Corrections

BIOCHEMISTRY

Correction for "Purified vitamin K epoxide reductase alone is sufficient for conversion of vitamin K epoxide to vitamin K and vitamin K to vitamin KH_2," by Pei-Hsuan Chu, Teng-Yi Huang, Jason Williams, and D. W. Stafford, which appeared in issue 51, December 19, 2006, of *Proc Natl Acad Sci USA* (103:19308–19313; first published December 12, 2006; 10.1073/pnas.0609401103).

The authors note that Fig. 5 and its corresponding legend appeared incorrectly. The corrected figure and its corrected legend appear below.

Also, the authors note that on page 19312, right column, 3rd full paragraph, line 5 "The reaction was carried out for 1 hr." should instead appear as "The reaction was carried out for 20 min."

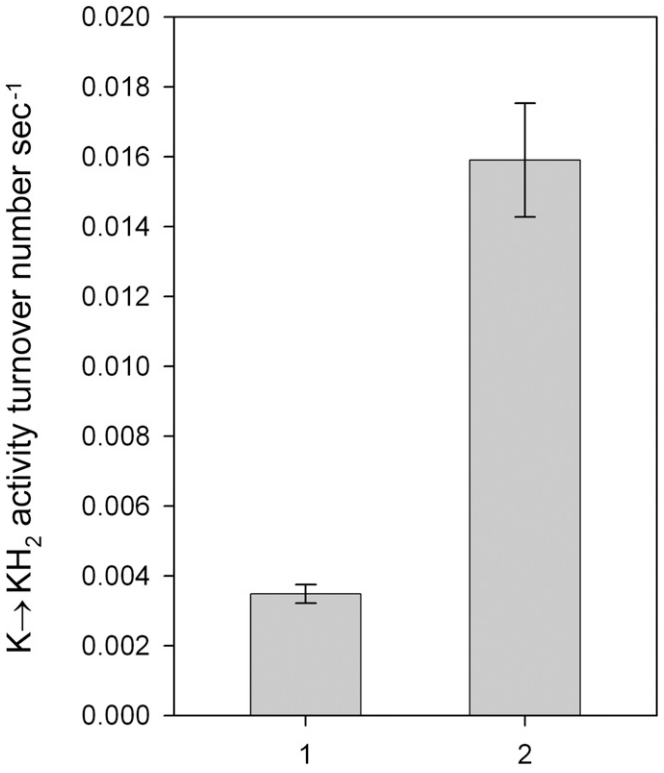

Fig. 5. Conversion of vitamin K to vitamin KH_2 by VKOR. The reaction was performed by using purified VKOR that had been dialyzed in the presence of THP. VKOR activity is represented as turnover number per second. Bar 1, DTT with elution buffer as background control; Bar2, purified VKOR after dialysis against buffer A with 4 mM THP. Data are represented as mean ± SD (n = 3).

www.pnas.org/cgi/doi/10.1073/pnas.1401722111

EVOLUTION

Correction for "The Burmese python genome reveals the molecular basis for extreme adaptation in snakes," by Todd A. Castoe, A. P. Jason de Koning, Kathryn T. Hall, Daren C. Card, Drew R. Schield, Matthew K. Fujita, Robert P. Ruggiero, Jack F. Degner, Juan M. Daza, Wanjun Gu, Jacobo Reyes-Velasco, Kyle J. Shaney, Jill M. Castoe, Samuel E. Fox, Alex W. Poole, Daniel Polanco, Jason Dobry, Michael W. Vandewege, Qing Li, Ryan K. Schott, Aurélie Kapusta, Patrick Minx, Cédric Feschotte, Peter Uetz, David A. Ray, Federico G. Hoffmann, Robert Bogden, Eric N. Smith, Belinda S. W. Chang, Freek J. Vonk, Nicholas R. Casewell, Christiaan V. Henkel, Michael K. Richardson, Stephen P. Mackessy, Anne M. Bronikowsi, Mark Yandell, Wesley C. Warren, Stephen M. Secor, and David D. Pollock, which appeared in issue 51, December 17, 2013, of *Proc Natl Acad Sci USA* (110:20645–20650; first published December 2, 2013; 10.1073/pnas.1314475110).

The authors note that the author name Anne M. Bronikowsi should instead appear as Anne M. Bronikowski. The corrected author line appears below. The online version has been corrected.

Todd A. Castoe, A. P. Jason de Koning, Kathryn T. Hall, Daren C. Card, Drew R. Schield, Matthew K. Fujita, Robert P. Ruggiero, Jack F. Degner, Juan M. Daza, Wanjun Gu, Jacobo Reyes-Velasco, Kyle J. Shaney, Jill M. Castoe, Samuel E. Fox, Alex W. Poole, Daniel Polanco, Jason Dobry, Michael W. Vandewege, Qing Li, Ryan K. Schott, Aurélie Kapusta, Patrick Minx, Cédric Feschotte, Peter Uetz, David A. Ray, Federico G. Hoffmann, Robert Bogden, Eric N. Smith, Belinda S. W. Chang, Freek J. Vonk, Nicholas R. Casewell, Christiaan V. Henkel, Michael K. Richardson, Stephen P. Mackessy, Anne M. Bronikowski, Mark Yandell, Wesley C. Warren, Stephen M. Secor, and David D. Pollock

www.pnas.org/cgi/doi/10.1073/pnas.1324133111

GENETICS

Correction for "Male-specific region of the bovine Y chromosome is gene rich with a high transcriptomic activity in testis development," by Ti-Cheng Chang, Yang Yang, Ernest F. Retzel, and Wan-Sheng Liu, which appeared in issue 30, July 23, 2013, of *Proc Natl Acad Sci USA* (110:12373–12378; first published July 10, 2013; 10.1073/pnas.1221104110).

The authors wish to note, "We have recently updated the data associated with our GenBank depositions to include age, tissue, and developmental stage of the bovine testis RNA-seq data. We have also deposited new data for information discussed in the *Supporting Information* of our article. We apologize for not providing this information at the time of publication. The updated accession numbers are as follows:

Accession Number	Sample	Developmental Stage	Tissue
SRX357350	Bos taurus 789_20D	postnatal-20 d	Testis
SRX357349	Bos taurus 789_20D	postnatal-20 d	Testis
SRX357348	Bos taurus 74_8M	puberty-8 mo	Testis
SRX357347	Bos taurus 74_8M	puberty-8 mo	Testis
SRX357346	Bos taurus 645_2Y	maturity-2 y	Testis
SRX357345	Bos taurus 645_2Y	maturity-2 y	Testis
SRX388838	Direct cDNA selection of the bovine Y chromosome		

"In addition, the project 'Transcriptome analysis of the bovine Y chromosome,' together with the bovine testis cDNA selection reads and assembled transcripts/ncRNAs (> 200 bp), were submitted to the Transcriptome Shotgun Assembly (TSA) database, www.ncbi.nlm.nih.gov/genbank/tsa (Bioproject accession no. PRJNA230872; reads accession no. SRX388838).

