City University of New York (CUNY)
CUNY Academic Works

Publications and Research | Queens College

2011

The genome of the green anole lizard and a comparative analysis with birds and mammals

Jessica Alföldi
Broad Institute of MIT and Harvard

Federica Di Palma
Broad Institute of MIT and Harvard

Manfred Grabherr
Broad Institute of MIT and Harvard

Christina Williams
North Carolina State University

Lesheng Kong
University of Oxford

See next page for additional authors

How does access to this work benefit you? Let us know!

More information about this work at: https://academicworks.cuny.edu/qc_pubs/88
Discover additional works at: https://academicworks.cuny.edu

This work is made publicly available by the City University of New York (CUNY).
Contact: AcademicWorks@cuny.edu

Authors

Jessica Alföldi, Federica Di Palma, Manfred Grabherr, Christina Williams, Lesheng Kong, Evan Mauceli, Pamela Russell, Craig B. Lowe, Richard Glor, Jacob D. Jaffe, David A. Ray, Stéphane Boissinot, Andrew M. Shedlock, Christopher Botka, Todd A. Castoe, John K. Colbourne, Matthew K. Fujita, Ricardo Godinez Moreno, Boudewijn F. ten Hallers, David Haussler, and Peter Novick

This article is available at CUNY Academic Works: https://academicworks.cuny.edu/qc_pubs/88

NIH Public Access
Author Manuscript
Nature. Author manuscript; available in PMC 2012 March 29.

Published in final edited form as:
Nature. ; 477(7366): 587–591. doi:10.1038/nature10390.

The genome of the green anole lizard and a comparative analysis with birds and mammals

Jessica Alföldi[1,*], **Federica Di Palma**[1,*], **Manfred Grabherr**[1], **Christina Williams**[2], **Lesheng Kong**[3], **Evan Mauceli**[1], **Pamela Russell**[1], **Craig B. Lowe**[4], **Richard Glor**[5], **Jacob D. Jaffe**[1], **David A. Ray**[6], **Stephane Boissinot**[7], **Andrew M. Shedlock**[8], **Christopher Botka**[9], **Todd A. Castoe**[10], **John K. Colbourne**[11], **Matthew K. Fujita**[3,12], **Ricardo Godinez Moreno**[12], **Boudewijn F. ten Hallers**[13], **David Haussler**[14], **Andreas Heger**[3], **David Heiman**[1], **Daniel E. Janes**[12], **Jeremy Johnson**[1], **Pieter J. de Jong**[13], **Maxim Y. Koriabine**[13], **Peter Novick**[7], **Chris L. Organ**[12], **Sally E. Peach**[1], **Steven Poe**[15], **David D. Pollock**[10,16], **Kevin de Queiroz**[17], **Thomas Sanger**[12], **Steve Searle**[18], **Jeremy D. Smith**[6], **Zachary Smith**[11], **Ross Swofford**[1], **Jason Turner-Maier**[1], **Juli Wade**[19], **Sarah Young**[1], **Amonida Zadissa**[18], **Genome Sequencing Platform and Whole Genome Assembly Team**[1], **Scott V. Edwards**[12], **Travis C. Glenn**[20], **Christopher J. Schneider**[21], **Jonathan B. Losos**[12], **Eric S. Lander**[1], **Matthew Breen**[2,22,23], **Chris P. Ponting**[3], and **Kerstin Lindblad-Toh**[1,24]

[1]Broad Institute of MIT and Harvard, Cambridge, MA

[2]Department of Molecular Biomedical Sciences, College of Veterinary Medicine, North Carolina State University, Raleigh, NC

[3]MRC Functional Genomics Unit, University of Oxford, Department of Physiology, Anatomy and Genetics, Oxford, UK

[4]Stanford University School of Medicine, Department of Developmental Biology, Stanford, CA

[5]University of Rochester, Rochester, NY

[6]Department of Biochemistry and Molecular Biology, Mississippi State University, Mississippi State, MS

[7]Department of Biology, Queens College, the City University of New York, New York, NY

Correspondence and requests for materials should be addressed to JA (jalfoldi@broadinstitute.org) and KLT (kersli@broadinstitute.org).

[*]These authors contributed equally to this work

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Author contributions

JA, FD and KLT planned and oversaw the project. GSP sequenced the genome. MG, DHe, SY, and WGAT assembled the genome. BFtH, MYK and PJJ constructed the BAC library. TCG and JW provided tissues for sequencing libraries and FISH analysis. MB, CW and DHe anchored the genome. TAC and DDP assembled the mitochondrial genome. JKC and ZS constructed the cDNA library. SS and AZ annotated the genome. LK, AH and CPP performed the gene repertoire analysis. TS aided egg protein experimental design. JDJ and SEP performed egg mass spectrometry. MG performed genome synteny analysis. EM performed segmental duplication analysis. CW and MB discovered the sex chromosomes and the pericentromeric inversions. PR performed the microchromosome and GC analysis. MKF and CPP participated in microchromosome and GC data interpretation. DAR constructed the repeat library. DAR, SB, PN, AMS, JDS, and CB performed the repeat analysis. MG, JBL, RG, SP, KdQ and RS participated in phylogeny data collection. SP, KdQ and RS participated in phylogeny analysis. All authors participated in data discussion and interpretation. JA, FD, CBL, RG, DAR, SVE, CJS, JBL, ESL, MB, CPP and KLT wrote the paper with input from other authors.

Author Information

The *A. carolinensis* whole-genome shotgun project has been deposited in GenBank under the project accession AAWZ00000000.2. All phylogeny sequence data can be found at http://purl.org/phylo/treebase/phylows/study/TB2:S11713. All animal experiments were approved by the MIT Committee for Animal Care. Reprints and permissions information is available at www.nature.com/reprints.

