PNAS RESEARCH ARTICLE | GENETICS

The genetic regulatory architecture and epigenomic basis for age-related changes in rattlesnake venom

Michael P. Hogan[a], Matthew L. Holding[a,b], Gunnar S. Nystrom[a], Timothy J. Colston[a,c], Daniel A. Bartlett[a], Andrew J. Mason[d,e], Schyler A. Ellsworth[a], Rhett M. Rautsaw[d,f,g], Kylie C. Lawrence[a], Jason L. Strickland[d,h], Bing He[a], Peter Fraser[a], Mark J. Margres[f], David M. Gilbert[i], H. Lisle Gibbs[e], Christopher L. Parkinson[d,j], and Darin R. Rokyta[a,1]

Edited by Nils Stenseth, Universitetet i Oslo, Oslo, Norway; received August 8, 2023; accepted March 13, 2024

Developmental phenotypic changes can evolve under selection imposed by age- and size-related ecological differences. Many of these changes occur through programmed alterations to gene expression patterns, but the molecular mechanisms and gene-regulatory networks underlying these adaptive changes remain poorly understood. Many venomous snakes, including the eastern diamondback rattlesnake (*Crotalus adamanteus*), undergo correlated changes in diet and venom expression as snakes grow larger with age, providing models for identifying mechanisms of timed expression changes that underlie adaptive life history traits. By combining a highly contiguous, chromosome-level genome assembly with measures of expression, chromatin accessibility, and histone modifications, we identified cis-regulatory elements and trans-regulatory factors controlling venom ontogeny in the venom glands of *C. adamanteus*. Ontogenetic expression changes were significantly correlated with epigenomic changes within genes, immediately adjacent to genes (e.g., promoters), and more distant from genes (e.g., enhancers). We identified 37 candidate transcription factors (TFs), with the vast majority being up-regulated in adults. The ontogenetic change is largely driven by an increase in the expression of TFs associated with growth signaling, transcriptional activation, and circadian rhythm/biological timing systems in adults with corresponding epigenomic changes near the differentially expressed venom genes. However, both expression activation and repression contributed to the composition of both adult and juvenile venoms, demonstrating the complexity and potential evolvability of gene regulation for this trait. Overall, given that age-based trait variation is common across the tree of life, we provide a framework for understanding gene-regulatory-network-driven life-history evolution more broadly.

venom | ontogeny | epigenomics | gene regulatory network | Crotalus

Significance

Many animals experience different ecologies as they mature, leading natural selection to favor different characteristics in adults and juveniles. Eastern diamondback rattlesnakes show major dietary shifts in adults as they become large enough to consume larger prey species, and their venom changes to accommodate this new diet. To determine the molecular basis for this change, we compared expression and epigenomic patterns between adult and juvenile venom glands and found co-option of regulatory systems involved in growth signaling, transcriptional activation, and circadian rhythm/biological timing. By connecting fitness-related life-history changes to their underlying mechanisms, our results will allow partitioning of selection among age-classified coding and regulatory sequences to determine the relative strengths of selection on venom in adults versus juveniles.

Complex phenotypes are produced and maintained through precise temporal, spatial, and quantitative regulation of the expression of the genes in the regulatory networks underlying them (1, 2). These expression patterns and the functional properties of the encoded proteins dictate observable phenotypes, and heritable variation in any of these characteristics can contribute to adaptive evolution. For nearly 20 y (3, 4), studies of the genetics of adaptation have focused on the relative evolutionary contributions of mutations changing protein-coding sequences directly and mutations influencing the expression of proteins by changing the underlying regulatory machinery. Both types of genetic variation contribute to adaptation (5, 6), but changes to gene-expression patterns have, in general, been found to explain a disproportionately higher amount of adaptive phenotypic variation (7–10). As our understanding of gene regulation has matured and technological advancements have improved the resolution with which we can characterize the organization and function of gene regulatory networks (11–13), focus is shifting from determining whether regulatory evolution is important to deciphering the genetic mechanisms, mutational processes, and adaptive pathways that make regulatory evolution so prevalent.

Temporal and spatial control of gene expression is determined by interactions among the genome, the epigenome, and the environment. In terms of heritable variation, transcription factors (TFs), their expression patterns, and the cis-regulatory elements (CREs) to which they bind, such as promoters and enhancers, are critical to setting expression patterns (14). These patterns, however, can be further modulated, as well as revealed, by epigenomic characteristics such as chromatin states, DNA alterations, and histone modifications (15). Whether one considers expression changes within an individual over its lifetime (16) or expression divergence among independently evolving lineages (17), these same underlying genetic factors are involved.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence may be addressed. Email: drokyta@bio.fsu.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2313440121/-/DCSupplemental.

Published April 5, 2024.

Understanding how and why expression evolution contributes to adaptation and speciation (18) requires uncovering the prevalent mechanisms contributing to expression changes that have evolved under selection resulting from the ecology of an organism. Ontogenetic expression changes are particularly suited to revealing mechanisms of expression regulation and evolution; control is accomplished using the same sets of alleles in the same genomes, allowing for direct observation of how expression is changed within a single genetic context. Just as regulatory control occurs in the same genetic background, selection over the lifetime of an individual is constrained to act on the same genome, despite specific alleles potentially having different fitness effects for different age classes (19). Therefore, ontogenetic phenotypic variation is limited to age-specific variation in gene expression (16), and characterizing this variation offers a unique opportunity to reveal the genetic mechanisms through which selection can tune expression to accommodate the distinct ecological niches experienced across an organism's lifespan.

Venom is essential for prey acquisition in venomous snakes, and gene expression within the venom glands dictates the venom phenotype (20). Venom composition can change dramatically as snakes mature, often with functional and medical consequences (21, 22). Although the tissue-specific transcriptional regulation of venom production (23–26) and ontogenetic venom shifts in rattlesnakes have been particularly well studied (27–31), the roles of individual TFs and CREs dictating age-specific expression of particular venom-gene paralogs remain unknown.

The eastern diamondback rattlesnake (*Crotalus adamanteus*), native to the southeastern United States, is the largest venomous snake in North America. As *C. adamanteus* grows, its prey preference shifts from small mice, to cotton rats, and eventually to rabbits (31, 32). Shifts in the main prey items are accompanied by major changes in venom composition (30, 31). In turn, functional toxicity of venom differs between juvenile and adult snakes, with juvenile venom being more potent to mice than adult venom (30), suggesting the ontogenetic shifts in venom composition are adaptive. The ontogenetic transition in venom composition is responsible for more proteomic differences than population-level geographic variation for this species (29), suggesting that selection along the life-history axis may be stronger than any geographically varying selective pressures on venoms. Furthermore, several notable cases of apparent venom paedomorphosis, whereby adults in one species retain juvenile characteristics from another, have been described in other rattlesnake species (33–35), suggesting that ontogenetic differences can be converted, or at least recapitulated, as species-level adaptive divergence.

To investigate the mechanisms that give rise to age-related shifts in rattlesnake venom composition, we generated a chromosome-level, highly contiguous genome assembly for *C. adamanteus* and compared venom gene expression, chromatin accessibility, and histone modifications related to active promoters and enhancers between adults and juveniles (*SI Appendix*, Figs. S1 and S2). To measure chromatin accessibility, we used Assay for Transposase-Accessible Chromatin using sequencing (ATAC-seq); greater than 75% of chromatin accessible regions represent enhancers and promoters (36). Previous work in rattlesnakes (26) showed that venom-gene regions have venom-gland-specific accessibility. To complement ATAC-seq, we also used Cleavage Under Targets and Release Using Nuclease (CUT&RUN) (37) to identify regions associated with modified histones, specifically targeting Histone 3 Lysine 27 acetylation (H3K27ac), a mark of active enhancers and promoters (38–40). Both ATAC-seq and CUT&RUN provide valuable insights into the chromatin landscape and the regulation of gene expression, but they detect different, but presumably correlated, signals.

We used two independent techniques alongside RNA-seq with identical statistical approaches to simultaneously identify changes in gene expression and epigenomic patterns with high confidence. With these data, we performed differential expression (DE) and differential accessibility/modification (DA) analyses and found genomic overlap between age-biased venom gene expression and age-biased venom gene accessibility and histone modifications. Guided by differentially expressed venom genes and TFs, we predicted high probability regulators and compared distributions of their DNA binding motifs between differentially accessible/modified venom regions to classify positive and negative regulators. We further evaluated the epigenomic relationship of putative ontogenetic regulators among all expressed TFs in the venom gland using protein-interaction-network map inferences. To confirm our results, we used comparative genomics to test our hypothesis that particular CREs control venom ontogeny, predicting that the loss of identified cis-regulatory elements would ablate differential expression in the western diamondback rattlesnake (*Crotalus atrox*), contributing to its substantially reduced ontogenetic venom shift.

Results

Complete Resolution of Venom Genomic Architecture. To determine the mechanisms responsible for the ontogenetic venom shift in *C. adamanteus* (30), we first generated a complete, chromosome-level genome assembly based on PacBio HiFi reads (*SI Appendix*, Table S1) and Hi-C data (*SI Appendix*, Table S2). Our final assembly included 381 contigs with a prescaffold *N*50 of 67.5 Mb (*SI Appendix*, Table S3). Hi-C scaffolding resulted in 44 scaffolds with an *N*50 of 208.9 Mb (Fig. 1*A*). Of these 44 scaffolds, 17 were determined to be contaminants of bacterial origin and were removed. The final consensus genome assembly consisted of 366 contigs in 27 scaffolds with a final genome size of 1.7 Gb. We identified scaffolds representing assemblies of all 17 autosomes and the Z sex chromosome (Fig. 1). Because our genome animal was female (*SI Appendix*, Table S4), we were also able to identify the complete W chromosome based on male–female read (*SI Appendix*, Table S5) ratio mapping as previously described (41). In addition to the complete mitochondrial genome, only seven small scaffolds (2,606 to 635,894 bp) remained unassigned to known chromosomes. We compared our chromosome recovery and identity with the male *Crotalus viridis* genome (23) and found strong overall agreement across all autosomes and the Z chromosome, with no loci matching to our W chromosome (Fig. 1*B*), as expected. We recovered both telomeres for chromosomes 2, 3, 5, 9, 10, 13, 16, and Z; we also recovered one telomere for chromosomes 6, 7, 11, 12, 14, 15, and 17. Putative centromeres were inferred for all chromosomes as the longest nontelomeric, high-order repeat regions (*SI Appendix*, Figs. S3–S7).

