa-13* is also expressed during the development of the cloaca (Roberts et al. 1995; Yokouchi et al. 1995), and snakes present cloacal scent glands (or anal glands) that are unique to and characteristic of all Serpentes, suggesting that they might have played an important role in the origin and successful radiation of the group (Greene 1997; Young et al. 1999). However, the fact that fish and snakes share a sequence signature is the strongest evidence that a limb-related signature actually exists in this gene (and that turtle, alligator, chick, and mouse share a sequence signature).

Results of the random forest analysis (Breiman 2001a; Breiman et al. 2007) suggest that the phenotypic states 'autopodium present' and 'autopodium absent' can be correctly classified based on exon 1 sequences of *Hoxa-13* with a relatively low error rate. Most notably in this classification is that, of the species lacking an autopodium, neither the five "fish" species (one shark, two basal actinopterygians, and two nontetrapod sarcopterygians) nor any of the 10 snake species were misclassified as having limbs. Similarly, none of the nonsquamate eutetrapods (*Xenopus*, chicken, alligator, turtle) have been misclassified. Eutetrapods is a noncladistic term used for tetrapod taxa with canonical limb development (frogs and amniotes), excluding urodeles and caecilians. This consistency of classification cannot be explained as a statistical or phylogenetic artifact, since the "fish" and the snakes are the most distantly related species in our taxon sample, and the eutetrapods are also a diverse sample. Classification ambiguity is concentrated on the lizards and the two amphibians in our sample other than the frog. Given the phylogenetic distribution of consistent classification, it is unlikely that the classification is artifactual, and by implication it is likely that the classification ambiguity among amphibians and lizards is a biologically significant signal. Indeed, the contrasts observed in the multiple analyses performed in the present study illuminate very interesting cases, such as the ones mentioned above.

The ambiguity in the classification of the lizard *Hoxa-13* sequences is particularly strong, as it is complementary in the two alignments. In the ClustalV alignment 4 of the 11 lizards with limbs are misclassified as being limbless, while in the ClustalX alignment 4 of the 6 limbless lizards are misclassified as having limbs. To assess what this classification ambiguity may mean, we need to remember two facts. First, the squamates have a longstanding and ongoing tendency to evolve a snakelike morphology. In a recent analysis, Wiens et al. (2006) estimated that in their taxon sample of 216 species, limb loss has occurred at least 25 times independently, which the authors consider an underestimate of the true frequency of limb loss. The limbless skink *Isopachys* is consistently classified as having a *Hoxa-13* sequence of a limbed species. In this case it is likely that the limb loss is very recent, since the closest relatives of *Isopachys*, for example, *Lipinia vitterga* and *Sphenomorphus indicus* (Honda et al. 2000), are fully limbed skinks. It is thus likely that the relaxed selection has not yet led to substitutions on limb-related amino acid residues. Second, some limbless lizards seem to be phylogenetically younger than the Serpentes clade. For instance, *Anniella* is limbless for less than 70 million years (Mya), and *Trogonophis* lost its limbs less than 50 Mya, according to the dated phylogeny of Wiens et al. (2006). Furthermore, there is evidence that digit and limb loss in squamates might not be irreversible (Greer 1992; Kearney and Stuart 2004; Kohlsdorf and Wagner 2006; Whiting et al. 2003; Brandley et al. 2008). It is plausible that many lizard lineages experienced total or partial loss of the autopodium in their ancestral lineages even though the current species may have fully formed limbs (e.g., *Bipes biporus* [Amphisbaenidae], *Scelotes mirus* [Scincidae], and *Bachia panoplia* and *B. scolecoides* [Gymnophthalmidae]). Although Squamata might have originated in the Permian-early Triassic (about 290-230 Mya), apparently most families of 'typical' snakes diverged in the Tertiary, some of them as early as 50-30 Mya (Kumazawa 2007). The origin of snakes is still controversial (e.g., Coates and Rutta 2000; Greene and Cundall 2000; Lee et al. 2000), and recent fossil discoveries support a scenario where some extinct snake species may have had 're-evolved' limbs (Apesteguía and Zaher 2006). However, it is a fact that none of the extant snakes known exhibit autopodium, which may explain why all the species studied here are correctly classified as limbless based on their exon 1 *Hoxa–13* amino sequence. In contrast, several lizard families are older than snakes (Kumazawa 2007), and groups that exhibit limb reduction usually enclose both limbed and limbless extant species (Wiens et al. 2006). Therefore, it is biologically plausible that the *Hoxa-13* gene sequence is ambiguous with respect to its association with the presence or absence of limbs in lizards, but not in snakes.

