

## Supplemental Information

### Inventory of Supplemental Materials:

**Figure S1**, related to Figure 1. Patterns of VISTA enhancer conservation across turtle, lizard, and snake genomes.

**Figure S2**, related to Figure 2. Enhancer overlaps and signal profiles H3K27ac enriched regions from mouse limb and genital tubercle.

**Figure S3**, related to Figure 3. Comparison of similarity scores for mouse Limb, Limb-GT, and GT enhancer orthologs in *Anolis carolinensis* vs. snake species.

**Figure S4**, related to Figure 4. Comparison of H3K27ac signal from orthologous enhancer regions from mouse and *Anolis carolinensis* forelimb, hindlimb, and genital tubercle or hemiphallus.

**Figure S5**, related to Figure 5. LacZ staining patterns of mouse embryos carrying HLEA and HLEB Hsp68LacZ transgenes.

**Figure S6**, related to Figure 6. LacZ positive E12.5 mouse embryos carrying squamate HLEB-Hsp68LacZ transgenes.

**Figure S7**, related to Figure 7. Location of HLEB deletion and phenotypes in *DeIA*<sup>-/-</sup> and *DeIB*<sup>-/-</sup> mice.

**Table S1**, related to Figure 1. Mouse genome (mm9) coordinates of VISTA forebrain-only, VISTA heart only, VISTA limb-only, and published limb enhancers.

**Table S2**, related to Figure 2. H3K27ac ChIP-Seq read counts, quality metrics, and data processing results.

**Table S3**, related to Figure 2. Mouse genome (mm9) coordinates of H3K27ac enriched regions and putative enhancers from embryonic mouse forelimbs, hindlimbs, genital tubercle, eye, and flank.

**Table S4**, related to Figure 2. Mouse genome (mm9) coordinates of H3K27ac limb, limb-GT, GT, eye, and flank K-means clusters. Also contains coordinates for limb specific, limb-GT specific, and GT specific H3K27ac enhancers.

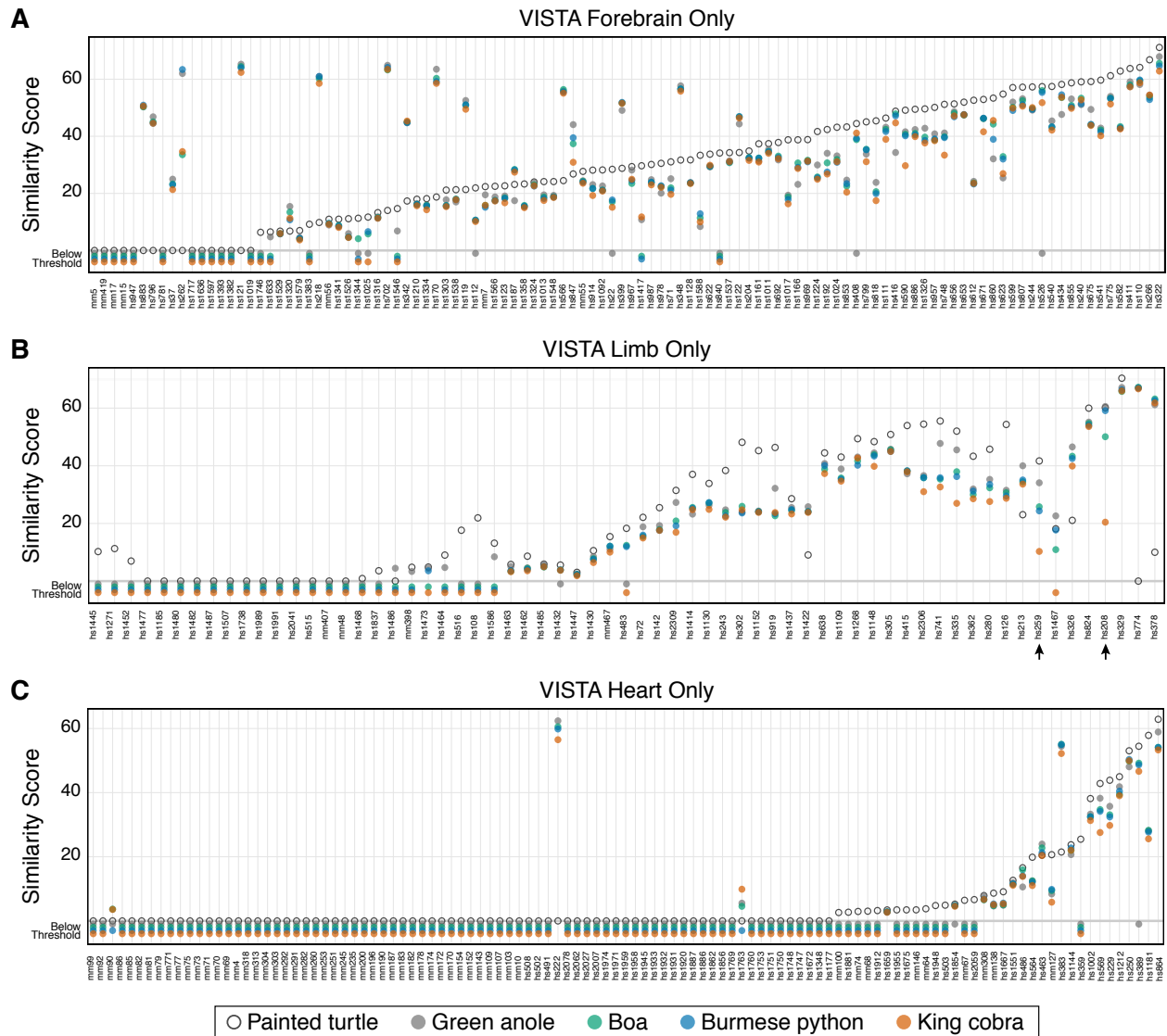
**Table S5**, related to Figure 3. Enriched sequence motifs in limb, limb-GT, and GT specific enhancers.

**Table S6**, related to Figure 4. *Anolis carolinensis* genome (anoCar2) coordinates of H3K27ac enriched regions from embryonic *Anolis* forelimbs, hindlimbs, hemiphallus.

