Quantity, Not Quality: Rapid Adaptation in a Polygenic Trait Proceeded Exclusively through Expression Differentiation

Mark J. Margres,*[†,1] Kenneth P. Wray,[‡,1] Alyssa T.B. Hassinger,[1] Micaiah J. Ward,[1] James J. McGivern,[1] Emily Moriarty Lemmon,[1] Alan R. Lemmon,[2] and Darin R. Rokyta[1]

[1]Department of Biological Science, Florida State University, Tallahassee, FL

[2]Department of Scientific Computing, Florida State University, Tallahassee, FL

[†]Present address: School of Biological Sciences, Washington State University, Pullman, WA

[‡]Present address: Department of Natural Resource Ecology and Management, Iowa State University, Ames, IA

***Corresponding author:** E-mail: mark.margres@wsu.edu.

Associate editor: Nicolas Vidal

Abstract

A trait's genomic architecture can affect the rate and mechanism of adaptation, and although many ecologically-important traits are polygenic, most studies connecting genotype, phenotype, and fitness in natural populations have focused on traits with relatively simple genetic bases. To understand the genetic basis of polygenic adaptation, we must integrate genomics, phenotypic data, ecology, and fitness effects for a genetically tractable, polygenic trait; snake venoms provide such a system for studying polygenic adaptation because of their genetic tractability and vital ecological role in feeding and defense. We used a venom transcriptome–proteome map, quantitative proteomics, genomics, and fitness assays in sympatric prey to construct a genotype–phenotype–fitness map for the venoms of an island–mainland pair of rattlesnake populations. Reciprocal fitness experiments demonstrated that each population was locally adapted to sympatric prey. We identified significant expression differentiation with little to no coding-sequence variation across populations, demonstrating that expression differentiation was exclusively the genetic basis of polygenic adaptation. Previous research on the genetics of adaptation, however, has largely been biased toward investigating protein-coding regions because of the complexity of gene regulation. Our results showed that biases at the molecular level can be in the opposite direction, highlighting the need for more systematic comparisons of different molecular mechanisms underlying rapid, adaptive evolution in polygenic traits.

Key words: **adaptation, expression, gene flow, fitness, polygenic.**

Introduction

The study of adaptive molecular evolution in natural populations has been limited by the difficulty of linking genetic and phenotypic variation to fitness effects. As such, most studies have used reverse genetic approaches to measure the functional effects of specific mutations in the laboratory (reviewed in Storz 2016). Many of the fundamental features of evolving systems, such as evolvability, epistasis (Rokyta et al. 2011), and pleiotropy (Hill and Zhang 2012), however, may be stronger determinants of evolutionary outcomes in natural populations than in the laboratory; artificial selection and breeding schemes are generally more simplistic than selection and demographic effects in natural settings. Therefore, although functionally characterizing the molecular basis of adaptation is necessary for understanding the predictability of evolution in nature (Rosenblum et al. 2010), the majority of our knowledge of adaptive molecular evolution is based on experimental evolutionary studies (Storz 2016).

The genetics of adaptation for particular traits with relatively simple genetic bases, however, have been described in natural populations, such as color variation in mice (Hoekstra et al. 2006) and lizards (Rosenblum et al. 2010), beak morphology in Darwin's finches (Lamichhaney et al. 2016), and eyespots in butterflies (Zhang and Reed 2016) among others; particular mutations have been functionally validated in some, but not all, of these cases (e.g., opsin genes in vertebrates Yokoyama et al. 2008 and toxin resistance in reptiles and other organisms Geffeney et al. 2005). The genomic basis of evolutionary change in polygenic traits, on the other hand, is poorly understood because the genomic architecture underlying these traits is often unknown despite many ecologically-important traits being the products of many loci (Wellenreuther and Hansson 2016).

