Supplementary Methods

Feeding Experiments

Prairie Rattlesnakes (*Crotalus viridis viridis*; Colorado: 17HP0974 to SPM) and Diamondback Watersnakes (*Nerodia rhombifer*; Mississippi Scientific Collecting Permit for 2015; No. 0508152 – to SMS) were wild-caught under state collecting permits. Animal care and tissue sampling was conducted under protocols approved by the Institutional Animal Care and Use Committee at the University of Alabama (14-06-0075) and the University of Northern Colorado (1701D-SM-S-20). Individuals from both species were sampled at three time points to target three distinct physiological states: fasted (30 days since last meal), 1 day post-feeding (DPF), and 4 DPF. Watersnakes were fed catfish filets and rattlesnake were fed adult mice in order to approximate their natural primary prey type (fish and mammals, respectively). Consumed meals were equal in mass to 25% of individual snake body mass. Snakes were humanely euthanized by severing the spinal cord immediately posterior to the head. Intestinal tissue was immediately extracted, snap frozen in liquid nitrogen, and stored at -80°C. Between three and four individual animals per species were sampled for each time point.

Transcriptomic Data Generation

Transcriptomic data for *C. viridis* and *N. rhombifer* were generated for this study. Transcriptomic data for the Burmese Python (*P. m. bivittatus*) was generated previously using protocols described in (Andrew et al. 2015; Andrew et al. 2017). Total RNA was extracted by placing ~50mg of snap-frozen tissue into 1mL of Trizol Reagent

(Invitrogen), mechanically lysing cells using a TissueLyzer for 10 minutes at 20 strokes/minute, and precipitating RNA from the aqueous phase using isopropanol. Individual Illumina mRNA-seq libraries were constructed using either a NEB Next RNAseq kit with poly-A selection, RNA fragmentation, cDNA synthesis, and indexed Illumina adapter ligation. RNAseq libraries were quantified using a BioAnalyzer (Agilent), pooled in equal molar ratios in various multiplex arrangements, and sequenced on an Illumina HiSeq 2500 using 100bp paired-end reads (Supplementary Table S1). Newly-generated transcriptomic data for *C. v. viridis* and *N. rhombifer* is archived at the NCBI Short Read Archive (NCBI: SRP200900) as well as previously-generated transcriptomic data for *P. m. bivittatus* (NCBI: SRP051827).

Identifying homologous genes between snakes and human

Downstream pathway and regulatory molecule analyses require gene expression data with human gene identifiers. To identify homologous genes between the snake species and human and ultimately assign human gene identifiers to snake genes for downstream analysis, OrthoMCL (Li et al. 2003) was run via the Orthomcl-pipeline (https://github.com/apetkau/orthomcl-pipeline) using species-specific protein fasta files as input. In cases where multiple isoforms were annotated for a given gene, the protein sequence corresponding to the longest coding sequence was used. For resulting OrthoMCL homolog groups containing a single gene from each of the four species, the identifier of the human gene was assigned to each snake gene within the group. For groups containing a single human gene identifier but more than one orthologous gene in one or more snake species, the human gene identifier was

assigned to all snake genes in the group. Lastly, for groups containing multiple human orthologs, each snake gene in the group was assigned the human identifier that produced the best one-way BLAST hit to that gene during the orthoMCL pipeline.

Quantifying and visualizing gene expression

Raw demultiplexed Illumina RNAseq reads were quality filtered and trimmed with Trimmomatic v. 0.32 (Bolger et al. 2014). All reads were mapped using STAR v.2.5.3a (Dobin et al. 2013) in basic two pass mode with --outFilterMultimapNmax set to 1 to exclude reads that mapped to multiple positions. Reads for P. m. bivittatus and C. v. viridis were mapped to their respective genome assemblies ((Castoe et al. 2013) and (Schield et al. 2019)); as no genome assembly is currently available for *N. rhombifer*, these reads were mapped to genome of the closely-related eastern gartersnake (*Thamnophis sirtalis*; (Perry et al. 2018)). Raw expression counts were determined using featureCounts v1.6.3 (Liao et al. 2013).

Count normalization and pairwise exact tests of differential expression between fasted versus 1DPF and 1DPF versus 4DPF were performed using DEseq2 v. 1.12.4 (Love et al. 2014). As transcriptomic data for *P. m. bivittatus* included both single-end 50bp and paired-end 120bp reads, read type was included as a factor in DEseq2 pairwise comparisons to account for any potential batch effects. Independent hypothesis weighting was applied to pairwise test results using the IHW package (Ignatiadis et al. 2016) with baseMean as a covariate.

The results of pairwise comparisons in each species were filtered to exclude genes that were not assigned a homologous gene identifier via OrthoMCL. We further filtered pairwise results to only include genes for which there was detectable expression (>= 1 raw expression count) in at least one time point in all three snake species.

Heatmaps were generated using the pHeatmap package (Kolde 2012) and alluvial plots of patterns of gene expression were generated using the ggalluvial package (Brunson 2018) in R (Team 2014).