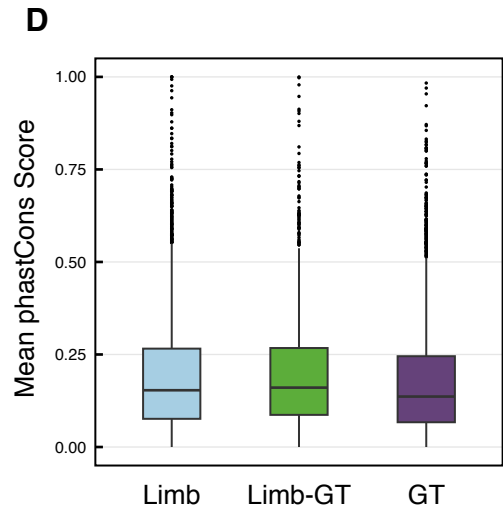
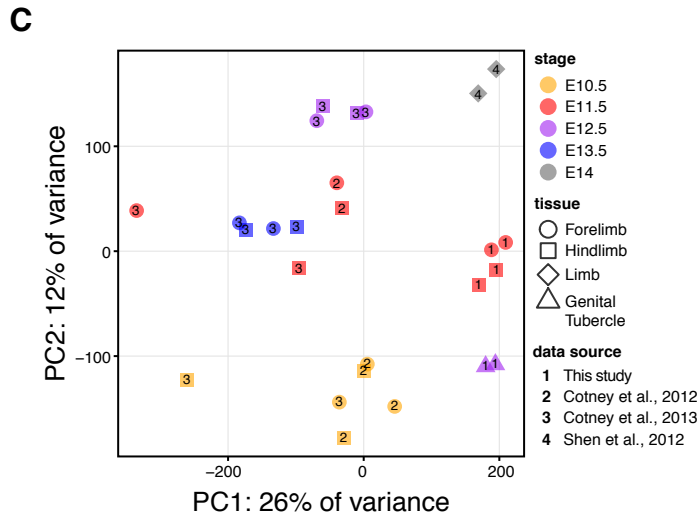
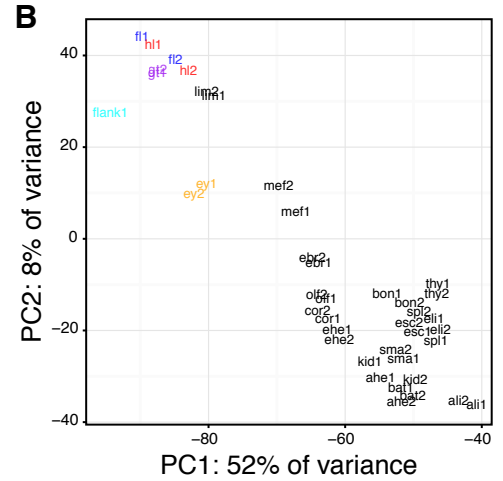
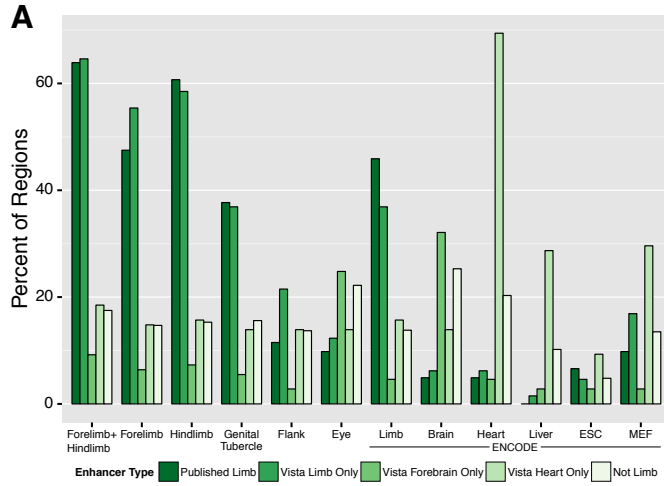
**Table S7**, related to Figure 7. The effects of hindlimb enhancer A (*DeIA*) and hindlimb enhancer B (*DeIB*) deletions on the length and width of the baculum.

### Supplementary Experimental Procedures

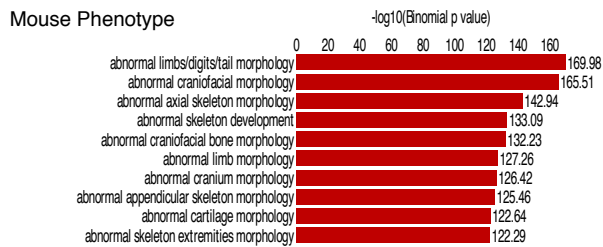
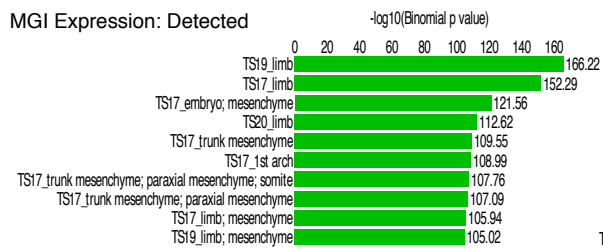
### Supplementary References



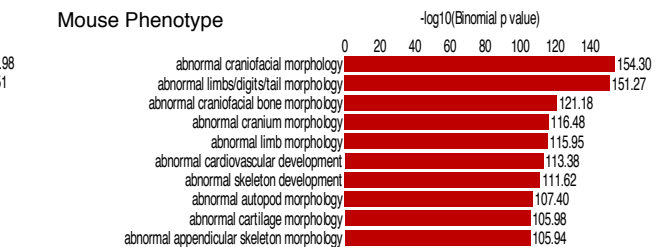
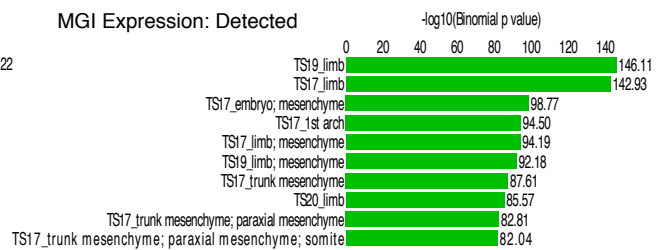
**Figure S1, related to Figure 1. Patterns of VISTA enhancer conservation across turtle, lizard, and snake genomes.** Similarity scores are shown for each VISTA Forebrain Only (A), VISTA Limb Only (B), and VISTA Heart Only (C) enhancer. An additional category, “Below Threshold”, described enhancers that were not detectable based on the alignment parameters used for the search. Black arrows in (B) point to two enhancers, hs208 and hs259, which have low conservation scores in king cobra due to gaps in the king cobra genome assembly. We PCR amplified hs208 and hs259 fragments from coral snake and corn snake and verified that these enhancer regions are not lost in other Colubroidea snakes (data not shown).