Most puzzling is the consistent misclassification of the caecilian *Gymnopis*, a limbless species calssified as limbed,

and *Ambystoma*, a fully limbed urodele classified as limbless. These results require a more extensive taxon sample to assess their significance (Kohlsdorf and Wagner, in preparation) but are potentially interesting. In particular, the "misclassification" of the urodele *Ambystoma* is interesting in the light of the highly derived mode of digit development of urodeles in general (Wagner et al. 1999; Stopper and Wagner 2005). Urodele digit development proceeds from anterior to posterior, while all other tetrapods develop their digits from posterior to anterior (Shubin and Alberch 1986). In urodeles *Hoxa-11* expression is not limited to the lower extremities, as in birds and mammals, but includes digits 3 and 4 of the hand (Wagner et al. 1999), and *Hoxd-11* expression is weak in the autopodium of *Ambystoma* (Torok et al. 1998), while it is strong and characteristic of the autopodium in other tetrapods including frogs (Stopper and Wagner 2005). The explanation of this derived limb development is controversial. One theory explains it as an adaptation to pond larva life (Wake and Marks 1993; Wake and Shubin 1994), while another theory explains it as a result of re-evolution of digits after the loss of most digits in the stem lineage of urodeles (Wagner et al. 1999). If a broader sample of urodele *Hoxa-13* sequences turns out to be consistently classified as typical for limbless forms, this would shed light on the origin of urodele limb development (Kohlsdorf and Wagner, in preparation).

Conclusion

The origin of limbs and the limb loss leave a consistent signal in the *Hoxa-13* sequence of a wide array of species ranging from sharks to birds, snakes, and turtles. This remarkable range of species correctly classified as either having or not having an autopodium, based on the *Hoxa-13* exon 1, suggests the existence of limb development-related amino acid motives in the Hoxa-13 protein. Ambiguity in the classification of *Hoxa-13* sequences is mostly focused on lizards, a large class of lineages with a high tendency toward limb loss (Wiens et al. 2006) and even a proven ability to regain digits after digit loss (Kohlsdorf and Wagner 2006; Brandley et al. 2008). We thus think that both the consistent classifications of most forms and the ambiguity of classification of the lizard sequence have biological significance.

Acknowledgments The authors thank K. Crow for help with lab molecular techniques and discussion of results, and Jacques Gauthier and Greg Watkins-Colwell (Yale Peabody Museum), Adalgisa Caccono (Yale University EEB), and David Wake (Museum of Vertebrate Zoology, University of California, Berkeley) for donation of tissue samples. We thank Drs. H. Greene, J. Gauthier, and P. Weldon for discussion of morphological patterns in snakes, Dr. K. Schwenk for discussion of phylogenetic relationships in Squamata, and M. Conte for assistance with some of the analyses. T.K. was supported by a CNPq/Brazil postdoctoral fellowship (PDE-201244/2003-9) and is supported by a FAPESP/Brazil grant (05/60140-4); G.P.W. is supported by a NSF/USA grant (IOB-0445971).

References

Apesteguía S, Zaher H (2006) A Cretaceous terrestrial snake with robust hindlimbs and a sacrum. Nature 440:1037–1040

Arbiza L, Dopazo J, Dopazo H (2006) Positive selection, relaxation, and acceleration in the evolution of the human and chimp genome. PloS Comput Biol 2(4):e38/0238-0300. doi:10.1371/journal.pcbi.0020038

Bejder L, Hall BK (2002) Limbs in whales and limblessness in other vertebrates: mechanisms of evolutionary and developmental transformation and loss. Evol Dev 4(6):445–458

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

Breiman L (1996) Bagging predictors. Mach Learn 24:123–140

Breiman L (2001a) Random forests. Mach Learn 45:5–32

Breiman L (2001b) Statistical modeling: the two cultures. Stat Sci 16:199–215

Breiman L, Friedman JH, Olshen RA, Stone CJ (1984) Classification and regression trees. Wadsworth and Brooks, Pacific Grove, CA

Breiman L, Cutler A, Liaw A, Wiener M (2007) Available at: http://cran.r-project.org/src/contib/Descriptions/randomForest.html