Proteomic Data Generation and Analysis

Proteins were extracted using T-PER Tissue Protein Extraction Reagent (Thermo Fisher, 78510) from multiple fasted, 1 DPF, and 4DPF proximal small intestine tissue samples for the python and watersnake (Supplementary Table S4). Proteins were quantified using a BCA assay, purified, and digested with trypsin. The dried pellet was resuspended in 50 mM NH_4CO_3. Following Kamal et al. 2018 (Kamal et al. 2018), proteins were reduced and alkylated, then digested with Trypsin (MS Grade) at a 1:50 enzyme/protein concentration for 16 h at 37 °C. Formic acid (pH < 3) was added to the resulting peptides for acidifying the sample. A C_{18} desalting column (ThermoFisher Scientific, IL, USA) was used for desalting the samples. After drying by speed vacuum, peptides were dissolved in 0.1% formic acid, and stored at -20°C.

Digested peptides were analyzed by nano-LC-MS/MS using a Velos Pro Dual-Pressure Linear Ion Trap Mass Spectrometer (ThermoFisher Scientific, MA) coupled to an UltiMate 3000 UHPLC (ThermoFisher Scientific, MA). Peptides were loaded onto the

analytical column and separated by reversed-phase chromatography using a 15-cm column (Acclaim PepMap RSLC) with an inner diameter of 75 µm and packed with 2 µm C_{18} particles (Thermo Fisher Scientific, MA). The peptide samples were eluted from the Nano column with multi-step gradients of 4-90% solvent B (A: 0.1% formic acid in water; B: 95% acetonitrile and 0.1% formic acid in water) over 70 min with a flow rate of 300 nL/min with a total run time of 90 min. The mass spectrometer was operated in positive ionization mode with nano spray voltage set at 2.50 kV and source temperature at 275°C. The three precursor ions with the most intense signal in a full MS scan were consecutively isolated and fragmented to acquire their corresponding MS2 scans. Full MS scans were performed with 1 micro scan at resolution of 3000, and a mass range of m/z 350-1500. Normalized collision energy (NCE) was set at 35%. Fragment ion spectra produced via high-energy collision-induced dissociation (CID) was acquired in the Linear Ion Trap with a resolution of 0.05 FWHM (full-width half maximum) with an Ultra Zoom-Scan between *m/z* 50-2000. A maximum injection volume of 5 µl was used during data acquisition with partial injection mode. The mass spectrometer was controlled in a data-dependent mode that toggled automatically between MS and MS/MS acquisition. MS/MS data acquisition and processing were performed by Xcalibur™ software (ThermoFisher Scientific, MA).

Spectra were searched using Proteome Discoverer software (ver. 2.0, ThermoFisher Scientific) against species-specific protein databases generated from the genome of the Burmese Python (Castoe et al. 2013) and the genome of the garter snake *Thamnophis sirtalis* (Perry et al. 2018), which was used as a reference set for the watersnake as it is the most closely-related snake species with a complete genome assembly and

annotation. The considerations in SEQUEST searches for normal peptides were used with carbamidomethylation of cysteine as the static modification and oxidation of methionine as the dynamic modification. Trypsin was indicated as the proteolytic enzyme with two missed cleavages. Peptide and fragment mass tolerance were set at ± 1.6 and 0.6 Da and precursor mass range of 350-5000 Da, and peptide charges were set excluding +1. SEQUEST results were filtered with the target PSM validator to improve the sensitivity and accuracy of the peptide identification. Using a decoy search strategy, target false discovery rates for peptide identification of all searches were < 1% with at least two peptides per proteins, and the results were strictly filtered by ΔCn (< 0.01), Xcorr (≥ 1.5) for peptides, and peptide spectral matches (PSMs) with high confidence (q-value of ≤ 0.05). Protein quantification was conducted using the total spectrum count of identified proteins. Additional criteria were applied to increase confidence that PSMs must be present in all three biological replicates samples. Protein identifiers from the rattlesnake and garter snake genomes were converted to orthologous python identifiers using reciprocal best BLASTp, followed by reciprocal best tBLASTx and, finally, stringent one-way BLASTp. Peptide spectrum matches (PSM) were normalized and analyzed using DEseq2 (Love et al. 2014) in R to identify proteins that exhibited significant changes in abundance between time points (Benjamini-Hochberg corrected p-value < 0.1). Differentially expressed proteins were characterized using GO term overrepresentation analysis against all protein coding genes using ClueGO (Bindea et al. 2009) and WebGestalt (Zhang et al. 2005).