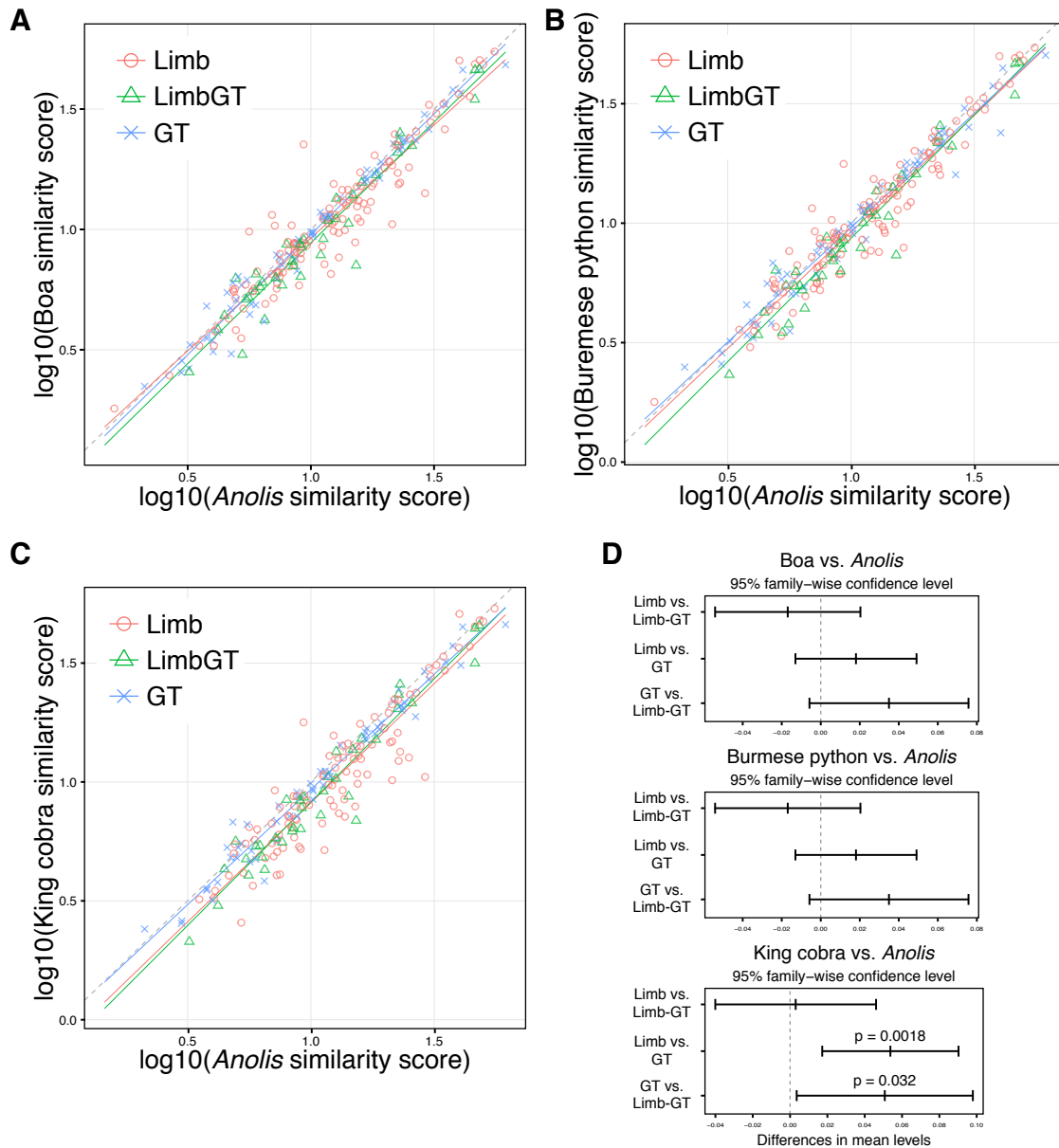
**E** Forelimb Enriched Regions



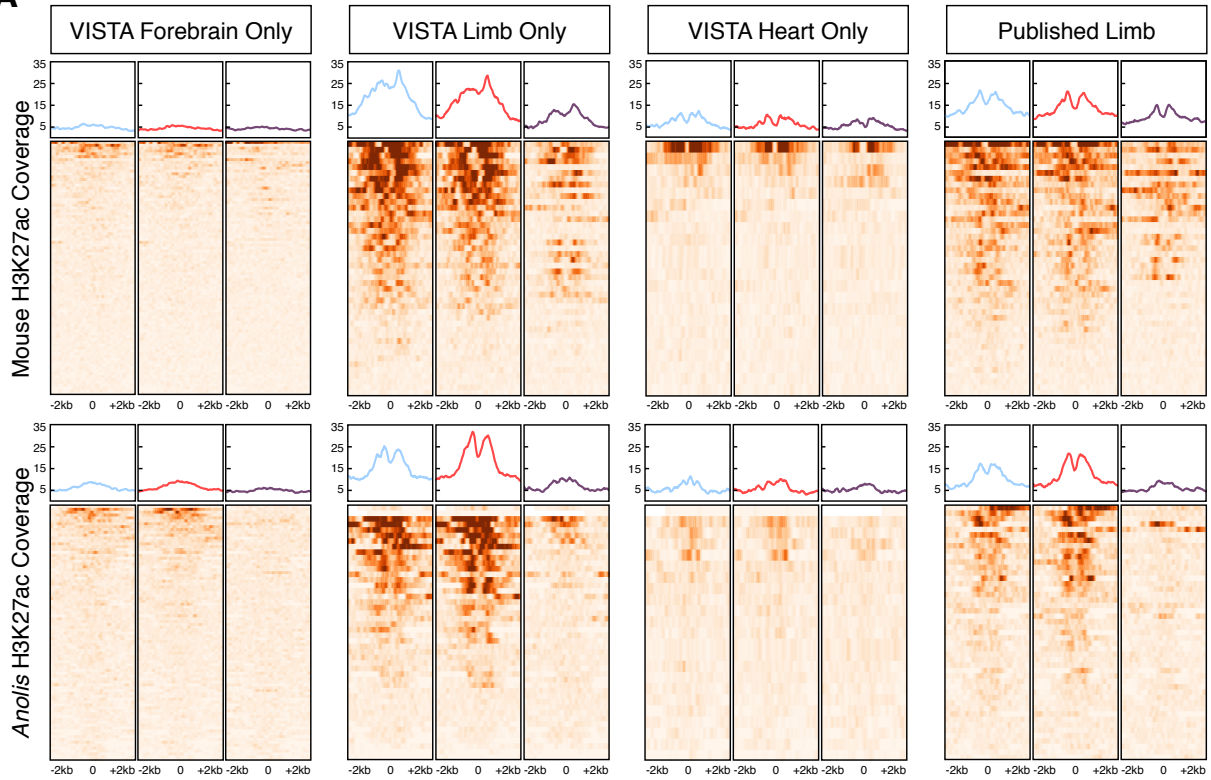
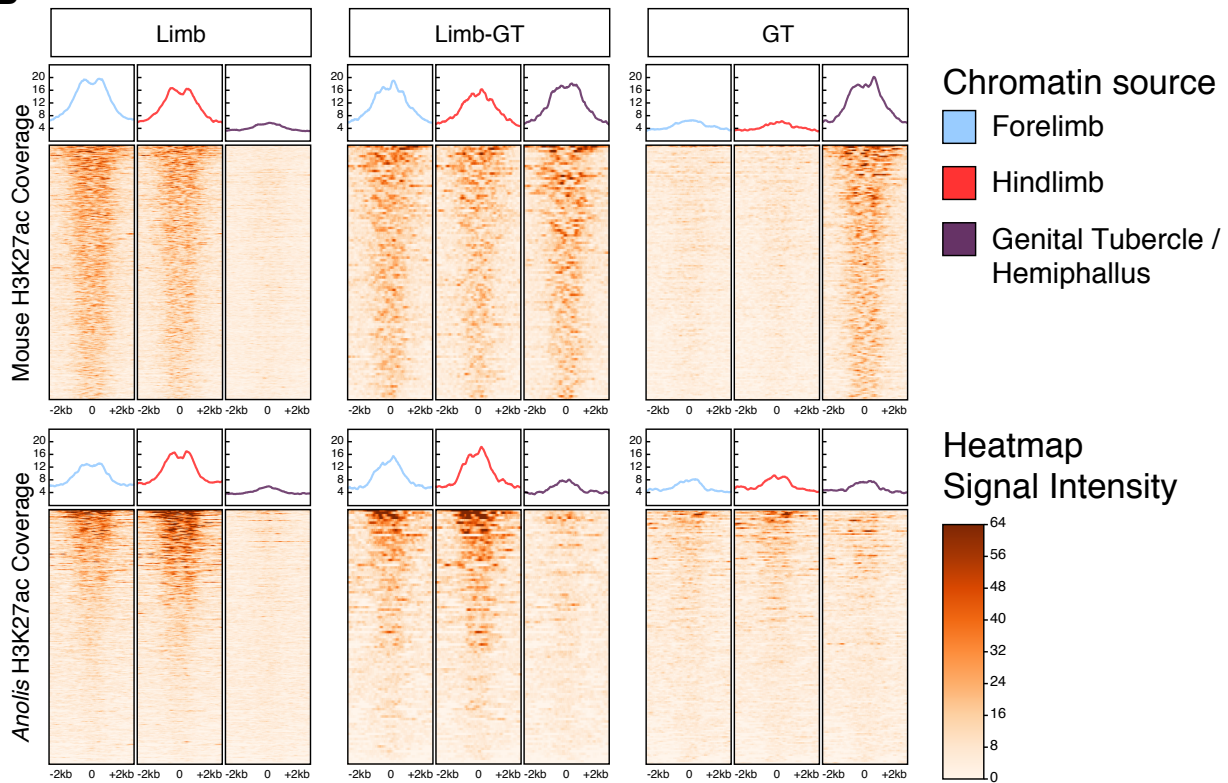
**F** Hindlimb Enriched Regions