"The assembled contigs (> 200 bp) have been deposited at DNA Data Bank of Japan/European Molecular Biology Laboratory/GenBank under the accession GAQO00000000. The version described in this paper is the first version, GAQO01000000."

www.pnas.org/cgi/doi/10.1073/pnas.1401739111

IMMUNOLOGY

Correction for "IRAK-1 bypasses priming and directly links TLRs to rapid NLRP3 inflammasome activation," by Keng-Mean Lin, Wei Hu, Ty Dale Troutman, Michelle Jennings, Travis Brewer, Xiaoxia Li, Sambit Nanda, Philip Cohen, James A. Thomas, and Chandrashekhar Pasare, which appeared in issue 2, January 14, 2014, of *Proc Natl Acad Sci USA* (111:775–780; first published December 30, 2013; 10.1073/pnas.1320294111).

The authors note that James A. Thomas should be included as a cocorresponding author. Correspondence can be addressed to him at james.thomas@bcm.edu.

Also, the authors note that they omitted references to articles by Juliana et al. and Fernandes-Alnemri et al. The complete references appear below.

34. Juliana C, et al. (2012) Non-transcriptional priming and deubiquitination regulate NLRP3 inflammasome activation. *J Biol Chem* 287(43):36617–36622.
35. Fernandes-Alnemri T, et al. (2013) Cutting edge: TLR signaling licenses IRAK1 for rapid activation of the NLRP3 inflammasome. *J Immunol* 191(8):3995–3999.

www.pnas.org/cgi/doi/10.1073/pnas.1401725111

The Burmese python genome reveals the molecular basis for extreme adaptation in snakes

Todd A. Castoe[a,b], A. P. Jason de Koning[a,c], Kathryn T. Hall[a], Daren C. Card[b], Drew R. Schield[b], Matthew K. Fujita[b], Robert P. Ruggiero[a], Jack F. Degner[d], Juan M. Daza[e], Wanjun Gu[f], Jacobo Reyes-Velasco[b], Kyle J. Shaney[b], Jill M. Castoe[a,b], Samuel E. Fox[g], Alex W. Poole[a], Daniel Polanco[a], Jason Dobry[h], Michael W. Vandewege[i], Qing Li[j], Ryan K. Schott[k], Aurélie Kapusta[j], Patrick Minx[l], Cédric Feschotte[j], Peter Uetz[m], David A. Ray[i,n], Federico G. Hoffmann[i,n], Robert Bogden[h], Eric N. Smith[b], Belinda S. W. Chang[k], Freek J. Vonk[o,p,q], Nicholas R. Casewell[q,r], Christiaan V. Henkel[p,s], Michael K. Richardson[p], Stephen P. Mackessy[t], Anne M. Bronikowski[u], Mark Yandell[j], Wesley C. Warren[l], Stephen M. Secor[v], and David D. Pollock[a,1]

[a]Department of Biochemistry and Molecular Genetics, University of Colorado School of Medicine, Aurora, CO 80045; [b]Department of Biology, University of Texas, Arlington, TX 76010; [c]Department of Biochemistry and Molecular Biology, University of Calgary and Alberta Children's Hospital Research Institute for Child and Maternal Health, Calgary, AB, Canada T2N 4N1; [d]Department of Human Genetics, University of Chicago, Chicago, IL 60637; [e]Instituto de Biologia, Universidad de Antiochia, Medellin, Colombia 05001000; [f]Key Laboratory of Child Development and Learning Science, Southeast University, Ministry of Education, Nanjing 210096, China; [g]Department of Biology, Linfield College, McMinnville, OR 97128; [h]Amplicon Express, Pullman, WA 99163; [i]Department of Biochemistry, Molecular Biology, Entomology and Plant Pathology, Mississippi State University, Mississippi State, MS 39762; [j]Department of Human Genetics, University of Utah School of Medicine, Salt Lake City, UT 84112; [k]Department of Ecology and Evolutionary Biology and Cell and Systems Biology, Centre for the Analysis of Genome Evolution and Function, University of Toronto, Toronto, ON, Canada M5S 3G5; [l]Genome Institute, Washington University School of Medicine, St. Louis, MO 63108; [m]Center for the Study of Biological Complexity, Virginia Commonwealth University, Richmond, VA 23284; [n]Department of Biological Sciences, Texas Tech University, Lubbock, TX 79409; [o]Naturalis Biodiversity Center, 2333 CR, Leiden, The Netherlands; [p]Institute of Biology Leiden, Leiden University, Sylvius Laboratory, Sylviusweg 72, 2300 RA, Leiden, The Netherlands; [q]Molecular Ecology and Evolution Group, School of Biological Sciences, Bangor University, Bangor LL57 2UW, United Kingdom; [r]Alistair Reid Venom Research Unit, Liverpool School of Tropical Medicine, Liverpool L3 5QA, United Kingdom; [s]ZF Screens, Bio Partner Center, 2333 CH, Leiden, The Netherlands; [t]School of Biological Sciences, University of Northern Colorado, Greeley, CO 80639; [u]Department of Ecology, Evolution, and Organismal Biology, Iowa State University, Ames, IA 50011; and [v]Department of Biological Sciences, University of Alabama, Tuscaloosa, AL 35487

Edited by David B. Wake, University of California, Berkeley, CA, and approved November 4, 2013 (received for review July 31, 2013)

Snakes possess many extreme morphological and physiological adaptations. Identification of the molecular basis of these traits can provide novel understanding for vertebrate biology and medicine. Here, we study snake biology using the genome sequence of the Burmese python (*Python molurus bivittatus*), a model of extreme physiological and metabolic adaptation. We compare the python and king cobra genomes along with genomic samples from other snakes and perform transcriptome analysis to gain insights into the extreme phenotypes of the python. We discovered rapid and massive transcriptional responses in multiple organ systems that occur on feeding and coordinate major changes in organ size and function. Intriguingly, the homologs of these genes in humans are associated with metabolism, development, and pathology. We also found that many snake metabolic genes have undergone positive selection, which together with the rapid evolution of mitochondrial proteins, provides evidence for extensive adaptive redesign of snake metabolic pathways. Additional evidence for molecular adaptation and gene family expansions and contractions is associated with major physiological and phenotypic adaptations in snakes; genes involved are related to cell cycle, development, lungs, eyes, heart, intestine, and skeletal structure, including GRB2-associated binding protein 1, SSH, WNT16, and bone morphogenetic protein 7. Finally, changes in repetitive DNA content, guanine-cytosine isochore structure, and nucleotide substitution rates indicate major shifts in the structure and evolution of snake genomes compared with other amniotes. Phenotypic and physiological novelty in snakes seems to be driven by system-wide coordination of protein adaptation, gene expression, and changes in the structure of the genome.