Bull JJ, Charnov EL (1985) On irreversible evolution. Evolution 39:1149–1155

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

Caldwell MW, Lee MSY (1997) A snake with legs from the marine Cretaceous of the Middle East. Nature 386:705–709

Carroll SB (2005) Endless forms most beautiful: the new science of evo devo and the making of the animal kingdom. W. W Norton, New York

Clark LA, Pregibon D (1993) Tree-based models. In: Chambers JM, Hastie TJ (eds) Statistical models in S. Chapman and Hall, London, pp 377–419

Coates M, Rutta M (2000) Nice snake, shame about the legs. TREE 15:503–507

Cummings MP, Myers DS (2004) Simple statistical models predict C-to-U edited sites in plant mitochondrial RNA. BMC Bioinformatics 5:132

Cummings MP, Segal MR (2004) Few amino acid positions in rpoB are associated with most of the rifampin resistance in Mycobacterium tuberculosis. BMC Bioinformatics 5:137

Davis AP, Witte DP, Hsieh-Li HM, Potter SS, Capecchi MR (1995) Absence of radius and ulna in mice lacking hoxa-11 and hoxd-11. Nature 375(6534):791–795

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

Duboule D (1994) Temporal colinearity and the phylotypic progression: a basis for the stability of a vertebrate Bauplan and the evolution of morphologies through heterochrony. Dev Suppl 1994:135–142

Favier B, Rijli FM, Fromental-Ramain C, Fraulob V, Chambon P, Dolle P (1996) Functional cooperation between the non-paralogous genes Hoxa-10 and Hoxd-11 in the developing forelimb and axial skeleton. Development 122(2):449–460

Felsenstein J (1985) Phylogenies and the comparative method. Am Nat 125:1–15

Fromental-Ramain C, Warot X, Messadecq N, LeMeur M, Dolle P, Chambon P (1996) *Hoxa-13* and *Hoxd-13* play a crucial role in the patterning of the limb autopod. Development 122:2997–3011

Fry BG, Vidal N, Norman JA, Vonk FJ, Scheib H, Ramjan FR, Kuruppu S, Fung K, Hedges SB, Richardson MK, Hodgson WC, Ignjatovic V, Summerhayes R, Kochva E (2006) Early evolution of the venom system in lizards and snakes. Nature 439(7076):584–588

Garland T Jr, Ives AR (2000) Using the past to predict the present: confidence intervals for regression equations in phylogenetic comparative methods. Am Nat 155:346–364

Garland T Jr, Harvey PH, Ives AR (1992) Procedures for the analysis of comparative data using phylogenetically independent contrasts. Syst Biol 41(1):18–32

Garland T Jr, Dickerman AW, Janis CM, Jones JA (1993) Phylogenetic analysis of covariance by computer simulation. Syst Biol 42(3):265–292

Garland T Jr, Midford PE, Ives AR (1999) An introduction to phylogenentically based statistical methods, with a new method for confidence intervals on ancestral values. Am Zool 39:374–388

Graur D, Li WH (2000) Fundamentals of molecular evolution. Sinauer Associates, Sunderland, MA

Greene HW (1997) Snakes—the evolution of mystery in nature. University of California Press, Berkeley

Greene HW, Cundall D (2000) Evolutionary biology: limbless tetrapods and snakes with legs. Science 287:1939–1941

Greer AE (1991) Limb reduction in squamates: identification of the lineages and discussion of the trends. J Herpetol 25:166–173

Greer AE (1992) Hyperphalangy in squamates: insights on the re-acquisition of primitive character states in limb-reduced lineages. J Herpetol 26:327–329

Hastie TJ, Tibshirani R, Friedman JH (2001) The elements of statistical learning. Springer, New York

Hérault Y, Beckers J, Gérard M, Duboule D (1999) *Hox* gene expression in limbs: colinearity by opposite regulatory controls. Dev Biol 208:157–165

Higgins DG, Bleasby AJ, Fuchs R (1992) CLUSTAL-V-improved software for multiple sequence alignment. Comp Appl Biosci 8:189–191

Honda M, Ota H, Kobayashi M, Nabhitabhata J, Yong HS, Hikida T (2000) Phylogenetic relationships, character evolution, and biogeography of the subfamily Lygosominae (Reptilia: Scincidae) inferred from mitochondrial DNA sequences. Mol Phylogenet Evol 15(3):452–461