We generated the first complete, paralog-resolved characterization of the venom genes of a snake genome, a prerequisite for studying the intricacies of venom-expression changes. We identified 21,841 genes in total (*SI Appendix*, Fig. S8), 134 of which were venom genes, and 1,566 of which were TFs. We found 124 of the 134 venom genes to occur in tandem arrays of duplicated genes (Fig. 1*C*), a pattern previously noted, but not fully resolved, for other venomous snake genomes (23, 26, 42, 43). We assembled all venom-gene arrays within single contigs in our primary assembly (Fig. 1*C*) except for the C-type lectin (CTL) array on chromosome 6, which had one

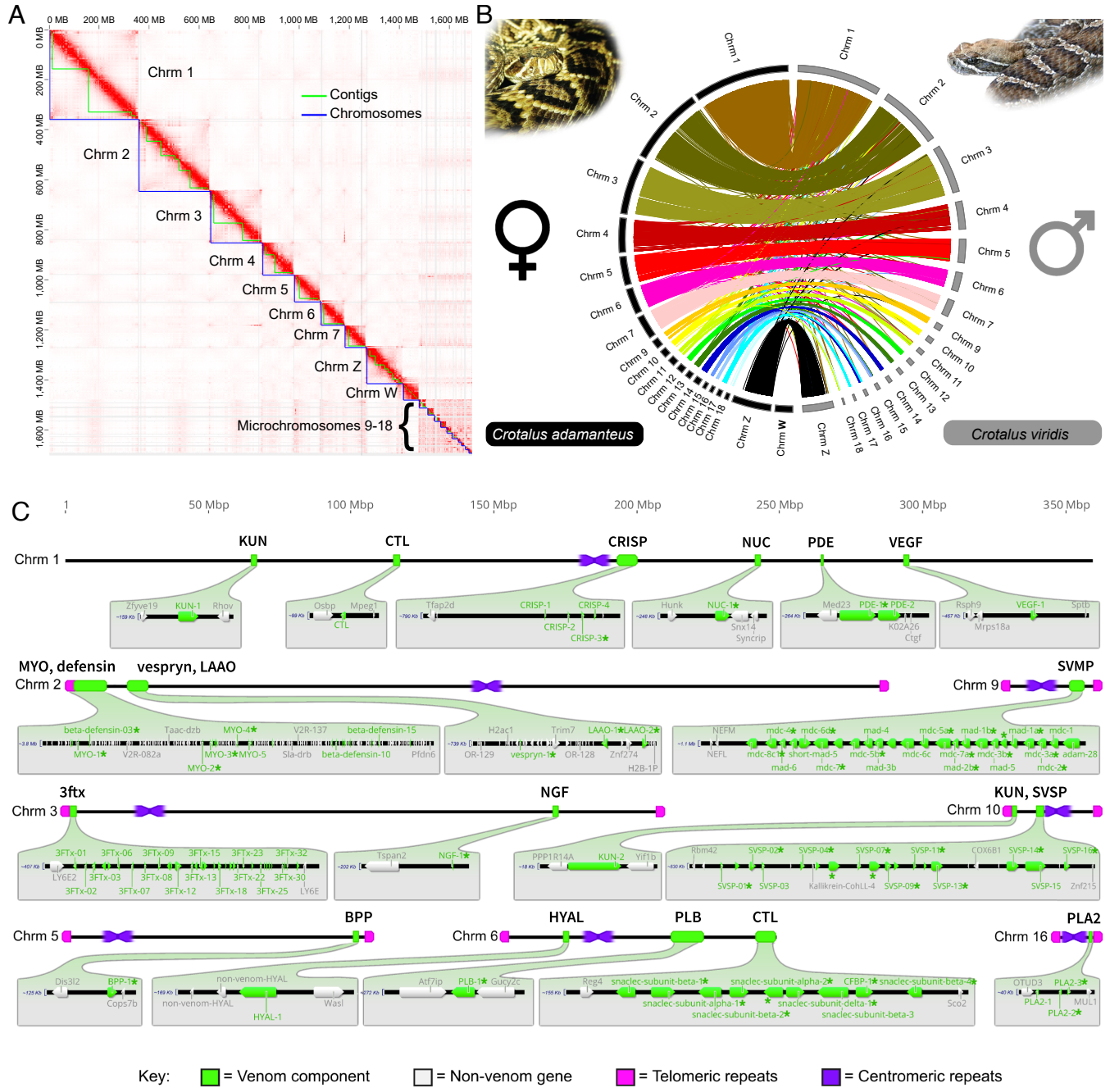

Fig. 1. Reference genome for *C. adamanteus* and the organization of venom genes. The final assembly consisted of 366 contigs and 27 scaffolds with a contig *N*50 of 67.533 Mb and a scaffold *N*50 of 208.942 Mb. (*A*) The highly contiguous primary assembly required little scaffolding to assemble all 17 autosomes, as well as both sex chromosomes. (*B*) We found broad agreement with the previously published genome for *C. viridis* (23) on the basis of locations of BUSCO loci. Lines connect BUSCO loci detected in both genomes. (*C*) Most chromosomes were assembled from telomere-to-telomere. All venom-gene arrays were assembled within single contigs in our primary assembly except for the C-type lectin (CTL) array on chromosome 6, which had one break. The CTL array structure was confirmed by comparison with other assemblies, where the break was fully resolved (*SI Appendix*, Table S3 and Fig. S9*D*). Blue numbers on the left end of the arrays indicate approximate sizes of the expanded region for each venom-gene array. Venom genes expressed at high levels (average TPM > 500) are indicated with an asterisk. Abbreviations: 3FTx—Three-finger toxin, BPP—Bradykinin-potentiating peptide, Chrm—Chromosome, CRISP—Cysteine-rich secretory protein, CTL—C-type lectin, HYAL—Hyaluronidase, KUN—Kunitz-type protease inhibitor, LAAO—L-amino acid oxidase, MYO—Myotoxin, NGF-Nerve growth factor, NUC—Nucleotidase, PDE—Phosphodiesterase, PLA2—Phospholipase A2, PLB—Phospholipase B, SVMP—Snake venom metalloproteinase, SVSP—Snake venom serine protease, VEGF—Vascular endothelial growth factor.

Hi-C contig splice between venom paralogs. The CTL array structure was confirmed by comparison with other assemblies of these data generated with Canu (44), Peregrine (45), and alternative versions of hifiasm (*SI Appendix*, Table S3), where the break was fully resolved (*SI Appendix*, Fig. S9*D*). We confirmed that the scaffolded CTL array did not interrupt any putative exons and proceeded with using the scaffolded hifiasm assembly for analyses. Our venom gene recovery mirrored previous genetic and proteomic characterizations for *C. adamanteus* (20), emphasizing paralogous gene families including 10 CTLs, 23 snake venom metalloproteinases (SVMPs), 16 snake venom serine proteinase (SVSPs), four cysteine-rich secretory proteins (CRISPs), three

phospholipase A2s (PLA2s), and five myotoxins (MYOs). We also detected unexpectedly high paralogy for three-finger toxins (3FTxs), with 32 putatively functional copies, despite only five paralogs being expressed in the venom glands (Dataset S1).

Identification of Differentially Expressed Genes across Age Classes. We identified age-biased differentially expressed (DE) venom and TF genes in the venom gland by comparing gene expression (*SI Appendix*, Table S6) between adults ($n = 8$) and juveniles ($n = 10$). We classified animals with a snout–vent length (SVL) greater than 1 m as adults and those less as juveniles (46). In total, we identified 24 out of 75 expressed venom-coding genes that were DE, with 18 biased toward adult expression (i.e., expressed higher in adults or lower in juveniles) and six biased toward juvenile expression (Fig. 2). The adult-biased genes include members of seven different toxin families, and the juvenile-biased genes included five SVMPs and a single CTL. The largest subadult (SVL = 880 mm) had an intermediate venom (*SI Appendix*, Fig. S9*A*), consistent with the gradual transition previously described for *C. adamanteus* (31). The largest class-level expression differences for venom genes based on total relative expression (Fig. 2*B*) correspond to toxin classes with differentially expressed tandem paralogs (e.g., CTLs, PLA2s, SVSPs, and SVMPs). One exception is a bradykinin-potentiating peptide (BPP), where a single gene on chromosome 5 is dramatically up-regulated in adult *C. adamanteus*, switching from a 1% average venom transcript contribution in juveniles to a 16% average contribution in adults. The overall proportion of expressed venom genes that were DE (32%) was significantly higher ($\chi^2 = 155.6$; $P < 0.00001$) than the proportion of nonvenom, non-TF genes that were DE (3.8%) in the venom glands (Fig. 2*D*).

We identified 1,108 putative TF genes expressed in *C. adamanteus* venom glands, 37 of which were DE with a $\log_2$-fold change > 1 and FDR < 0.05, including 26 biased toward adults and 11 biased toward juveniles (Fig. 2). Of the significant TFs identified from our DE analysis, 30 represent previously unrecognized venom regulators, including Fam13c, Irf8, NFIL3, Six4, Etv1, Bhlhe40, and Klf15. The remaining seven DE TFs match top candidate venom regulators identified from the *C. viridis* genome (25), including Nr4a1, Fos, Ddit3, Barx2, and Grhl3. We also identified 28 additional TF matches that were DE with a less stringent statistical cutoff (FDR < 0.05 but with no restrictions on $\log_2$-fold change), including PitX2, Ehf, Creb3l1, Elf3. In contrast to the venom-gene results, the overall proportion of expressed TF genes that were DE (3.3%) was slightly lower, but not significantly so ($\chi^2 = 0.6$; $P = 0.43$), compared to the proportion of nonvenom, non-TF coding genes that were DE (3.8%) in the venom glands. The full gene-expression results are provided in Dataset S1.

Epigenomic Patterns Predict Ontogenetic Expression Changes. We performed venom-gland ATAC-seq (*SI Appendix*, Tables S7 and S8) and H3K27ac CUT&RUN (*SI Appendix*, Table S9) for adults ($n = 4$ biological and $n = 15$ technical replicates for ATAC-seq; $n = 2$ biological and $n = 4$ technical replicates for CUT&RUN) and juveniles ($n = 4$ biological and $n = 19$ technical replicates for ATAC-seq; $n = 3$ biological and $n = 6$ technical replicates for CUT&RUN) to test for