comparative genomics | transposable elements | systems biology | transcriptomics | physiological remodeling

Biological research increasingly incorporates nontraditional model organisms, particularly model organisms with extreme phenotypes that can provide novel insight into vertebrate and human biology. Snakes are one such model, and they exhibit many extreme phenotypes that can be viewed as innovative evolutionary experiments capable of illuminating key aspects of vertebrate gene function and systems biology (1–5). The evolutionary origin of snakes involved extensive morphological and physiological adaptations, including limb loss, functional loss of one lung, and elongation of the trunk, skeleton, and organs (6). They evolved suites of radical adaptations to consume extremely large prey relative to their body size, including the evolution a kinetic skull and diverse venom proteins (7–9). They also evolved the ability to extensively remodel their organs and physiology on feeding

Significance

The molecular basis of morphological and physiological adaptations in snakes is largely unknown. Here, we study these phenotypes using the genome of the Burmese python (*Python molurus bivittatus*), a model for extreme phenotypic plasticity and metabolic adaptation. We discovered massive rapid changes in gene expression that coordinate major changes in organ size and function after feeding. Many significantly responsive genes are associated with metabolism, development, and mammalian diseases. A striking number of genes experienced positive selection in ancestral snakes. Such genes were related to metabolism, development, lungs, eyes, heart, kidney, and skeletal structure—all highly modified features in snakes. Snake phenotypic novelty seems to be driven by the system-wide coordination of protein adaptation, gene expression, and changes in genome structure.

Author contributions: T.A.C. and D.D.P. designed research; T.A.C., A.P.J.d.K., K.T.H., D.C.C., D.R.S., M.K.F., R.P.R., J.F.D., J.M.D., W.G., J.R.-V., K.J.S., J.M.C., S.E.F., A.W.P., D.P., M.W.V., Q.L., R.K.S., A.K., P.M., P.U., E.N.S., B.S.W.C., F.J.V., N.R.C., C.V.H., M.K.R., S.P.M., M.Y., W.C.W., and S.M.S. performed research; T.A.C., J.D., P.M., R.B., E.N.S., A.M.B., W.C.W., S.M.S., and D.D.P. contributed new reagents/analytic tools; T.A.C., A.P.J.d.K., K.T.H., D.C.C., D.R.S., M.K.F., R.P.R., J.F.D., J.M.D., W.G., J.R.-V., K.J.S., S.E.F., M.W.V., Q.L., R.K.S., A.K., P.M., C.F., D.A.R., F.G.H., B.S.W.C., F.J.V., N.R.C., C.V.H., M.K.R., S.P.M., M.Y., W.C.W., S.M.S., and D.D.P. analyzed data; and T.A.C. and D.D.P. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The sequence reported in this paper has been deposited in the GenBank database (accession no. AEQU00000000).

[1]To whom correspondence should be addressed. E-mail: david.pollock@ucdenver.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1314475110/-/DCSupplemental.

(10, 11), while enduring fluctuations in metabolic rates that are among the most extreme of any vertebrate (11). At the molecular level, previous research has shown that they have undergone an unprecedented degree of evolutionary redesign and accelerated adaptive evolution across multiple mitochondrial proteins (1, 12). We hypothesized that the extreme changes in the mitochondrial proteins were likely to extend to metabolic genes in the nuclear genome. We also hypothesized that other genes and gene networks associated with extreme physiological and phenotypic adaptations in snakes may also have undergone significant evolutionary changes.

To gain insight into the genomic basis of these phenotypes, we sequenced and annotated the genome of a female Burmese python (*Python molurus bivittatus*). The python is the most extreme

