Article

Accumulation of transposable elements in *Hox* gene clusters during adaptive radiation of A*nolis* lizards

Nathalie Feiner[1,2]*

1 Department of Biology, Lund University, Sölvegatan 37, 223 62 Lund, Sweden

2 Department of Zoology, University of Oxford, Oxford OX1 3PS, UK

* Nathalie.feiner@biol.lu.se

Running head: Transposable elements in *Anolis* lizards

Keywords: adaptive radiation, *Anolis* lizards, *Hox* gene cluster, transposable elements, speciation

Data archival location: GenBank database (accession numbers LN899470-LN899543)

Abstract

Transposable elements (TEs) are DNA sequences that can insert elsewhere in the genome and modify genome structure and gene regulation. The role of transposable elements in evolution is contentious. One hypothesis posits that TE activity generates genomic incompatibilities that can cause reproductive isolation between incipient species. This predicts that TEs will accumulate during speciation events. Here I tested the prediction that extant lineages with a relatively high rate of speciation have a high number of TEs in their genomes. I sequenced and analysed the TE content of a marker genomic region (*Hox* clusters) in *Anolis* lizards, a classic case of an adaptive radiation. Unlike other vertebrates, including closely related lizards, *Anolis* lizards have high numbers of TEs in their *Hox* clusters, genomic regions that regulate development of the morphological adaptations that characterize habitat specialists in these lizards. Following a burst of TE activity in the lineage leading to extant *Anolis*, TEs have continued to accumulate during or after speciation events, resulting in a positive relationship between TE density and lineage speciation rate. These results are consistent with the prediction that TE activity contributes to adaptive radiation by promoting speciation. Although there was no evidence that TE density *per se* is associated with ecological morphology, the activity of TEs in *Hox* clusters could have been a rich source for phenotypic variation that may have facilitated the rapid parallel morphological adaptation to microhabitats seen in extant *Anolis* lizards.

1. Introduction

Transposable elements are mobile genetic units that have the potential to multiply and populate the genome [1,2]. Unlike point mutations, their activity can lead to deletions, insertions and chromosomal rearrangements (e.g. through unequal crossing over), and thus impact gross genomic architecture [3]. While this may often be deleterious to individuals, the ability of TEs to bring about fast genomic changes might at times contribute to adaptation and diversification [4-8]. For example, TE activity may accelerate speciation by causing genomic incompatibilities between incipient species [9,10]. This hypothesis predicts that lineages with abundant TEs will show a higher rate of diversification than lineages with fewer TEs. This is consistent with the observation that some of the most iconic adaptive radiations, such as *Heliconius* butterflies [11,12], African cichlids [13-15] or angiosperms [16] show unusually high numbers of TEs in their genomes. However, evidence for a positive relationship between TE abundance and the rate of speciation is lacking. Insertion of TEs may also contribute to phenotypic diversification of species through its effects on gene regulation, for example by modification of the epigenome or rewiring of gene networks through shuffling of regulatory elements [17-20]. Indeed, a number of evolutionary innovations, such as egg-spots in cichlids and the mammalian placenta, appear to have originated through the co-option of newly inserted TEs [21-23].

Anolis lizards are a remarkable example of an adaptive radiation. Following colonization of the Caribbean islands starting around 40 mya [24], species repeatedly diversified into habitat specialists (i.e., ecomorphs [24,25]). On some islands, numerous species belonging to up to six different ecomorphs originated from a single ancestor, whereas only a small number of species evolved on other islands. Ecomorphs are characterized by changes in body shape, including modifications of limbs [24], which are likely to be caused by changes in gene regulation [26].

Rapid speciation and morphological diversification make *Anolis* lizards a prime candidate for TE-driven adaptive radiation [15]. Recent evidence also points towards an important role for TEs in the evolution of the *Anolis* genome. Firstly, the genome of the green anole, *Anolis carolinensis*, contains a very high number of relatively young, active elements, in contrast to bird and mammalian genomes that are rich in remnants of old, extinct TEs [27,28]. Thus, proliferation of TEs may have contributed to genomic incompatibilities that enabled formation of distinct species during the adaptive radiation.

Secondly, relative to mammal and bird genomes, TEs appear to be enriched in developmental genes that control morphogenesis [29]. Specifically, in contrast to the highly condensed clusters of *Hox* genes in other vertebrates, *A. carolinensis Hox* clusters are interspersed by a large number of TEs, a feature that may impact on the temporal and spatial expression patterns during morphogenesis [29]. Indeed, a relationship between *Hox* gene expression and limb length has recently been demonstrated experimentally in the mouse [30-32]. Consequently, TE insertions in *Hox* clusters may have contributed to the repeated evolutionary diversification of body and limb forms that today are represented in the recognized ecomorphs of *Anolis* lizards.

To test if TE abundance is positively associated with speciation, I characterized the TE content in the *Hox* gene clusters of 20 species of *Anolis* lizards and ten additional lepidosaurian reptiles (figure 1a). I sequenced posterior parts of the *HoxA* and *–D* clusters (figure 1b and 1c), which are reliable candidate marker regions for TE density and vital for development, in particular limb morphogenesis [33,34]. TE contents were compared among species, and to genomic regions outside the *Hox* cluster of the green anole. The results show that proliferation of TEs is associated with speciation, suggesting that TE activity has played a role in the adaptive diversification of *Anolis* lizards.

2. Material and methods

(a) PCR amplification and DNA sequencing

Genomic DNA was extracted from 20 *Anolis* species and three other lepidosaurians (electronic supplementary material, table S1) using the Qiagen DNeasy Blood & Tissue kit. *Anolis* species were sampled across the phylogeny, and representatives of diverse geographical origins (Greater and Lesser Antillean islands and North America) and of diverse ecomorphs (crown-giant, grass-bush, trunk, trunk-crown and trunk-ground) were chosen. 10 ng of DNA were used as template in long-range PCRs with the Phusion Green High-Fidelity DNA Polymerase (Thermo Fisher Scientific). All primers were designed based on the *Anolis carolinensis* genome assembly ([28]; electronic supplementary material, table S2). To increase PCR specificity in more distantly related species (e.g. non-*Anolis* lepidosaurians), primer binding sites were chosen in the coding region of *Hox* genes and nested PCRs were performed.

The long intergenic region between *HoxA13* and *–A11* (23.5 kb) was bridged using a conserved non-coding element as a primer binding site (see figure 1c). PCR bands were cut from agarose gels and purified with QIAquick Gel Extraction Kit, or QIAEX II Gel Extraction Kit for fragments above 10 kb length. The obtained fragments were pooled per species in equimolar amounts and libraries were prepared with the Nextera XT and indexing kit (Illumina). These 23 libraries were sequenced on a MiSeq platform using a 300 bp paired end protocol.

Quality control of the output of the MiSeq run was performed with the FastQC software (URL: http://www.bioinformatics.babraham.ac.uk/projects/fastqc). Low quality reads were removed and a sliding window approach was used to trim low quality bases at the ends of the reads using Trimmomatic ([35]; quality threshold 30; minimum length 36). Paired end reads were merged into contigs using FLASH software [36]. These contigs were used as input for the SPAdes assembler tool (version 3.1.1 with standard settings; [37]). In cases where the expected fragments were not retrieved in this initial assembly, a subsampling strategy was applied to compensate for the high coverage and the assembly process was repeated. *Anolis carolinensis* intergenic fragments were used as queries in blastn searches against each species' SPAdes assembly to assign identities to the individual fragments. A total of 82 genomic fragments for *Anolis* and 7 genomic fragments for other lepidosaurians were obtained (electronic supplementary material, table S1) and deposited in the GenBank database (accession numbers LN899470-LN899543). Note that the two PCR fragments between *HoxA13* and *–A11* were submitted as one concatenated fragment, but analysed as two separate fragments.