The genetic basis of a trait has specific implications regarding the mechanism underlying adaptive phenotypic change, and differences in genomic architecture can directly influence the rate and mechanism of adaptive evolution (Rosenblum et al. 2010; Lourenco et al. 2013). A relatively simple genetic basis may facilitate rapid adaptation through novel, large-effect mutations (e.g., Hoekstra et al. 2006; Lamichhaney et al. 2016), but could also reduce evolvability, whereas adaptation in polygenic traits may be more likely through many small-effect mutations from standing genetic variation (Daub et al. 2013). To fully understand the

Article

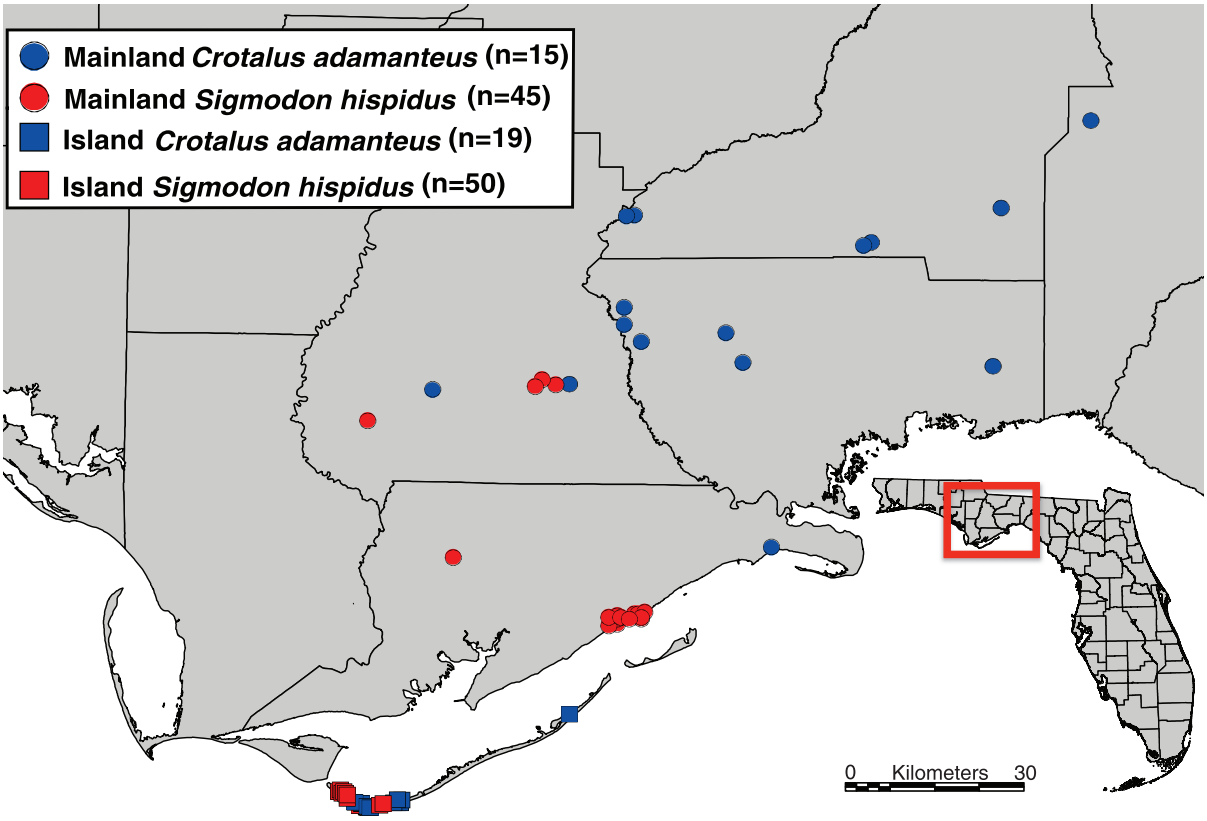

FIG. 1. Sampling of *Crotalus adamanteus* and *Sigmodon hispidus*. We collected venom and/or tissue samples from 15 mainland and 19 island adult *C. adamanteus*. We live-trapped 45 mainland and 50 island *S. hispidus*; these individuals were brought back to the lab for toxicity assays.

genotype–phenotype–fitness connection for complex traits in natural populations, we need to integrate genomics, phenotypic data, ecology, population-level experiments, and fitness effects for a genetically tractable, polygenic trait (Wellenreuther and Hansson 2016).