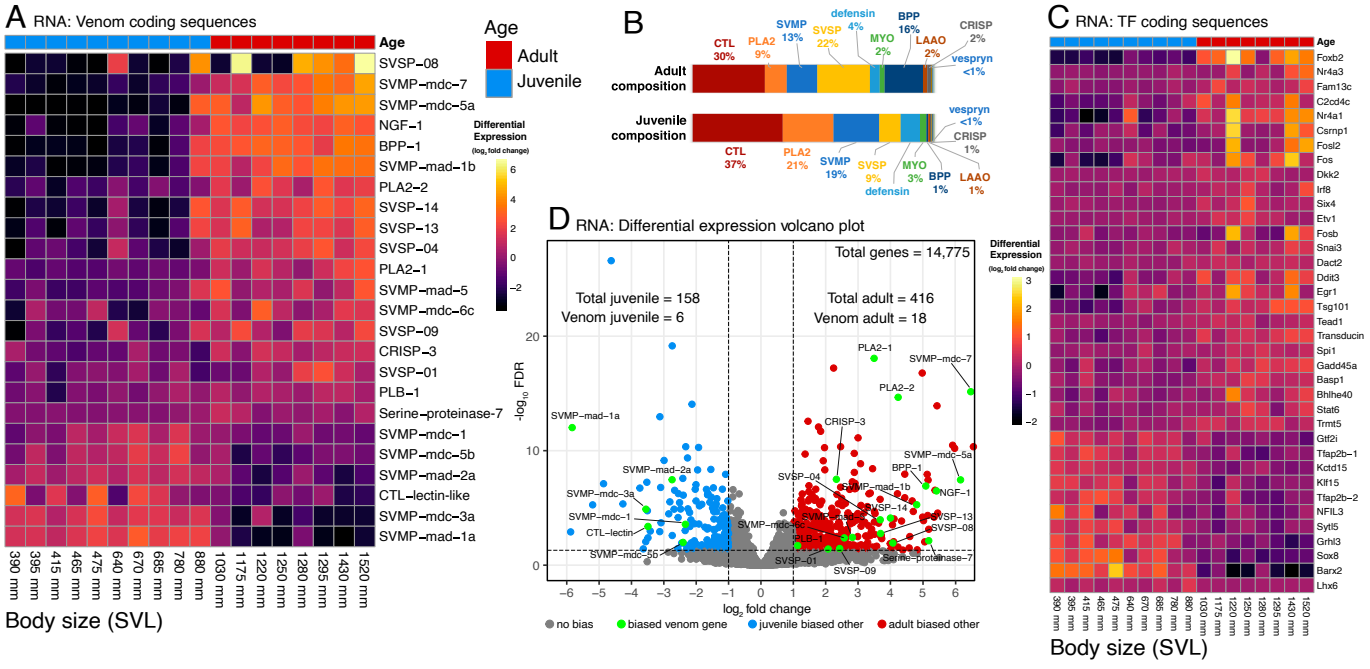

Fig. 2. Differential venom and transcription factor (TF) gene expression between adults and juveniles. (*A*) We identified 24 differentially expressed venom genes ($\log_2$-fold change > 1 and FDR < 0.05) using snout–vent length (SVL; adults >1,000 mm) as our proxy for age. The vast majority of age-biased venom genes showed higher expression in adults relative to juveniles. The largest subadult (SVL = 880 mm) showed an intermediate pattern, consistent with a gradual transition (31). Heatmap cells are colored based on regularized log (rlog) count differences per gene. (*B*) The largest class-level expression differences for venom genes correspond to toxin classes with differentially expressed paralogs (e.g., CTLs, PLA2s, SVSPs, SVMPs). One exception is BPP, where a single gene on chromosome 5 is up-regulated in adults (*SI Appendix*, Fig. S17). Stacked bars represent venom composition based on average transcripts per million (TPM) for juveniles and adults. (*C*) We identified 37 differentially expressed putative TFs, with the vast majority being adult biased. (*D*) A complete visualization of venom gland expression shows drastic shifts in venom compared to the majority of nonvenom genes. Abbreviations: BPP—Bradykinin-potentiating peptide, CRISP—Cysteine-rich secretory protein, CTL—C-type lectin, LAAO—L-amino acid oxidase, MYO—Myotoxin, NGF-Nerve growth factor, PLA2—Phospholipase A2, PLB—Phospholipase B, SVMP—Snake venom metalloproteinase, SVSP—Snake venom serine protease.

differential chromatin accessibility and to quantify differences in H3K27ac modifications, respectively. We mirrored our DE analysis to detect age-biased differential accessibility and histone modifications (DA) from ATAC-seq and CUT&RUN data and found significant signals for age-biased epigenomic patterns with both approaches (Figs. 3 and 4). Read coverage distributions and peak calls (*SI Appendix*, Fig. S11) for ATAC-seq CUT&RUN for all venom genes are provided in *SI Appendix*, Figs. S12–S19. Both ATAC-seq and CUT&RUN revealed differential epigenomic patterns (DA) flanking DE venom genes, consistent with enhancer/promoter mediated expression mechanisms postulated previously for snake venom systems (25). However, we also observed comparable DA within the gene regions, primarily overlapping with the introns of SVMPs and SVSPs. We found that almost all the juvenile-biased venom accessibility is explained by the two primary SVMP genes (mdc-3a and mad-1a) responsible for juvenile expression bias. Of the eleven juvenile-biased venom regions detected with ATAC-seq, four overlap with these two genes, three flank them, one overlaps with the nonbiased venom gene PDE-1, one flanks the nonbiased venom gene HYAL-1, and one flanks the nonbiased venom gene MYO-3. Of the 22 juvenile-biased venom regions detected as DA with CUT&RUN, 18 of these appear to be associated with the same two biased SVMPs, including eight overlapping and nine flanking these genes.

We next tested for an overall association between venom gene expression and epigenomic ontogenetic patterns. We detected a significant positive correlation between $\log_2$-fold change in gene expression with both $\log_2$-fold changes in chromatin accessibility and histone modifications at multiple genomic scales (Fig. 5*B*). Differentially accessible ATAC-seq peaks exhibited a strong positive correlation with DE venom genes for peaks within venom genes ($P < 0.0001$; $R^2 = 0.68$), 1 kb flanking venom genes ($P < 0.0001$; $R^2 = 0.80$), and 1 to 10 kb flanking venom genes ($P < 0.0001$; $R^2 = 0.61$); ATAC-seq peaks from 10 to 100 kb distance of a venom gene did not show a significant correlation with venom gene DE ($P = 0.23$; $R^2 = 0.17$; Fig. 5*B*). Similarly, DA CUT&RUN peaks exhibited significant positive correlations with DE venom genes at all four levels ($P < 0.0001$; $0.52 \leq R^2 \leq 0.69$; Fig. 5*B*) with the strongest

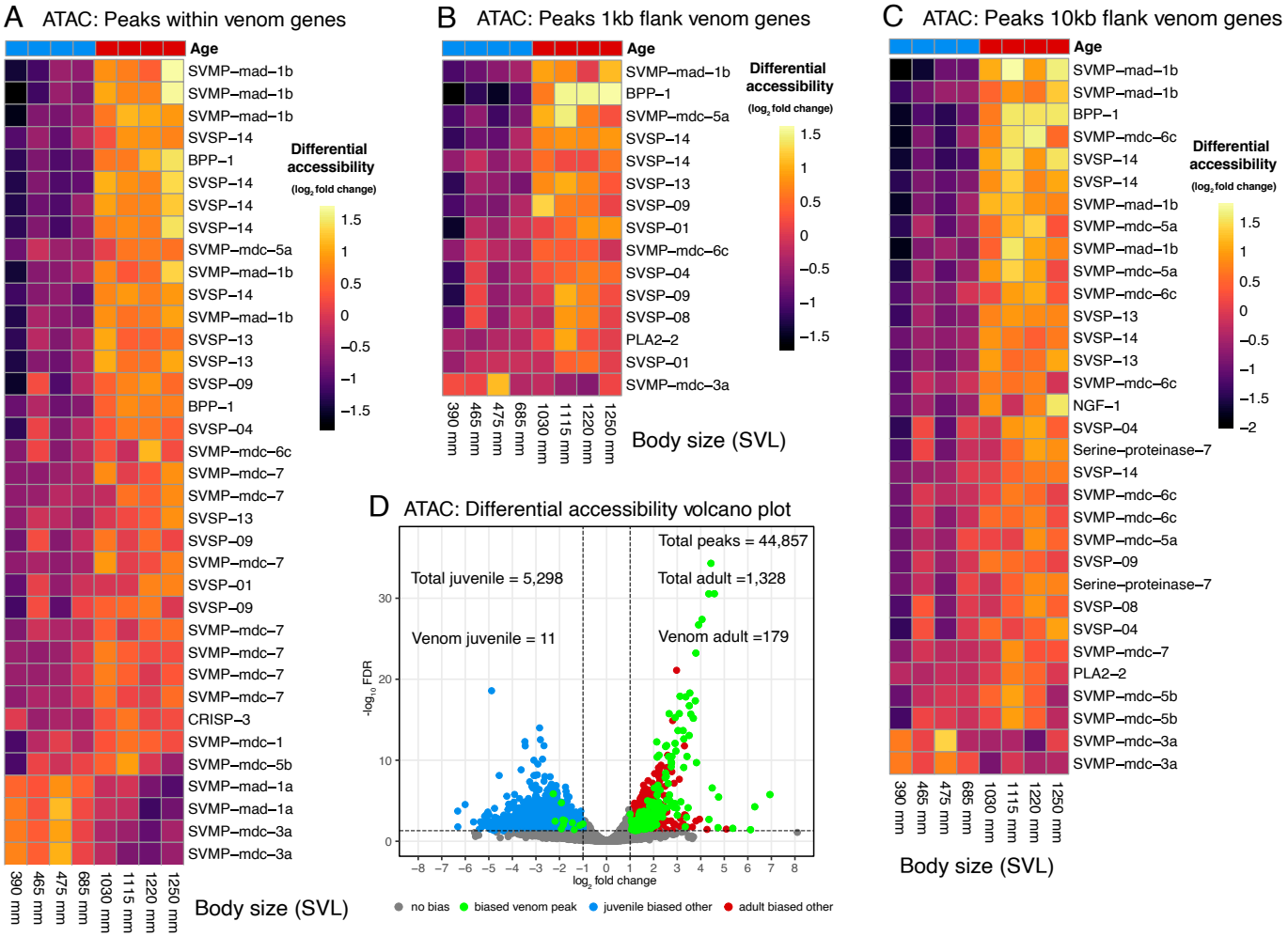

Fig. 3. ATAC-seq reveals differential chromatin accessibility between adults and juveniles. (*A–C*) In parallel with gene-expression patterns, the vast majority of age-biased peaks showed higher accessibility in adults relative to juveniles. We used the same cutoff for differential accessibility as differential expression ($\log_2$-fold change > 1 and FDR < 0.05). To illustrate the spatial relation of accessible regions and venom genes, we split results based on spatial proximity to the nearest venom gene: (*A*) within gene margins, (*B*) flanking gene to 1 kb, and (*C*) flanking between 1 and 10 kb. We represented peaks by the name of the closest gene, with many venom genes having multiple associated peaks. Heatmap cells are colored based on regularized log (rlog) count differences per peak. For clarity, only peaks associated with DE venom genes are represented in the heatmaps shown. (*D*) A complete visualization of venom-gland chromatin accessibility from ATAC-seq suggests that the adult bias is more centered around venom than the juvenile bias. Read coverage distributions for all venom genes are provided in *SI Appendix*, Figs. S12–S19.