(b) Retrieval of *Hox* intergenic sequences of non-*Anolis* lepidosaurians

In addition to the sequences obtained by long-range PCRs, intergenic regions were downloaded from the NCBI database for the following species: the slow worm *Anguis fragilis* (GenBank accession no. GU320307 for *HoxD12-D10*), the golden gecko *Gekko ulikovskii* (GenBank accession no. GU320308 for *Hox*D12-D11-D10), the corn snake *Pantherophis guttatus* (GenBank accession no. GU320305 for *HoxA13-A10*, and GU320304.1 for *HoxD12-D10*), and the tuatara *Sphenodon punctatus* (GenBank accession no. GU320310 for *HoxD12-D10*). *Hox* intergenic regions were curated from publicly available genome assemblies of the Burmese Python *Python molurus bivittatus* ([38], scaffold 1009 for

HoxA cluster and scaffold 548 for *HoxD* cluster), the king cobra *Ophiophagus Hannah* ([39]; scaffold 260 for *HoxA* cluster and scaffold 55 for *HoxD* cluster) and the central bearded dragon *Pogona vitticeps* ([40]; scaffold 3 for *HoxA* cluster and scaffold 430 for *HoxD* cluster).

(c) TE annotation

Hox intergenic regions obtained in long-range PCRs and from public databases were aligned using the mLAGAN software [41] resulting in multiple sequence alignments of 12.5 kb average length. The Censor tool with default parameters (against the vertebrate Repbase library, version 20.06, 3 June 2014) was used to annotate TEs on these *Hox* intergenic regions [42]. Reptilian entries of the vertebrate Repbase library are highly biased towards *Anolis carolinensis*, and therefore I complemented the annotation with a screen of a snake-specific repeat library (pit viper and Burmese python RepeatModeler library; available from URL: http://www.snakegenomics.org). All predicted TEs were filtered and only TEs with at least 50 bp length, 75% identity, and a score of >500 to a Repbase entry were further investigated. For the intra-genomic analysis of TE content in the *A. carolinensis* genome, TE annotations were obtained from RepeatMasker open-4.0.5, Repeat Library 20140131 [43,44].

(d) TE density and fragment length statistical analysis

Only *Hox* intergenic regions that were obtained in full-length were included in this analysis. Sequenced fragments that did not overlap with the open reading frame of flanking *Hox* genes, or sequences that were disrupted by stretches of 'N's, were excluded. TE density was calculated as the number of TEs per 1 kb. In order to be able to compare metrics between the six fragments, the fragment length in bp and the number of TEs were transformed, such that the longest fragment and the maximum number of TEs were both scaled to 1. To test if patterns of TE proliferation differed from that of other repetitive sequences I identified microsatellites in the *Hox* clusters using the software ATRHunter [45]. The number of repeats was summed for each *Anolis* species and for each of the nine microsatellites that were identified, and the maximum number was scaled to 1.

(e) Mapping of *A. carolinensis* transcriptomes against *Hox* intergenic regions

I downloaded publicly available organ-specific transcriptomes of *A. carolinensis* from the NCBI sra archive (two developmental stages and eight adult organs, accession IDs: embryo stage 38, SRR389085; embryo stage 28, SRR389086; brain, SRR540258; ovary, SRR543709; dewlap, SRR543711; heart, SRR540256; skeletal muscle, SRR391650; liver, SRR391651; lungs, SRR391654; adrenal glands, SRR495267). Raw reads of each transcriptomic dataset were mapped against intergenic regions of the posterior *HoxA* and *–D* clusters using the alignment tool bowtie2 [46]. Results were visualized in the Integrative Genomics Viewer [47] with the annotated TEs as input GFF file (electronic supplementary material, figure S1).

(f) Phylogenetic comparative analysis

I employed the method described by Freckleton *et al.* [48] to test systematically for the relationship between TE density and diversification rate, and TE density and limb length within the *Anolis* clade. This method is particularly useful for relatively small data sets as it allows the inference of speciation nodes to be performed on the complete phylogeny on anoles, while the phylogenetic generalized least squares (pgls) analysis is conducted on a pruned dataset containing only the focal species with available trait data. Although more sophisticated methods for testing predictors of the rate of speciation are available [49], these require a more complete taxonomic sampling and are therefore not suitable to these data [50,51]. Phenotypic measures (snout-vent and long bone length) and islands of origin of *Anolis* species were obtained from the literature [24,52,53]. Geographic area estimates were obtained from URL: http://www.worldatlas.com for Greater Antillean islands and from [54] for Lesser Antillean islands. As a measure of net speciation rate I used 'node depth', which is the number of speciation nodes linking a given extant species to the last common ancestor of *Anolis* [48]. I based this analysis on the phylogenetic tree published by Gamble *et al.* [55] which contains 216 out of 387 described species [53]. The species *A. forresti* was not part of this dataset, and it was added to the phylogenetic tree as sister species to *A. wattsi* [56]. Although it should be noted here that estimating rates of extinction is complex and prone to large errors [57-59], a previous study suggested that the diversification patterns of the genus

Anolis are characterized by low levels of extinctions [60]. I therefore conducted the analyses assuming equal and constant extinction rates, which means that the number of speciation nodes will be proportional to the rate of speciation (see also Discussion).

I accounted for phylogenetic non-independence using a pgls approach [61,62] in the R package caper (PGLS code; [63]) with a variance-covariance matrix inferred from the phylogenetic tree published by Gamble *et al.* [55] pruned to contain only *Anolis* species with available trait data. To ensure robustness, the results of these pgls analyses were replicated using two alternative phylogenetic datasets of *Anolis*: the dataset of Poe [64] containing 174 *Anolis* species, which is based on a combination of morphological and molecular markers (concatenated from chromosome, allozyme, immunological and DNA sequence data), and the molecular dataset of Nicholson *et al.* [53] containing 189 taxa (electronic supplementary material, table S3 and S4). For all models, diagnostic plots were examined to check for normal distribution of errors and heteroscedasticity. To test if there is a relationship between the rate of changes in limb length and TE density, as may be predicted if TE activity stimulate phenotypic variation, a phylogenetically independent contrast analysis [65] was performed. Limb contrast was calculated as the absolute difference of the limb measure between a pair of species, where the limb measure was defined as the mean of humerus, femur, radius and tibia divided by the snout-vent length of the species. This limb contrast was determined for every node of the pruned phylogenetic tree containing the 16 *Anolis* species for which both TE density and limb data are available. Limb contrast was regressed on the reconstructed ancestral TE density (mean of descendent nodes or species; electronic supplementary material, figure S2).

3. Results

(a) *Anolis Hox* clusters have high TE densities

Given that the structure of the *Anolis Hox* clusters is so strikingly different from all other vertebrate *Hox* clusters [29], I first asked if high TE density is restricted to *Hox* clusters, or if it is a general characteristic of the *Anolis* genome. I confined my analysis to the *HoxB* and *–C* clusters as the current release of the *A. carolinensis* genome does not provide chromosome-wide information for the *HoxA* and *–D* clusters

[28]. Using a 500 kb sliding window approach, I found that the absolute numbers of TEs are significantly higher in the 500 kb windows containing *HoxB* and *–C* clusters compared to the remainder of the chromosomes (figure 2; except for centromeric and telomeric regions that show a characteristic TE landscape [20,66]). In addition, the composition of major classes of TEs within *Hox* clusters is different compared to the genomic vicinity of the *HoxB* and *–C* clusters as well as in a genome-wide comparison, with an overrepresentation of simple repeats and DNA transposons, and an underrepresentation of LTRs, SINEs and LINEs (electronic supplementary material, figure S3).