The authors declare no competing financial interests.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

[8]Biology Department and Graduate Program in Marine Genomics, College of Charleston, Charleston, SC

[9]Harvard Medical School, Boston, MA

[10]Department of Biochemistry and Molecular Genetics, University of Colorado School of Medicine, Aurora, CO

[11]The Center for Genomics and Bioinformatics, Indiana University, Bloomington, IN

[12]Department of Organismic and Evolutionary Biology and Museum of Comparative Zoology, Harvard University, Cambridge, MA

[13]Children's Hospital Oakland Research Institute, Oakland, CA

[14]Center for Biomolecular Science and Engineering, University of California, Santa Cruz, CA

[15]Department of Biology, University of New Mexico, Albuquerque, NM

[16]Program in Computational Bioscience, University of Colorado School of Medicine, Aurora, CO

[17]National Museum of Natural History, Smithsonian Institution, Washington, DC

[18]Wellcome Trust Sanger Institute, Hinxton, UK

[19]Departments of Psychology and Zoology, Program in Neuroscience, Michigan State University, East Lansing, MI

[20]Department of Environmental Health Science and Georgia Genomics Facility, University of Georgia, Athens, GA

[21]Biology Department, Boston University, Boston, MA

[22]Center for Comparative Medicine and Translational Research, North Carolina State University, Raleigh, NC

[23]University of North Carolina Lineberger Comprehensive Cancer Center, Chapel Hill, NC

[24]Science for Life Laboratory Uppsala, Department of Medical Biochemistry and Microbiology, Uppsala University, Uppsala, Sweden

Abstract

The evolution of the amniotic egg was one of the great evolutionary innovations in the history of life, freeing vertebrates from an obligatory connection to water and thus permitting the conquest of terrestrial environments[1]. Among amniotes, genome sequences are available for mammals[2] and birds[3–5], but not for non-avian reptiles. Here we report the genome sequence of the North American green anole lizard, *Anolis carolinensis*. We find that *A. carolinensis* microchromosomes are highly syntenic with chicken microchromosomes, yet do not exhibit the high GC and low repeat content that are characteristic of avian microchromosomes[3]. Also, *A. carolinensis* mobile elements are very young and diverse – more so than in any other sequenced amniote genome. This lizard genome's GC content is also unusual in its homogeneity, unlike the regionally variable GC content found in mammals and birds[6]. We describe and assign sequence to the previously unknown *A. carolinensis* X chromosome. Comparative gene analysis shows that amniote egg proteins have evolved significantly more rapidly than other proteins. An anole phylogeny resolves basal branches to illuminate the history of their repeated adaptive radiations.

The amniote lineage divided into the ancestral lineages of mammals and reptiles ~320 million years ago (MYA). Today, the surviving members of those lineages are mammals, comprising ~ 4,500 species, and reptiles, containing ~17,000 species. Within the reptiles, the two major clades diverged ~280 MYA -the lepidosaurs, which contains lizards (including

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

snakes) and the tuatara; and the archosaurs, containing crocodilians and birds (the position of turtles remains unclear)[7]. For simplicity, we will refer here to lepidosaurs as lizards(Figure 1).

The study of the major genomic events that accompanied the transition to a fully terrestrial life cycle has been assisted by the sequencing of several mammal [2] and three bird genomes[3–5]. The genome of the lizard *Anolis carolinensis* thus fills an important gap in the coverage of amniotes, splitting the long branch between mammals and birds and allowing more robust evolutionary analysis of amniote genomes.

For instance, almost all reptilian genomes contain microchromosomes, but these have only been studied at a sequence level in birds [3,8], raising the question as to whether the avian microchromosomes' peculiar sequence features are universal across reptilian microchromosomes[9]. Another example is the study of sex chromosome evolution. Nearly all placental and marsupial mammals share homologous sex chromosomes (XY)[10] and all birds share ZW sex chromosomes. However, lizards exhibit either genetic or temperature-dependent sex determination[11]. Characterization of lizard sex chromosomes would allow the study of previously unknown sex chromosomes and comparison of independent sex chromosome systems in closely related species.

Anolis lizards comprise a spectacularly diverse clade of ~400 described species distributed throughout the Neotropics. These lizards have radiated, often convergently, into a variety of ecological niches with attendant morphological adaptations, providing one of the best examples of adaptive radiation. In particular, their diversification into multiple replicate niches on diverse Caribbean islands via interspecific competition and natural selection has been documented in detail[12]. *Anolis carolinensis* is the only anole native to the USA and can be found from Florida and Texas up to North Carolina. We chose this species for genome sequencing because it is widely used as a reptile model for experimental ecology, behavior, physiology, endocrinology, epizootics and, increasingly, genomics.

The green anole genome was sequenced and assembled (AnoCar 2.0) using DNA from a female *Anolis carolinensis* (Supplementary Tables 1–4). Fluorescence in-situ hybridization (FISH) of 405 Bacterial Artificial Chromosome (BAC) clones (from a male) allowed the assembly scaffolds to be anchored to chromosomes (Supplementary Table 5 and Supplementary Figure 1). The *Anoliscarolinensis* genome has been reported to have a karyotype of n=18 chromosomes, comprising six pairs of large macrochromosomes and 12 pairs of small microchromosomes [13]. The draft genome sequence is 1.78 Gb in size (see Table 1 for assembly statistics)and represents an intermediate between genome assemblies of birds (0.9–1.3 Gb) and mammals (2.0–3.6 Gb).