To identify the genomic basis of rapid adaptation for a polygenic trait, we constructed a genotype–phenotype–fitness map for the venoms of an island–mainland pair of eastern diamondback rattlesnake (*Crotalus adamanteus*) populations (fig. 1); the island arose $<$ 5,000 years ago (Lopez and Rink 2007). We used the previously published venom transcriptome–proteome map (Margres et al. 2014, 2015, 2016a) along with target enrichment, quantitative proteomics, and fitness assays to identify the genetic basis of adaptive venom divergence. Because venom genes are not expressed in tissues other than venom glands (but see Casewell et al. 2012), we expect no significant pleiotropic constraints/crossphenotype associations on mutations affecting toxin protein sequences. Because venoms are secretions, changes in toxin-gene expression levels directly alter protein amounts in the venom and thereby directly influence venom efficacy (Casewell et al., 2011; Margres et al., 2016b). These unique characteristics of venom allow the straightforward ascertainment of the relative roles of expression versus coding-sequence changes in adaptive evolution for a trait with little inherent bias toward either class of mutation. We characterized the functional effects of the identified sequence and expression variation by conducting reciprocal fitness comparisons in sympatric, wild populations of the primary prey species of *C. adamanteus* in north Florida, the hispid cotton rat (*Sigmodon hispidus*); *S. hispidus* has been shown to comprise ~57% of the diet, with the next most abundant species comprising only ~14% (Means 2017). Quantifying total fitness for natural, vertebrate populations is extremely difficult, but foraging success is undoubtedly a critical fitness component. We, therefore, used venom toxicity in natural, sympatric prey as a proxy for fitness because of venom's direct role in feeding and defense. Venom is metabolically costly to produce (McCue 2006), and increasing toxicity reduces the amount of venom expended per envenomation event, increasing the feeding efficiency of the snake. Therefore, it is predicted that selection will operate to increase venom economy by reducing the amount of venom required to subdue and incapacitate a specific prey item. Our joint genomic, proteomic, and functional approach enabled us to examine both the pattern and process of local adaptation for a polygenic trait in natural populations.

Results and Discussion

Reciprocal Toxicity Assays Identified Local Adaptation

To determine whether either population of *C. adamanteus* was adapted to local prey, we conducted reciprocal venom toxicity (i.e., fitness) comparisons in island and mainland

Table 1. Reciprocal Toxicity Assays.

	Island Prey		Mainland Prey	
	LD_{50}	ED_{50}	LD_{50}	ED_{50}
Island venom	**78.40 (39.50–155.59)**	**68.24 (80.53)**	**214.48 (165.82–277.41)**	**244.45 (86.37)**
Mainland venom	**105.96 (88.51–126.85)**	**95.83 (75.68)**	**70.43 (41.51–119.49)**	**60.07 (48.76)**

NOTE.—Values represent the median lethal dose (LD_{50}) or the median effective dose (ED_{50}) for each venom in mg/kg; lower values represent a more toxic (i.e., effective) venom. LD_{50} estimates were calculated using the Trimmed–Spearman Karber method and log-doses; 95% confidence intervals are shown in parentheses. ED_{50} estimates were calculated using the *drm* function from the *drc* R package; standard errors are shown in parentheses, and log doses were not used for ED_{50} estimates.

Table 2. Significance Tests for Differences in Median Effective Doses (ED_{50}s).

	Home-versus-away	Local-versus-foreign
Island venom	**0.28 (−0.29 to 0.95)***	**0.71 (−1.27 to 2.70)**
Mainland venom	**0.63 (−0.77 to 2.02)**	**0.25 (−0.18 to 0.67)***

NOTE.—Values represent the ratio of ED_{50}s for each comparison. 95% confidence intervals are shown in parentheses, and ED_{50}s were considered significantly different if the confidence intervals did not contain 1 as outlined in Wheeler et al. (2006).
*Significance.