On average, the TE density of the six *Hox* intergenic regions, measured in numbers of TEs per 1 kb, is 23 times higher in *Anolis* lizards compared to other lepidosaurians (figure 3a; t-test: p-value < 0.001). For example, no TEs were found in the basilisk *Basiliscus plumifrons*, a very close relative of *Anolis* ([67,68]; electronic supplementary material, figure S4). Within *Anolis*, fragment length increases with the number of TEs, showing that TE insertion is responsible for the evolutionary expansion of *Hox* clusters (figure 3b, electronic supplementary material, figure S5). In contrast, non-*Anolis* lepidosauria show no correlation between fragment length and TE number (figure 3b, electronic supplementary material, figure S5).

(b) *Anolis* TE densities are positively correlated with speciation rates

A reconstruction of the phylogenetic timing of TE insertions revealed that the accumulation of TEs occurred in parallel with the *Anolis* radiation (electronic supplementary material, figure S6) allowing us to test if lineages that have undergone more speciation events have a higher density of TEs, as predicted if TE activity causes genomic incompatibilities [9,10]. Controlling for island area, TE density was higher in extant lineages with a high propensity to speciate (table 1 and figure 4). To rule out the possibility that TEs accumulate in the *Hox* clusters as a result of neutral population genetic processes, I tested the same statistical model with the number of microsatellite repeats instead of TE density as predictor of speciation nodes. A total of nine microsatellites located in the posterior *HoxA* and *–D* clusters were analysed (electronic supplementary material, figure S4), with no significant correlation between repeats and speciation nodes (electronic supplementary material, table S4).

219 The unusually high number of TEs in the *Anolis Hox* clusters compared to other vertebrates has
220 also been discussed as a driver of morphological diversification [69]. A functional impact of TEs in the
221 *Hox* clusters should be associated with transcriptional activity of at least some TEs. Mapping of organ-
222 specific transcriptomes to the *Hox* intergenic regions analysed in this study revealed that most of the
223 annotated TEs (81%) belong to families that are transcriptionally active (electronic supplementary
224 material, figure S1 and table S5). The most conspicuous difference between *Anolis* ecomorphs is their
225 limb length relative to body size, which has evolved repeatedly on different islands. Given that the
226 effects of expansion of intergenic regions on *Hox* gene regulation may very well be conserved, we might
227 predict that similarity in morphology would be accompanied by similarity in the density of TEs in *Hox*
228 clusters. However, this was not the case (table 1). Furthermore, using phylogenetically independent
229 contrasts, there was no evidence that differences in relative limb length between closely related species
230 was associated with a historical proliferation of TEs (electronic supplementary material, figure S2).
231

232 4. Discussion

233 Lineages with a high abundance of TEs may be more likely to undergo adaptive radiation if TE activity
234 promotes genomic incompatibility between incipient species and phenotypic innovation. The results
235 here provide support for a role of TEs in the adaptive radiation of *Anolis* lizards by demonstrating that
236 *Hox* gene clusters are remarkably rich in active TEs and that the density of TEs is positively associated
237 with rates of speciation within this genus.
238

239 Transposable elements and the radiation of *Anolis* lizards

240 Previous data showed that the *Hox* clusters of the green anole, *A. carolinensis*, are richer in TEs than
241 *Hox* clusters of model species of mammals, birds, and fish [29]. The comparative analysis of TE
242 densities of *Hox* intergenic regions in this study extends these data to 20 *Anolis* species and 10 other
243 lepidosaurians. This reveals that *Anolis* is highly unusual also within lizards and that there was a rapid
244 increase in TE numbers in the lineage leading to extant *Anolis* rather than a gradual taxonomic trend
245 within lepidosaurians.

Importantly, the comparative analysis showed that TEs continued to accumulate during or following speciation events within *Anolis*, and that at least some TE families whose members are located within *Anolis Hox* clusters remain transcriptionally active. This positive relationship between TE density and the number of speciation nodes leading up to extant species is consistent with the hypothesis that TE activity causes genomic changes that result in reproductive incompatibilities and hence reproductive isolation [4,7,9,70]. This estimate of diversification is potentially confounded by lineage differences in the rate of extinction. However, if lineage-specific rates of extinction would explain these results, lineages with high rates of extinction must have fewer TEs. This would be highly unexpected since bursts of TE activity most commonly cause genomic instabilities [71] and hence loss of fitness [72]. Thus, although more data within and across species would be necessary to rule out extinction as a contributor to the observed patterns, the results are consistent with the main prediction of the hypothesis that TE activity promotes speciation [9,10].

Rapid reproductive isolation through re-organisation of genome structure may be particularly likely to facilitate adaptive radiation when ecological conditions limit the ability for reproductive isolation and phenotypic divergence in allopatry. *Anolis* speciation is associated with radiation into ecomorphs, and extant species on the Greater Antillean islands are indeed mostly found in sympatry since the ecomorph habitats are highly connected [24]. This may point towards a scenario where the processes of speciation and adaptation are intertwined [73]. TEs may facilitate such adaptive ecological speciation because their activity can produce both phenotypic variation and structural changes in the genome that cause reproductive isolation through pre- (e.g., sexual signals) or post-zygotic (e.g., genomic incompatibility) mechanisms. A similar role for TEs in speciation has also been suggested in cichlid fish [13-15] and *Heliconius* butterflies [11,12], but to my knowledge this is the first evidence that the number of speciation events is positively correlated with an estimate of TE abundance.

An alternative explanation for a positive relationship between TE density and the number of speciation events is that lineages with high rates of speciation exhibit population dynamics that promote the accumulation of neutral or even slightly deleterious TEs. The principles of TE accumulation are similar to those of other sequence-based variants that may accumulate under low effective population sizes and weak natural selection [74,75]. Although neutral processes cannot be completely ruled out as

causes of the positive relationship between speciation rate and TE density in *Hox* clusters, the fact that microsatellite repeats located in the *Hox* clusters did not accumulate with the number of speciation nodes, and showed the expected pattern from neutral evolution, argues against this. In addition, low effective population sizes are unlikely to apply to Caribbean *Anolis* lizards as they are characterized by generally high abundances [24] and estimates of historical effective population sizes suggest this has been the case throughout the lineage's history [76]. Environmental factors such as temperature or humidity can potentially promote TE activity as demonstrated in model organisms exposed to biotic and abiotic stressors [77-79]. However, this cannot solely explain the patterns of TE abundance within *Anolis* since it is unlikely that lineages with high rates of speciation also systematically experience more stressful environmental conditions.