We find that few chromosomal rearrangements occurred in the 280 million years since anole and chicken diverged, as had been hinted at by previous comparisons using *Xenopus* and chicken[14]. There are 259 syntenic blocks (defined as consecutive syntenic anchors that are consistent in order, orientation, and spacing, at a resolution of 1 Mb) between lizard and chicken (Supplementary Table 6 and Supplementary Figure 2). Interestingly, 19 out of 22 anchored chicken chromosomes are each syntenic to a single *A. carolinensis* chromosome over their entire lengths (Figure 2a); by contrast, only 6 (of 23) human chromosomes are syntenic to a single opossum chromosome over their entire lengths, even though the species diverged only 148 million years ago [15]. Segmental duplications follow trends seen in other amniote genomes (Supplementary Note, Supplementary Table 7 and Supplementary Figure 3).

Approximately 30% of the *A. carolinensis* genome is composed of mobile elements (ME), which comprise a much wider variety of active repeat families than is seen for either bird[3] or

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

mammalian[16] genomes. The most active classes are long interspersed (LINE) elements (27%) and short interspersed (SINE) elements (16%) [17] (Supplementary Table 8). The majority of LINE repeats belong to five groups (L1, L2, CR1, RTE, and R4)and appear to be recent insertions based on their sequence similarity (divergence ranges from 0.00–0.76%[18]). This contrasts with observations of mammalian genomes, where only a single family of LINEs, L1, has predominated over tens of millions of years. The DNA transposons comprise at least 68 families belonging to five superfamilies: hAT, Chapaev, Maverick, Tc/Mariner, and Helitron[19]. As with retrotransposons, the majority of DNA transposon families appear to be relatively young in contrast to the extremely few recently active DNA transposons found in other amniote genomes (Supplementary Table 9). Overall, *A. carolinensis* MEs feature significantly higher GC content (43.5%, $p<10^{-20}$) than the genome-wide average of 40.3%. In addition to mobile elements, *A. carolinensis* exhibits a high density (3.5%)of tandem repeats, with length and frequency distributions similar to those of human microsatellite DNA[16]. We now know that amniote genomes come in at least three types: mammalian genomes are enriched for L1 elements and have a high degree of ME accumulation, bird genomes are repeat-poor with very little ME activity, while the lizard genome contains an extremely wide diversity of active ME families, but has a low rate of accumulation, which is reminescent of the ME profile of teleostean fish[20].

Most reptile genomes contain microchromosomes, but the numbers vary among species; the *A. carolinensis* genome contains 12 pairs of microchromosomes[13], whereas the chicken genome contains 28 pairs. Bird microchromosomes have very distinctive properties compared to bird macrochromosomes, such as higher GC and lower repeat contents [3], whereas lizard microchromosomes do not exhibit these features (Figure 2b). Remarkably, all sequence anchored to microchromosomes in *A. carolinensis* also aligns to microchromosomes in the chicken genome, and all but one *A. carolinensis* microchromosome are each syntenic to only a single corresponding chicken microchromosome (Figure 2a). Microchromosomes conserved between *A. carolinensis* and chicken thus could have arisen in the reptile ancestor, whereas the remaining chicken microchromosomes could be derived in the bird lineage. Alternatively, the remaining chicken microchromosomes could have been present in the reptile ancestor but fused to form macrochromosomes in the lizard lineage.

The *Anoliscarolinensis* genome has surprisingly little regional variation of GC content, substantially less than previously observed for birds and mammals; it is the only amniotic genome known whose nucleotide composition is as homogenous as the frog genome[6] (Supplementary Figures 4–5). Figure 3 illustrates how local GC content is evolutionarily conserved between human chromosome 14 and chicken chromosome 5, but to a much lesser degree with *A. carolinensis* chromosome 1. Since all sequenced amniote genomes other than *A. carolinensis* contain these homologous varying levels of GC content (“isochores”)[21], the ancestral amniote GC heterogeneity is likely to have eroded towards homogeneity in this lizard’s lineage. It has been proposed that isochores with high GC content are a consequence of higher rates of GC-biased gene conversion in regions of higher recombination [3]. The greater GC homogeneity in the anole genome may thus reflect more uniform recombination rates, or else a substantially reduced bias towards GC during the resolution of gene conversion events in the *A. carolinensis* lineage (for a discussion, see[6]).

Both temperature-dependent sex determination and XY genetic sex determination have been found in Iguania[11]. Within the genus *Anolis*, there are species with heteromorphic XY chromosomes (including those with multiple X and Y chromosomes), and others with entirely homomorphic chromosomes[13]. *Anolis carolinensis* is known to have genotypic sex determination [22], but the form of its sex chromosomes (ZW or XY) has thus far been unknown due to a lack of obviously heteromorphic chromosomes.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

In-depth examination of male and female cells using FISH allowed us to identify the microchromosome previously designated as 'b' as the *Anolis carolinensis* X chromosome; it is present in two copies in females and one in males. This chromosome is syntenic to chicken microchromosome 15. Eleven BACs assigned to two scaffolds, #154 (3.3 Mb) and chrUn0090 (1.8 Mb), hybridize via FISH to the p arms of the two X chromosomes in females, and hybridize to the p arm of the single X chromosome in males (Figure 4, Supplementary Figure 1). *Anolis carolinensis* thereby shows a pattern representative of a male heterogametic system of genotypic sex determination. We have not identified the Y chromosome, but we hypothesize that *A. carolinensis* possesses both X and Y chromosomes, as both male and female cells contain the same number of chromosomes.