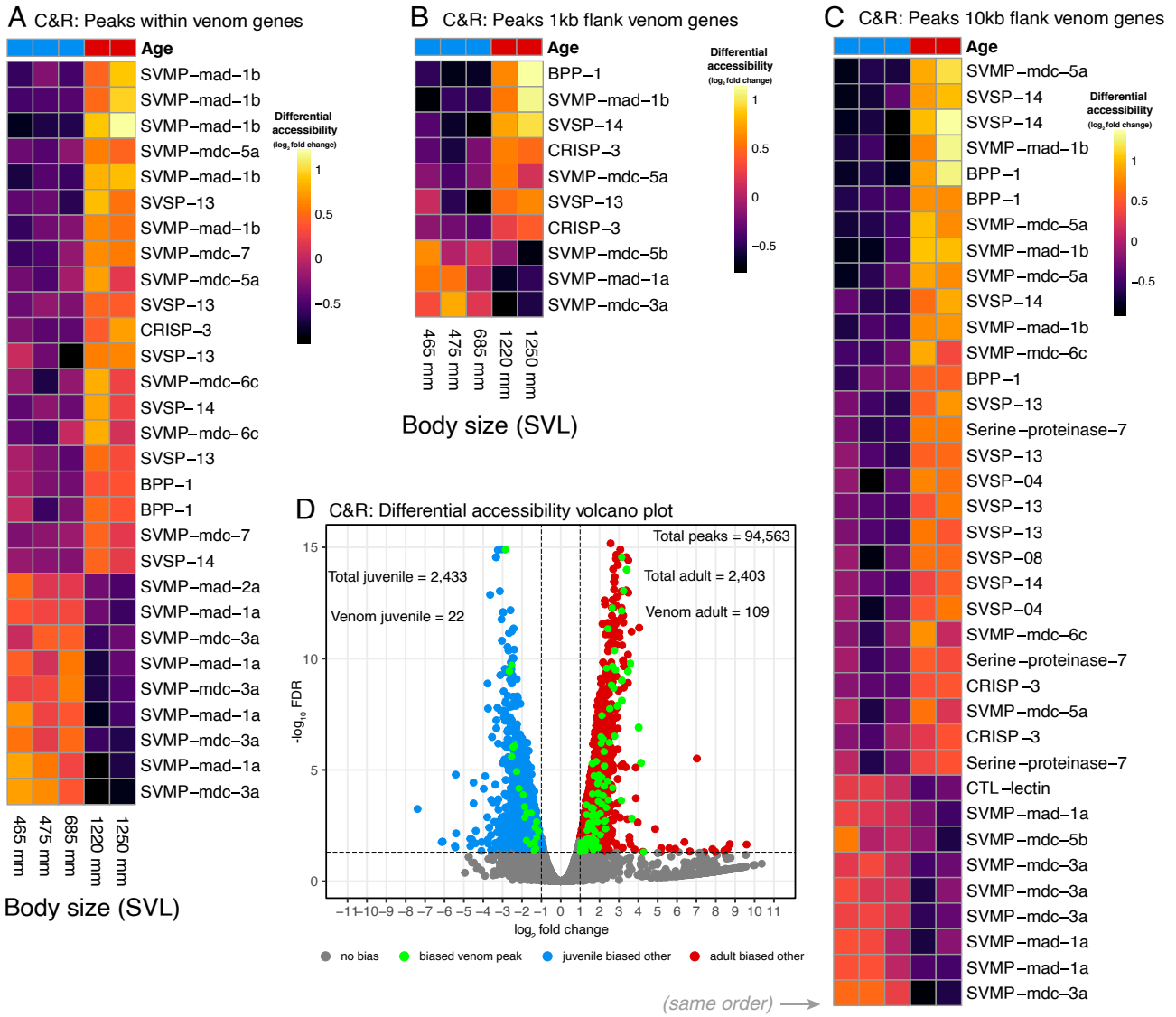

Fig. 4. CUT&RUN reveals differential patterns of H3K27ac histone modification between adults and juveniles. (*A*–*C*) We see general agreement with the ATAC-seq results, where differential H3K27ac modifications occur around differentially expressed venom genes. We analyzed and reported patterns from CUT&RUN following our ATAC-seq results (Fig. 3). For clarity, only peaks associated with DE venom genes are represented in the heatmaps shown. We represented peaks by the name of the closest gene, with many venom genes having multiple associated peaks. (*D*) A visualization of venom-gland H3K27ac patterns from CUT&RUN suggests that the venom accessibility bias is more comparable between adults and juveniles when compared to our ATAC-seq patterns (Fig. 3*D*). Read coverage distributions for all venom genes are provided in *SI Appendix*, Figs. S12–S19.

association between histone modifications and venom gene DE occurring within venom genes ($R^2 = 0.68$) and 1 to 10 kb flanking venom genes ($R^2 = 0.69$).

Because the adult-biased DA was much more complex, and the relative sizes of DA regions were not uniform, we implemented a proportional nucleotide-count *z*-score approach and rated all DE-DA overlap scenarios for venom ontogeny based on these scores (Fig. 5). Sites within epigenomic peaks can be classified on the basis of their location relative to an expressed gene (i.e., within the gene, within 1 kb of the gene, or 1 to 10 kb from the gene), as well as by whether they showed a signal for DA (i.e., not biased, juvenile biased, or adult biased). We excluded peaks from 10 to 100 kb flanking venom genes because gene assignments were less reliable at this distance. The genes with which these sites associate can also be classified on the basis of their DE patterns (i.e., not biased, juvenile biased, or adult biased). If, for example, adult-biased gene expression were driven by changes to accessibility of promoters, we would expect to see a significant enrichment for sites in adult-biased hbixATAC-seq peaks within 1 kb of adult-biased genes. We identified juvenile-biased DA within juvenile-biased venom genes as the most supported DE-DA overlap scenario for both ATAC-seq and CUT&RUN on the basis of *z*-scores. This pattern was followed by the other forms of age-matching DE-DA overlap scenarios (adult-adult, juvenile-juvenile, nonbiased-nonbiased), with slightly different *z*-score rating orders between ATAC-seq and CUT&RUN (Fig. 5). Mismatching bias DE-DA overlap scenarios (juvenile-adult, nonbiased-juvenile, nonbiased-adult) were the least supported scenarios, with nonbiased accessibility within adult-biased genes having the least support for ATAC-seq, and nonbiased differential H3K27ac histone modifications flanking adult-biased genes between 1 and 10 kb having the least support for CUT&RUN. All DE-DA overlap test calculations are included in Dataset S2.

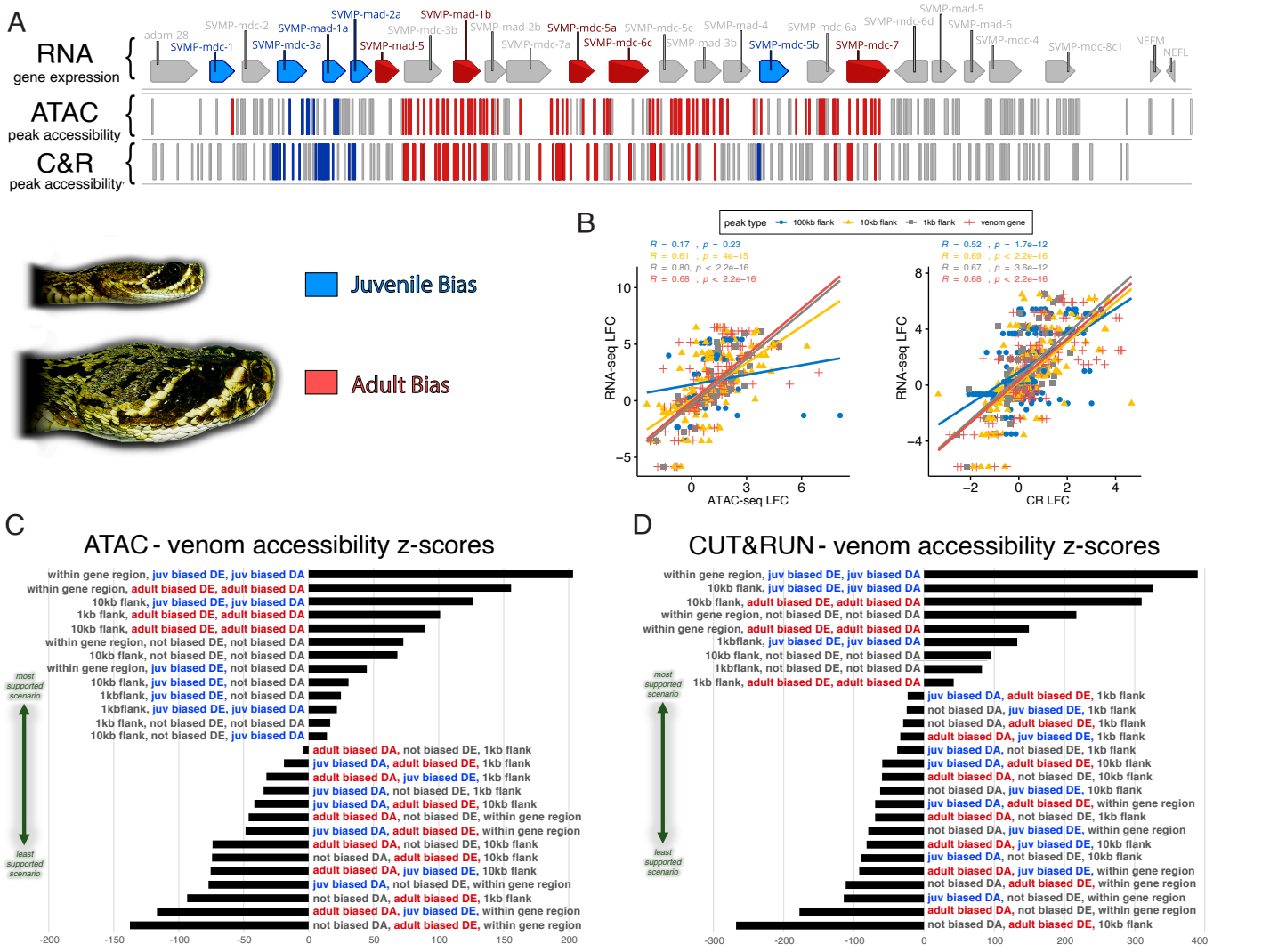

Fig. 5. Overlapping ontogenetic biases between venom gene expression and chromatin accessibility and differential H3K27ac histone modifications in *C. adamanteus*. Differential gene expression (DE) via RNA-seq and differential peak accessibility and histone modifications (DA) via ATAC-seq and CUT&RUN between adults (colored red) and juveniles (colored blue). Both DE and DA analyses utilized the same significance threshold ($\log_2$-fold change > 1 and FDR < 0.05). (*A*) SVMP gene array and accessible peaks colored by significant age bias highlighting DE and DA overlap, with DE gene regions typically overlapping or flanked by matching DA peaks. The array is oriented with adam28 on the *Left* and NEFL on the *Right*. (*B*) The association between changes in expression patterns of venom genes and changes in chromatin accessibility and histone modifications (H3K27ac). We performed linear regression analyses of the $\log_2$-fold changes (LFCs) in expression and epigenomic patterns between adults and juveniles. Peaks were binned according to their relationship to their nearest venom gene in the following four categories: 1) within the gene, 2) outside the gene but within 1 kb, 3) from 1 to 10kb from the gene, and 4) from 10 to 100 kb from the gene. We found significant positive correlations for all categories for both ATAC-seq and CUT&RUN, except for ATAC-seq peaks 10 to 100 kb from their nearest venom gene. We found the same pattern of significance when using only the peak with the largest magnitude of change for each gene for each peak category to address RNA-seq data nonindependence. (*C*) Proportional *z*-scores of DE–DA bias agreement from ATAC-seq, ordered from most supported scenario to least. Positive values represent DE–DA overlap scenarios occurring more than all other venom peak scenarios, and negative values represent DE–DA overlap scenarios occurring less than all other venom peak scenarios. The highest *z*-scores reflect a strong preference for matching ontogenetic DE–DA combinations. Peaks are classified based on positional relation to genes, ontogenetic DE bias of the gene, and ontogenetic DA bias of the peak. Proportions were calculated using total bp counts for each venom peak classification compared to all other venom peaks. (*D*) Proportional *z*-scores of DE-DA bias agreement from CUT&RUN, mirroring the layout shown in panel C for ATAC-seq. The CUT&RUN *z*-scores show overall agreement with ATAC-seq, with the highest *z*-scores depicting a strong preference for matching ontogenetic DE-DA combinations.

Motif-Based Identification of Ontogeny-Related Transcription Factors. From the translated coding sequences of differentially expressed TF genes (FDR < 0.05), we successfully assigned DNA binding motifs for 31 of 65 searched TFs based on predicted matches from the JASPAR CORE motif database (47). We next identified these motif sequences in genomic regions overlapping with accessible venom peaks using TFBStools (48) and tested for enrichment by comparing whether the total number of base pairs occupied by each motif was proportionally higher or lower for age-biased peaks. Comparisons between motif distributions spanning ATAC-seq DA venom regions with matching venom gene expression biases (i.e., adult-adult versus juvenile-juvenile) identified 15 TFs with ontogenetic motif enrichment using a 90% match cutoff and 13 using a 95% cutoff, where enrichment reflects significance (P <0.05) from the nucleotide-proportional *z*-tests comparing motif distributions between the two age groups. By comparing motif distributions spanning CUT&RUN DA venom regions with matching expression biases, we identified 14 TFs with ontogenetic motif enrichment using the 90% cutoff (including nine overlapping with ATAC-seq) and 10 TFs using the 95% cutoff (including three overlapping with ATAC-seq). We identified

our top TF candidates (Table 1) by applying an additional LFC> 1 cutoff.