Transposable elements and *Anolis Hox* clusters

Although TE density across the genome in the green anole is high relative to many other vertebrates [27], *Hox* clusters appear to be particularly TE rich. This might be explained by biased TE insertion due to chromatin characteristics (facultative heterochromatin; [80]). *Hox* genes are expressed during early stages of development, and hence, the open chromatin configuration can potentially attract TEs, possibly even in germ cells [81,82]. Nevertheless, a high TE density in *Hox* clusters is highly unexpected given that the condensed, TE poor, *Hox* clusters observed in other vertebrates have been hypothesized to be crucial for the temporal and spatial collinearity that specifies positional identities along the embryonic head-to-tail axis [83,84]. The structural organization of *Hox* clusters should be under strong purifying selection [85]. That the TE composition of *Hox* clusters differs from the genomic vicinity, and genome-wide, suggests that TE insertions are non-random or that selection has contributed to shape the TE landscape of *Hox* clusters. Accordingly, the abundant occurrence of TEs in the *Anolis Hox* clusters could indicate that they contribute functionally to development, for example by modification of histones, insertion of novel regulatory elements, or positional effects [20,86]. Interestingly, the *Anolis Hox* clusters are thus reminiscent of invertebrate *Hox* clusters which appear to be generally less constrained and accumulate more TEs than their jawed vertebrate counterparts [87,88].

A particular interesting hypothesis is that TE-induced changes in the regulation of *Hox* genes have contributed to the morphological variation that characterizes the adaptive radiation of *Anolis* [29]. Indeed, the *HoxA* and -*D* clusters are involved in regulating limb growth [89], a key feature of *Anolis* ecomorphs [24]. This may suggest that species convergent in limb length would also show convergence in intergenic accumulation of TEs. There was, however, no evidence for this, nor for the prediction that proliferation of TEs should be associated with the rate of change in limb morphology, although statistical power to detect these relationships was low. These results do not, of course, rule out that specific TEs could be responsible for changes in gene expression underlying the different ecomorphs. However, rigorous tests of these hypotheses would require knowledge about the developmental genetic regulation of morphological differences between ecomorphs that is currently lacking.

In summary, this study supports the prediction that TE proliferation is positively associated with speciation in the adaptive radiation of *Anolis* lizards. Whether or not TEs also have contributed to adaptive divergence between species remains unknown, but the accumulation and transcriptional activity of TEs in genes that control morphological development suggests that this hypothesis warrants further study.

Acknowledgments

I thank Jason Kolbe, Roger Thorpe and Tobias Uller for providing tissue samples of *Anolis* lizards, Juliane Cohen for technical assistance, Tobias Uller, Alex Pigot, Adam Hargreaves, Ferdinand Marletaz, Jordi Paps and Peter Holland for discussions and Tobias Uller, Charlie Cornwallis and Peter Holland for comments on the manuscript. I am grateful to Peter Holland for hosting me in his lab, where most of this work was conducted. I thank the High-Throughput Genomics Group at the Wellcome Trust Centre for Human Genetics for the generation of sequencing data.

Funding statement

This research was supported by the John Fell fund, the Humboldt Foundation and the Wenner-Gren Foundation.

References

1. McClintock B. 1951 Chromosome organization and genic expression. *Cold Spring Harbor symposia on quantitative biology* **16**, 13-47.
2. McClintock B. 1984 The significance of responses of the genome to challenge. *Science* **226**(4676), 792-801.
3. Fedoroff NV. 2012 Presidential address. Transposable elements, epigenetics, and genome evolution. *Science* **338**(6108), 758-767.
4. Flavell R. 1982 Sequence amplification, deletion and rearrangement: major sources of variation during species divergence. In *Genome Evolution* (eds. Dover G., Flavell R.), pp. 301-321. London, Academic Press.
5. Georgiev GP, Kramerov DA, Ryskov AP, Skryabin KG, Lukanidin EM. 1983 Dispersed repetitive sequences in eukaryotic genomes and their possible biological significance. *Cold Spring Harbor symposia on quantitative biology* **47 Pt 2**, 1109-1121.
6. McDonald JF. 1983 The molecular basis of adaptation: a critical review of relevant ideas and observations. *Annu Rev Ecol Syst* **14**, 77-102.
7. Sylvanen M. 1984 The evolutionary implications of mobile genetic elements. *Annu Rev Genet* **18**, 271–293.
8. Oliver KR, Greene WK. 2012 Transposable elements and viruses as factors in adaptation and evolution: an expansion and strengthening of the TE-Thrust hypothesis. *Ecology and evolution* **2**(11), 2912-2933. (doi:10.1002/ece3.400).
9. Zeh DW, Zeh JA, Ishida Y. 2009 Transposable elements and an epigenetic basis for punctuated equilibria. *BioEssays : news and reviews in molecular, cellular and developmental biology* **31**(7), 715-726. (doi:10.1002/bies.200900026).
10. Rebollo R, Horard B, Hubert B, Vieira C. 2010 Jumping genes and epigenetics: Towards new species. *Gene* **454**(1-2), 1-7. (doi:10.1016/j.gene.2010.01.003).

11. Lavoie CA, Platt RN, 2nd, Novick PA, Counterman BA, Ray DA. 2013 Transposable element evolution in Heliconius suggests genome diversity within Lepidoptera. *Mobile DNA* **4**(1), 21. (doi:10.1186/1759-8753-4-21).

12. Consortium THG. 2012 Butterfly genome reveals promiscuous exchange of mimicry adaptations among species. *Nature* **487**(7405), 94-98. (doi:http://www.nature.com/nature/journal/v487/n7405/abs/nature11041.html#supplementary-information).

13. Brawand D, Wagner CE, Li YI, Malinsky M, Keller I, Fan S, Simakov O, Ng AY, Lim ZW, Bezault E, et al. 2014 The genomic substrate for adaptive radiation in African cichlid fish. *Nature* **513**(7518), 375-381. (doi:10.1038/nature13726).

14. Fan S, Meyer A. 2014 Evolution of genomic structural variation and genomic architecture in the adaptive radiations of African cichlid fishes. *Frontiers in genetics* **5**, 163. (doi:10.3389/fgene.2014.00163).

15. Berner D, Salzburger W. 2015 The genomics of organismal diversification illuminated by adaptive radiations. *Trends in genetics : TIG* **31**(9), 491-499. (doi:10.1016/j.tig.2015.07.002).

16. Oliver KR, McComb JA, Greene WK. 2013 Transposable elements: powerful contributors to angiosperm evolution and diversity. *Genome biology and evolution* **5**(10), 1886-1901. (doi:10.1093/gbe/evt141).

17. Feschotte C. 2008 Transposable elements and the evolution of regulatory networks. *Nature reviews Genetics* **9**(5), 397-405. (doi:10.1038/nrg2337).

18. Sundaram V, Cheng Y, Ma Z, Li D, Xing X, Edge P, Snyder MP, Wang T. 2014 Widespread contribution of transposable elements to the innovation of gene regulatory networks. *Genome research* **24**(12), 1963-1976. (doi:10.1101/gr.168872.113).

19. Cowley M, Oakey RJ. 2013 Transposable elements re-wire and fine-tune the transcriptome. *Plos Genet* **9**(1), e1003234. (doi:10.1371/journal.pgen.1003234).