The 5.1 Mb of sequence assigned to the X chromosome contain 62 protein-coding genes (Supplementary Table 10); Gene Ontology (GO) terms associated with these genes show no significant enrichment. It is very likely that there is more X chromosome sequence that is currently labeled as unanchored scaffolds in the AnoCar 2.0 assembly. Identification of the *A. carolinensis* sex determination gene will require considerable functional biology, but we note that the chicken sex determination gene *DMRT1* is located on *A. carolinensis* chromosome 2, and that *SOX3* (the X chromosome paralog of the therian mammal sex determination gene *SRY*) is located on an unanchored *A. carolinensis* scaffold and are thus unlikely to be the *A. carolinensis* sex determination gene.

All ten *A. carolinensis* individuals (originating from South Carolina and Tennessee) used for FISH mapping showed large pericentromeric inversions in one or more of chromosomes 1–4, with no correlation between different chromosomal inversions or with the sex of the lizard(see Supplementary Note, Supplementary Table 11, and Supplementary Figure 6).

A total of 17,472 protein-coding genes and 2,924 RNA genes were predicted from the *Anolis carolinensis* genome assembly (Ensembl release 56, Sept. 2009). We built a phylogeny for all *A. carolinensis* genes and their homologues in eight other vertebrate species(human, mouse, dog, opossum, platypus, chicken, zebra finch and pufferfish), allowing us to identify a conservative set of 3994 one-to-one orthologs, i.e. genes that have not been duplicated or deleted in any of these vertebrates since their last common ancestor. These gene phylogenies were also used to identify genes that arose by duplication in the lizard lineage after the split with the avian lineage and, separately, those that were lost in the mammalian lineage after the mammal-reptile split(Figure 1, Supplementary Note, Supplementary Figure 7, Supplementary Table 12).

We found 11 *A. carolinensis* opsin genes that have no mammalian orthologs (but have orthologs in invertebrates, fishes and frog), and thus appear to have been lost during mammalian evolution (Supplementary Table 13). The large repertoire of opsins may contribute to the excellent color vision of anoles, including the ability to see in the ultraviolet range, and also may contribute to their hyperdiversity by allowing the evolution of diverse, species-specific coloration of the dewlap, which plays an important role in sexual selection and species recognition[12]. Similarly, olfactory receptor and beta-keratin genes are highly duplicated in *A. carolinensis* (Supplementary Note, Supplementary Figure 9).

Many reptiles, including green anoles, differ from placental mammals in being oviparous (laying eggs). Vivipary in placental mammals is a derived state, reflected in their loss of some egg-related genes. We used mass spectrometry to identify proteins present in the immature *A. carolinensis* egg, as most egg proteins are produced in the mother's body and then transported into the immature egg. We found that in contrast with mammals, reptiles have lineage-specific gene duplications, including vitellogenins (VTGs), apovitellenin-1, ovomucin-alpha and three homologs of ovocalyxin-36, a chicken eggshell matrix protein.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Our results show rapid evolution of egg protein genes among amniotes. Specifically, we found proteins from 276 *A. carolinensis* genes in immature *A. carolinensis* eggs (Supplementary Tables 14–15), of which only 50 have been confirmed to be present in chicken eggs by mass spectrometry[23–24]. These genes include VTGs, a lysozyme, vitelline membrane outer layer protein 1(*VMO1*) paralogues, protease inhibitors, natterin, and nothepsin. By aligning genes that are one-to-one orthologs in *Anolis carolinensis* and chicken, we found that egg proteins evolve significantly more rapidly than non-egg proteins (mean *dN/dS* values of 0.186 and 0.135, respectively; $p = 1.2\times10^{-5}$), which reflects reduced purifying selection and/or more frequent episodes of adaptive evolution.

Using multiple vertebrate genome sequences, we identified three *VMO1* paralogs (which we name α, β and γ)that we infer to have been present in the last common ancestor of all reptiles and mammals. While at least one of *VMO1*-α, *VMO1*-β, and *VMO1*-γ have been lost in all other amniote genomes, the *A. carolinensis* genome contains representatives of all three paralogs. Moreover, the *A. carolinensis*-specific *VMO1*-α family has grown to 13 members and has experienced positive selection of amino acid substitutions within a negatively charged, likely substrate-binding cavity, changes which, presumably, modify its lysozyme-like transferase activity(Supplementary Note, Supplementary Figure 8, Supplementary Tables 16–17) [40].

The extensive and active repeat repertoire of *Anolis carolinensis* has allowed us to discover the origin of several mammalian conserved elements. Through the process of exaptation (a major change in function of a sequence during evolution), certain MEs that were active in the amniote ancestor have become conserved, and presumably functional, in mammals, while remaining active MEs in *A. carolinensis*. The origin of these conserved mammalian sequences in MEs was not recognizable without comparison to a distant and repeat-rich genome sequence[25]. We identified 96 such exapted elements (see Supplementary Table 18) in the human genome tracing back to MEs present in the amniote ancestor that are still present in *A. carolinensis*, particularly the CR1, L2 and gypsy families. While most exapted elements are non-coding and likely serve a regulatory function, we also identified a protein-coding exon that was exapted from an L2-like LINE, now constituting exon 2 in a mammal-specific N-terminal region of the *MIER1* (Mesoderm Induction Early Response 1) protein. This exon is highly conserved across 29 mammals and therefore likely represents a mammalian innovation since the amniote ancestor.

GO terms associated with the transcription start site closest to each exapted element in the human genome show enrichment for neurodevelopmental genes(see Methods), with “ephrin receptor binding”, “nervous system development” and “synaptic transmission” being strongly enriched (all p-values $< 5\times10^{-3}$). These enrichments are consistent with adaptive changes in neurodevelopment occurring during the emergence of mammals.