Joint analyses of motif enrichment, DA, and DE yielded predictions of specific TF roles for regulating venom production in rattlesnakes. If expression increased when a TF motif became accessible, we considered it a putative expression enhancer. If expression decreased, we considered it a putative repressor. We found that seven TFs likely enhance venom gene expression in adults, five likely repress expression in adults, two likely enhance expression in juveniles, and two likely repress expression in juveniles (Table 1; Dataset S3). By comparing motif enrichment across DA venom regions, we independently predicted general regulatory functions which correctly matched prior characterizations for several top candidate TFs in other systems: Foxb2 is a transcriptional activator (49), Sox8 is an early-life positive transcriptional regulator (50), Tfab2 is an early life growth repressor (51), Klf15 inhibits RNA polymerase II binding (52) and interacts directly with PPAR-family genes driving circadian rhythmicity (53), and Bhlhe40 is a late-life negative transcriptional regulator (54, 55).

Interaction Network and the Epigenomics of Venom Ontogeny. To assess our candidate venom ontogeny regulators in a broader gene-regulatory context, we modeled a protein-interaction-network map for 204 TFs expressed in rattlesnake venom glands (DESeq2 baseMean ≥ 30; *SI Appendix*, Fig. S20). Integrating our venom ontogeny findings into the protein-interaction network allowed us to identify several key clusters of TFs representing general regulatory or developmental pathways that are likely important for regulating venom production. One cluster, which includes the SMAD subunits, Runx1, Irf8, Tfap2b, and the AP-1 activation subunits, is involved in growth signaling and may play a role in the development and maturation of venom glands. Another cluster, which includes the RNA polymerase II subunits, Egr1, Nr4a1, Ppard, and Irf6 is involved in transcriptional activation and may be important for turning venom genes on and off. A third cluster, which includes internal time-keeping components such as CLOCK, Arntl, and two of our top candidate venom ontogeny TFs, Bhlhe40 and Klf15, is involved in maintaining the biological clock and controls timing interactions.

Comparative Genomics Establishes the Roles of Bhlhe40 and Klf15 in Venom Ontogeny. To test our hypothesis that specific *cis*-regulators control venom ontogeny and that a loss of these factors would ablate age-biased differential expression, we compared venom gene expression and motif distributions with the western diamondback rattlesnake (*C. atrox*; *SI Appendix*, Table S11), a species that shows a substantially diminished ontogenetic venom shift (*SI Appendix*, Figs. S22 and S23). For *C. atrox*, the SVMP gene array on chromosome 9 was fully resolved by Giorgianni et al. (42). We generated venom gland RNA-seq data for nine adult and six juvenile *C. atrox* (*SI Appendix*, Table S12) and detected significant DE for conserved venom orthologs across the two species (*SI Appendix*, Figs. S22 and S23). Only nine venom genes were differentially expressed across adults and juveniles for *C. atrox*, and these were comprised entirely of SVMP genes (*SI Appendix*, Fig. S22*A*). Additionally, eight of the nine biased SVMPs were biased toward juveniles in *C. atrox*, presenting a stark contrast to the adult-skewed pattern in *C. adamanteus*. Venom in *C. adamanteus* appears to transition from a simpler SVMP makeup in juveniles to a more complex SVMP makeup in adults, but venom in *C. atrox* transitions from a more complex SVMP makeup in juveniles to a simpler SVMP makeup in adults (*SI Appendix*, Fig. S23). We also determined that the only age-biased TFs shared between species were juvenile biased, including Sox8, Crebl1, Sytl5, and Barx2 (*SI Appendix*, Fig. S22).

To understand the overall lack of TF expression bias overlap, we conducted a motif comparison between the two species using the same 31 TF motifs evaluated earlier in *C. adamanteus* (Fig. 6). Because only the SVMP gene array and no epigenomic data were available for *C. atrox*, we modeled motifs spanning the entire SVMP gene array in both species (*SI Appendix*, Figs. S24–S34) and compared orthologous regions (*SI Appendix*, Table S13). Binding motifs can be lost or gained through point mutations and indels (e.g., Fig. 6), leading to altered expression patterns and

Table 1. Top candidate venom ontogeny regulators and supporting evidence from expression bias and sequence-based motif inference

Transcription factor	Expression bias	Motif enrichment 90%	Motif enrichment 95%	Predicted expression effect	Generalized functional context	HGNC #
Foxb2	Adult	C&R	–	Enhance adult	Activator; DNA binding; forkhead family	23315
Nr4a1	Adult	C&R	ATAC	Enhance adult	Stress-response; hormone-retinoid receptor	7980
Fosl2	Adult	–	Both	Repress adult	Leucine zipper; AP-1 complex	3798
Fos	Adult	–	Both	Repress adult	Leucine zipper; AP-1 complex	3798
Etv1	Adult	Both	C&R	Enhance adult	Translocations; ETS family	3490
Bhlhe40	Adult	Both	ATAC	Repress adult	Circadian rhythm repressor; block E-box binding	1046
Egr1	Adult	Both	C&R	Enhance adult	Growth; zinc-finger family	3238
Snai3	Adult	C&R	–	Enhance adult	Growth; snail family	18411
Spi1	Adult	ATAC	ATAC	Enhance adult	Activator; ETS family	11241
Fosb	Adult	–	ATAC	Repress adult	Leucine zipper; AP-1 complex	3798
Nr4a3	Adult	C&R	ATAC	Enhance adult	Activator; hormone-retinoid receptor	7982
Tead1	Adult	Both	C&R	Repress adult	Activator and repressor; TEA family	11714
Barx2	Juvenile	C&R	C&R	Enhance juvenile	Growth; homeobox family	956
Sox8	Juvenile	ATAC	ATAC	Enhance juvenile	Embryonic activator; SOX family	11203
Tfap2b	Juvenile	Both	Both	Repress juvenile	Embryonic repressor; AP-2 family	11743
Klf15	Juvenile	Both	C&R	Repress juvenile	DNA binding; circadian rhythm repressor	14536

This subset includes 16 differentially expressed TFs with motif enrichment detected in differentially accessible venom peaks flanking or overlapping with differentially expressed toxin genes. Motif enrichment support via TFBStools is shown for two separate runs, one using a 90% minimum sequence identity match and the other using 95%. Functional predictions relating to age-biased venom gene expression were assigned based on motif bias agreement or disagreement with TF gene DE, venom gene DE, and venom peak DA.

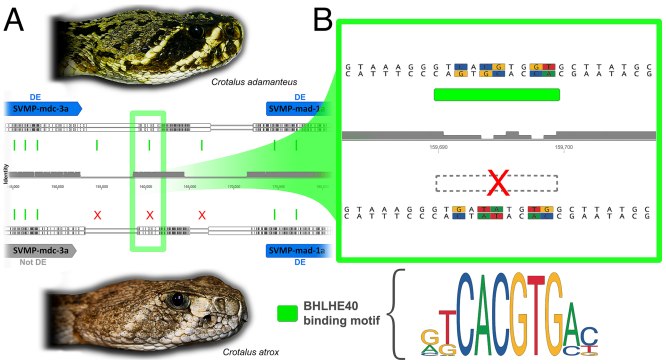

Fig. 6. Presence and absence of Bhlhe40 binding motifs in two rattlesnake species. (*A*) *Crotalus atrox* does not show a drastic ontogenetic shift in venom phenotype as does *C. adamanteus* (*SI Appendix*, Figs. S22 and S23). Comparing binding motif presence and absence between the two species supports our hypothesis that Bhlhe40 functions as a transcriptional repressor for SVMP gene mdc-3a in adult *C. adamanteus*. (*B*) An example of apparent Bhlhe40 motif loss/gain by means of five nucleotide changes.

changes to ontogenetic control. Proportional motif contributions surrounding SVMP mdc-1a were found to be significantly higher in *C. adamanteus* for Etv1, Bhlhe40, Egr1, and Tead1 compared to *C. atrox*. Although both species shared the juvenile expression bias for mdc-1a, the shift is much larger in *C. adamanteus* ($\log_2$-fold change = −5.8) than in *C. atrox* ($\log_2$-fold change = −3.8). We also found motif contributions surrounding SVMP gene mdc-7 are significantly higher in *C. adamanteus* for Etv1, Egr1, Spi1, Tfap2b, and Klf15 compared to *C. atrox*. The ontogenetic expression bias for mdc-7 is opposite between the two species, with the *C. adamanteus* ortholog biased toward adults ($\log_2$-fold change = 6.5) and the *C. atrox* version biased toward juveniles ($\log_2$-fold change = −2.4). For our last motif comparison, we compared proportional abundances across the entire SVMP array and found significantly higher motif contributions of Foxb2, Etv1, Bhlhe40, Egr1, Tfap2b, and Klf15 in *C. adamanteus*. The Sox8 motif is more abundant in the *C. atrox* SVMP array than the *C. adamanteus* array ($z = -2.01$), which together with the juvenile expression bias suggest this particular TF has a higher importance in regulating *C. atrox* SVMPs than in *C. adamanteus*. The high density of TF-binding motifs near large venom-gene arrays such as the SVMPs (*SI Appendix*, Figs. S24–S34) is consistent with their broader regions of accessibility with corresponding H3K27ac enrichment (e.g., *SI Appendix*, Fig. S14) relative to nonvenom regions (e.g., *SI Appendix*, Fig. S12), suggesting that these regions are actively being bound by TFs.

Discussion

Our work substantially expands our understanding of snake venom-gene expression by focusing on the previously unexplored basis for changes to expression leading to functional venom differences between adult and juvenile snakes (30). Previous work examined baseline control of venom-gene expression in adult rattlesnakes from the species *Crotalus viridis* (23, 25), as well as the housekeeping genes involved in venom production on a broader taxonomic scale (56). We confirm key regulators through matched detection, including TFs involved in the MAPK/ERK growth signaling pathway (e.g., Ehf and Fos), the unfolded protein response (e.g., Ddit3, Creb3l1), and several biological timing regulators (e.g., CLOCK). Our genome-wide TF scanning using DeepTFactor (57) also predicted putative TF genes that have no detectable homology with previously described TFs, suggesting that novel regulatory mechanisms may be employed in the regulation of venom-gene expression. Finally, our strongest candidate venom ontogeny TFs include seven adult expression enhancers, five adult expression repressors, two juvenile expression enhancers, and two juvenile expression repressors (Table 1). Both repression and activation therefore play major roles in regulating expression ontogeny. We found no signal for a master ontogeny switch for rattlesnake venom, but instead subtle and varied control structures.