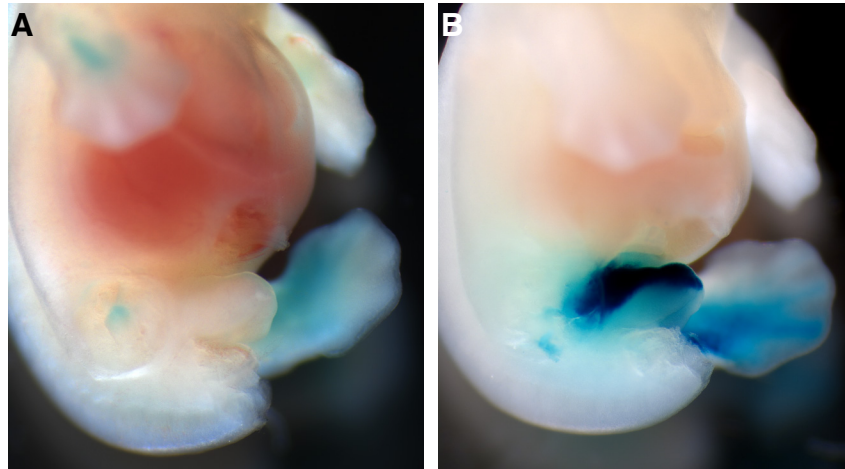
**Figure S2, related to Figure 2. H3K27ac enriched regions from mouse limb and genital tubercle share similar enhancer overlaps and signal profiles. (A)** Percent of H3K27ac enriched regions from combined E11.5 forelimb+hindlimb data, E11.5 forelimb, E11.5 hindlimb, E12.5 genital tubercle, E11.5 flank, and E10.5 eye generated by this study, and additional ENCODE tissues that overlap published limb, VISTA Limb-only, VISTA Forebrain-only, or VISTA Heart-only enhancers. The “Not Limb” category is comprised of positive VISTA enhancers that have no limb activity. **(B)** Principal component analysis (PCA) of H3K27ac signal at shared enhancer regions from E11.5 forelimb, E11.5 hindlimb, E12.5 genital tubercle, E11.5 flank, E10.5 eye generated by this study and 16 additional ENCODE tissues. Forelimb, hindlimb, and genital tubercle signal profiles group with ENCODE E14.5 limb. Abbreviations: forelimb (fl), hindlimb (hl), genital tubercle (gt), eye (ey), adult heart (ahe), adult liver (ali), adult brown adipose tissue (bat), adult bone marrow (bon), adult cortical plate (cor), embryonic E14.5 brain (ebr), embryonic E14.5 heart (ehe), embryonic E14.5 liver (eli), embryonic stem cell (esc), adult kidney (kid), embryonic E14.5 limb (lim), mouse embryonic fibroblast (mef), adult olfactory bulb (olf), adult small intestine (sma), adult spleen (spl), and adult thymus (thy). Numbers following abbreviation designate replicate number. **(C)** PCA of limb and GT H3K27ac signal at putative limb enhancers identified by combining data from this study, Cotney et al., 2012, Cotney et al., 2013, and mouse ENCODE E14.5 limb from Shen et al., 2012. PC1 is associated with technical variation between datasets. **(D)** Distribution of phastCons scores for H3K27ac regions in Limb, Limb-GT, and GT clusters. **(E & F)** The top ten MGI Expression and Mouse Phenotype enriched terms for forelimb **(E)** and hindlimb **(F)** H3K27ac regions generated using GREAT (McLean et al., 2010).



**Figure S3, related to Figure 3. Comparisons of snake and *Anolis* similarity scores for Limb, Limb-GT, and GT enhancers.** Boa (A), Burmese python (B), and King cobra (C) similarity scores compared to *Anolis carolinensis* scores for mouse enhancers. Linear regressions for each enhancer type are indicated by colored lines; gray dashed line has a slope of one through the origin. Points below gray dashed line indicate higher conservation in *Anolis*. ANOVA indicated significant difference in similarity scores only between King cobra and *Anolis*. (D) Pairwise comparisons using Tukey-HSD of similarity scores shows significant differences between King cobra and *Anolis* for GT vs. Limb ( $p=0.0018$ ) and GT vs. Limb-GT ( $p=0.032$ ), indicating a decrease in sequence conservation for Limb and Limb-GT compared to GT enhancer orthologs in the King cobra genome compared to *Anolis* orthologs.