Anolis lizards are a textbook case of adaptive radiation, having diversified independently on each island in the Greater Antilles, and throughout the Neotropics, producing a wide variety of ecologically and morphologically differentiated species, with as many as 15 found at a single locality[12]. Although anoles are widely used as a model system for phylogenetic comparative studies, it has been difficult to determine the evolutionary relationships among major anole clades due to rapid evolutionary radiations associated with access to new dimensions of ecological opportunity. Successfully resolving the relatively short branching events associated with such a radiation requires a wealth of data from loci evolving at an appropriate rate.

We used the genome sequence of *Anolis carolinensis* to develop a new phylogenomic dataset comprised of 20 kb of sequence data sampled from across the genomes of 93 species

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

of anoles (Supplementary Tables 19–20). Analyses of this dataset infer a well-supported phylogeny that reinforces and clarifies the adaptive and biogeographic history of anoles (Figure 5, details in Supplementary Figure 10). First, our phylogenomic analysis reaffirms previous molecular and morphological studies suggesting that similar anole habitat specialists have evolved independently on each of the four large Greater Antillean islands. Second, our analyses suggest a complex biogeographic scenario involving a limited number of dispersal events between islands and extensive *in situ* diversification within islands. The closest relatives of *Anolis* occur on the mainland, and the phylogeny confirms the existence of two colonizations, one into the southern Lesser Antilles and the second producing the diverse adaptive radiations throughout the rest of the Caribbean. Within this latter clade, anoles initially diversified primarily on the two larger Greater Antillean islands (though Puerto Rico also seems to have been involved) before subsequently undergoing secondary radiations on all of the islands and eventually returning to the mainland, where this back-colonization has produced an extensive evolutionary radiation. The phylogeny also indicates that very few inter-island dispersal events occurred in Greater Antillean evolution. Rather, the Greater Antillean faunas, renowned for the extent to which the same ecomorphs are found on each island, are primarily the result of convergent evolution[26].

In conclusion, the genome sequence of *Anolis carolinensis* allows a deeper understanding of amniote evolution. Filling this important reptilian node with a sequenced genome has revealed derived states in each major amniote branch and has helped to illuminate the amniote ancestor. However, the tree of sequenced reptilian genomes is still extremely sparse, and the sequencing of additional non-avian reptiles would be necessary to fully understand how typical *A. carolinensis* and the sequenced bird genomes are of the entire reptile clade.

In addition to the utility of the *A. carolinensis* genome sequence as a representative of non-avian reptiles, *Anolis* species are a unique resource for the study of adaptive radiation and convergent evolution. With their invasions of and subsequent radiations on Caribbean islands, anoles provide a terrestrial analog to stickleback and cichlid fish, which underwent adaptive evolution in separate aquatic environments. Just as genomic research in sticklebacks has deepened the study of aquatic ecological speciation, a large-scale genomic phylogenetic survey of the Caribbean anoles would be an outstanding opportunity for detailed study of adaptive evolution in a land animal[27], in particular since anole genomes contain large numbers of active MEs that we speculate could form substrates for exaptation of novel regulatory elements.

Methods

Appear in the online supplement.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Generation of the *Anolis carolinensis* sequence at the Broad Institute of MIT and Harvard was supported by grants from the National Human Genome Research Institute (NHGRI). We would also like to thank the David and Lucile Packard Foundation for their early support of anole genomics, R. Andrews for her advice on lizard egg biology, C. Hickman and B. Temple and the Herpetology group at the University of Georgia's Savannah River Ecology Lab for assistance with sample collection, and L. Gaffney and L. Virnoche for assistance with figure and text preparation.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