Inferences of canonical pathway and upstream regulatory molecule activity

To infer canonical pathways and regulatory molecules that may be driving observed patterns of gene expression in these three species, and specifically those that may be driving regenerative growth, we employed a multi-level comparative approach to analyze particular subsets of differentially expressed genes ($p < 0.05$) using Core Analysis in Ingenuity Pathway Analysis software (Krämer et al. 2013). First, we used all differentially expressed genes from pairwise analyses as input. Second, to distinguish between mechanisms that are uniquely shared between regenerating species and those that are shared between all three species, we compared inferences generated from genes that were differentially expressed in both the python and rattlesnake to inferences generated from genes that were differentially expressed in all three species. To compare evidence of pathway and regulatory activity inferred from gene expression and protein abundance, we also used Core Analysis in IPA to analyze proteins that demonstrated significant changes in abundance between Fasted and 4DPF ($p <0.1$). Networks of overlapping genes between pathways was generated using the GGally (Barret Schloerke 2018) and network (Butts 2008; Handcock et al. 2008) packages in R (Fig. 3). In these networks, pathways were connected if at least 50% of the genes underlying the prediction of one of the pathway overlap with the genes underlying the prediction of the other pathway. Pathways that did not exhibit this overlap with at least one additional pathway are shown as unconnected nodes. Networks were manually annotated to group pathways based on similar biological function.

Supplementary Table 1. Sample and sequencing information for rattlesnake and watersnake small intestine gene expression data generated for this study.

Species	Time point	Animal ID	Instrument	cDNA Prep Kit	SRA Accession
Prairie Rattlesnake (*Crotalus viridis viridis*)	Fasted	CV1	HiSeq	NEB Next	SAMN12003898
	Fasted	CV2	HiSeq	NEB Next	SAMN12003899
	Fasted	CV4	HiSeq	NEB Next	SAMN12003900
	Fasted	CV7	HiSeq	NEB Next	SAMN12003901
	1DPF	CV3	HiSeq	NEB Next	SAMN12003902
	1DPF	CV6A	HiSeq	NEB Next	SAMN12003903
	1DPF	CV8	HiSeq	NEB Next	SAMN12003904
	4DPF	CV9	HiSeq	NEB Next	SAMN12003905
	4DPF	CV10	HiSeq	NEB Next	SAMN12003906
	4DPF	CV11	HiSeq	NEB Next	SAMN12003907
	4DPF	CV12	HiSeq	NEB Next	SAMN12003908
Diamondback Watersnake (*Nerodia rhombifer*)	Fasted	NR1316	HiSeq	NEB Next	SAMN12003909
	Fasted	NR1317	HiSeq	NEB Next	SAMN12003910
	Fasted	NR1464	HiSeq	NEB Next	SAMN12003911
	Fasted	NR1416	HiSeq	NEB Next	SAMN12003912
	1DPF	NR1357	HiSeq	NEB Next	SAMN12003913
	1DPF	NR1436	HiSeq	NEB Next	SAMN12003914
	1DPF	NR1388	HiSeq	NEB Next	SAMN12003915
	1DPF	NR1331	HiSeq	NEB Next	SAMN12003916
	4DPF	NR1442	HiSeq	NEB Next	SAMN12003917
	4DPF	NR1354	HiSeq	NEB Next	SAMN12003918
	4DPF	NR1327	HiSeq	NEB Next	SAMN12003919
	4DPF	NR1324	HiSeq	NEB Next	SAMN12003920

Supplementary Table 2. Read mapping statistics for all samples used in the study.