S. hispidus populations. We found that both snake populations were adapted to local prey (tables 1, 2, and fig. 2); each venom satisfied local-versus-foreign (local population has higher fitness than population from another habitat) and home-versus-away (home population has higher fitness at home than away) criteria (fig. 2). Island venom was significantly more effective on island prey than mainland prey, indicative of a fitness trade-off, and mainland venom was significantly more effective than island venom on mainland prey (table 2). Venoms were not significantly different in effect within island prey (most likely due to a limited sample size; see Materials and Methods), but the reciprocal pattern was consistent with local adaptation (fig. 2).

Local adaptation is the result of ongoing or very recent selection, and we do not expect a single phenotype to be superior in all environments. What we do expect, however, is for each population to exhibit fitness trade-offs because of genotype-by-environment interactions. Consistent with these expectations, we found that each venom phenotype exhibited an allopatric cost (table 2 and fig. 2). This dynamic was much more pronounced in the island population, potentially reflecting a greater degree of dietary specialization relative to mainland individuals; *S. hispidus* was the only prey species trapped on the island, but a number of prey species were trapped on the mainland (e.g., *Peromyscus* and *Neotoma*; Means 2017). The mainland population, therefore, may feed on a broader range of prey species and possess a more generalist venom phenotype, whereas the island population may feed solely (or at least predominantly) on *S. hispidus* and possess an extremely specialized venom phenotype.

Although we identified significant toxicity differences between island and mainland *C. adamanteus*, we needed to identify the underlying process that produced this pattern. We next used genomic and proteomic approaches to identify the causal genetic mechanism underlying the identified adaptive phenotypic variation.

Significant Expression Differentiation in Low-Expression Venom Proteins

To determine if island ($n = 19$) and mainland ($n = 11$) populations exhibited significant differentiation in toxin expression, we used reversed-phase high-performance liquid chromatography (RP-HPLC) to compare venom expression levels and detected significant differentiation in toxin expression ($P = 0.006$), consistent with previous studies (Margres et al. 2015, 2016a). Margres et al. (2016a) compared these two populations and found that the expression levels of abundant venom protein peaks (i.e., expressed > geometric mean) were conserved with significant expression differentiation restricted to low-expression protein peaks (i.e., expressed < geometric mean). These expression differences were hypothesized to be the result of stronger stabilizing selection and more limiting physiological constraints on high-expression proteins, reducing the amount of standing expression variation in abundant relative to less abundant toxins. Here, because differentially expressed RP-HPLC peaks contained multiple proteins (Margres et al. 2016a), we could not resolve which particular toxin(s) exhibited the significant differentiation; any number, but not necessarily all, of the venom proteins detected in a given peak could be differentially expressed.

To address this lack of resolution, we performed quantitative mass spectrometry on the venom samples of five individuals from each population to identify which specific toxins were most differentially expressed across populations (fig. 3). Consistent with previous work (Margres et al. 2016a), high-expression venom proteins (as defined above) were largely conserved ($R^2 = 0.802$), and low-expression venom proteins exhibited substantial intra and interpopulation variation ($R^2 = 0.458$; fig. 3*A*). We next calculated residuals for each venom protein and identified candidate toxins in the upper and lower 2.5% of the residual distribution; these toxins represent the most differentially expressed proteins in the mass spectrometry data set and, therefore, should be major contributors to the significant expression and toxicity differences identified above. We identified two proteins that were more highly expressed in the island population (snake venom type III metalloproteinases (SVMPIII) 4 and 5), and two proteins that were more highly expressed in the mainland population (myotoxin 2, phosphodiesterase 1; fig. 3*B*). The five largest residuals, indicating an upregulation in the island population, all belonged to the SVMP gene family, suggesting that this