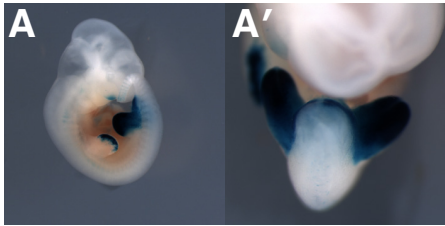
**A****B**

**Figure S4, related to Figure 4. Comparison of H3K27ac signal from mouse and *Anolis carolinensis* forelimb, hindlimb, and genital tubercle or hemiphallus tissues.** For each enhancer category, a profile plot of the average H3K27ac enrichment across all orthologous regions is shown above a heatmap of H3K27ac signal at individual enhancer regions for mouse forelimb, hindlimb, and genital tubercle and *Anolis* forelimb, hindlimb, and hemiphallus. **(A)** For the VISTA Forebrain Only, VISTA Limb Only, VISTA Heart Only, and Published Limb enhancers, signal intensities for each tissue are similar in mouse and *Anolis*, with forelimb and hindlimb samples showing enrichment in both VISTA Limb Only and the Published Limb enhancer datasets. **(B)** H3K27ac signal at shared regions from the Limb, Limb-GT, and GT enhancers identified in this study. Profile plots and heatmaps generated with deepTools (Ramirez et al., 2014).



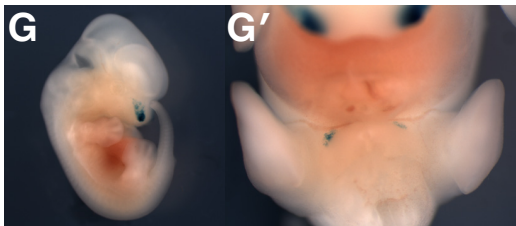
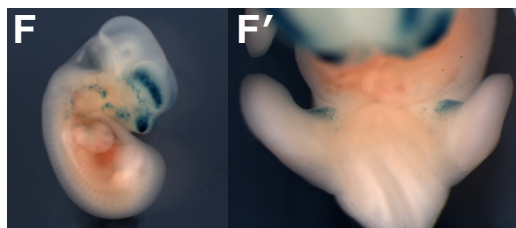
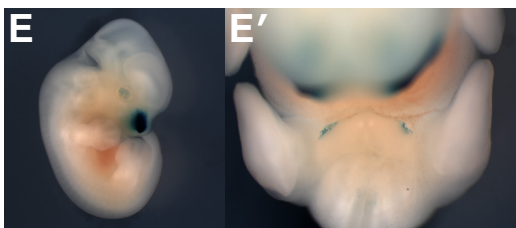
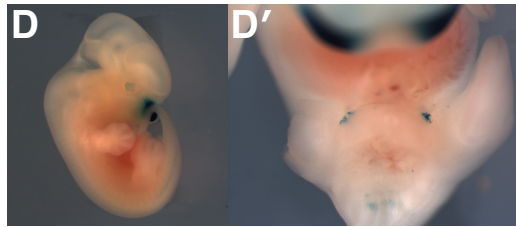
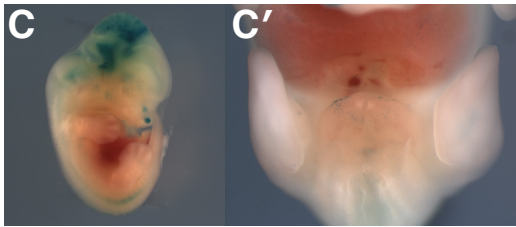
**Figure S5, related to Figure 5. Mouse embryos carrying *HLEA* and *HLEB Hsp68LacZ* transgenes display different staining patterns in the genital tubercle (GT). (A) No staining is observed in the GT of *HLEA Hsp68LacZ* embryos at E12.5. (B) *HLEB Hsp68LacZ* embryos display intense staining in the dorsal half of the GT.**





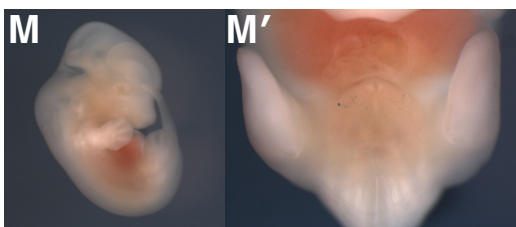
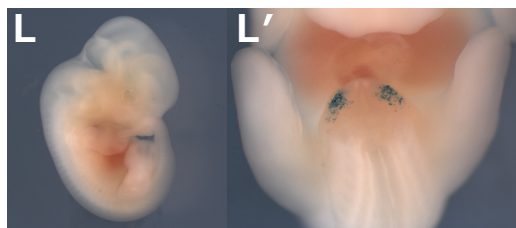
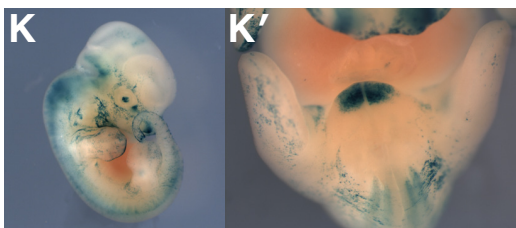
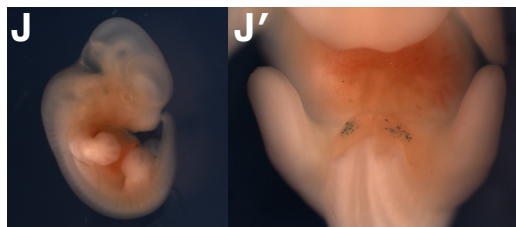
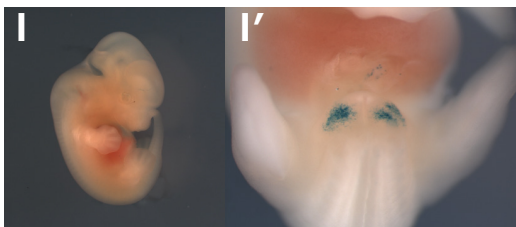
**B** *Anolis carolinensis* HLEB

| Hindlimb | Genital Tubercle | Nasal Stripe |
|----------|------------------|--------------|
| 2/2      | 2/2              | 0/2          |