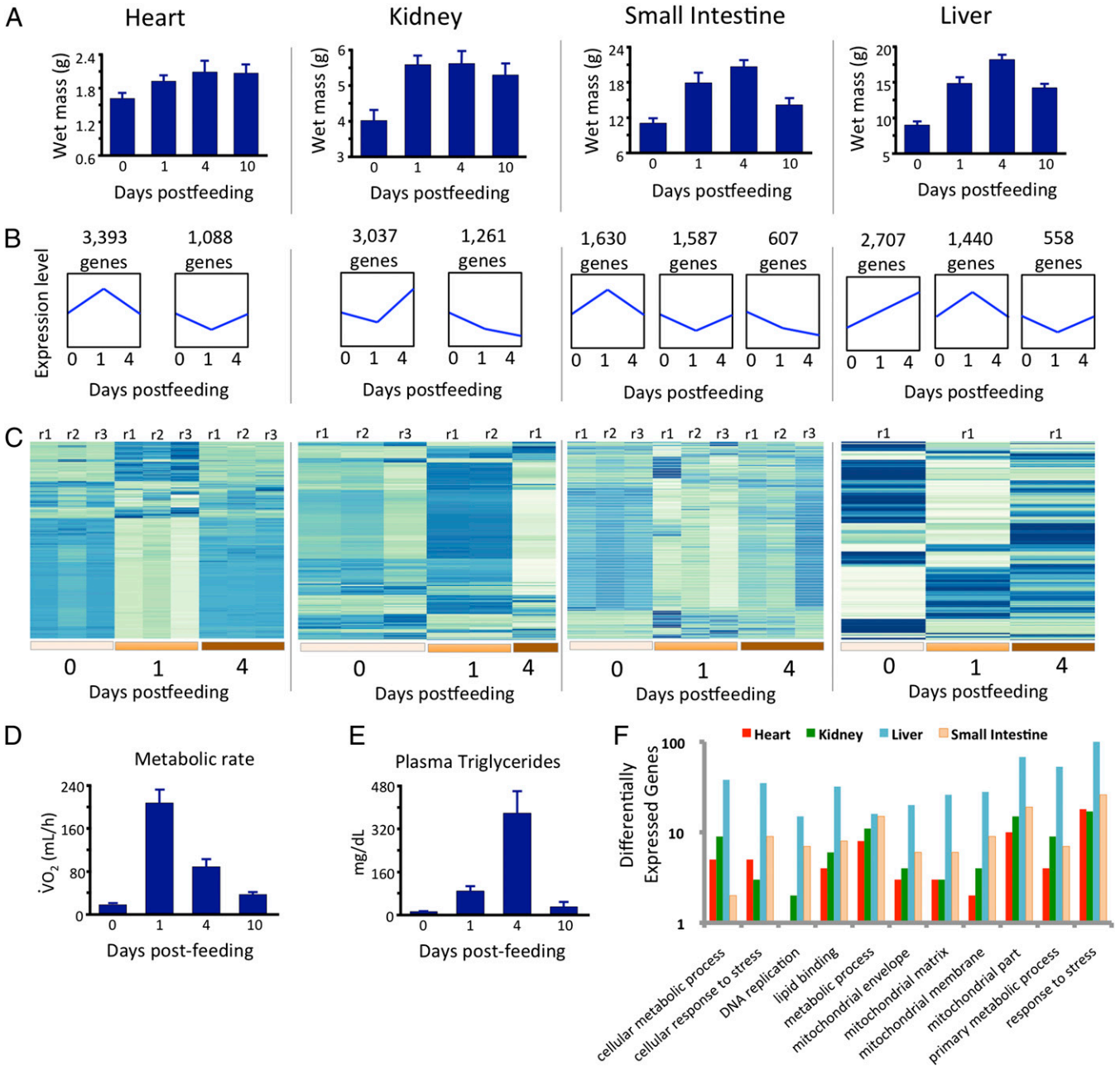

Fig. 1. Coordinated shifts in physiology and gene expression across organs in pythons after feeding. (*A*) Rapid increases in organ mass that accompany feeding in the python. (*B*) Generalized trends in gene expression that are significantly overrepresented in each organ across time points before and after feeding in the python. Results are based on cluster analysis of gene expression profiles and identification of statistically overpopulated profiles. The numbers of genes clustered within each profile are shown above profiles, and trends in the relative magnitude of gene expression for each set of genes are shown in blue. (*C*) Heat maps of normalized gene expression levels for the top 300 significantly differentially expressed genes across time points shown for different tissues and time points before and after feeding. Low expression levels are indicated by pale colors, and high expression levels are indicated by darker blue. Most time points per tissue have replicates that are indicated at the tops of profiles [labeled by replicate (r) number]; every column per tissue indicates a different individual. (*D*) Changes in oxygen consumption indicating changes in oxidative metabolism in fasted and fed pythons. (*E*) Changes in plasma triglyceride levels in fasted and fed pythons. (*F*) Numbers of significantly differentially expressed genes between 0 and 1 DPF in select GO categories in different tissues.

vertebrate model for studying physiological remodeling, rapid organ growth, and metabolic fluctuations after feeding (10, 11). Within 2–3 d after feeding, Burmese pythons can experience 35–100% increases in the mass of major organs, including the heart, liver, small intestine, and kidneys (Fig. 1*A*) (10, 13). On completion of digestion, each of these phenotypes is reversed; physiological functions are down-regulated, and tissues atrophy in a matter of days (Fig. 1*A*) (14).

To complement the genome sequencing results, we also studied transcriptional responses to feeding in four python organs (heart, kidney, liver, and small intestine) for multiple time points. We also evaluated evidence for positive selection in protein coding genes on the lineages leading to the python, cobra, or their common ancestor. In addition, we analyzed changes in functionally important multigene families, such as vomeronasal and olfactory receptors, and opsin genes. Finally, we studied changes in aspects of snake genome structure [repeat content, microsatellite structure, and guanine-cytosine (GC) content] and rates of molecular evolution.

Results and Discussion

The python genome was sequenced using a hybrid approach (combining Illumina and 454 reads), and it is available at National Center for Biotechnology Information: Bioproject PRJNA61234 (GenBank accession no. AEQU00000000). The scaffolded python genome assembly Pmo2.0 is 1.44 Gbp (including gaps), which happened to be the same as the genome size estimated for the related species *P. reticulatus* (1.44 Gbp) (15). This assembly (Pmo2.0) has an N50 contig size of 10.7 kb and a scaffold size of 207.5 kb (*SI Appendix*, Tables S1 and S2). Transcriptomic data were used to annotate this genome to provide robust gene models (*SI Appendix*, Tables S3–S8). For comparative genomic analysis, we analyzed the python genome in conjunction with the genome of the king cobra, *Ophiophagus hannah* (8). The repetitive contents of the python and king cobra genomes estimated using the de novo *P* clouds method (16, 17) were similar (python = 59.4%, cobra = 60.4%). Annotation of readily identifiable repeats using a standard consensus repeat element reference library-based approach (18) also found similar repetitive content in the python (31.8%) and cobra (35.2%) genomes (*SI Appendix*, Table S6). These percentages are only slightly lower than the percentages for humans (~67% for *P* clouds and 45% for library-based methods) (16), although the snake genomes are around one-half as large.

We studied transcriptional responses to feeding in four organs (heart, kidney, liver, and small intestine) for multiple time points before [0 d postfed (DPF)] and after (1 and 4 DPF) feeding (Fig. 1 *A*–*D*). These responses involve thousands of genes and large changes in gene expression that are tightly coordinated with the extreme and rapid changes in organ size and performance after feeding (Fig. 1 and *SI Appendix*, Figs. S1–S17 and Tables S9 and S10). The genes with significant expression responses are functionally diverse and involved in metabolism, chromatin remodeling, growth and development, and human pathologies (Dataset S1). Given the postfeeding increases in oxidative metabolism (Fig. 1*D*), plasma lipid levels (Fig. 1*E*), and organ size (Fig. 1*A*) (10, 11, 14) in the python, we expected genes associated with metabolism, lipids, mitochondria, and DNA replication to be highly regulated between 0 and 1 DPF. As predicted, many differentially expressed genes are associated with these functional groups [based on gene ontology (GO) (19) term analyses], with the one exception being the lack of genes in the heart associated with DNA replication (Fig. 1*F*). This lack is consistent with previous findings that the python heart experiences hypertrophy (cell growth) rather than hyperplasia (cell division) during postfeeding growth (20, 21). To elucidate core shared aspects of this response, we identified 20 genes that are significantly differentially expressed in all four organs between 0 and 1 DPF (Fig. 2*A*). Functions of these genes include chromatin remodeling, mitochondrial function, development, translation, and glycosylation, and there are organ-specific patterns in the direction and magnitude of expression changes, with the heart tending to be the most unique (Fig. 2*B* and Dataset S1).

Based on previous work showing extraordinary selection in snake mitochondrial genomes (1, 22, 23), we hypothesized that many nuclear-encoded metabolic genes might show evidence of positive selection (particularly on the ancestral snake branch), and we were curious if positive selection might partially explain other unique physiological and phenotypic features of snakes. To detect selection on protein coding genes, we assembled orthologous gene alignments for 7,442 genes from the python and cobra along with all other tetrapod species in Ensembl (24). We used branch site codon models to detect genes that experienced positive selection on the lineages leading to the python, cobra, or their common ancestor (Fig. 3). We inferred positive selection in 516 genes on the ancestral snake lineage, 174 genes on the cobra, and 82 genes on the python at a P value < 0.001 (Dataset S2). To link these gene sets to phenotypes, we identified mouse KO phenotypes (25) and GO (19) terms that were statistically enriched on different snake lineages (Dataset S3 and *SI Appendix*, *Supplementary Methods*). A number of functionally enriched categories of positively selected genes in snakes is readily