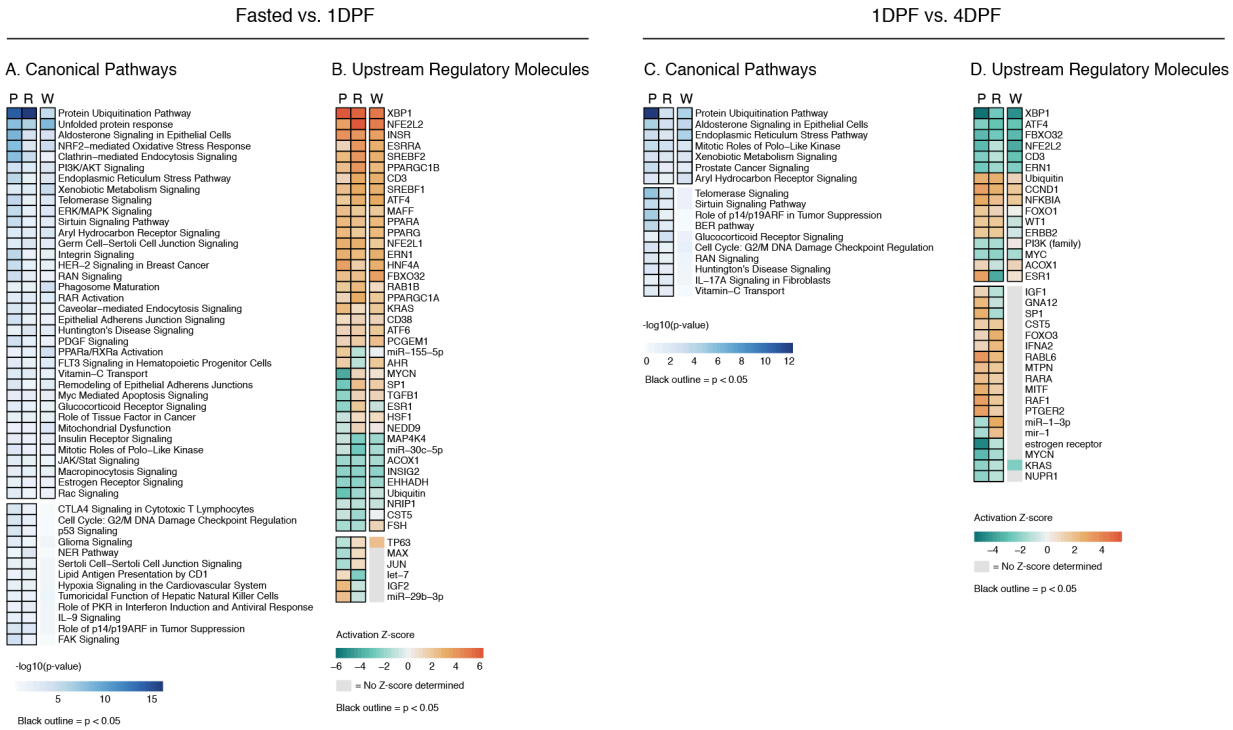
Species	Time Point	Sample ID	Number Input Reads	Number Uniquely Mapped Reads	% Uniquely Mapped Reads
Python	Fasted	U25	11,140,509	7,845,947	70.43%
		AI6a	3,409,861	2,898,451	85.00%
		AI6b	992,739	871,455	87.78%
		AJ6a	5,154,838	4,529,326	87.87%
		AJ6b	2,060,920	1,837,885	89.18%
		AJ6c	2,643,193	2,414,192	91.34%
	1DPF	W20	2,043,673	1,960,430	95.93%
		V43	5,656,165	4,193,004	74.13%
		Z14a	4,085,590	3,680,246	90.08%
		Z14b	1,678,087	1,527,872	91.05%
		Z14c	2,401,791	2,183,797	90.92%
	4DPF	Y18a	3,560,391	3,107,561	87.28%
		Y18b	1,364,726	1,203,620	88.19%
		Y18c	2,216,650	1,949,233	87.94%
		Y23a	2,971,652	2,551,637	85.87%
		Y23b	1,145,443	1,001,545	87.44%
		Y5a	3,689,468	3,256,951	88.28%
		Y5b	1,449,679	1,292,496	89.16%
Rattlesnake	Fasted	CV2	12,730,320	9,974,214	78.35%
		CV4	11,271,037	8,782,204	77.92%
		CV7	8,708,873	6,947,966	79.78%
		CV1	1,915,910	1,071,479	55.93%
	1DPF	CV6A	8,556,459	7,335,838	85.73%
		CV8	8,248,596	6,768,378	82.05%
		CV3	3,010,519	1,273,622	42.31%
	4DPF	CV11	9,610,524	8,174,577	85.06%
		CV12	8,926,893	7,465,208	83.63%
		CV9	11,594,658	10,156,595	87.60%
		CV10	2,318,138	1,607,507	69.34%
Watersnake	Fasted	NR1316	9,885,620	6,590,139	66.66%
		NR1317	9,239,753	6,384,714	69.10%
		NR1416	6,598,368	4,319,166	65.46%
		NR1464	11,832,553	7,814,975	66.05%
	1DPF	NR1331	6,486,641	3,561,965	54.91%
		NR1357	10,446,473	7,218,437	69.10%
		NR1388	9,144,085	5,465,528	59.77%
		NR1436	7,363,598	5,135,712	69.74%
	4DPF	NR1324	5,338,747	3,349,379	62.74%
		NR1327	19,008,775	8,448,405	44.44%
		NR1354	5,420,532	3,474,894	64.11%
		NR1442	8,042,516	5,501,847	68.41%

Supplementary Table 3. Relevant statistics pertaining to homolog assignment via OrthoMCL. The values in parentheses represent the number of groups containing a single human ID and multiple human IDs, respectively, in the homolog groups that were not full 1-to-1. See Supplementary Methods for additional details.