**H** Burmese python HLEB

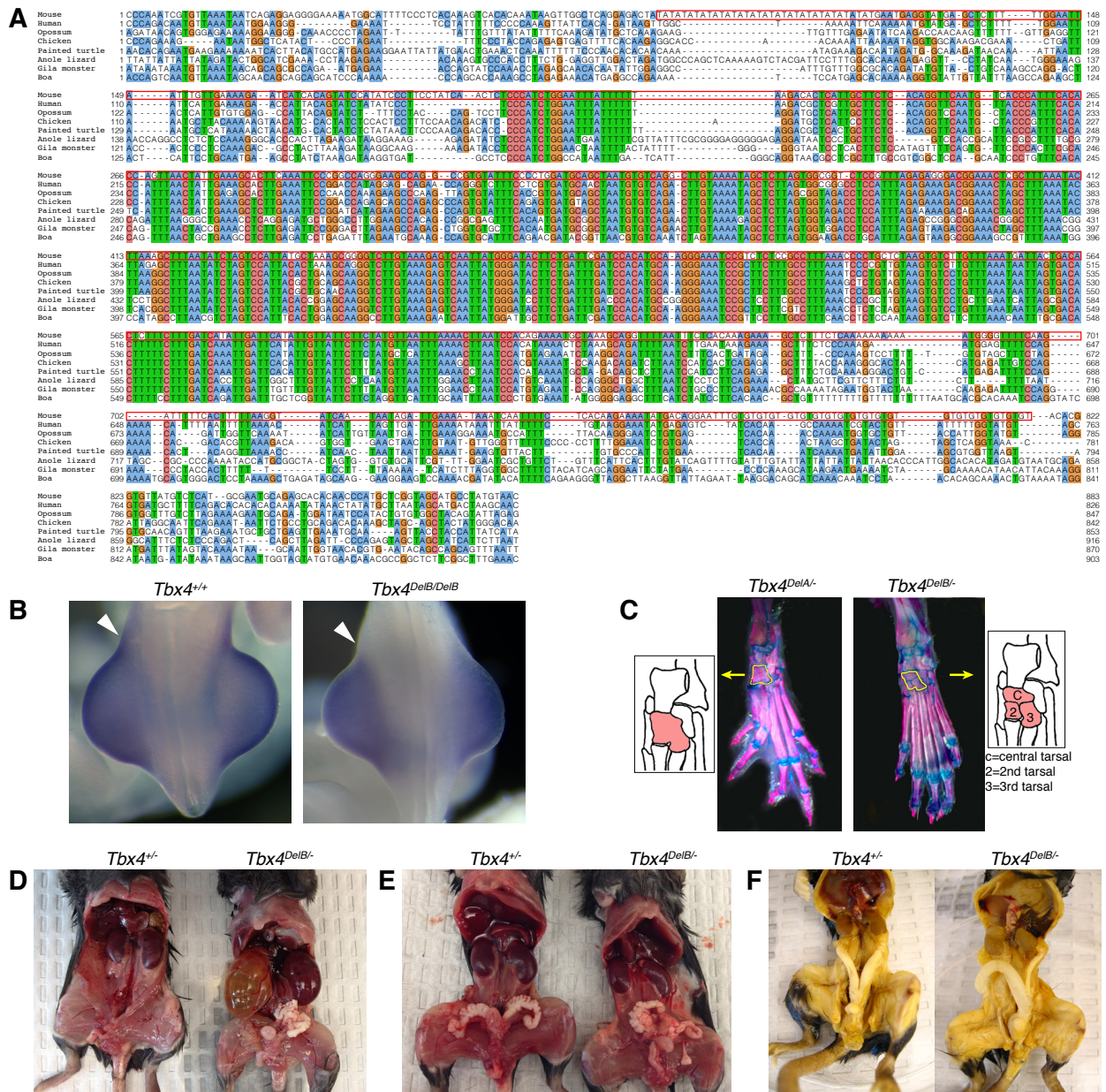
| Hindlimb | Genital Tubercle | Nasal Stripe |
|----------|------------------|--------------|
| 0/10     | 5/10             | 6/10         |



**N** King cobra HLEB

| Hindlimb | Genital Tubercle | Nasal Stripe |
|----------|------------------|--------------|
| 0/12     | 6/12             | 0/12         |

**Figure S6, related to Figure 6. LacZ positive E12.5 mouse embryos carrying squamate HLEB-Hsp68LacZ transgenes\*.** (A & A') Side and ventral views of a transgenic mouse embryo carrying the *Anolis carolinensis* HLEB-Hsp68LacZ transgene. (B) Number of *Anolis carolinensis* HLEB-Hsp68LacZ transgenic embryos with LacZ staining in hindlimb, genital tubercle, and nasal region over the total number of transgenic embryos obtained. (C-G and C'-G') Side and ventral views of transgenic mouse embryos carrying the *Burmese python* HLEB-Hsp68LacZ transgene. (H) Number of *Burmese python* HLEB-Hsp68LacZ transgenic embryos with LacZ staining in hindlimb, genital tubercle, and nasal region over the total number of transgenic embryos obtained. (I-M and I'-M') Side and ventral views of transgenic mouse embryos carrying the *King cobra* HLEB-Hsp68LacZ transgene. (N) Number of *King cobra* HLEB-Hsp68LacZ transgenic embryos with LacZ staining in hindlimb, genital tubercle, and nasal region over the total number of transgenic embryos obtained. \*LacZ positive embryos pictured in Figure 6 are not shown.



**Figure S7**, related to **Figure 7**. **Location of HLEB deletion and phenotypes in *DeIA*<sup>-/-</sup> and *DeIB*<sup>-/-</sup> mice.** (A) Multi-species alignment of HLEB region from mammals, bird, turtle, and squamates. The region deleted from the mouse *DeIB* allele is outlined in red and encompasses all mammal-squamate conserved sequences. (B) Whole-mount mRNA in situ hybridization for *Tbx4* on E10.5 wild-type and *DeIB/DeIB* embryos. A modest decrease in *Tbx4* expression is apparent in *DeIB/DeIB* embryos and is most noticeable in the anterior most location of the hindlimb bud (white arrowheads). (C) Fusion between the tarsal bones and fusions at the base of the 1st and 2nd metatarsals occur in *DeIA*<sup>-/-</sup>, but not *DeIB*<sup>-/-</sup> animals. (D) *DeIB*<sup>-/-</sup> mouse presenting with severe hydronephrosis of the right kidney, and hyperplasia in the left kidney. (E) *DeIB*<sup>-/-</sup> male with malformed seminal vesicles. (F) *DeIB*<sup>-/-</sup> female with underdeveloped left uterine

horn. Incidence of *De/B/-* kidney phenotypes: Kidney agenesis 3 of 12, kidney hypoplasia/hydronephrosis 2 of 12, malformed kidney 1 of 12; Incidence of reproductive tissue phenotypes in *De/B/-* males: Seminal vesicle overgrowth 6 of 12, seminal vesicle multiple horns 8 of 12, Distended bulbourethral gland 8 of 12, Absent bulbourethral gland 1 of 12; Incidence of reproductive tissue phenotypes in *De/B/-* females: Bifurcated vagina 8 of 10, Imperforate vagina 2 of 10, Incomplete vaginal canal 2 of 10, Distended uterine horns, Underdeveloped uterine horn 2 of 10.