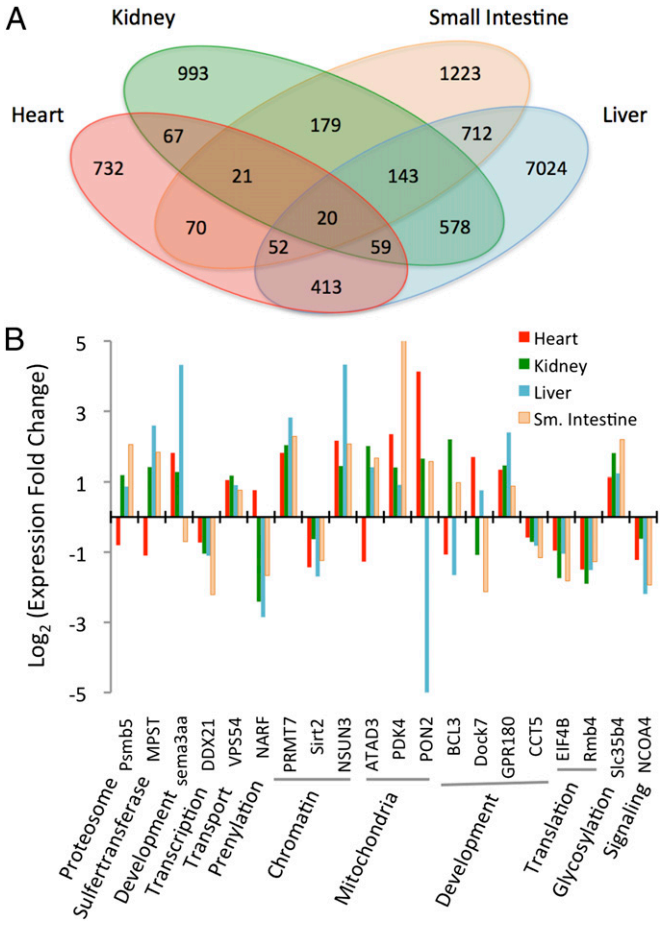

Fig. 2. Differentially expressed genes between fasting and 1 DPF across tissues in the python. (*A*) Numbers of genes significantly differentially expressed between 0 and 1 DPF in four python tissues and the overlap in these gene sets among tissues. (*B*) Fold expression changes between 0 and 1 DPF for 20 genes differentially expressed in all four tissues and broad functional classifications of these genes.

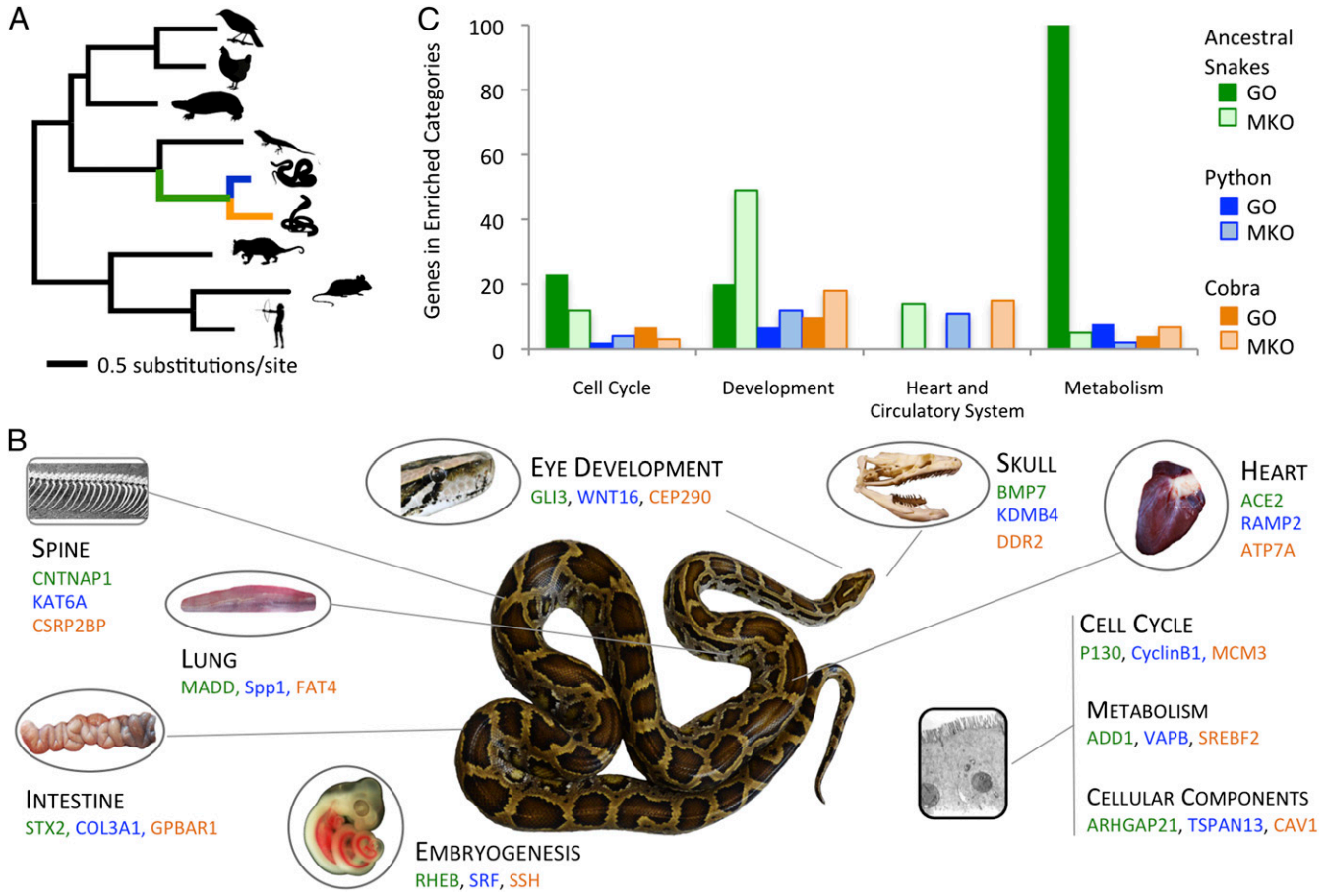

Fig. 3. Functional categories of genes under positive selection related to the extreme biology of snakes. (*A*) Phylogenetic tree of amniotes with branch lengths estimated by maximum likelihood analysis of aligned Ensembl 1:1 orthologs from a subset of taxa analyzed for positive selection. Branches representing snake lineages are colored green for ancestral snake lineage, blue for ancestral python lineage, and orange for ancestral cobra lineage. (*B*) Examples of genes that have experienced positive selection ($P < 0.001$) on snake lineages and are related to prominent phenotypic or cellular traits of snakes (colors correspond to the branches in *A*). Genes are grouped with phenotypic characteristics based on GO and mouse KO phenotype (MKO) terms associated with these genes (Datasets S2 and S3), although no claim is made that genes listed directly explain the snake phenotypes that they are associated with per se or that specific genes shown were selected based on their relative prominence in other literature. (*C*) Functional clusters of GO and MKO terms that are statistically enriched ($P < 0.05$) for positively selected genes ($P < 0.001$). Numbers on the *y* axis represent the combined numbers of genes in clustered enriched categories.

interpretable in light of the unique aspects of snake physiology and morphology (Fig. 3, Dataset S3, and *SI Appendix*, Fig. S17). Genes under positive selection include genes that are functionally related to development, the cardiovascular system, signaling pathways, cell cycle control, and lipid and protein metabolism (Fig. 3). We also find a high level of correspondence between enriched categories of differentially expressed genes involved in organ remodeling (Figs. 1 and 2) and genes that have experienced positive selection in snakes, including genes involved in the cell cycle, development, the heart and circulatory system, and metabolism (Fig. 3*B*).