By splitting our putative cis-regulatory elements (CREs) into classes on the basis of their proximity to their nearest venom gene, we defined the relationships between differential expression and differential epigenomic patterns and established the importance of intragenic CREs in venom-gene expression. Our strongest signals for epigenomic control of expression levels for both ATAC-seq and H3K27ac CUT&RUN fell within the gene bodies of venom genes themselves (Fig. 5); CREs within introns of venom genes have been previously noted in snakes (58) and are generally associated with tissue-specific expression (59). Although this signal for chromatin accessibility may, in part, result simply from high levels of expression and the resulting high occupancy of RNA polymerase II, the consistency across techniques suggests this pattern reflects true CREs within the gene bodies. Such an organization may facilitate the evolution of venom genes into tandemly arrayed duplicates with high paralogy (Fig. 1). If required CREs are contained within gene bodies, the duplication of the gene itself is certain to include its necessary CREs. We also showed strong support for the involvement of promoters (i.e., 1 kb flanking genes) and CREs outside the genes bodies (1 to 10 kb flanking genes) in the control of venom ontogeny. The latter results assume that CREs can be assigned to genes on the basis of linear genomic proximity rather than three-dimensional proximity, which is not an optimal assignment criterion (60), and further work using techniques such as promoter-capture Hi-C (61) or HiChIP (62) will be required to fully resolve these relationships.

By comparing the *C. adamanteus* regulatory architecture with that of *C. atrox*, we found evidence that the absence of specific CREs ablates age-biased differential expression in orthologous SVMPs (Fig. 6). Predicted binding affinity of 10 candidate TFs differed significantly between the two species, including two noteworthy negative regulators, Bhlhe40 and Klf15. Unlike the other candidate TFs, the protein–protein interactions of Bhlhe40 and Klf15 directly overlap with one another, with both linked to the same key biological timing factors, CLOCK and Arntl (*SI Appendix*, Fig. S20). In *C. adamanteus*, Bhlhe40 is expressed higher in adults, and Klf15 is expressed higher in juveniles, while neither are DE in *C. atrox*. Since Bhlhe40 and Klf15 were only age-biased in *C. adamanteus*, we expect selection to favor enrichment of the corresponding CREs (i.e., binding motifs) around age-biased venom genes in *C. adamanteus* compared to *C. atrox*, which was confirmed to be the case. We found significantly more Bhlhe40 and Klf15 binding motifs associated with SVMPs in *C. adamanteus* than *C. atrox*, and the distribution of these CREs was consistent with repressive activity in *C. adamanteus*. Prior functional characterizations reported Bhlhe40 as an important regulatory repressor (54, 55), which is consistent with our interpretation of this TF repressing venom gene expression in rattlesnakes (Table 1). These findings, in particular, yield a set of highly supported genetic regulatory targets for future hypothesis testing. For example, the roles of these TFs in ontogeny in *C. adamanteus* or other pitviper species could be functionally tested by means of RNA interference (RNAi) (63) targeting these TFs. Analyses similar to ours in

additional rattlesnake species could not only allow estimation of the timing of regulatory component co-option for venom ontogeny, but also begin to relate the genetic architecture of ontogenetic regulation to across-species variation in the nature of the ontogenetic change and venom composition in general.

The ontogenetic shift in venom composition and venom-gland gene expression for *C. adamanteus* was found to be one of overall increasing complexity in adults relative to juveniles. Five SVMPs and a single CTL highly expressed early in life were down-regulated in adults, along with the concomitant upregulation in adults of 18 venom genes, including five SVMPs and seven SVSPs (Fig. 2). Previous work (64, 65) linked increasingly complex venom to increasingly diverse diets across rattlesnake species, but with an explicit focus on adult venoms and adult diets. As *C. adamanteus* grows larger, its diet grows more phylogenetically diverse. Whereas small snakes feed solely on mice and rats (Rodentia: Cricetidae), larger snakes add squirrels and chipmunks (Rodentia: Sciuridae), and eventually rabbits (Lagomorpha: Leporidae) to their diets. Hence, we show here how an increasingly diverse diet (31) is mirrored by an increasingly complex venom as these snakes grow, underpinned by a complex but tractable suite of regulatory changes. The remarkable levels of functional, medically significant variation in snake venoms among species may emerge in part from initial divergent selection within snake lifetimes linked to life-history change.

Beneath the overall pattern of increasing venom complexity with age lies a sophisticated regulatory architecture involving both juvenile- and adult-biased upregulation (Fig. 2). Expression changes for each age class are controlled by both expression suppression and enhancement (Table 1), often mediated by multiple CREs for each affected gene (Figs. 3 and 4). Previous work by Schonour et al. (31) showed proteomically that for both *C. adamanteus* and *Crotalus horridus*, the ontogentic change is multifaceted, with some venom components changing gradually and others changing abruptly. Such patterns suggest diverse gene-regulatory mechanisms are involved, a pattern confirmed and elaborated upon by our results.

Regulatory complexity may be critical to these snakes if, like for *C. adamanteus* (29), selection associated with life-history patterns is as strong as, or even stronger than, selection across geography. Variation in the toxins involved in rattlesnake ontogenetic venom changes appears to be widespread; we showed, for example, significant differences in the ontogenetic change between *C. adamanteus* and *C. atrox*, although these differences were largely restricted to changes in expression among SVMP paralogs. Other rattlesnake species emphasize other toxin classes in their ontogentic changes. For example, *Crotalus molossus* showed a strong bias toward myotoxins and SVSPs in juveniles (66). Such variation likely reflects a modular regulatory architecture (67) with the capacity for mutation to affect changes to individual genes, correlated sets of genes, or even the presence of an ontogenetic change altogether.

We have revealed the requisite regulatory structure, including both associated TFs and CREs, that may lend evolutionary lability to this trait. In addition, because populations within species often vary along the same compositional axes as adults and juveniles of other species (33–35), this regulatory control may be subject to co-option as a means of phenotypic integration, enacting correlated changes among venom components that have already been screened by natural selection (29). As a result, venom evolution may be predictable, and the ontogenetic regulatory network identified here may allow us to explore the repeatability of the genetic architecture underlying adaptive phenotypic convergence among populations. For example, given that some islands in the southeastern United States inhabited by *C. adamanteus* lack the large prey items (e.g., rabbits) typical of adult diets (68, 69), we may expect suppression of the adult venom phenotype through, for instance, Bhlhe40 regulation. Although putative venom paedomorphism has been described in pitvipers previously (33–35, 70), the mechanisms underlying such changes were entirely unknown. Some subspecies of the South American rattlesnake (*Crotalus durissus*) do not undergo ontogenetic changes, retaining their juvenile phenotypes, whereas other subspecies show dramatic venom shifts (34). The common lancehead (*Bothrops atrox*), the leading cause of reptile-induced fatalities in South America, shows venom ontogeny in Colombia and Venezuela, but paedomorphic venoms in Brazil (70). More generally, paedomorphism has been implicated in major evolutionary transitions in, for example, birds (71), mammals (72), and salamanders (73). Our results not only enable predictions about mechanisms of adaptive divergence across ages, populations, and species, but also, by partitioning coding and regulatory sequences by age class, will allow for direct sequence-based comparisons of patterns of variation resulting from life-history-based selection on expression versus protein sequences.

Materials and Methods

Genome Sequencing and Assembly. High-molecular-weight (HMW) DNA was extracted from blood stored in 95% ethanol at −80 °C using phenol–chloroform extraction from an adult female *C. adamanteus* (DRR0105) from the Apalachicola National Forest in the Florida panhandle (*SI Appendix*, Table S4 and Fig. S1). Two flowcells of PacBio HiFi reads were generated using PacBio Sequel II sequencers. Hi-C data were generated from blood collected from an adult female *C. adamanteus* (MM0114), following previously described procedures (61). HiFi reads from both flowcells were assembled using hifiasm (74) v0.13 (r308), followed by purge_dups (75) v1.2.5. For Hi-C scaffolding, we used the juicer pipeline v1.6 (76), 3d-DNA (77), and Juicebox Assembly Tools v1.11.08 (78). Additional details are provided in *SI Appendix*.

RNA-seq. Total mRNA was extracted using the NEBNext Poly(A) mRNA Magnetic Isolation Module (New England Biolabs), and libraries were prepared for 150 nucleotide paired-end Illumina sequencing using NEBNext Ultra RNA-seq kits. Libraries were visualized and assessed for size and quality using a Bioanalyzer with a High-sensitivity DNA Kit (Aligent Technologies) and amplifiable concentrations were determined using a NGS Library Quantification Kit (KAPA Biosystems) performed at the Florida State University Biological Core Facilities. Individual libraries were pooled and sequenced with 150 nucleotide paired-end reads at the Florida State University College of Medicine Translational Science Laboratory on an Illumina HiSeq 2500 or NovaSeq 6000 or at Clemson University on an Illumina NextSeq (*SI Appendix*, Tables S6, S12, and S14). For some individuals, left and right glands were processed separately using unique indexes and later combined for all analyses. Read quantity and quality was assessed with FastQC v0.11.9. We performed adapter and quality trimming using Trim Galore! v0.6.6 (79) with settings –length 50 -q 10 –stringency 1 -e 0.1. Reads were aligned to the reference genome using hisat2 v2.2.1 (80) with parameters –no-unal –max-intronlen 25000 –dta and sorted using samtools (81) v1.12.

ATAC-seq. Venom-gland tissues were collected and used immediately for nuclei isolation and library preparation for either standard ATAC-seq (26) or Omni-ATAC-seq (82), with some modifications. Approximately 30 to 100 mg of venom-gland tissue was sheared with a razor blade and washed twice with ice-cold phosphate-buffered saline. For standard ATAC-seq samples, nuclei extraction and tagmentation were performed exactly as described (26). For Omni-ATAC-seq, the supernatant was discarded, samples were resuspended in ice-cold ATAC-resuspension buffer (10 mM Tris-HCL, 10 mM NaCl, 3 mM MgCl2, 0.01% digitonin, 0.1% Tween-20, and 0.1% NP40), and nuclei were extracted

as previously described (82). Nuclei were counted using hemocytometer chips, and approximately 50,000 to 60,000 intact nuclei were transposed using the Illumina Tagment DNA TDE1 Enzyme and Buffer Kit. Following tagmentation, the reaction was cleaned with a Zymo DNA Clean and Concentrator-5 kit. Transposed DNA was PCR amplified for five to six cycles. A qPCR side reaction was performed to determine the number of PCR cycles needed for effective amplification of each sample without oversaturation. The PCRs were purified with Agencourt AMPure XP beads (Beckman Coulter) and quantified using a Qubit Fluorometer (ThermoFisher Scientific). Library traces were quality checked using a D1000 ScreenTape and Tapestation (Agilent). Amplifiable concentrations were determined via KAPA PCR before being sequenced on an Illumina NovaSeq 6000; sequencing depths are summarized in *SI Appendix*, Tables S7 and S8.

Read quantity and quality was assessed with FastQC v0.11.9. One sample (four replicates) was sequenced with single-end 101 nt Illumina reads (*SI Appendix*, Table S8). Reads were trimmed with Cutadapt v3.4 (83) with options -e 0.1 -O 3 -m 20. Reads were aligned with bowtie2 (84) v2.4.2 with options –no-unal –very-sensitive and sorted using samtools (81) v1.12. Duplicate removal was performed using Picard. All remaining samples were sequenced with paired-end 150 nt Illumina reads (*SI Appendix*, Table S7). Reads were trimmed with Cutadapt v3.4 (83) with options -e 0.1 -O 3 -m 20. Reads were aligned with bowtie2 v2.4.2 (84) with options –no-unal -X 2000 –very-sensitive and sorted using samtools (81) v1.12. Duplicate removal was performed again using Picard. Peaks were called with MACS2 v2.2.7.1 (85). Our peak filtering and classification protocols are described in detail in *SI Appendix*. Our filtering strategy was based on the approach previously described (86) and is outlined in *SI Appendix*, Fig. S11.