379 20. Slotkin RK, Martienssen R. 2007 Transposable elements and the epigenetic regulation of the
380 genome. *Nature reviews Genetics* **8**(4), 272-285. (doi:10.1038/nrg2072).
381 21. Lynch VJ, Nnamani MC, Kapusta A, Brayer K, Plaza SL, Mazur EC, Emera D, Sheikh SZ, Grutzner
382 F, Bauersachs S, et al. 2015 Ancient transposable elements transformed the uterine regulatory
383 landscape and transcriptome during the evolution of mammalian pregnancy. *Cell reports* **10**(4),
384 551-561. (doi:10.1016/j.celrep.2014.12.052).
385 22. Emera D, Wagner GP. 2012 Transposable element recruitments in the mammalian placenta:
386 impacts and mechanisms. *Briefings in functional genomics* **11**(4), 267-276.
387 (doi:10.1093/bfgp/els013).
388 23. Santos ME, Braasch I, Boileau N, Meyer BS, Sauteur L, Bohne A, Belting HG, Affolter M,
389 Salzburger W. 2014 The evolution of cichlid fish egg-spots is linked with a cis-regulatory change.
390 *Nature communications* **5**, 5149. (doi:10.1038/ncomms6149).
391 24. Losos JB. 2009 *Lizards in an Evolutionary Tree: Ecology and Adaptive Radiation of Anoles*.
392 Berkeley, CA, University of California Press; 1-507 p.
393 25. Williams E. 1972 The Origin of Faunas. Evolution of Lizard Congeners in a Complex Island
394 Fauna: A Trial Analysis. *Evolutionary Biology* (6), 47-89. (doi:10.1007/978-1-4684-9063-3_3).
395 26. Wray GA. 2007 The evolutionary significance of cis-regulatory mutations. *Nature reviews*
396 *Genetics* **8**(3), 206-216. (doi:10.1038/nrg2063).
397 27. Tollis M, Boissinot S. 2011 The transposable element profile of the anolis genome: How a lizard
398 can provide insights into the evolution of vertebrate genome size and structure. *Mobile genetic*
399 *elements* **1**(2), 107-111. (doi:10.4161/mge.1.2.17733).
400 28. Alfoldi J, Di Palma F, Grabherr M, Williams C, Kong LS, Mauceli E, Russell P, Lowe CB, Glor RE,
401 Jaffe JD, et al. 2011 The genome of the green anole lizard and a comparative analysis with birds
402 and mammals. *Nature* **477**(7366), 587-591. (doi:Doi 10.1038/Nature10390).

29. Di-Poï N, Montoya-Burgos JI, Duboule D. 2009 Atypical relaxation of structural constraints in Hox gene clusters of the green anole lizard. *Genome research* **19**(4), 602-610. (doi:10.1101/gr.087932.108).

30. Neufeld SJ, Wang F, Cobb J. 2014 Genetic interactions between Shox2 and Hox genes during the regional growth and development of the mouse limb. *Genetics* **198**(3), 1117-1126. (doi:10.1534/genetics.114.167460).

31. Gonzalez-Martin MC, Mallo M, Ros MA. 2014 Long bone development requires a threshold of Hox function. *Developmental biology* **392**(2), 454-465. (doi:10.1016/j.ydbio.2014.06.004).

32. Pineault KM, Swinehart IT, Garthus KN, Ho E, Yao Q, Schipani E, Kozloff KM, Wellik DM. 2015 Hox11 genes regulate postnatal longitudinal bone growth and growth plate proliferation. *Biology open*. (doi:10.1242/bio.012500).

33. Andrey G, Montavon T, Mascrez B, Gonzalez F, Noordermeer D, Leleu M, Trono D, Spitz F, Duboule D. 2013 A switch between topological domains underlies HoxD genes collinearity in mouse limbs. *Science* **340**(6137), 1234167. (doi:10.1126/science.1234167).

34. Morgan BA, Tabin CJ. 1993 The role of Hox genes in limb development. *Progress in clinical and biological research* **383A**, 1-9.

35. Bolger AM, Lohse M, Usadel B. 2014 Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**(15), 2114-2120. (doi:10.1093/bioinformatics/btu170).

36. Magoc T, Salzberg SL. 2011 FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* **27**(21), 2957-2963. (doi:10.1093/bioinformatics/btr507).

37. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, et al. 2012 SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *Journal of computational biology : a journal of computational molecular cell biology* **19**(5), 455-477. (doi:10.1089/cmb.2012.0021).

38. Castoe TA, de Koning AP, Hall KT, Card DC, Schield DR, Fujita MK, Ruggiero RP, Degner JF, Daza JM, Gu W, et al. 2013 The Burmese python genome reveals the molecular basis for extreme

adaptation in snakes. *Proceedings of the National Academy of Sciences of the United States of America* **110**(51), 20645-20650. (doi:10.1073/pnas.1314475110).

39. Vonk FJ, Casewell NR, Henkel CV, Heimberg AM, Jansen HJ, McCleary RJ, Kerkkamp HM, Vos RA, Guerreiro I, Calvete JJ, et al. 2013 The king cobra genome reveals dynamic gene evolution and adaptation in the snake venom system. *Proceedings of the National Academy of Sciences of the United States of America* **110**(51), 20651-20656. (doi:10.1073/pnas.1314702110).

40. Georges A, Li Q, Lian J, O'Meally D, Deakin J, Wang Z, Zhang P, Fujita M, Patel HR, Holleley CE, et al. 2015 High-coverage sequencing and annotated assembly of the genome of the Australian dragon lizard Pogona vitticeps. *GigaScience* **4**, 45. (doi:10.1186/s13742-015-0085-2).

41. Brudno M, Do CB, Cooper GM, Kim MF, Davydov E, Program NCS, Green ED, Sidow A, Batzoglou S. 2003 LAGAN and Multi-LAGAN: efficient tools for large-scale multiple alignment of genomic DNA. *Genome research* **13**(4), 721-731. (doi:10.1101/gr.926603).

42. Kohany O, Gentles AJ, Hankus L, Jurka J. 2006 Annotation, submission and screening of repetitive elements in Repbase: RepbaseSubmitter and Censor. *BMC bioinformatics* **7**, 474. (doi:10.1186/1471-2105-7-474).

43. Smit AFA, Hubley R, Green P. 2013 Repeatmasker Open-4.0. http://wwwrepeatmaskerorg.

44. Jurka J. 2000 Repbase update: a database and an electronic journal of repetitive elements. *Trends in genetics : TIG* **16**(9), 418-420.

45. Wexler Y, Yakhini Z, Kashi Y, Geiger D. 2005 Finding approximate tandem repeats in genomic sequences. *Journal of computational biology : a journal of computational molecular cell biology* **12**(7), 928-942. (doi:10.1089/cmb.2005.12.928).

46. Langmead B, Salzberg SL. 2012 Fast gapped-read alignment with Bowtie 2. *Nature methods* **9**(4), 357-359. (doi:10.1038/nmeth.1923).

47. Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP. 2011 Integrative genomics viewer. *Nature biotechnology* **29**(1), 24-26. (doi:10.1038/nbt.1754).

48. Freckleton RP, Phillimore AB, Pagel M. 2008 Relating traits to diversification: a simple test. *The American naturalist* **172**(1), 102-115. (doi:10.1086/588076).

49. FitzJohn RG. 2010 Quantitative traits and diversification. *Systematic biology* **59**(6), 619-633. (doi:10.1093/sysbio/syq053).

50. Cooper N, Thomas GH, FitzJohn RG. 2016 Shedding light on the 'dark side' of phylogenetic comparative methods. *Methods in Ecology and Evolution* **7**(6), 693--699. (doi:10.1111/2041-210X.12533).

51. Rabosky DL, Goldberg EE. 2015 Model inadequacy and mistaken inferences of trait-dependent speciation. *Systematic biology* **64**(2), 340-355. (doi:10.1093/sysbio/syu131).

52. Mahler DL, Revell LJ, Glor RE, Losos JB. 2010 Ecological opportunity and the rate of morphological evolution in the diversification of Greater Antillean anoles. *Evolution* **64**(9), 2731-2745. (doi:10.1111/j.1558-5646.2010.01026.x).