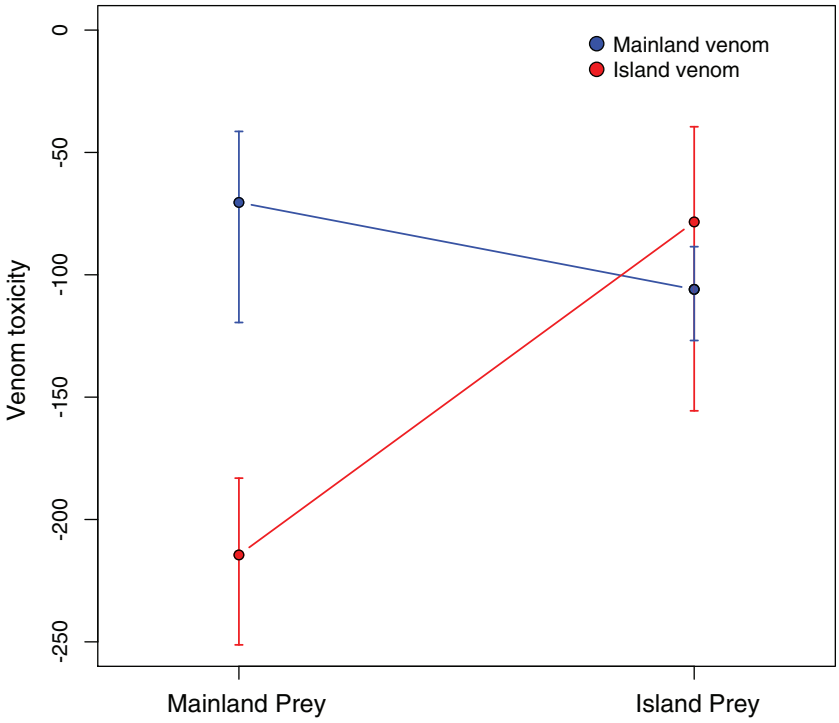

Fig. 2. Reciprocal toxicity assays demonstrated that island and mainland *Crotalus adamanteus* populations were locally adapted. Island and mainland venoms satisfied local-versus-foreign and home-versus-away criteria. Venom toxicity (LD_{50}), our proxy for fitness, is measured in mg/kg and shown on the *y*-axis. Because a lower LD_{50} represents a more toxic or effective venom, we plotted inverse values to allow fitness to increase on the *y*-axis. Error bars represent 95% confidence intervals. The number of *Sigmodon hispidus* used per assay was as follows: island venom injected into island prey ($n = 26$), island venom injected into mainland prey ($n = 18$), mainland venom injected into island prey ($n = 24$), and mainland venom injected into mainland prey ($n = 27$). Abbreviations: LD_{50}, median lethal dose.