CUT&RUN. Venom-gland tissues were collected and used immediately for nuclei isolation and library preparation. For nuclei isolation, venom-gland tissue was suspended in ice-cold lysis buffer (10 mM Tris-HCL, 10 mM NaCl, 3 mM MgCl2, and 0.1% IGEPAL-360) and then placed in a petri-dish on ice and sliced repeatedly with a razor blade or 21g needle followed by mixing through pipetting up and down with a p200 pipet tip to break up the tissue. Cells were counted using a hemocytometer, and intact nuclei were DAPI (4,6-diamidino-2-phenylindole) stained and visualized under a fluorescent microscope. For each library, approximately 50,000 cells containing intact nuclei were isolated and bound to ConA magnetic beads. CUT&RUN was performed as described (87) using anti-H3K27ac (cat no. ab4729) antibody and pA-MNase supplied by the Henikoff Lab. Library DNA was prepared using the NEBNext Ultra kit (#E6040S) and sequenced as 150 bp paired-end reads on an Illumina NovaSeq 6000.

Read quantity and quality was assessed with FastQC v0.11.9. Reads were trimmed with Cutadapt v3.4 (83) with options -e 0.1 -O 3 -m 20. Reads were aligned with bowtie2 (84) v2.4.2 with options –local –very-sensitive-local –no-unal –no-mixed –no-discordant -I 10 -X 700, sorted using samtools (81) v1.12., and duplicate removal was performed using Picard. Sequencing depth is summarized in *SI Appendix*, Table S9. Peaks were called MACS2 v2.2.7.1 (85). Our peak filtering and classification protocols are described in detail in *SI Appendix*. Our filtering strategy is outlined in *SI Appendix*, Fig. S11.

Differential Expression, Accessibility, and Histone Modification Analyses. We performed parallel analyses of differential expression (DE), differential accessibility, and differential H3K27ac modifications (DA) using our RNA-seq, ATAC-seq, and CUT&RUN data, respectively. We estimated read counts on genes/peaks with HTSeq-count v0.13.5 (88), and then compared adults to juveniles using DESeq2 v1.32.0 (89). Our threshold for significance was $\log_2$-fold change > 1 and FDR < 0.05, and significant genes/peaks were classified as adult or juvenile biased depending on whether the $\log_2$-fold change value was positive or negative. Additional details are provided in *SI Appendix*. The full DE/DA test results are included in Dataset S1.

Additional Analyses. Details of our other statistical and bioinformatic analyses are provided in *SI Appendix*.

Permits and Protocols. All animals were collected under the following permits: Florida, USA, permits LSSC-13-00004A, LSSC-13-00004B, and LSSC-13-00004C; New Mexico, USA, permit NMDGF 3563; Oklahoma, USA, permit ODWC 7413; Arizona, USA, permit SP628489; Texas, USA, permit SPR-0713-098; and Mexico permit OFICIO NUM. SGPA/DGVS/01090/17. All animal procedures were performed under active IACUC protocols: Florida State University protocols 0924, 1230, 1333, 1529, and 1836; University of Central Florida protocols 13-17 W and 16-17 W; Clemson University protocol 2017-067.

Data, Materials, and Software Availability. The genome assembly and data generated for this project have been deposited at DDBJ/ENA/GenBank under the accessions JAOTOJ000000000 and BioProjects PRJNA868880 (90), PRJNA88989 (91), and PRJNA667573 (92). Individual SRA accessions for all samples are listed in *SI Appendix*, Tables S1–S2, S5–S9, S12, and S14.

ACKNOWLEDGMENTS. We thank Kenneth P. Wray, Alyssa Hassinger, Emilie Broussard, Laura E. Koffinas, Rachel Saul, Alex Oliver, and Simone Gable for assistance in collecting and processing animals and Dr. Juan Miguel Borja-Jiménez for facilitating collecting and permits in Mexico. We thank Jack Facente for training and Sean B. Carroll, Noah L. Dowell, and Matt W. Giorgianni for providing unpublished data. Funding for this work was provided by the NSF (NSF DEB 1638902 to D.R.R., NSF DEB 1638879 and NSF DEB 1822417 to C.L.P., and DEB 1638872 to H.L.G.) with additional support from the Clemson University Genomics and Bioinformatics Facility, which receives support from the College of Science and two Institutional Development Awards (IDeA) from the National Institute of General Medical Sciences of the NIH under grant numbers P20GM146584 and P20GM139767.

Author affiliations: [a]Department of Biological Science, Florida State University, Tallahassee, FL 32306; [b]Life Sciences Institute, University of Michigan, Ann Arbor, MI 48109; [c]Department of Biology, University of Puerto Rico at Mayagüez, Mayagüez, PR 00681; [d]Department of Biological Sciences, Clemson University, Clemson, SC 29634; [e]Department of Evolution, Ecology and Organismal Biology, The Ohio State University, Columbus, OH 43210; [f]Department of Integrative Biology, University of South Florida, Tampa, FL 33620; [g]School of Biological Sciences, Washington State University, Pullman, WA 99164; [h]Department of Biology, University of South Alabama, Mobile, AL 36688; [i]Laboratory of Chromosome Replication and Epigenome Regulation, San Diego Biomedical Research Institute, San Diego, CA 92121; and [j]Department of Forestry and Environmental Conservation, Clemson University, Clemson, SC 29634

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

1. M. Levine, E. H. Davidson, Gene regulatory networks for development. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 4936–4942 (2005).
2. D. H. Erwin, E. H. Davidson, The evolution of hierarchical gene regulatory networks. *Nat. Rev. Genet.* **10**, 141–148 (2009).
3. H. E. Hoekstra, J. A. Coyne, The locus of evolution: Evo devo and the genetics of adaptation. *Evolution* **61**, 995–1016 (2007).
4. S. B. Carroll, Evo-devo and an expanding evolutionary synthesis: A genetic theory of morphological evolution. *Cell* **134**, 25–36 (2008).
5. M. Barrier, R. H. Robichaux, M. D. Purugganan, Accelerated regulatory gene evolution in an adaptive radiation. *Proc. Natl. Acad. Sci. U.S.A.* **98**, 10208–10213 (2001).
6. H. E. Hoekstra, R. J. Hirschmann, R. A. Bundey, P. A. Insel, J. P. Crossland, A single amino acid mutation contributes to adaptive beach mouse color pattern. *Science* **313**, 101–104 (2006).
7. N. Gompel, B. Prud'homme, P. J. Wittkopp, V. A. Kassner, S. B. Carroll, Chance caught on the wing: Cis-regulatory evolution and the origin of pigment patterns in *Drosophila*. *Nature* **433**, 481–487 (2005).
8. H. B. Fraser, Gene expression drives local adaptation in humans. *Genom. Res.* **23**, 1089–1096 (2013).
9. M. Konczal, W. Babik, J. Radwan, E. T. Sadowska, P. Koteja, Initial molecular-level response to artificial selection for increased aerobic metabolism occurs primarily through changes in gene expression. *Mol. Biol. Evol.* **32**, 1461–1473 (2015).
10. M. J. Margres, A. T. Bigelow, E. M. Lemmon, A. R. Lemmon, D. R. Rokyta, Selection to increase expression, not sequence diversity, precedes gene family origin and expansion in rattlesnake venom. *Genetics* **206**, 1569–1580 (2017).
11. A. Rhie *et al.*, Towards complete and error-free genome assemblies of all vertebrate species. *Nature* **592**, 737–746 (2021).
12. S. Fu *et al.*, Differential analysis of chromatin accessibility and histone modifications for predicting mouse developmental enhancers. *Nucleic Acids Res.* **46**, 11184–11201 (2018).
13. A. Gershman *et al.*, Epigenetic patterns in a complete human genome. *Science* **376**, eabj5089 (2022).
14. M. S. Halfon, Perspectives on gene regulatory network evolution. *Trend. Genet.* **33**, 436–447 (2017).