Ihaka R, Gentleman R (1996) R: a language for data analysis and graphics. J Comput Graph Stat 5:299–314

Innis JW, Margulies EH, Kardia S (2002) Integrative biology and the developing limb bud. Evol Dev 4(5):378–389

Innis JW, Mortlock D, Chen Z, Ludwig M, Williams ME, Williams TM, Doyle CD, Shao Z, Glynn M, Mikulic D, Lehmann K, Mundlos S, Utsch B (2004) Polyalanine expansion in *Hoxa-13*: three new affected families and the molecular consequences in a mouse model. Hum Mol Genet 13(22):2841–2851

Kearney M, Stuart BL (2004) Repeated evolution of limblessness and digging heads in worm lizards revealed by DNA from old bones. Proc Roy Soc Lond B 271:1677–1683

Knosp WM, Scott V, Bächgubger GO, Stadler HS (2004) *HOXA13* regulates the expression of bone morphogenetic proteins 2 and 7 to control distal limb morphogenesis. Development 131:4581–4592

Kohlsdorf T, Wagner GP (2006) Evidence for the reversibility of digit loss: a phylogenetic study of limb evolution in *Bachia* (Gymnophthalmidae, Squamata). Evolution 60(9):1896–1912

Kumazawa Y (2007) Mitochondrial genomes from major lizard families suggest their phylogenetic relationships and ancient radiations. Gene 388:19–26

Lande R (1978) Evolutionary mechanisms of limb loss in tetrapods. Evolution 32:73–92

Lee MSY, Scanlon JD, Caldwell MW (2000) Snake origins. Science 288:1343–1345

Lynch VJ, Roth JJ, Takahashi K, Dunn C, Nonaka DF, Stopper G, Wagner GP (2004) Adaptive evolution of Hoxa-11 and Hoxa-13 at the origin of the uterus in mammals. Proc Roy Soc Lond B 271:2201–2207

Maddison WP, Maddison DR (1992) MacClade: analysis of phylogeny and character evolution. Sinauer Associates, Sunderland, MA

Maddison WP, Maddison DR (2004) Available at: http://mesquiteproject.org

Marshall CR, Raff EC, Raff RA (1994) Dollo's law and the death and resurrection of genes. Proc Natl Acad Sci USA 91:12283–12287

Metscher BD, Takahashi K, Crow K, Amemiya C, Nonaka DF, Wagner GP (2005) Expression of *Hoxa-11* and *Hoxa-13* in the pectoral fin of a basal ray-finned fish, *Polyodon spathula*: implications for the origin of tetrapod limbs. Evol Dev 7(3):186–195

Mortlock DP, Innis JW (1997) Mutation of *HOXA-13* in hand-foot-genital syndrome. Nat Genet 15:179–180

Mortlock DP, Sateesh P, Innis JW (2000) Evolution of N-terminal sequences of the vertebrate *Hoxa-13* protein. Mammal Gen 11:151–158

Nelson CE, Morgan BA, Burke AC, Laufer E, DiMambro E, Murtaugh LC, Gonzales E, Tessarollo L, Parada LF, Tabin C (1996) Analysis of *Hox* gene expression in the chick limb bud. Development 122:1449–1466

Pagel MD (1992) A method for the analysis of comparative data. J Theor Biol 156:431–442

Pond SLK, Frost SDW, Muse SV (2005) HyPhy: hypothesis testing using phylogenies. Bioinformatics 21(5):676–679. doi:10.1093/bioinformatics/bti079

Post LC, Innis JW (1999) Altered *Hox* expression and increased cell death distinguish *Hypodactyly* from *Hoxa13* null mice. Int J Dev Biol 43:287–294

Pough FH, Andrews RM, Cadle JE, Crump ML, Savitzki AH, Wells KD (1998) Herpetology. Prentice-Hall, Upper Saddle River, NJ

Rage JC (2001) Fossil history. In: Seigel RA, Collins JT, Novak SS (eds) Snakes—ecology and evolutionary biology. Blackburn Press, Caldwell, NJ

Rage JC, Escuillié F (2000) Un nouveau serpent bipède du Cénomanien (Crétacé) Implications phylétiques. CR Acad Sci Paris 330:513–520

Rechavi MR, Huchon D (2000) RRTree: relative-rate tests between groups of sequences on a phylogenetic tree. Bioinformatics 16:296–297