References

1. Sumida, SS.; Martin, KLM. Amniote Origins: Completing the Transition to Land. Academic Press; 1997.
2. Lindblad-Toh K, et al. A high-resolution map of evolutionary constraint in the human genome based on 29 eutherian mammals. Nature. (submitted).
3. Hillier LW, et al. Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. Nature. 2004; 432:695–716. [PubMed: 15592404]
4. Warren WC, et al. The genome of a songbird. Nature. 464:757–762. [PubMed: 20360741]
5. Dalloul RA, et al. Multi-platform next-generation sequencing of the domestic turkey (Meleagris gallopavo): genome assembly and analysis. PLoS biology. :8.
6. Fujita MK, Edwards SV, Ponting CP. Unexpected homogeneity in GC content and divergent isochore structure in the *Anolis* genome. Genome Biol Evol. (in press).
7. Benton MJ, Donoghue PC. Paleontological evidence to date the tree of life. Molecular biology and evolution. 2007; 24:26–53. [PubMed: 17047029]
8. Axelsson E, Webster MT, Smith NGC, Burt DW, Ellegren H. Comparison of the chicken and turkey genomes reveals a higher rate of nucleotide divergence on microchromosomes than macrochromosomes. Genome Research. 2005; 15:120–125. [PubMed: 15590944]
9. Burt DW. Origin and evolution of avian microchromosomes. Cytogenetic and genome research. 2002; 96:97–112. [PubMed: 12438785]
10. Potrzebowski L, et al. Chromosomal gene movements reflect the recent origin and biology of therian sex chromosomes. PLoS biology. 2008; 6:e80. 08-PLBI-RA-0019 [pii]. 10.1371/journal.pbio.0060080 [PubMed: 18384235]
11. Organ CL, Janes DE. Evolution of sex chromosomes in Sauropsida. Integrative and Comparative Biology. 2008; 48:512–519. [PubMed: 21669812]
12. Losos, JB. Lizards in an Evolutionary Tree: Ecology and Adaptive Radiation of Anoles. University of California Press; 2009.
13. Gorman, GC. Cytotaxonomy and Vertebrate Evolution. Chiarelli, AB.; Capanna, E., editors. Vol. Ch 12. Academic Press; 1973. p. 349-424.
14. Hellsten U, et al. The genome of the Western clawed frog Xenopus tropicalis. Science. 2010; 328:633–636. 328/5978/633 [pii]. 10.1126/science.1183670 [PubMed: 20431018]
15. Bininda-Emonds OR, et al. The delayed rise of present-day mammals. Nature. 2007; 446:507–512. [PubMed: 17392779]
16. Shedlock AM, et al. Phylogenomics of nonavian reptiles and the structure of the ancestral amniote genome. Proceedings of the National Academy of Sciences of the United States of America. 2007; 104:2767–2772. [PubMed: 17307883]
17. Kordis D. Transposable elements in reptilian and avian (sauropsida) genomes. Cytogenetic and genome research. 2009; 127:94–111. [PubMed: 20215725]
18. Novick PA, Basta H, Floumanhaft M, McClure MA, Boissinot S. The evolutionary dynamics of autonomous non-LTR retrotransposons in the lizard Anolis carolinensis shows more similarity to fish than mammals. Molecular biology and evolution. 2009; 26:1811–1822. [PubMed: 19420048]
19. Novick PA. The evolution and diversity of DNA transposons in the genome of the lizard Anolis carolinensis. Genome Biology and Evolution. 2010 In Press.
20. Furano AV, Duvernell DD, Boissinot S. L1 (LINE-1) retrotransposon diversity differs dramatically between mammals and fish. Trends Genet. 2004; 20:9–14. S0168952503003214 [pii]. [PubMed: 14698614]
21. Duret L, Eyre-Walker A, Galtier N. A new perspective on isochore evolution. Gene. 2006; 385:71–74. S0378-1119(06)00479-3 [pii]. 10.1016/j.gene.2006.04.030 [PubMed: 16971063]
22. Viets BE, Ewert MA, Talent LG, Nelson CE. Sex-Determining Mechanisms in Squamate Reptiles. Journal of Experimental Zoology. 1994; 270:45–46.
23. Mann K. The chicken egg white proteome. Proteomics. 2007; 7:3558–3568. [PubMed: 17722208]
24. Farinazzo A, et al. Chicken egg yolk cytoplasmic proteome, mined via combinatorial peptide ligand libraries. Journal of chromatography. 2009; 1216:1241–1252. [PubMed: 19081572]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

25. Lowe CB, Bejerano G, Salama SR, Haussler D. Endangered species hold clues to human evolution. J Hered. 2010; 101:437–447. esq016 [pii]. 10.1093/jhered/esq016 [PubMed: 20332163]

26. Williams, EE. Lizard Ecology: Studies of a Model Organism. Huey, RB.; Pianka, ER.; Schoener, TW., editors. Harvard University Press; 1983.

27. Colosimo PF, et al. Widespread parallel evolution in sticklebacks by repeated fixation of Ectodysplasin alleles. Science (New York, NY). 2005; 307:1928–1933. 307/5717/1928 [pii]. 10.1126/science.1107239

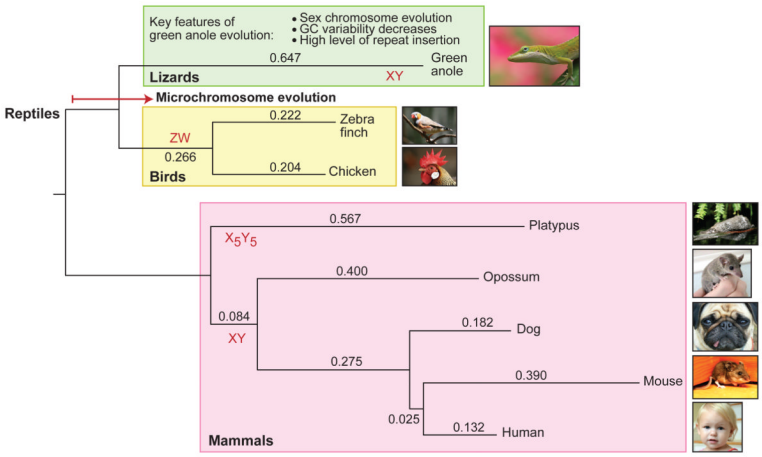


Figure 1. Amniote phylogeny based on protein synonymous sites showing major features of amniote evolution

Major features of lizard evolution including homogenization of GC content, high sex chromosome turnover and high levels of repeat insertion are highlighted. Sex chromosome inventions are indicated in red. The X-axis is proportional to dS.

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

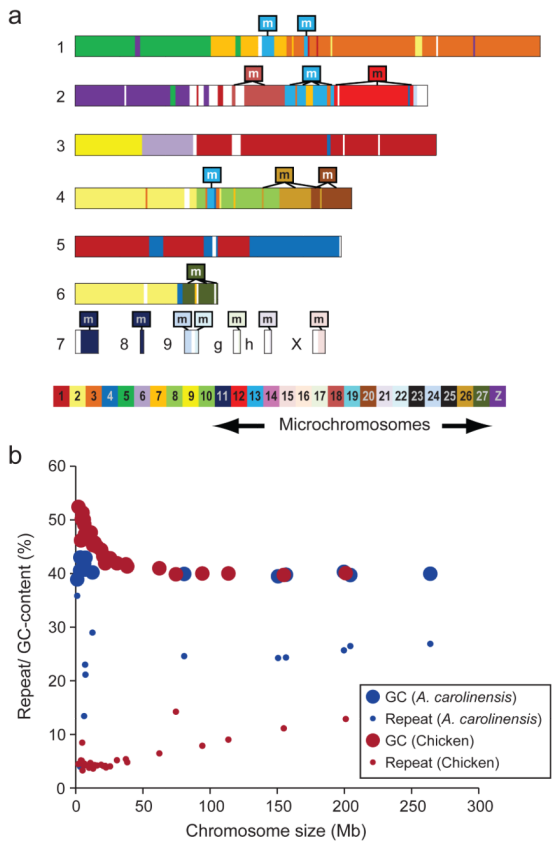


Figure 2. *Anolis carolinensis*-chicken synteny map reveals synteny of reptile microchromosomes but dissimilar GC and repeat content