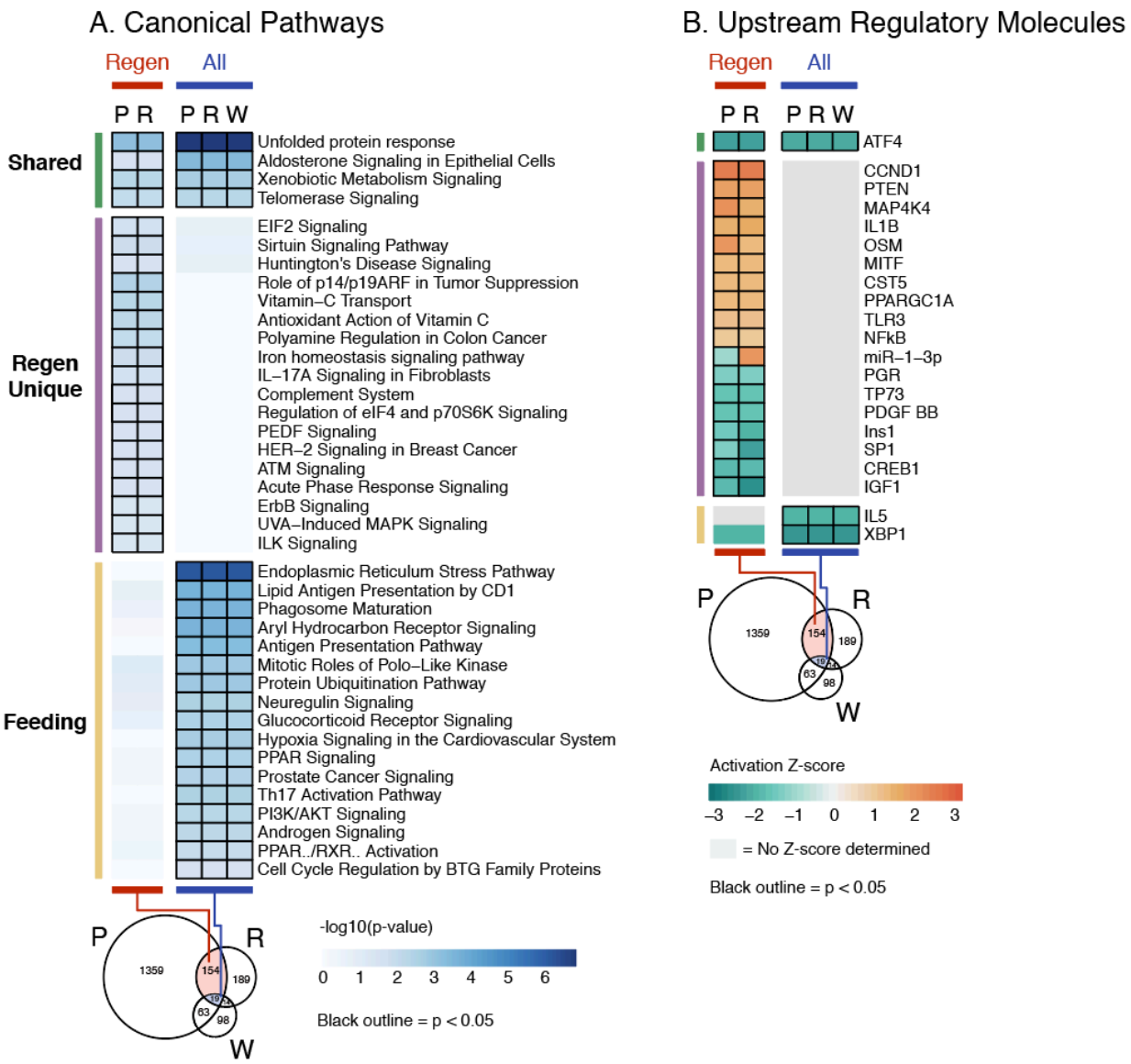
Total input genes per species:		
Human	20,182	
Python	17,985	
Rattlesnake	17,480	
Gartersnake	17,524	
Total OrthoMCL homolog groups per species:		
Human	17,464	
Python	16,721	
Rattlesnake	14,513	
Gartersnake	15,623	
OrthoMCL homolog groups containing all species:		
Total	9,553	
Full 1-to-1	7,151	
Not full 1-to-1	2,402	(832; 1,570)

Supplementary Table 4. Python and watersnake small intestine samples used to generate label-free quantitative proteomics data for this study. See (Andrew et al. 2015; Andrew et al. 2017) for additional information on python tissues.