53. Nicholson KE, Crother BI, Guyer C, Savage JM. 2012 It is time for a new classification of anoles (Squamata: Dactyloidae). *Zootaxa* **3477**, 1-108.

54. Algar AC, Mahler DL. 2015 Area, climate heterogeneity, and the response of climate niches to ecological opportunity in island radiations of Anolis lizards. *Global Ecology and Biogeography*, n/a-n/a. (doi:10.1111/geb.12327).

55. Gamble T, Geneva AJ, Glor RE, Zarkower D. 2014 Anolis sex chromosomes are derived from a single ancestral pair. *Evolution* **68**(4), 1027-1041. (doi:10.1111/evo.12328).

56. Lazell JD. 1972 The Anoles (Sauria, Iguanidae) of the Lesser Antilles. *Bulletin of The Museum of Comparative Zoology* **143**.

57. Rabosky DL. 2010 Extinction rates should not be estimated from molecular phylogenies. *Evolution* **64**(6), 1816-1824. (doi:10.1111/j.1558-5646.2009.00926.x).

58. Quental TB, Marshall CR. 2010 Diversity dynamics: molecular phylogenies need the fossil record. *Trends Ecol Evol* **25**(8), 434-441. (doi:10.1016/j.tree.2010.05.002).

59. Pyron RA, Burbrink FT. 2013 Phylogenetic estimates of speciation and extinction rates for testing ecological and evolutionary hypotheses. *Trends in ecology & evolution* **28**(12), 729-736. (doi:10.1016/j.tree.2013.09.007).

60. Rabosky DL, Glor RE. 2010 Equilibrium speciation dynamics in a model adaptive radiation of island lizards. *Proceedings of the National Academy of Sciences of the United States of America* **107**(51), 22178-22183. (doi:10.1073/pnas.1007606107).

61. Martins E, Hansen T. 1997 Phylogenies and the Comparative Method: A General Approach to Incorporating Phylogenetic Information into the Analysis of Interspecific Data. *The American naturalist* **149**(4), 646-667. (doi:citeulike-article-id:701018).

62. Pagel M. 1997 Inferring evolutionary processes from phylogenies. *Zoologica Scripta* **26**(4), 331-348. (doi:citeulike-article-id:4249120
doi: 10.1111/j.1463-6409.1997.tb00423.x).

63. Freckleton RP, Harvey PH, Pagel M. 2002 Phylogenetic analysis and comparative data: a test and review of evidence. *The American naturalist* **160**(6), 712-726. (doi:10.1086/343873).

64. Poe S. 2004 Phylogeny of Anoles. *Herpetological Monographs* **18**, 37-89.

65. Felsenstein J. 1985 Phylogenies and the Comparative Method. *The American naturalist* **125**(1), 1-15. (doi:doi:10.1086/284325).

66. DeBaryshe PG, Pardue ML. 2011 Differential maintenance of DNA sequences in telomeric and centromeric heterochromatin. *Genetics* **187**(1), 51-60. (doi:10.1534/genetics.110.122994).

67. Streicher JW, Schulte JA, 2nd, Wiens JJ. 2016 How Should Genes and Taxa be Sampled for Phylogenomic Analyses with Missing Data? An Empirical Study in Iguanian Lizards. *Systematic biology* **65**(1), 128-145. (doi:10.1093/sysbio/syv058).

68. Pyron RA, Burbrink FT, Wiens JJ. 2013 A phylogeny and revised classification of Squamata, including 4161 species of lizards and snakes. *BMC evolutionary biology* **13**, 93. (doi:10.1186/1471-2148-13-93).

69. Di-Poï N, Montoya-Burgos JI, Miller H, Pourquie O, Milinkovitch MC, Duboule D. 2010 Changes in Hox genes' structure and function during the evolution of the squamate body plan. *Nature* **464**(7285), 99-103. (doi:10.1038/nature08789).

70. Belyayev A. 2014 Bursts of transposable elements as an evolutionary driving force. *Journal of evolutionary biology* **27**(12), 2573-2584. (doi:10.1111/jeb.12513).

71. Ayarpadikannan S, Kim HS. 2014 The impact of transposable elements in genome evolution and genetic instability and their implications in various diseases. *Genomics & informatics* **12**(3), 98-104. (doi:10.5808/GI.2014.12.3.98).

72. Pasyukova EG, Nuzhdin SV, Morozova TV, Mackay TF. 2004 Accumulation of transposable elements in the genome of Drosophila melanogaster is associated with a decrease in fitness. *The Journal of heredity* **95**(4), 284-290. (doi:10.1093/jhered/esh050).

73. Losos JB. 2004 Adaptation and speciation in Greater Antillean anoles. In *Adaptive Speciation* (eds. Dieckmann U., Doebeli M., Metz J.A.J., Tautz D.). Cambridge, UK, Cambridge University Press.

74. Charlesworth B. 1985 The population genetics of transposable elements. In *Population Genetics and Molecular Evolution* (eds. Ohta T., Aoki K.), pp. 213-232. New York, Springer-Verlag.

75. Lynch M. 2007 *The origins of genome architecture*. Sunderland, USA, Sinauer Associates.

76. Munoz MM, Crawford NG, McGreevy TJ, Jr., Messana NJ, Tarvin RD, Revell LJ, Zandvliet RM, Hopwood JM, Mock E, Schneider AL, et al. 2013 Divergence in coloration and ecological speciation in the Anolis marmoratus species complex. *Molecular ecology* **22**(10), 2668-2682. (doi:10.1111/mec.12295).

77. Mourier T, Nielsen LP, Hansen AJ, Willerslev E. 2014 Transposable elements in cancer as a by-product of stress-induced evolvability. *Frontiers in genetics* **5**, 156. (doi:10.3389/fgene.2014.00156).

78. Capy P, Gasperi G, Biemont C, Bazin C. 2000 Stress and transposable elements: co-evolution or useful parasites? *Heredity* **85 (Pt 2)**, 101-106.

79. Guerreiro MPG. 2012 What makes transposable elements move in the Drosophila genome? *Heredity* **108**(5), 461-468. (doi:10.1038/hdy.2011.89).

80. Trojer P, Reinberg D. 2007 Facultative heterochromatin: is there a distinctive molecular signature? *Molecular cell* **28**(1), 1-13. (doi:10.1016/j.molcel.2007.09.011).

81. Zhang Y, Mager DL. 2012 Gene properties and chromatin state influence the accumulation of transposable elements in genes. *PloS one* **7**(1), e30158. (doi:10.1371/journal.pone.0030158).

82. Fontanillas P, Hartl DL, Reuter M. 2007 Genome organization and gene expression shape the transposable element distribution in the Drosophila melanogaster euchromatin. *Plos Genet* **3**(11), e210. (doi:10.1371/journal.pgen.0030210).

83. Duboule D, Dolle P. 1989 The structural and functional organization of the murine HOX gene family resembles that of Drosophila homeotic genes. *The EMBO journal* **8**(5), 1497-1505.

84. Graham A, Papalopulu N, Krumlauf R. 1989 The murine and Drosophila homeobox gene complexes have common features of organization and expression. *Cell* **57**(3), 367-378.

85. Almirantis Y, Provata A, Papageorgiou S. 2013 Evolutionary constraints favor a biophysical model explaining hox gene collinearity. *Current genomics* **14**(4), 279-288. (doi:10.2174/13892029113149990003).