The ancestral snake lineage shows significant enrichment of positively selected genes in GO and mouse knock out (KO) phenotype categories related to metabolism, eye structure, spine and skull shape, and embryonic patterning mechanisms contributing to somite formation and left–right asymmetry (Fig. 3 and Dataset S3). The high number of positively selected nuclear genes associated with metabolism on this lineage corresponds well with previous studies indicating substantial mitochondrial protein selection also occurring early in snake evolution (1, 12). Other enriched categories correspond well with the major phenotypic shifts, including limb loss, trunk elongation and skeletal changes, associated with the shift to a fossorial lifestyle. On the cobra lineage, we found enrichment of categories related to heart, lung, and neuronal development (Fig. 3 and Dataset S3). This lineage includes the ancestor of colubroid snakes, many of which (like the cobra) are highly active foragers, and categories of positively selected genes may be related to this shift in natural history. The python lineage showed enrichment in categories regulating heart and blood vessel development, hematopoiesis, and cell cycle regulation, which correspond well with the python's extreme postfeeding response, and potentially, the use of constriction to subdue prey in members of this lineage. These enriched categories in the python include the angiogenic PDGF pathway, signaling through the cytoskeletal regulator ρ-GTPase, the TGF-β/bone morphogenetic protein signaling pathway, and categories associated with bone strength and growth (Dataset S3). Many of the positively selected genes also have prominent medical significance. For example, angiotensin 1 converting enzyme and endothelin 1 are important therapeutic targets for cardiovascular disease (26, 27). Similarly, GRB2-associated binding protein 1, which integrates receptor tyrosine kinase signaling, is known to contribute to breast cancer, melanomas, and childhood leukemia (28).

The extreme phenotypes that characterize snakes can also be linked to changes in multigene families. A prominent feature of

snakes is a long forked tongue used to enhance chemoreception. Genes encoding vomeronasal receptors, olfactory receptors, and ephrin-like receptors all show major expansions in the ancestral snake lineage as well as the cobra and python, indicating that expansions in these gene families may also contribute to enhanced chemoreception in snakes (*SI Appendix*, Figs. S18–S20). It is hypothesized that ancestral snakes were fossorial (6), which reduced selection for light perception. Supporting this fossorial snake ancestor hypothesis, we found that 10 visual and nonvisual opsin genes were lost in snakes but are otherwise present in squamates, including *RH2*, *SWS2*, *PIN*, *PPIN*, *PARIE*, *MEL1*, *NEUR2*, *NEUR3*, *TMT2*, and *TMTa* (*SI Appendix*, Fig. S21 and Table S11). The absence of these genes was verified using tBlastx searches for these opsins in the cobra and python annotated gene sets as well as genome assemblies and cDNA assemblies.

As with the mitochondrial genome (1, 22, 23), snake nuclear genomes have evolved substantial changes in structure and patterns of molecular evolution relative to other vertebrate genomes. To compare repeat element content among partially sequenced and fully assembled snake genomes, we used a single multispecies combined repeat element library for analysis of repeat content (*SI Appendix*, *Supplementary Methods*). Despite low variance in snake genome size (*SI Appendix*, Fig. S21), repeat library-based annotations of genome samples from 10 additional snake species show surprisingly high variance in genomic repeat content but a relatively constant diversity of repeat types (Fig. 4 and *SI Appendix*, Tables S5–S8) (29). Exceptions are the peculiar families of snake1 (L3) CR1 LINEs that tend to contain microsatellite repeats at their 3′ end (30), which we find to have expanded almost exclusively in colubroid snakes, such as the King Cobra, the Western Diamondback Rattlesnake and Copperhead, and the Garter snake (Fig. 4*B*). An unexpected result from the analysis of the *Anolis* lizard genome was that it lacked the GC isochore structure present in mammalian and avian genomes (31, 32). The GC isochore structures of snakes, however, are intermediate between the lack of isochores in *Anolis* and the clear isochore structure in turtles, birds, and mammals (Fig. 4). These differences in isochore structure raise the intriguing possibility that snakes reevolved GC isochore structure or that the *Anolis* (or an ancestral squamate) lineage lost GC isochore structure. Trends in GC content at third codon position across amniotes indicate a shift to higher AT content in snakes and based on equilibrium GC content at third codon position calculations, a continued erosion of GC content in king cobra and an increase of GC content in python (*SI Appendix*, Fig. S24). We also used whole-genome pairwise alignments among *Anolis*, python, and king cobra to examine how GC content varies among squamates, and we found that both snake species had similar GC profiles to each other but lower GC content than lizard when comparing across aligned genomic sites (*SI Appendix*, Fig. S25). This inference of dynamic GC content evolution is consistent with a shift in isochore structure in reptile genomes.

Rates of molecular evolution also differ substantially across reptile lineages (33). Although turtle genes evolve slowly compared with other sequenced amniotes (34), we find snake genes have evolved rapidly compared to other amniotes (Fig. 3*A*). Based on a subset of 10,000 aligned codons sampled from orthologous gene alignments, snake lineages experienced relatively high rates of evolution compared with the turtle and other amniote lineages (*SI Appendix*, Fig. S26). Analysis of the full set of 62,817 fourfold degenerate third codon positions from the orthologous gene set indicates that snake neutral substitution rates are also accelerated relative to other reptilian lineages (*SI Appendix*, Fig. S27). Additional analyses of 44 nuclear genes for >150 squamate reptile species (from a previous phylogenetic study) (35) indicate accelerated neutral evolution in the ancestral lineages of squamate reptiles, snakes, and colubroid snakes (*SI Appendix*, Fig. S28).