15. G. Cavalli, E. Heard, Advances in epigenetics link genetics to the environment and disease. *Nature* **571**, 489–499 (2019).
16. R. K. Schott *et al*., Transcriptomic evidence for visual adaptation during the aquatic to terrestrial metamorphosis in leopard frogs. *BMC Biol.* **20**, 138 (2022).
17. A. Abzhanov, M. Protas, R. Grant, P. R. Grant, C. J. Tabin, Bmp4 and morphological variation of beaks in Darwin's finches. *Science* **305**, 1462–1465 (2004).
18. K. L. Mack, M. W. Nachman, Gene regulation and speciation. *Trend. Genet.* **33**, 68–80 (2017).
19. M. Kawajiri *et al*., Ontogenetic stage-specific quantitative trait loci contribute to divergence in developmental trajectories of sexually dimorphic fins between medaka populations. *Mol. Ecol.* **23**, 5258–5275 (2014).
20. M. J. Margres *et al*., Linking the transcriptome and proteome to characterize the venom of the eastern diamondback rattlesnake (*Crotalus adamanteus*). *J. Proteom.* **96**, 145–158 (2014).
21. S. P. Mackessy, N. M. Sixberry, W. H. Heyborne, T. Fritts, Venom of the brown treesnake, *Boiga irregularis*: Ontogenetic shifts and taxa-specific toxicity. *Toxicon* **47**, 537–548 (2006).
22. C. M. Modahl, A. K. Mukherjee, S. P. Mackessy, An analysis of venom ontogeny and prey-specific toxicity in the monocled cobra (*Naja kaouthia*). *Toxicon* **119**, 8–20 (2016).
23. D. R. Schield *et al*., The origins and evolution of chromosomes, dosage compensation, and mechanisms underlying venom regulation in snakes. *Genome Res.* **29**, 590–601 (2019).
24. B. W. Perry, D. R. Schield, A. K. Westfall, S. P. Mackessy, T. A. Castoe, Physiological demands and signaling associated with snake venom production and storage illustrated by transcriptional analyses of venom glands. *Sci. Rep.* **10**, 18083 (2020).
25. B. Perry *et al*., Snake venom gene expression is coordinated by novel regulatory architecture and the integration of multiple co-opted vertebrate pathways. *Genome Res.* **32**, 1058–1073 (2022).
26. M. J. Margres *et al*., The tiger rattlesnake genome reveals a complex genotype underlying a simple venom phenotype. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2014634118 (2021).
27. J. Durban *et al*., Integrated omics profiling indicates that miRNAs are modulators of the ontogenetic venom composition shift in the Central American rattlesnake. *Crotalus simus simus. BMC Genom.* **14**, 234 (2013).
28. K. P. Wray, M. J. Margres, M. Seavy, D. R. Rokyta, Early significant ontogenetic changes in snake venoms. *Toxicon* **96**, 74–81 (2015).
29. M. J. Margres *et al*., Phenotypic integration in the feeding system of the eastern diamondback rattlesnake (*Crotalus adamanteus*). *Mol. Ecol.* **24**, 3405–3420 (2015).
30. D. R. Rokyta, M. J. Margres, M. J. Ward, E. E. Sanchez, The genetics of venom ontogeny in the eastern diamondback rattlesnake (*Crotalus adamanteus*). *PeerJ* **5**, e3249 (2017).
31. R. B. Schonour *et al*., Gradual and discrete ontogenetic shifts in rattlesnake venom composition and assessment of hormonal and ecological correlates. *Toxins* **12** (2020).
32. D. B. Means, "Food and prey acquisition" in *Diamonds in the Rough: Natural History of the Eastern Diamondback Rattlesnake* (Tall Timbers Press, Tallahassee, FL, 2017), pp. 173–198.
33. S. P. Mackessy, K. Williams, K. G. Ashton, Ontogenetic variation in venom composition and diet of *Crotalus oreganus concolor*: A case of venom paedomorphosis? *Copeia* **2003**, 769–782 (2003).
34. J. J. Calvete *et al*., Snake venomics of the central american rattlesnake *Crotalus simus* simus and the South American *Crotalus durissus* complex points to neurotoxicity as an adaptive paedomorphic trend along *Crotalus* dispersal in South America. *J. Proteome Res.* **9**, 528–544 (2010).
35. D. R. Rokyta, K. P. Wray, J. J. McGivern, M. J. Margres, The transcriptomic and proteomic basis for the evolution of a novel venom phenotype within the timber rattlesnake (*Crotalus horridus*). *Toxicon* **98**, 34–48 (2015).
36. F. Yan, D. R. Powell, D. J. Curtis, N. C. Wong, From reads to insight: A hitchhiker's guide to ATAC-seq data analysis. *Genome Biol.* **21**, 22 (2020).
37. P. J. Skene, S. Henikoff, An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites. *eLife* **6**, e21856 (2017).
38. M. P. Creyghton *et al*., Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 21931–21936 (2010).
39. D. Pasini *et al*., Characterization of an antagonistic switch between histone H3 lysine 27 methylation and acetylation in the transcriptional regulation of Polycomb group target genes. *Nucleic Acids Res.* **38**, 4958–4969 (2010).
40. H. Kang *et al*., Dynamic regulation of histone modifications and long-range chromosomal interactions during postmitotic transcriptional reactivation. *Genes Dev.* **34**, 1–18 (2020).
41. M. P. Hogan *et al*., The chemosensory repertoire of the eastern diamondback rattlesnake (*Crotalus adamanteus*) reveals complementary genetics of olfactory and vomeronasal-type receptors. *J. Mol. Evol.* **89**, 313–328 (2021).
42. M. Giorgianni *et al*., The origin and diversification of a novel protein family in venomous snakes. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 201920011 (2020).
43. D. D. Almeida *et al*., Tracking the recruitment and evolution of snake toxins using the evolutionary context provided by the *Bothrops jararaca* genome. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2015159118 (2021).
44. S. Nurk *et al*., HiCanu: Accurate assembly of segmental duplications, satellites, and allelic variants from high-fidelity long reads. *Genome Res.* **30**, 1291–1305 (2020).
45. C. Chin, A. Khalak, Human genome assembly in 100 minutes. bioRxiv [Preprint] (2019). https://www.biorxiv.org/content/10.1101/705616v1 (Accessed 17 July 2021).
46. J. L. Waldron, S. M. Welch, S. H. Bennett, W. G. Kalinowsky, T. A. Mousseau, Life history constraints contribute to the vulnerability of a declining North American rattlesnake. *Biol. Conserv.* **159**, 530–538 (2013).
47. O. Fornes *et al*., JASPAR 2020: Update of the open-access database of transcription factor binding profiles. *Nucleic Acids Res.* **48**, D87–D92 (2020).
48. G. Tan, B. Lenhard, TFBSTools: An R/bioconductor package for transcription factor binding site analysis. *Bioinformatics* **32**, 1555–1556 (2016).
49. L. Moparthi, G. Pizzolato, S. Koch, Wnt activator FOXB2 drives the neuroendocrine differentiation of prostate cancer. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 22189–22195 (2019).
50. S. Kimura *et al*., Sox8 is essential for M cell maturation to accelerate IgA response at the early stage after weaning in mice. *J. Exp. Med.* **216**, 831–846 (2019).
51. N. Zainolabidin, S. P. Kamath, A. R. Thanawalla, A. I. Chen, Distinct activities of Tfap2A and Tfap2B in the specification of GABAergic interneurons in the developing cerebellum. *Front. Mol. Neurosci.* **10** (2017).
52. S. Ray, J. W. Pollard, KLF15 negatively regulates estrogen-induced epithelial cell proliferation by inhibition of DNA replication licensing. *Proc. Natl. Acad. Sci. U.S.A.* **109**, E1334–E1343 (2012).
53. L. Fan *et al*., Transcription factors KLF15 and PPARδ cooperatively orchestrate genome-wide regulation of lipid metabolism in skeletal muscle. *J. Biol. Chem.* **298**, 101926 (2022).
54. R. Rauschmeier *et al*., Bhlhe40 and Bhlhe41 transcription factors regulate alveolar macrophage self-renewal and identity. *EMBO J.* **38**, e101233 (2019).
55. R. Rauschmeier *et al*., Bhlhe40 function in activated B and TFH cells restrains the GC reaction and prevents lymphomagenesis. *J. Exp. Med.* **219**, e20211406 (2021).
56. A. Barua, A. S. Mikheyev, An ancient, conserved gene regulatory network led to the rise of oral venom systems. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2021311118 (2021).
57. X. R. Bao, Z. Yi-Heng, D. J. Yu, "DeepTF: Accurate prediction of transcription factor binding sites by combining multi-scale convolution and long short-term memory neural network in Intelligence Science and Big Data Engineering" in *Big Data and Machine Learning* (Springer International Publishing, Cham, 2019), pp. 126–138.
58. G. Zancolli, N. R. Casewell, Venom systems as models for studying the origin and regulation of evolutionary novelties. *Mol. Biol. Evol.* **37**, 2777–2790 (2020).
59. B. Borsari *et al*., Enhancers with tissue-specific activity are enriched in intronic regions. *Genome Res.* **31**, 1325–1336 (2021).
60. S. Schoenfelder, P. Fraser, Long-range enhancer-promoter contacts in gene expression control. *Nat. Rev. Genet.* **20**, 437–455 (2019).
61. S. Schoenfelder, B. M. Javierre, M. Furlan-Magaril, S. W. Wingett, P. Fraser, Promoter capture Hi-C: High-resolution, genome-wide profiling of promoter interactions. *J. Vis. Exp.* **136** (2018).
62. M. R. Mumbach *et al*., HiChIP: Efficient and sensitive analysis of protein-directed genome architecture. *Nat. Methods* **13**, 919–922 (2016).
63. A. Walton, M. J. Sheehan, A. L. Toth, Going wild for functional genomics: RNA interference as a tool to study gene-behavior associations in diverse species and ecological contexts. *Horm. Behav.* **124**, 104774 (2020).
64. M. L. Holding *et al*., Phylogenetically diverse diets favor more complex venoms in North American pitvipers. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2015579118 (2021).
65. A. J. Mason *et al*., Venom gene sequence diversity and expression jointly shape diet adaptation in pitvipers. *Mol. Biol. Evol.* **39**, msac082 (2022).
66. M. Borja *et al*., Ontogenetic change in the venom of Mexican black-tailed rattlesnakes (*Crotalus molossus nigrescens*). *Toxins* **10**, 501 (2018).
67. A. J. Mason *et al*., Trait differentiation and modular toxin expression in palm-pitvipers. *BMC Genom.* **21**, 147 (2020).
68. M. J. Margres *et al*., Quantity, not quality: Rapid adaptation in a polygenic trait proceeded exclusively through expression differentiation. *Mol. Biol. Evol.* **34**, 3099–3110 (2017).
69. D. B. Means, "Morphology" in *Diamonds in the Rough: Natural History of the Eastern Diamondback Rattlesnake* (Tall Timbers Press, Tallahassee, FL, 2017), pp. 35–61.
70. J. J. Calvete *et al*., Snake population venomics and antivenomics of *Bothrops atrox*: Paedomorphism along its transamazonian dispersal and implications of geographic venom variability on snakebite management. *J. Proteom.* **74**, 510–527 (2011).
71. B. A. S. Bhullar *et al*., Birds have paedomorphic dinosaur skulls. *Nature* **487**, 223–226 (2012).
72. H. E. White *et al*., Pedomorphosis in the ancestry of marsupial mammals. *Curr. Biol.* **33**, 2136–2150 (2023).
73. J. J. Wiens, R. M. Bonett, P. T. Chippindale, Ontogeny discombobulates phylogeny: Paedomorphosis and higher-level salamander relationships. *Syst. Biol.* **54**, 91–110 (2005).
74. H. Cheng, G. T. Concepcion, X. Feng, H. Zhang, H. Li, Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm. *Nat. Methods* **18**, 170–175 (2021).
75. D. Guan *et al*., Identifying and removing haplotypic duplication in primary genome assemblies. *Bioinformatics* **36**, 2896–2898 (2020).
76. N. C. Durand *et al*., Juicer provides a one-click system for analyzing loop-resolution Hi-C experiments. *Cell Syst.* **3**, 95–98 (2016).
77. O. Dudchenko *et al*., De novo assembly of the *Aedes aegypti* genome using Hi-C yields chromosome-length scaffolds. *Science* **356**, 92–95 (2017).
78. O. Dudchenko *et al*., The Juicebox Assembly Tools module facilitates de novo assembly of mammalian genomes with chromosome-length scaffolds for under $1000. BioRxiv [Preprint] (2018). https://www.biorxiv.org/content/10.1101/254797v1 (Accessed 3 October 2022).
79. F. Krueger, T. Galore, A wrapper tool around Cutadapt and FastQC to consistently apply quality and adapter trimming to FastQ files, with some extra functionality for MspI-digested RRBS-type (Reduced Representation Buisulfite-Seq) libraries (2015).
80. D. Kim, J. M. Paggi, C. Park, C. Bennett, S. L. Salzberg, Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol.* **37**, 907–915 (2019).
81. H. Li *et al*., The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
82. M. R. Corces *et al*., An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. *Nat. Methods* **14**, 959–962 (2017).
83. M. Martin, Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. J.* **17**, 10–12 (2011).
84. B. Langmead, S. L. Salzberg, Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
85. Y. Zhang *et al*., Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).
86. T. B. Sackton *et al*., Convergent regulatory evolution and loss of flight in paleognathous birds. *Science* **364**, 74–78 (2019).
87. D. Janssens, S. Henikoff, CUT&RUN: Targeted in situ genome-wide profiling with high efficiency for low cell numbers. Protocols.io **N/A** (2019).
88. G. H. Putri, S. Anders, P. T. Pyl, J. E. Pimanda, F. Zanini, Analysing high-throughput sequencing data in python with HTSeq 2.0. *Bioinformatics* **38**, 2943–2945 (2022).
89. M. I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
90. M. P. Hogan *et al*., Data from "Eastern Diamondback Rattlesnake chromosome-resolution genome." NCBI BioProject. https://www.ncbi.nlm.nih.gov/bioproject/PRJNA868880. Deposited 12 August 2022.
91. D. R. Rokyta *et al*., Data from "Snake venom-gland transcriptomics." NCBI BioProject. https://www.ncbi.nlm.nih.gov/bioproject/PRJNA88989. Deposited 8 August 2022.
92. M. P. Hogan *et al*., Data from "Crotalus adamanteus isolate:Cadam-KW1264 Genome sequencing and assembly." NCBI BioProject. https://www.ncbi.nlm.nih.gov/bioproject/PRJNA667573/. Deposited 21 May 2021.