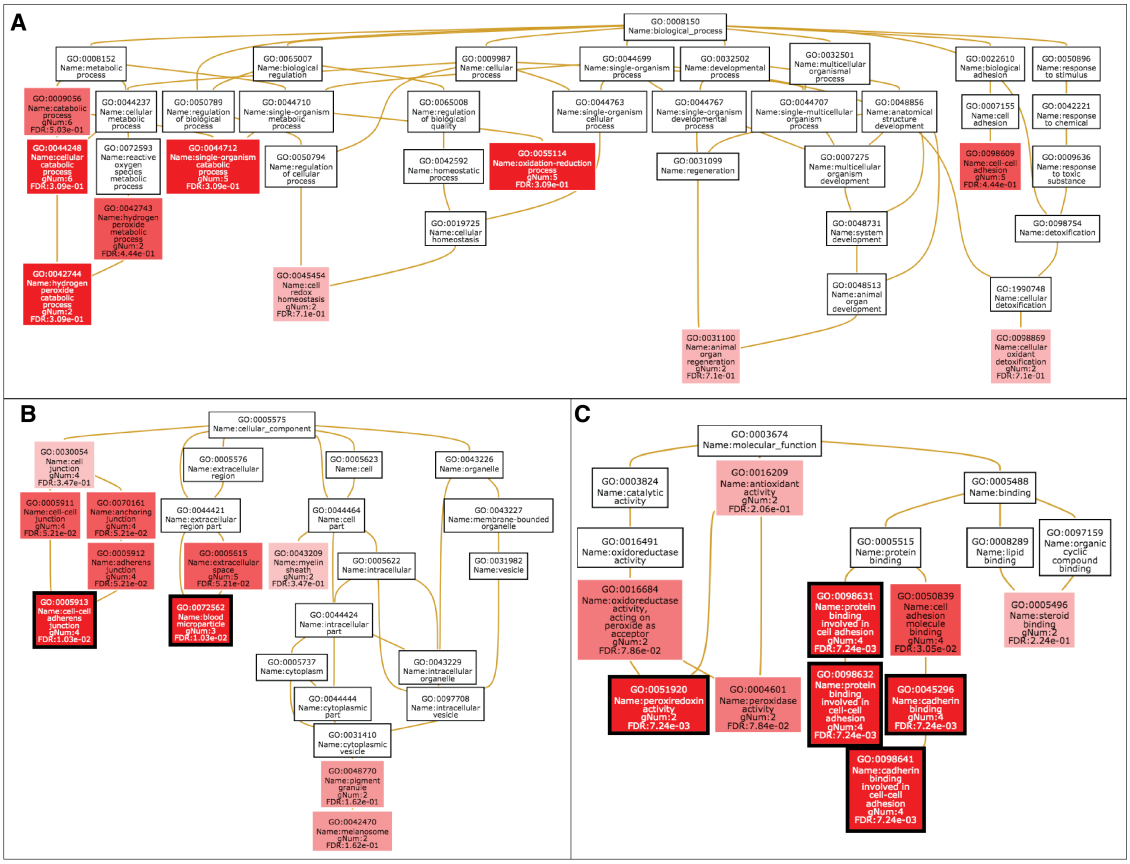
Species	Time point	Animal ID
Burmese Python (*Python molurus bivittatus*)	Fasted	AI6
	Fasted	AJ6
	Fasted	U25
	1DPF	V43
	1DPF	S6
	1DPF	W20
	4DPF	Y5
	4DPF	Y18
Diamondback Watersnake (*Nerodia rhombifer*)	Fasted	NR1317
	Fasted	NR1464
	Fasted	NR1416
	1DPF	NR1436
	1DPF	NR1388
	1DPF	NR1331
	4DPF	NR1442
	4DPF	NR1354
	4DPF	NR1324



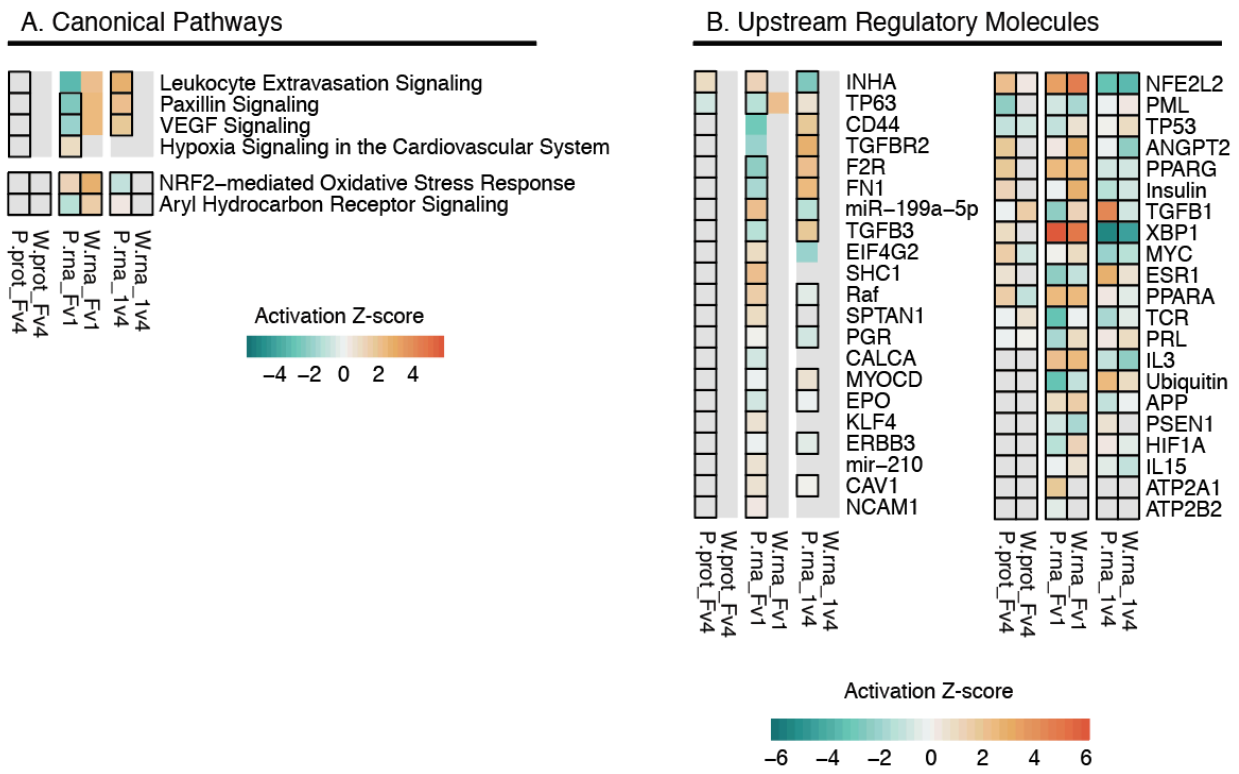
Supplementary Figure 1. Canonical pathway and upstream regulatory molecule activation inferences based on comparisons of all differentially expressed genes. A) Canonical pathways enrichment of differentially expressed genes between fasted and 1DPF. Black outlines denote a $p < 0.05$. B) Predicted upstream regulatory molecule activity between fasted and 1DPF. Cells with a black outline indicate significant enrichment and predicted activity ($p < 0.05$ and $|z| > 1$). C) Canonical pathways enrichment of differentially expressed genes between 1DPFand 4DPF. Black outlines denote a $p < 0.05$. D) Predicted upstream regulatory molecule activity between 1DPFand 4DPF. Cells with a black outline indicate significant enrichment and predicted activity ($p < 0.05$ and $|z| > 1$). For each heatmap, only pathways/regulatory molecules that are either significant in all three species or just the two regenerating species (python and rattlesnake) are shown.