The comparative systems genomics approach that we have taken to study snake genomes has provided hundreds of candidate genes to study the process by which genes and gene networks coevolve to produce phenotypic diversity in vertebrates. The degree to which the physiological, morphological, and metabolic changes in the snakes coincide with molecular changes is remarkable. It has been hypothesized that major morphological changes are primarily driven by changes in gene expression (36). Snakes provide an alternative vertebrate model system, in which the extensive system-wide evolutionary coordination of protein adaptation, gene expression, and changes in the structure and organization of the genome itself seem to have driven phenotypic novelty. Although there are examples of such types of changes in other vertebrates (37), we expect that the genomic changes seen in snakes and documented here are exceptional in their number and magnitude. There have been sufficient vertebrate

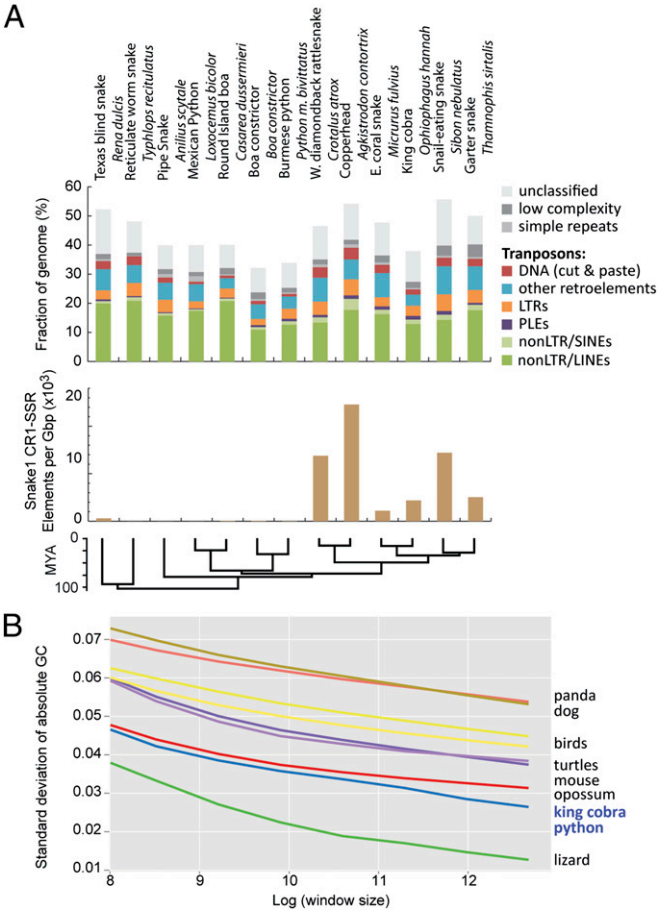

Fig. 4. Variation and uniqueness of snake genome content and structure. (*A*) Amounts and types of readily identified repeat elements in snake complete and sampled genomes. Estimates in *Top* and *Middle* show abundance of genomic repeat elements across 10 snake species based on sampled genome sequences, except for the python and cobra, which are based on complete genomes. *Bottom* shows genomic density of snake1 CR1 LINE elements for subfamilies that tend to contain microsatellite repeats at their 3′ tails in snake genomes and genome samples. (*B*) Evidence for shifts in genomic GC isochore structure in squamate reptiles. The *y* axis is the standard variation of GC content when examining the genome at nonoverlapping window sizes (from 5- to 320-kb windows log-transformed on the *x* axis). Larger values indicate greater among-window GC content heterogeneity. For example, at all spatial scales, mammals have the greatest GC heterogeneity, and squamate reptiles have the least GC heterogeneity. The right side of the graph (GC SD at a window size of 320 kb) shows GC heterogeneity at a spatial scale on the order of isochore structure; the low GC heterogeneity of squamate reptiles indicates a reduced representation of GC-rich isochores compared with the other taxa. LTR, long terminal repeat; PLE, Penelope-like element; SINE, short interspersed nuclear element.

mitochondrial genomes sampled to make a convincing case that the adaptive changes observed in snake mitochondrial genomes are truly exceptional (1, 22, 23), and among the 60+ currently published vertebrate nuclear genomes, the degree of change observed in snakes is exceptional as well. The python genome, together with the genome of the king cobra (8), will accelerate understanding of the genomic features that underlie the phenotypic uniqueness of snakes.

Materials and Methods

Burmese Python Genome Sequencing. All animal procedures were conducted with registered Institutional Animal Care and Use Committee (University of Texas, Arlington, TX) protocols. Python genome sequencing libraries were constructed from a single *P. molurus bivittatus* female obtained commercially. We created and sequenced multiple whole-genome shotgun libraries on an Illumina Genome Analyzer IIx, an Illumina HiSeq 2000, and a 454 FLX sequencer. In total, we generated 73.8 Gbp (~49× genome coverage) for python genome assembly and scaffolding, which included data generated for a previous draft assembly (38). The genome was assembled using *SOAPdenovo* (39) and *Newbler* (Roche), and alternate assemblies were merged (40). Details are provided in *SI Appendix, Supplementary Methods*.

Annotation and Gene Prediction. The genome assembly was annotated using the MAKER annotation pipeline (41). All available RNAseq data from multiple organs and fasted/fed time points were used in developing gene model predictions, and repeat element libraries from both complete snake genomes and the snake genome sampling were used to annotate repeat elements in the python (*SI Appendix, Supplementary Methods* and *SI Appendix*, Tables S5–S10). The annotated genome assembly is available under the National Center for Biotechnology Information Bioproject PRJNA61234 (GenBank accession no. AEQU00000000).

Comparative and Evolutionary Analyses. Ortholog sets were assembled by addition of our python and cobra gene coding DNA sequences (CDSs) sets to the Ensembl Compara v70 1:1 vertebrate ortholog set (24, 42). Analyses included filtering likely gene duplicates in snakes, quality control steps, and alignment. Alignments were analyzed using a maximum likelihood branch site test for sites that were uniquely positively selected on the python, cobra, or ancestral snake lineages (*SI Appendix, Supplementary Methods*).

ACKNOWLEDGMENTS. We thank Carl Franklin for donation of the specimen used for genome sequencing. We thank Roche 454 for contributing 454 sequencing used in python genome assembly, Roger Winer for running *Newbler* assemblies, WestGrid and Compute/Calcul Canada (A.P.J.d.K.) for providing extra computing resources, and two anonymous reviews for constructive comments. The project was supported by setup funds from the University of Texas (to T.A.C.) and the University of Colorado School of Medicine (to D.D.P.); Roche-454 Proof of Concept grants (to T.A.C. and D.D.P.); National Science Foundation Grants IOS 0922528 (to A.M.B.) and IOS 0466139 (to S.M.S.); and National Institutes of Health Grants R01GM083127 (to D.D.P.) and R01GM097251 (to D.D.P.).