(A) Very few rearrangements have occurred in the 280 million years since *Anolis carolinensis* and chickens diverged. *A. carolinensis* microchromosomes are exclusively syntenic to chicken microchromosomes. Horizontal colored bars depict the six *A. carolinensis* macrochromosomes (1–6) and the six (of 12) *A. carolinensis* microchromosomes that have sequence anchored to them that is syntenic to the chicken genome (7, 8, 9, X, LGg, LGh). Chromosomes that could be ordered by size were assigned a number; the smaller microchromosomes that could not be distinguished by size were assigned a lower case letter. Each color corresponds to a different chicken chromosome as indicated in the legend. Any part of an *A. carolinensis* chromosome that is syntenic to a chicken microchromosome is indicated by a lower case m. **(B)** Chicken microchromosomes have both higher GC content and lower repeat content than chicken macrochromosomes, whereas *A. carolinensis* chromosomes do not vary in GC or repeat content by chromosome size. Large circles designate the GC % of each chromosome in the chicken and lizard genomes with greater than 100 kb of sequence anchored to it. Small circles designate the percentage of the genome made up of repetitive sequence of each chromosome in the chicken (blue circles) and lizard (red circles) genomes.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

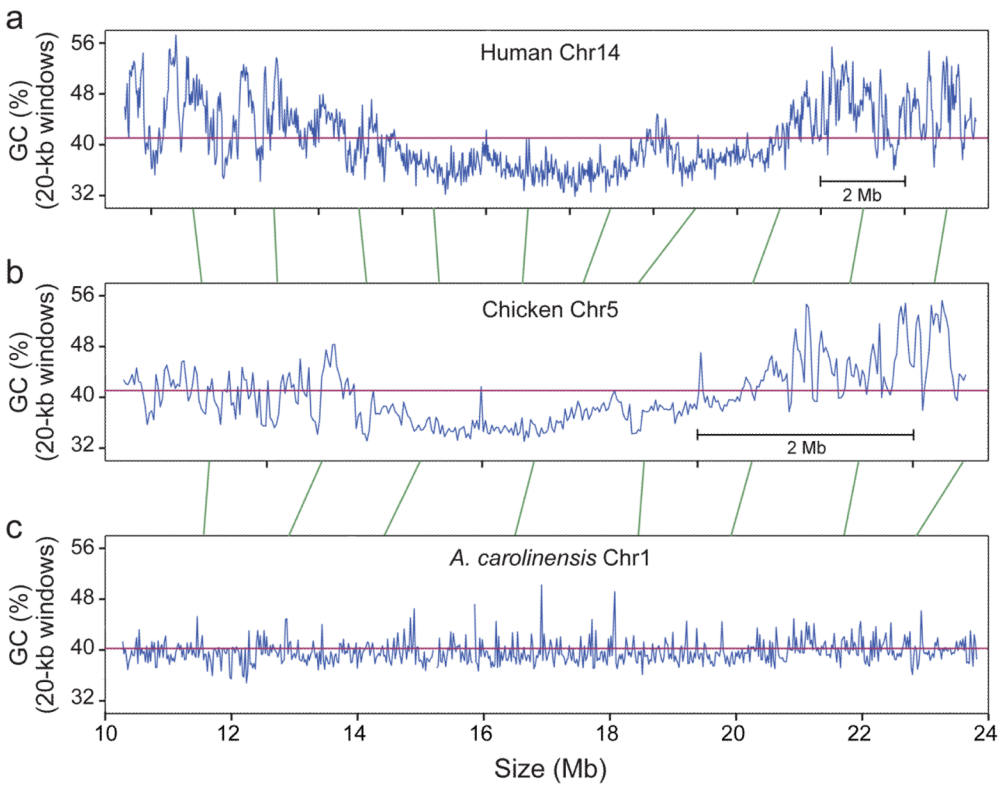


Figure 3. The *Anolis carolinensis* genome lacks isochores

The *A. carolinensis* genome shows only very local variation in GC content, unlike the human and chicken genomes, which also show larger trends in GC variation, sometimes called ‘isochores’. Syntenic regions of human chromosome 14, chicken chromosome 5, and *A. carolinensis* chromosome 1 are shown. The human and chicken regions are inverted and rearranged to align with the *A. carolinensis* region. Blue lines depict GC % in 20 kb windows. The purple line designates the genome average. Green lines represent examples of syntenic anchors between the three genomes.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