86. Cordaux R, Batzer MA. 2009 The impact of retrotransposons on human genome evolution. *Nature reviews Genetics* **10**(10), 691-703. (doi:10.1038/nrg2640).

87. Fried C, Prohaska SJ, Stadler PF. 2004 Exclusion of repetitive DNA elements from gnathostome Hox clusters. *Journal of experimental zoology Part B, Molecular and developmental evolution* **302**(2), 165-173. (doi:10.1002/jez.b.20007).

88. Wagner GP, Amemiya C, Ruddle F. 2003 Hox cluster duplications and the opportunity for evolutionary novelties. *Proceedings of the National Academy of Sciences of the United States of America* **100**(25), 14603-14606. (doi:10.1073/pnas.2536656100).

555 89. Fromental-Ramain C, Warot X, Messadecq N, LeMeur M, Dolle P, Chambon P. 1996 Hoxa-13
556 and Hoxd-13 play a crucial role in the patterning of the limb autopod. *Development* **122**(10),
557 2997-3011.
558
559

560 # Table

561 **Table 1.** Regression coefficients with standard errors (β ± SE) and statistics (t-value) and their
562 associated p-values for the complete multiple regression model with the length of the long bones and
563 speciation nodes as response variables.

	Trait	β ± SE	t-value	p-value
Humerus*				
	TE density	-0.023 ± 0.119	-0.193	0.850
	svl*	**1.010 ± 0.065**	**15.449**	**> 0.001**
Femur*				
	TE density	-0.099 ± 0.138	-0.719	0.485
	svl*	**0.991 ± 0.098**	**10.098**	**> 0.001**
Radius*				
	TE density	-0.019 ± 0.180	-0.108	0.916
	svl*	**1.014 ± 0.099**	**10.260**	**> 0.001**
Tibia*				
	TE density	-0.082 ± 0.180	-0.454	0.657
	svl*	**0.981 ± 0.099**	**9.889**	**> 0.001**
Speciation nodes*				
	TE density	**0.684 ± 0.223**	**3.065**	**0.007**
	island area*	**-0.054 ± 0.023**	**-2.351**	**0.031**

Variables presenting a significant partial regression coefficient are highlighted in bold. *log transformed; Abbreviation: svl, snout-vent length; full statistical model used in pgls: length of long bone ~ TE density + svl and speciation nodes ~ TE density + island area, resp., including a variance-covariance matrix accounting for phylogenetic relatedness.

564

Figure Legends

Figure 1. Taxonomic overview of the dataset and *Hox* cluster structure of the green anole. (*a*) Schematic tree of the species analysed in this study. Phylogenetic relationships are derived from previous studies [55,68]. Animal pictures are licensed from URL: https://stock.adobe.com. (*b*) The location of *Hox* genes along the four *Hox* clusters of the green anole. The size of coding regions and the intergenic distance between *Hox* genes are drawn to scale. (*c*) Intergenic regions targeted in this study. The actual lengths of the fragments amplified by long-range PCRs are given in kilo bases. Note the position of the conserved non-coding element (CNE) between *HoxA13* and *–A11* used as primer binding site.

Figure 2. Comparison of *Hox* clusters versus chromosome-wide TE content in the green anole. (*a* and *b*) The number of TEs per 500 kb is plotted along the genomic regions containing the *HoxC* and *–B* cluster (chromosome 2 and 6, resp.). As expected, TE densities are high in the telomeric regions and absent from the centromeric regions (highlighted with grey bars). (*c* and *d*) Frequency plot of TE densities of all 500 kb genomic windows. The genomic windows of the *HoxB* and *–C* clusters are shown in red and are significantly higher in TE densities than expected under a normal distribution. Dashed lines show the 95% confidence intervals of the distribution.

Figure 3. Relationship between intergenic length and TE content. (*a*) Boxplot showing that species of the genus *Anolis* show higher TE density as other non-*Anolis* lepidosaurians. The TE density, measured as the number of TEs per 1 kb, differs significantly between *Anolis* and non-*Anolis* species (t-test p-value < 0.01). (*b*) A linear phylogenetic model established a significant relationship between number of TEs and intergenic length for *Anolis* (slope, 0.23; p-value < 0.001; 95% confidence interval shaded in grey), but not for other lepidosaurians (slope, -1.15; p-value 0.6). Plotted are the mean values per species of the normalized number of TEs and the intergenic length (maximum value per fragment was

591 scaled to 1). For a plot of individual intergenic fragments and number of TEs per species, see electronic
592 supplementary material, figure S5.
593
594 **Figure 4.** Relationship of diversification rates and density of TEs in *Anolis* lizards. (*a*) Distribution of the
595 20 *Anolis* species analysed in this study, highlighted in red (phylogeny from Gamble *et al.* [55]). For an
596 enlarged version of this phylogeny, see electronic supplementary material, figure S7. (*b*) Correlation
597 between the number of nodes, as a measure of diversification rate, and the density of TEs (slope,
598 0.684; p-value, 0.007; 95% confidence interval shaded in grey). The full statistical model includes
599 phylogenetic relationships among species and the geographic area of the islands where the lizards
600 originated (see table 1 and Results). The maximum likelihood estimation of phylogenetic signal, the λ
601 parameter, was 0.53; however the estimate was not significantly different from zero (p-value, 0.81) or
602 one (p-value, 0.46).
603