Roberts DJ, Johnson RL, Burke AC, Nelson CE, Morgan BA, Tabin C (1995) Sonic hedgehog is an endodermal signal inducing Bmp-4 and Hox genes during induction and regionalization of the chick hindgut. Development 121:3163–3174

Robinson M, Gouy M, Gautier C, Mouchiroud D (1998) Sensitivity of the relative-rate test to taxonomic sampling. Mol Biol Evol 15(9):1091–1098

Segal MR, Cummings MP, Hubbard AE (2001) Relating genotype to phenotype: analysis of peptide binding data. Biometrics 57:632–643

Shubin NH, Alberch P (1986) A morphogenetic approach to the origin and basic organization of the tetrapod limb. Evol Biol 20:319–387

Stadler HS, Higgins KM, Capecchi MR (2001) Loss of Eph-receptor expression correlates with loss of cell adhesion and

chondrogenic capacity in *Hoxa-13* mutant limbs. Development 128:4177–4188

Stopper GS, Wagner GP (2005) Of chicken wings and frog legs: a somrgasbord of evolutionary variation in mechanisms of tetrapod limb development. Dev Biol 288:21–39

Tchernov E, Rieppel O, Zaher H, Polcyn MJ, Jacobs LL (2000) A fossil snake with limbs. Science 287:2010–2012

Teotónio H, Rose MR (2001) Perspective: reverse evolution. Evolution 55(4):653–660

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 24:4876–4882

Torok MA, Gardiner DM, Shubin NH, Bryant SV (1998) Expression of *HoxD* genes in developing and regenerating axolotl limbs. Dev Biol 200:225–233

Veraksa A, Campo MD, McGinnis W (2000) Developmental patterning genes and their conserved functions: from model organisms to humans. Mol Genet Metab 69:85–100

Vidal N, Hedges BS (2005) The phylogeny of squamate reptiles (lizards, snakes, and amphisbaenians) inferred from nine nuclear protein-coding genes. CR Biol 328:1000–1008

Wagner GP, Lynch VJ (2008) The gene regulatory logic of transcription factor evolution. Tree 23(7):377–385

Wagner GP, Khan PA, Blanco MJ, Misof G, Liversage RA (1999) Evolution of *Hoxa-11* expression in amphibians: is the urodele autopodium an innovation? Am Zool 39:686–694

Wagner GP, Takahashi K, Lynch VJ, Prohaska SJ, Fried C, Stadler PF, Amemiya C (2005) Molecular evolution of duplicated ray finned fish HoxA clusters: increased synonymous substitution rate and asymmetrical co-divergence of coding and noncoding sequences. J Mol Evol 60:665–676

Wake DB, Marks SB (1993) Development and evolution of plethodontid salamanders—a review of prior studies and a prospectus for future research. Herpetologica 49:194–203

Wake DB, Shubin N (1994) Urodele limb development in relation to phylogeny and life history. J Morphol 220:407–408

West-Eberhard M (2003) Developmental plasticity and evolution. Oxford University Press, Oxford, UK

Whiting AS, Bauer AM, Sites JW (2003) Phylogenetic relationships and limb loss in sub-Saharan African scincine lizards (Squamata: Scincidae). Mol Phylogenet Evol 29:582–598

Wiens JJ, Brandley MC, Reeder TW (2006) Why does a trait evolve multiple times within a clade? Repeated evolution of snakelike body form in squamate reptiles. Evolution 60:123–141

Wright S (1968) Evolution and the genetics of populations, vol I. Genetic and biometric foundations, University of Chicago Press, Chicago

Yang Z (1997) PAML: a program package for phylogenetic analysis by maximum likelihood. Comp Appl Biosci 13:555–556

Yokouchi Y, Nakazato S, Yamamoto M, Goto Y, Kameda T, Iba H, Kuroiwa A (1995) Misexpression of *Hoxa–13* induces cartilage homeotic transformation and changes cell adhesiveness in chick limb buds. Gen Dev 9:2509–2522

Young BA, Marsit C, Meltzer K (1999) Comparative morphology of the cloacal scent gland in snakes (Serpentes: Reptilia). Anat Rec 256(2):127–138

Zákány J, Duboule D (1999) *Hox* genes in digit development and evolution. Cell Tissue Res 296:19–25

Zhang J, Nielsen R, Yang Z (2005) Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. Mol Biol Evol 22(12):2472–2479