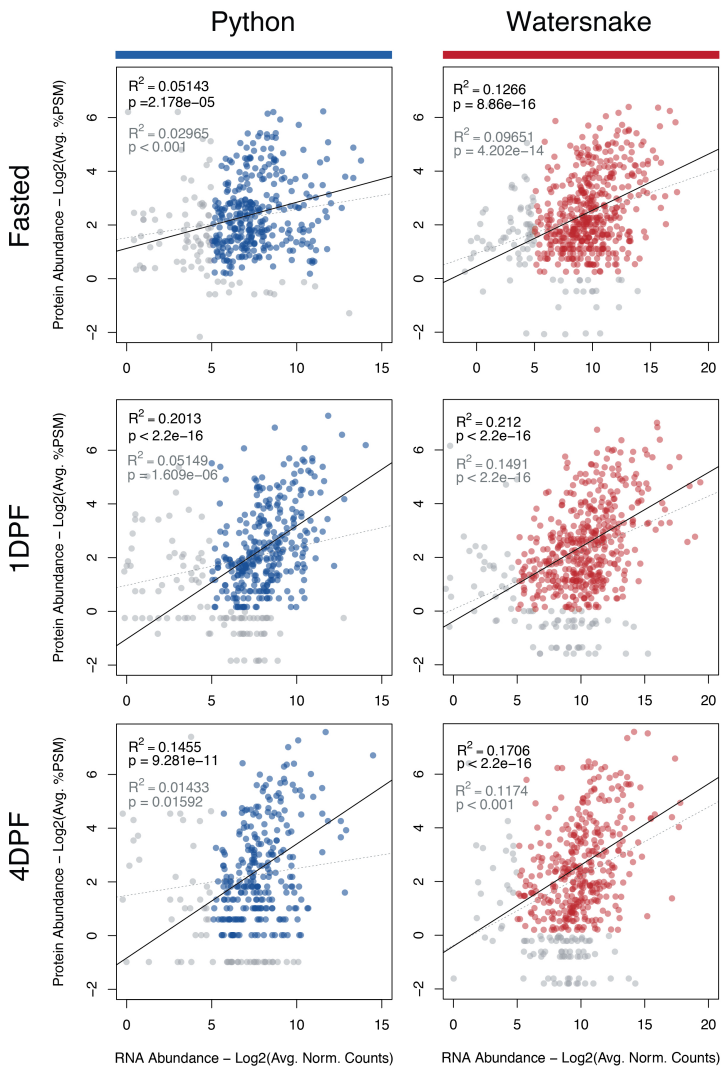
Supplementary Figure 2. Canonical pathway and upstream regulatory molecule activation inferences based on comparisons of 1DPF versus 4DPF RNAseq data. A) Canonical pathways enrichment of differentially expressed genes between fasted and 1DPF based on genes shared uniquely between the the two regenerating species ("Regen," left column) and genes shared between all three species ("All," right column). Black outlines denote a $p < 0.05$. B) Predicted upstream regulatory molecule activity based on Regen and All gene sets. Cells with a black outline indicate significant enrichment and predicted activity ($p < 0.05$ and $|z| > 1$).



Supplementary Figure 3. GO term overrepresentation for the 12 proteins that exhibited significant changes in abundance between fasting and feeding in both the python and watersnake. A) Biological process. **B**) Cellular component. **C**) Molecular function. Boxes with black outline denote significant overrepresentation (FDR < 0.05).



Supplementary Figure 7. Canonical pathway and upstream regulatory molecule activation inferences based on proteomic data. Comparison of predicted activity of A) canonical pathways and B) upstream regulatory molecules based on proteins with significant changes in abundance between fasted and 4DPF and genes significantly differentially expressed between fasted vs. 1DPF and 1DPF vs. 4DPF, filtered to only show pathways and molecules that are significant from both protein and RNA data but only in python, or significant based on both protein and RNA data in both python and watersnake. Cells with a black outline indicate significant predicted activity ($p < 0.05$).