Figures

Figure 1

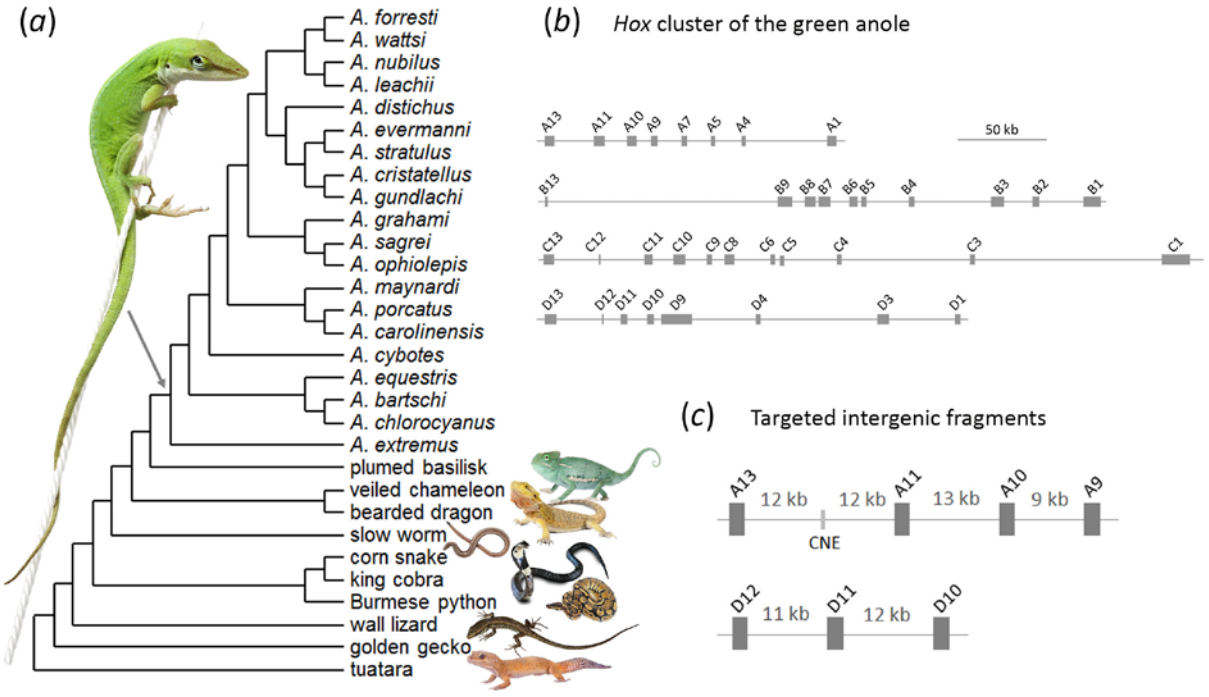
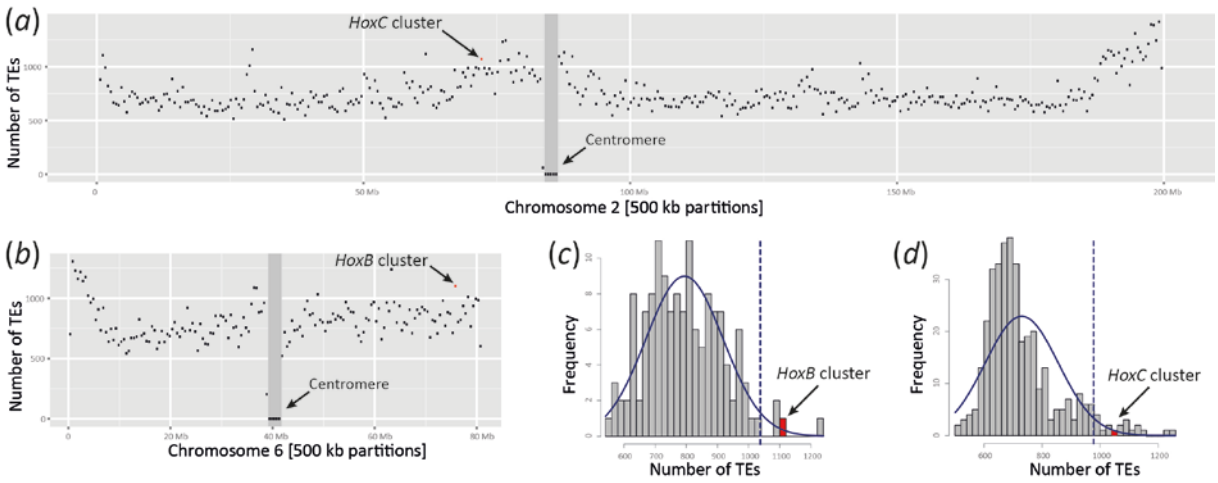
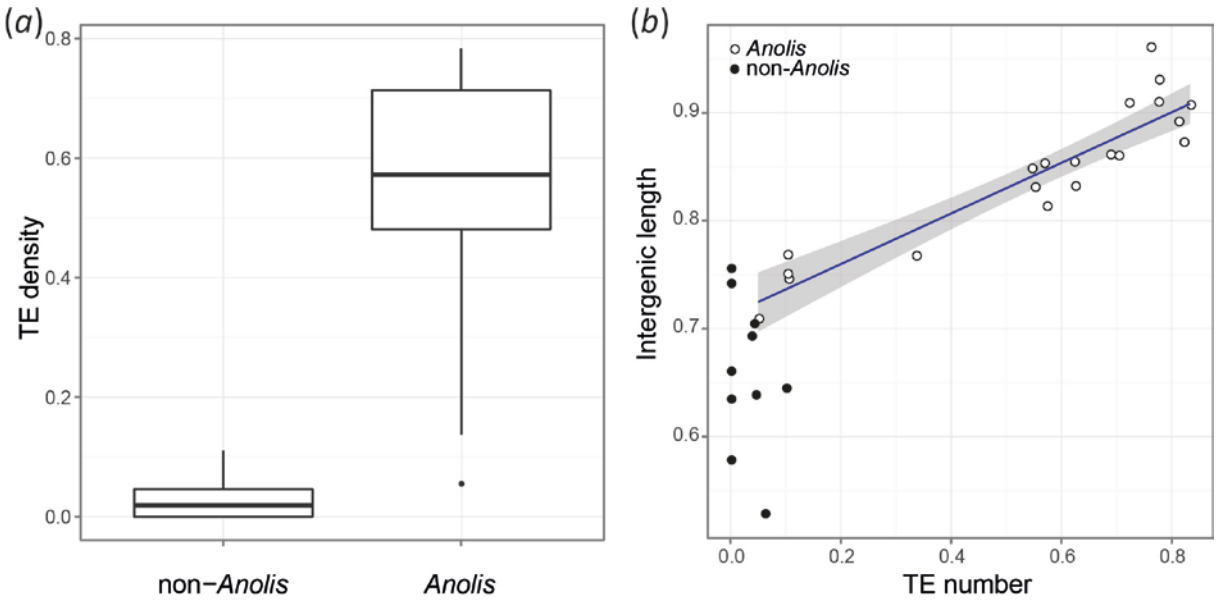


Figure 2



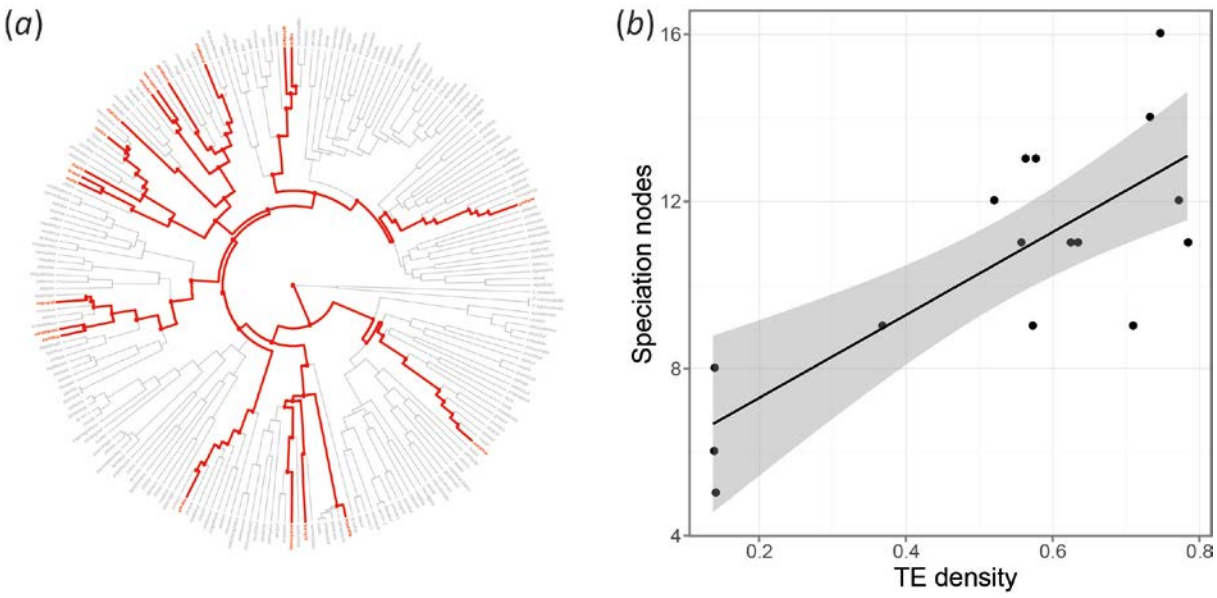
613 Figure 3



614

615

616 Figure 4



617