1. Castoe TA, Jiang ZJ, Gu W, Wang ZO, Pollock DD (2008) Adaptive evolution and functional redesign of core metabolic proteins in snakes. *PLoS One* 3(5):e2201.
2. Gomez C, et al. (2008) Control of segment number in vertebrate embryos. *Nature* 454(7202):335–339.
3. Cohn MJ, Tickle C (1999) Developmental basis of limblessness and axial patterning in snakes. *Nature* 399(6735):474–479.
4. Di-Poï N, et al. (2010) Changes in Hox genes' structure and function during the evolution of the squamate body plan. *Nature* 464(7285):99–103.
5. Vonk FJ, Richardson MK (2008) Developmental biology: Serpent clocks tick faster. *Nature* 454(7202):282–283.
6. Vidal N, Hedges SB (2004) Molecular evidence for a terrestrial origin of snakes. *Proc Biol Sci* 271(Suppl 4):S226–S229.
7. Casewell NR, Huttley GA, Wüster W (2012) Dynamic evolution of venom proteins in squamate reptiles. *Nat Commun* 3:1066.
8. Vonk FJ, et al. (2013) The king cobra genome reveals dynamic gene evolution and adaptation in the snake venom system. *Proc Natl Acad Sci USA* 110:20651–20656.
9. Mackessy SP (2002) Biochemistry and pharmacology of colubrid snake venoms. *J Toxicol* 21(1-2):43–83.
10. Secor SM, Diamond J (1995) Adaptive responses to feeding in Burmese pythons: Pay before pumping. *J Exp Biol* 198(Pt 6):1313–1325.
11. Secor SM, Diamond J (1998) A vertebrate model of extreme physiological regulation. *Nature* 395(6703):659–662.
12. Castoe TA, et al. (2009) Evidence for an ancient adaptive episode of convergent molecular evolution. *Proc Natl Acad Sci USA* 106(22):8986–8991.
13. Cox CL, Secor SM (2008) Matched regulation of gastrointestinal performance in the Burmese python, *Python molurus*. *J Exp Biol* 211(Pt 7):1131–1140.
14. Secor SM (2008) Digestive physiology of the Burmese python: Broad regulation of integrated performance. *J Exp Biol* 211(Pt 24):3767–3774.
15. De Smet WHO (1981) The nuclear Feulgen-DNA content of the vertebrates (especially reptiles), as measured by fluorescence cytophotometry, with notes on the cell and chromosome size. *Acta Zool et Pathologica Antverpiensia* 76(1):119–167.
16. de Koning APJ, Gu W, Castoe TA, Batzer MA, Pollock DD (2011) Repetitive elements may comprise over two-thirds of the human genome. *PLoS Genet* 7(12):e1002384.
17. Gu W, Castoe TA, Hedges DJ, Batzer MA, Pollock DD (2008) Identification of repeat structure in large genomes using repeat probability clouds. *Anal Biochem* 380(1):77–83.
18. Smit AFA, Hubley R, Green P (2004) *RepeatMasker Open-3.0*. Available at http://www.repeatmasker.org. Accessed December 1, 2012.
19. Ashburner M, et al. (2000) Gene ontology: Tool for the unification of biology. *Nat Genet* 25(1):25–29.
20. Andersen JB, Rourke BC, Caiozzo VJ, Bennett AF, Hicks JW (2005) Physiology: Postprandial cardiac hypertrophy in pythons. *Nature* 434(7029):37–38.
21. Riquelme CA, et al. (2011) Fatty acids identified in the Burmese python promote beneficial cardiac growth. *Science* 334(6055):528–531.
22. Jiang ZJ, et al. (2007) Comparative mitochondrial genomics of snakes: Extraordinary substitution rate dynamics and functionality of the duplicate control region. *BMC Evol Biol* 7:123.
23. Castoe TA, et al. (2009) Dynamic nucleotide mutation gradients and control region usage in squamate reptile mitochondrial genomes. *Cytogenet Genome Res* 127(2–4):112–127.
24. Flicek P, et al. (2013) Ensembl 2013. *Nucleic Acids Res* 41(Database Issue):D48–D55.
25. Eppig JT, et al. (2012) The Mouse Genome Database (MGD): Comprehensive resource for genetics and genomics of the laboratory mouse. *Nucleic Acids Res* 40(Database Issue):D881–D886.
26. Lang CC, Struthers AD (2013) Targeting the renin-angiotensin-aldosterone system in heart failure. *Nat Rev Cardiol* 10(3):125–134.
27. Kaoukis A, et al. (2013) The role of endothelin system in cardiovascular disease and the potential therapeutic perspectives of its inhibition. *Curr Top Med Chem* 13(2):95–114.
28. Ortiz-Padilla C, et al. (2013) Functional characterization of cancer-associated Gab1 mutations. *Oncogene* 32(21):2696–2702.
29. Shedlock AM, et al. (2007) Phylogenomics of nonavian reptiles and the structure of the ancestral amniote genome. *Proc Natl Acad Sci USA* 104(8):2767–2772.
30. Castoe TA, et al. (2011) Discovery of highly divergent repeat landscapes in snake genomes using high-throughput sequencing. *Genome Biol Evol* 3:641–653.
31. Fujita MK, Edwards SV, Ponting CP (2011) The Anolis lizard genome: An amniote genome without isochores. *Genome Biol Evol* 3:974–984.
32. Alföldi J, et al. (2011) The genome of the green anole lizard and a comparative analysis with birds and mammals. *Nature* 477(7366):587–591.
33. Tzika AC, Helaers R, Schramm G, Milinkovitch MC (2011) Reptilian-transcriptome v1.0, a glimpse in the brain transcriptome of five divergent Sauropsida lineages and the phylogenetic position of turtles. *Evodevo* 2(1):19.
34. Shaffer HB, et al. (2013) The western painted turtle genome, a model for the evolution of extreme physiological adaptations in a slowly evolving lineage. *Genome Biol* 14(3):R28.
35. Wiens JJ, et al. (2012) Resolving the phylogeny of lizards and snakes (Squamata) with extensive sampling of genes and species. *Biol Lett* 8(6):1043–1046.
36. Carroll SB (2005) Evolution at two levels: On genes and form. *PLoS Biol* 3(7):e245.
37. Wan QH, et al. (2013) Genome analysis and signature discovery for diving and sensory properties of the endangered Chinese alligator. *Cell Res* 23(9):1091–1105.
38. Castoe TA, et al. (2011) Sequencing the genome of the Burmese python (*Python molurus bivittatus*) as a model for studying extreme adaptations in snakes. *Genome Biol* 12(7):406.
39. Li R, et al. (2010) *De novo* assembly of human genomes with massively parallel short read sequencing. *Genome Res* 20(2):265–272.
40. Yao G, et al. (2012) Graph accordance of next-generation sequence assemblies. *Bioinformatics* 28(1):13–16.
41. Cantarel BL, et al. (2008) MAKER: An easy-to-use annotation pipeline designed for emerging model organism genomes. *Genome Res* 18(1):188–196.
42. Vilella AJ, et al. (2009) EnsemblCompara GeneTrees: Complete, duplication-aware phylogenetic trees in vertebrates. *Genome Res* 19(2):327–335.