**Table S1, related to Figure 1. Mouse genome (mm9) coordinates of VISTA forebrain-only, VISTA heart only, VISTA limb-only, and published limb enhancers.** Table listing the chromosome, start, end, and enhancer name for VISTA enhancers downloaded 24 January 2014 and published limb enhancers compiled from the literature as of 31 July 2014. References for the published limb enhancers are also included.

**Table S2, related to Figure 2. H3K27ac ChIP-Seq read counts, quality metrics, and data processing results.**

| Tissue                    | Replicate | Total Reads | Aligned Reads | NSC <sup>a</sup> | RSC <sup>b</sup> | Enriched Regions | Merged Regions | Combined Regions |
|---------------------------|-----------|-------------|---------------|------------------|------------------|------------------|----------------|------------------|
| Mouse Forelimb            | ChIP 1    | 48,918,246  | 27,450,425    | 1.04             | 1.01             | 40,303           | 18,404         | 14,071           |
|                           | ChIP 2    | 63,450,153  | 45,173,555    | 1.05             | 1.51             | 45,170           | 23,909         |                  |
|                           | Input     | 49,495,329  | 30,772,420    |                  |                  |                  |                |                  |
| Mouse Hindlimb            | ChIP 1    | 56,879,507  | 34,546,836    | 1.04             | 1.01             | 35,166           | 16,148         | 13,317           |
|                           | ChIP 2    | 57,348,471  | 41,470,307    | 1.05             | 1.36             | 44,951           | 23,300         |                  |
|                           | Input     | 55,651,083  | 32,727,378    |                  |                  |                  |                |                  |
| Mouse Genital Tubercle    | ChIP 1    | 32,834,034  | 24,242,884    | 1.04             | 1.55             | 34,531           | 19,231         | 15,718           |
|                           | ChIP 2    | 31,985,957  | 23,929,270    | 1.06             | 1.85             | 38,867           | 21,250         |                  |
|                           | Input     | 34,529,785  | 23,544,688    |                  |                  |                  |                |                  |
| Mouse Eye                 | ChIP 1    | 62,171,724  | 44,150,100    | 1.03             | 1.58             | 34,382           | 19,263         | 11,549           |
|                           | Input 1   | 62,230,024  | 42,451,834    |                  |                  |                  |                |                  |
|                           | ChIP 2    | 64,813,891  | 45,749,616    | 1.02             | 1.54             | 22,947           | 13,899         |                  |
|                           | Input 2   | 61,179,959  | 41,547,551    |                  |                  |                  |                |                  |
| Mouse Flank               | ChIP      | 62,318,253  | 44,255,495    | 1.01             | 0.95             | 19,261           | 11,875         |                  |
|                           | Input     | 64,330,413  | 44,800,173    |                  |                  |                  |                |                  |
| <i>Anolis</i> Forelimb    | ChIP      | 68,259,468  | 42,467,374    | 1.06             | 1.13             | 47,573           | 28,528         |                  |
|                           | Input     | 69,008,637  | 41,292,208    |                  |                  |                  |                |                  |
| <i>Anolis</i> Hindlimb    | ChIP      | 84,899,388  | 55,848,522    | 1.08             | 1.23             | 53,464           | 31,553         |                  |
|                           | Input     | 68,802,300  | 36,517,301    |                  |                  |                  |                |                  |
| <i>Anolis</i> Hemiphallus | ChIP      | 62,957,995  | 37,187,831    | 1.02             | 0.86             | 23,906           | 15,381         |                  |
|                           | Input     | 69,105,710  | 42,013,212    |                  |                  |                  |                |                  |

a. Normalized strand cross-correlation coefficient; values less than 1.05 indicate low signal to noise ratio in the dataset.

b. Relative strand cross-correlation coefficient; values less than 0.8 indicate low signal to noise ratio in the dataset.

**Table S3, related to Figure 2. Mouse genome (mm9) coordinates of H3K27ac enriched regions and putative enhancers from embryonic mouse forelimbs, hindlimbs, genital tubercle, eye, and flank.** Table listing the chromosome, start, end, name, MACS2 score, and strand of enriched regions for all forelimb, hindlimb, genital tubercle, eye, and flank replicates. Putative forelimb, hindlimb, genital tubercle, eye, and flank enhancers also include associated genes determined by GREAT (McLean et al., 2010).

**Table S4, related to Figure 2. Mouse genome (mm9) coordinates of H3K27ac limb, limb-GT, GT, eye, and flank K-means clusters. Also contains coordinates for limb specific, limb-GT specific, and GT specific H3K27ac enhancers.** Table listing the chromosome, start, end, name, score, and strand of limb, limb-GT, eye, and flank enhancers grouped by kmeans clustering, as well as filtered subsets of limb specific, limb-GT specific, and GT specific enhancers. All regions also include associated genes determined by GREAT (McLean et al., 2010).

**Table S5, related to Figure 3. Enriched sequence motifs in limb, limb-GT, and GT specific enhancers.** Table listing the HOMER de novo motif and known motif search output for limb, limb-GT, and GT specific enhancers.

**Table S6, related to Figure 4. *Anolis carolinensis* genome (anoCar2) coordinates of H3K27ac enriched regions from embryonic *Anolis* forelimbs, hindlimbs, hemiphallus.** Table listing the chromosome, start, end, name, MACS2 score, and strand of enriched regions identified from *Anolis* forelimb, hindlimb, and hemiphallus H3K27ac CHIP-Seq datasets.

**Table S7, related to Figure 7. The effects of hindlimb enhancer A (*DeIA*) and hindlimb enhancer B (*DeIB*) deletions on the length and width of the baculum.**

| Genotype         | n  | Length                     |         |                              | Width                      |         |                              |
|------------------|----|----------------------------|---------|------------------------------|----------------------------|---------|------------------------------|
|                  |    | Mean $\pm$ SD <sup>a</sup> | % of Wt | <i>p</i> -value <sup>b</sup> | Mean $\pm$ SD <sup>a</sup> | % of Wt | <i>p</i> -value <sup>b</sup> |
| Wildtype (Wt)    | 18 | .384 $\pm$ .011            |         |                              | .140 $\pm$ .007            |         |                              |
| <i>DeIA/DeIA</i> | 12 | .390 $\pm$ .010            | 101.7   | n.s.                         | .140 $\pm$ .007            | 100.1   | n.s.                         |
| <i>DeIB/DeIB</i> | 13 | .368 $\pm$ .011            | 96.5    | .00074                       | .135 $\pm$ .005            | 96.3    | .029                         |
| Wt/Null          | 10 | .364 $\pm$ .012            | 94.9    | .00044                       | .142 $\pm$ .005            | 101.1   | n.s.                         |
| <i>DeIA/Null</i> | 13 | .378 $\pm$ .012            | 98.6    | n.s.                         | .141 $\pm$ .005            | 100.6   | n.s.                         |
| <i>DeIB/Null</i> | 11 | .332 $\pm$ .012            | 86.7    | 2.67 x 10 <sup>-10</sup>     | .103 $\pm$ .004            | 7.35    | 6.76 x 10 <sup>-13</sup>     |

<sup>a</sup> Measurement means  $\pm$  standard deviation are normalized by dividing by the length of the humerus and presented as a ratio.