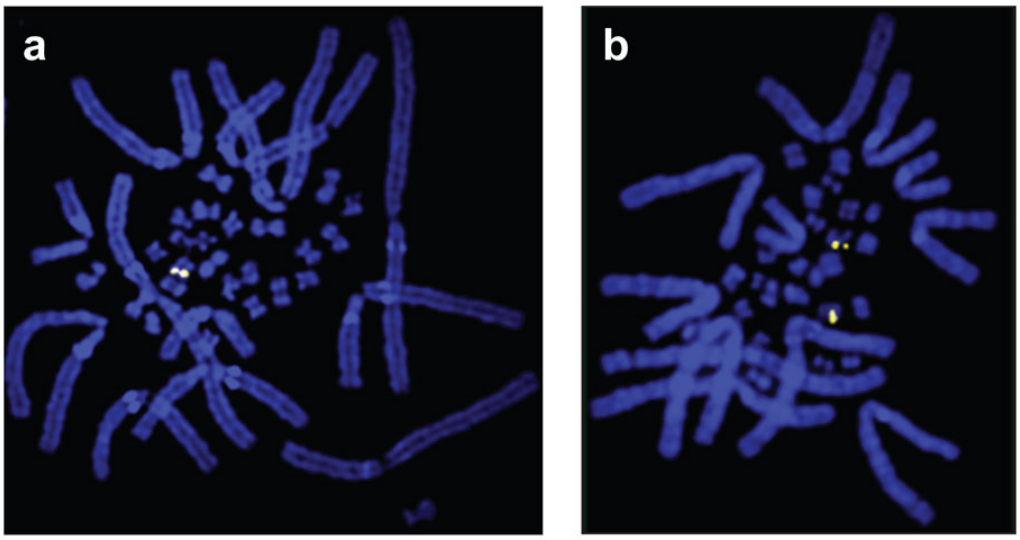


Figure 4. The *Anolis carolinensis* genome contains a newly discovered X chromosome

The X chromosome, a microchromosome, is found in **(A)** one copy in male *Anolis carolinensis* and in **(B)** two copies in females. The BAC 206M13 (CHORI-318 BAC library) is hybridized to the p arm of the X chromosome using FISH in both male and female metaphase spreads. 206M13 and ten other BACs showed this sex-specific pattern in cells derived from five male and five female individuals.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

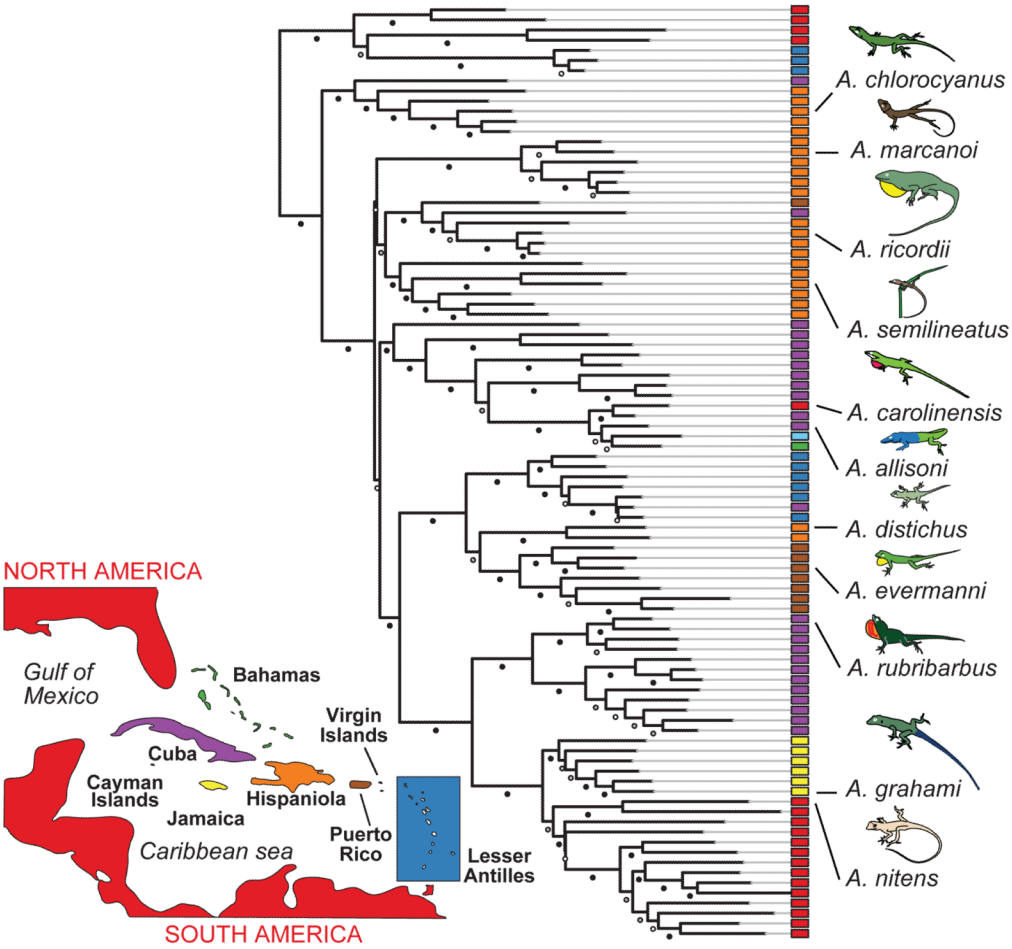


Figure 5. A phylogeny of 93 *Anolis* species clarifies the biogeographic history of anoles

Anolis ecomorphs derive from convergent evolution and not from frequent inter-island migration. Using conserved primer pairs distributed across the genome of *Anolis carolinensis*, we obtain sequences from 46 genomically diverse loci evolving at a range of evolutionary rates and representing both protein-coding and non-coding regions. Maximum likelihood analyses of this new dataset of 20 kb aligned nucleotides infer nearly all previously established anole relationships while also partially resolving the basal relationships that have plagued previous studies. White circles indicate bootstrap values <70; gray circles: 70<bs<95; black circles bs >95.