Supplementary Figure 8. Correlation of RNA and protein abundance in the python and Watersnake. Grey R^2 values, p-values, and trend lines correspond to all data points (grey and colored), while black values and trend lines correspond to only data points in which average log2 RNA and protein abundance are greater than 5 and 0, respectively (colored points).

Supplementary References

Andrew, A. L., D. C. Card, R. P. Ruggiero, D. R. Schield, R. H. Adams, D. D. Pollock, S. M. Secor, and T. A. Castoe. 2015. Rapid changes in gene expression direct rapid shifts in intestinal form and function in the Burmese python after feeding. Physiol Genomics 47:147-157.

Andrew, A. L., B. W. Perry, D. C. Card, D. R. Schield, R. P. Ruggiero, S. E. McGaugh, A. Choudhary, S. M. Secor, and T. A. Castoe. 2017. Growth and stress response mechanisms underlying post-feeding regenerative organ growth in the Burmese python. BMC genomics 18:338.

Barret Schloerke, J. C., Di Cook, Francois Briatte, Moritz Marbach, Edwin Thoen, Amos Elberg, Joseph Larmarange. 2018. GGally: Extension to 'ggplot2'.

Bindea, G., B. Mlecnik, H. Hackl, P. Charoentong, M. Tosolini, A. Kirilovsky, W.-H. Fridman, F. Pagès, Z. Trajanoski, and J. Galon. 2009. ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. Bioinformatics 25:1091-1093.

Bolger, A. M., M. Lohse, and B. Usadel. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114-2120.

Brunson, J. C. 2018. ggalluvial: Alluvial Diagrams in 'ggplot2'.

Butts, C. T. 2008. network: a Package for Managing Relational Data in R. Journal of Statistical Software 24:1-36.

Castoe, T. A., A. P. de Koning, K. T. Hall, D. C. Card, D. R. Schield, M. K. Fujita, R. P. Ruggiero, J. F. Degner, J. M. Daza, W. Gu, J. Reyes-Velasco, K. J. Shaney, J. M. Castoe, S. E. Fox, A. W. Poole, D. Polanco, J. Dobry, M. W. Vandewege, Q. Li, R. K. Schott, A. Kapusta, P. Minx, C. Feschotte, P. Uetz, D. A. Ray, F. G. Hoffmann, R. Bogden, E. N. Smith, B. S. Chang, F. J. Vonk, N. R. Casewell, C. V. Henkel, M. K. Richardson, S. P. Mackessy, A. M. Bronikowski, M. Yandell, W. C. Warren, S. M. Secor, and D. D. Pollock. 2013. The Burmese python genome reveals the molecular basis for extreme adaptation in snakes. Proc Natl Acad Sci U S A 110:20645-20650.

Dobin, A., C. A. Davis, F. Schlesinger, J. Drenkow, C. Zaleski, S. Jha, P. Batut, M. Chaisson, and T. R. Gingeras. 2013. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29:15-21.

Handcock, M. S., D. R. Hunter, C. T. Butts, S. M. Goodreau, and M. Morris. 2008. statnet: Software tools for the representation, visualization, analysis and simulation of network data. Journal of statistical software 24:1548.

Ignatiadis, N., B. Klaus, J. B. Zaugg, and W. Huber. 2016. Data-driven hypothesis weighting increases detection power in genome-scale multiple testing. Nature methods 13:577.

Kamal, A. H. M., M. B. Fessler, and S. M. Chowdhury. 2018. Comparative and network-based proteomic analysis of low dose ethanol-and lipopolysaccharide-induced macrophages. PloS one 13:e0193104.

Kolde, R. 2012. Pheatmap: pretty heatmaps. R package version 61.

Krämer, A., J. Green, J. Pollard Jr, and S. Tugendreich. 2013. Causal analysis approaches in ingenuity pathway analysis. Bioinformatics 30:523-530.

Li, L., C. J. Stoeckert, and D. S. Roos. 2003. OrthoMCL: identification of ortholog groups for eukaryotic genomes. Genome research 13:2178-2189.
Liao, Y., G. K. Smyth, and W. Shi. 2013. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 30:923-930.
Love, M. I., W. Huber, and S. Anders. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome biology 15:550.
Perry, B. W., D. C. Card, J. W. McGlothlin, G. I. M. Pasquesi, R. H. Adams, D. R. Schield, N. R. Hales, A. B. Corbin, J. P. Demuth, and F. G. Hoffmann. 2018. Molecular adaptations for sensing and securing prey and insight into amniote genome diversity from the garter snake genome. Genome biology and evolution 10:2110-2129.
Schield, D. R., D. C. Card, N. R. Hales, B. W. Perry, G. M. Pasquesi, H. Blackmon, R. H. Adams, A. B. Corbin, C. F. Smith, and B. Ramesh. 2019. The origins and evolution of chromosomes, dosage compensation, and mechanisms underlying venom regulation in snakes. Genome Research.
Team, R. C. 2014. R: A language and environment for statistical computing.
Zhang, B., S. Kirov, and J. Snoddy. 2005. WebGestalt: an integrated system for exploring gene sets in various biological contexts. Nucleic acids research 33:W741-W748.