<sup>b</sup> *p*-value indicates a significant difference from the wildtype measurements using Student's t-test.

## Supplemental Experimental Procedures

### **Mouse strains**

The *DeiB* allele was produced through homologous recombination using an F1 hybrid (C57BL/6 x 129SvEv) ES cell line from inGenious Targeting Laboratory (Ronkonkoma, NY). The targeting construct (pLIZR) was created using C57BL/6J genomic DNA and included a 7.5kb 5' homology arm (mm10: chr11:85,955,841-85,963,352) and a 7.2kb 3' homology arm (mm10: chr11:85,964,085-85,971,324). A 732 bp region of HLEB (mm10: chr11:85,963,353-85,964,084) was replaced with orthologous sequences from *Anolis* species and a Neo selection cassette from plasmid PL451 (Liu, 2003). The entire 18.1kb targeting construct was resequenced to ensure that no unintended mutations were introduced during gene targeting. ES cells were electroporated with the linearized targeting construct by inGenious Targeting Laboratory. An initial PCR screen revealed that 152 of 288 ES clones were homologously targeted. Southern blotting with 5' and 3' external probes was used to verify homologously targeted clones. The 732 bp deletion removed two small microsatellites that both differ in size between C57BL/6 and 129SvEV. This allowed us to determine whether the 129 or B6 allele was targeted in each ES cell clones. Two ES cells clones with targeted B6 alleles were injected into blastocysts and one of these yielded chimeras that transmitted the targeted allele when crossed to wild-type C57BL/6J females. The final *DeiB* allele was generated by crossing the targeted allele to the *Ella-Cre* deleter strain (JAX # 003724). This removed the *Anolis* sequences and the Neo selection cassette, leaving behind a *LoxP* and an *FRT* site. Mouse phenotypes were analyze after backcrossing to the equivalent of N4 on a C57BL/6J genetic background.

The *Anolis HLEB Hsp68LacZ* transgene was created by PCR amplifying a 980bp region from *Anolis carolinensis* genomic DNA. This region encompasses all mammal-squamate conserved sequences (AnoCar2.0: chrUn\_GL343932:29,865-30,844). To ensure we captured the full enhancer activity of the snake HLEB orthologs we amplified a 5.5kb region from king cobra genomic DNA and an orthologous 4.8kb region from Burmese python DNA to create the *king cobra HLEB Hsp68LacZ* and *Burmese python HLEB Hsp68LacZ* transgenes, respectively. The conserved HLEB core is located within the central portion of these snake PCR fragments. All transgenes were created by cloning PCR fragments into the *NotI* site of *p5'-Not-HspLacZ* (DiLeone et al., 1998). Transgene constructs were linearized and injected into fertilized mouse oocytes by Cyagen Biosciences (Santa Clara, CA) to produce transient F0 transgenic embryos.

HLEB Primer sequences for *Anolis carolinensis*, cobra and python:

ACAR HLEB-NotI-F1 5'-AATTAGCGGCCGCACAAAGTGCCACCTTTCTGG-3'  
ACAR HLEB-NotI-R1 5'-AATAAGCGGCCGCATGACAGCTGATGTGTGCCAA-3'

Cobra HLEB Region-NotI-F2 5'-  
TTATTGCGGCCGCTAGAAGCTTGCTTCTTCAAGTCAGG-3'

Cobra HLEB Region-NotI-R2 5'-  
TTATTGCGGCCGCAACCTTTGTATGTGAAGTTTGG-3'



Python HLEB Region-NotI-F3 5'-  
ATTATGCGGCCCGCCAATGTGGACTAGAAAGCTTGCTTC-3'

Python HLEB Region-NotI-R3 5'-  
ATTATGCGGCCCGCTGAACAGCAGAGGTCGTCGAGAC-3'

### **Allele-Specific Gene Expression Analysis**

Relative levels of expression from wild-type and enhancer deletions alleles were assessed using melting curve analysis with FRET probes as reported (Jeong et al., 2007). *Tbx4* FRET probe pair follows: 5'-ACATCAAAGTGGGGCTGCATGAGAA/36-FAM/-3' and 5'-/5Cy5/AGCTGTGGAAGAAGTTCCACGAGGCAGG/3Phos/-3'

### **Generation of Gila Monster *Tbx4* Genomic Sequence**

Nylon filters spotted with clones from a Gila monster (*Heloderma suspectum*) BAC library were obtained from the Joint Genome Institute (Walnut Creek, CA; Wang et al., 2006). *Tbx4* containing BACs were identified through hybridization with *Tbx4* overgo probes using an established protocol (Han et al., 2000). Three overlapping *Tbx4* BACs (clones 248\_M\_11, 297\_I\_15, and 386\_L\_11) were purified and then sequenced on a Roche 454 FLX sequencing platform at the University of Georgia Genomics Facility. BAC sequence reads were assembled with MIRA (<http://sourceforge.net/projects/mira-assembly>) and Newbler (<http://454.com/contact-us/software-request.asp>) and the results were combined to generate a consensus assembly of all three BACs. PCR was used to resolve assembly gaps.

## **Supplemental References**

- Han, C.S., Sutherland, R.D., Jewett, P.B., Campbell, M.L., Meincke, L.J., Tesmer, J.G., Mundt, M.O., Fawcett, J.J., Kim, U.J., Deaven, L.L., Doggett, N.A. (2000). Construction of a BAC contig map of chromosome 16q by two-dimensional overgo hybridization. *Genome Res.* 10, 714-721.
- Jeong, S., Hahn, Y., Rong, Q., and Pfeifer, K. (2007). Accurate quantitation of allele-specific expression patterns by analysis of DNA melting. *Genome Res.* 17, 1093–1100.
- Liu, P. (2003). A Highly Efficient Recombineering-Based Method for Generating Conditional Knockout Mutations. *Genome Res* 13, 476–484.
- Wang, Z., Miyake, T., Edwards, S.V., and Amemiya, C.T. (2006). Tuatara (*Sphenodon*) genomics: BAC library construction, sequence survey, and application to the DMRT gene family. *J. Hered.* 97, 541-548.