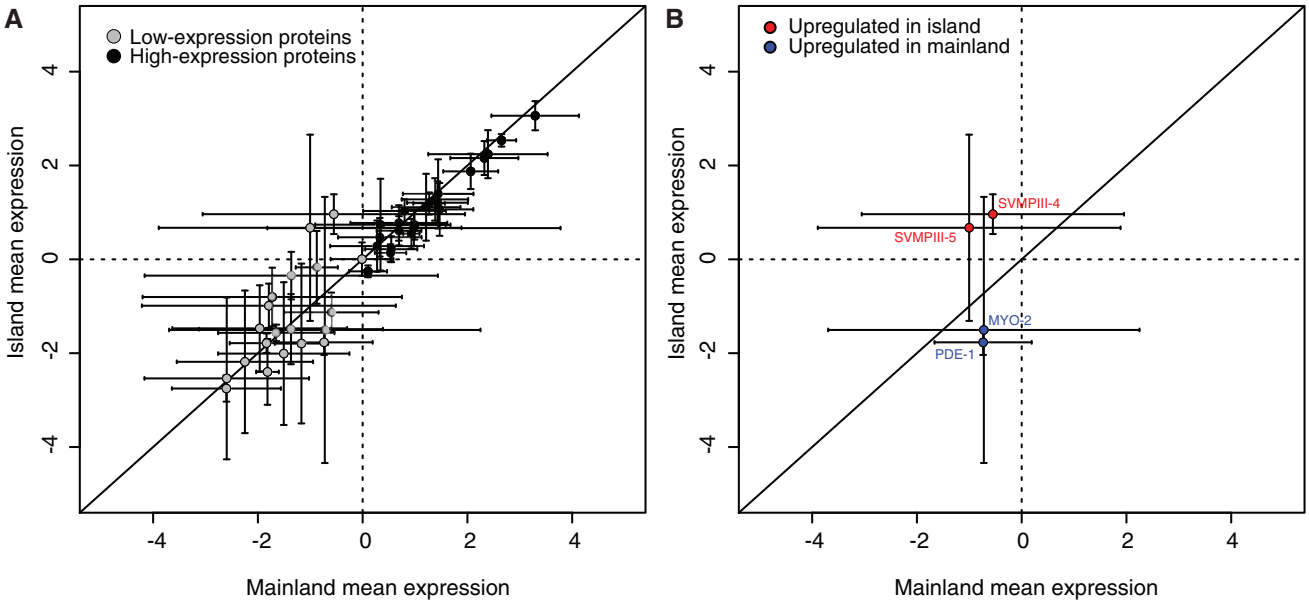

Fig. 3. Toxin expression differentiation between island and mainland venoms. (*A*) We performed quantitative mass spectrometry on the venom samples of five individuals from each population to identify which loci were most differentially expressed. Consistent with previous results, we found that high-expression venom proteins were largely conserved ($R^2 = 0.802$), and low-expression venom proteins were differentially expressed across populations ($R^2 = 0.458$). (*B*) A plot of the four outlier loci; outliers were identified in the upper and lower 2.5% of the residual distribution and represent the most differentially expressed toxins. Expression values were centered logratio transformed prior to plotting. Bars indicate standard deviation, the solid line indicates a perfect agreement (slope = 1), dashed lines indicate the origin (the geometric mean), and proteins less than these values were considered low-expression proteins. Abbreviations: MYO, myotoxin; PDE, phosphodiesterase; SVMPIII, snake venom type III metalloproteinase.

concerted regulation may be the result of parallel regulatory changes.

Our results again indicated that the expression levels of highly expressed proteins evolve under considerable constraints, and rapid, adaptive divergence in expression may be restricted to low-expression proteins over ecological timescales (Margres et al. 2016a). Because we identified significant expression differentiation, we next used target enrichment to sequence all toxin genes to assess the relative roles of expression and coding-sequence changes in polygenic adaptation.

Paucity of Sequence Divergence Indicates Expression Differentiation Was the Genetic Basis of Polygenic Adaptation

To identify potential adaptive mutations in toxin coding-regions, we used target enrichment to sequence all 59 toxin loci identified in the venom-gland transcriptome (Rokyta et al. 2017) and >1,000 neutral loci for ten individuals from each population; these neutral loci were comprised of non-toxin exons identified in the venom-gland transcriptome (Margres et al. 2015), anchored loci (Lemmon et al. 2012; Ruane et al. 2015), and anonymous loci (see Materials and Methods for details). We identified 160 single nucleotide polymorphisms (SNPs) in 41 toxin loci and 4,063 SNPs in 1,077 neutral loci; 18 toxin loci did not possess variable sites. We calculated mean Weir and Cockerham's F_{ST} estimates for all locus data sets (toxin, nontoxin, anchored, and short and long anonymous; see Materials and Methods) and found that mean neutral divergence ($F_{ST} = 0.023–0.033$) exceeded mean toxin divergence ($F_{ST} = 0.013$) for all four classes of neutral loci (table 3).

These estimates and comparisons, however, were based on all loci, and we did not expect all toxin loci to be evolving under strong, directional selection, especially over a timescale of only a few thousand years. To determine which toxin loci exhibited the most differentiation, we compared the per-locus toxin F_{ST} estimates ($n = 41$) to the distribution of the per-locus neutral F_{ST} estimates ($n = 1,077$); the neutral loci provided a null distribution, and we identified toxins as being significantly differentiated between populations if the locus exceeded the 95% percentile of the null distribution. F_{ST}-based approaches for measuring differentiation between parental (i.e., mainland) and derived (i.e., island) populations have been shown to effectively capture derived polymorphisms due to selection (Innan and Kim 2008). We identified three toxin loci that exceeded the 95th percentile of the neutral distribution: Nerve growth factor 1 (NGF-1), SVMPII-4, and vascular endothelial growth factor 1 (VEGF-1; fig. 4A). Nonsynonymous changes were identified in SVMPII-4 and VEGF-1. For SVMPII-4, the island population only possessed the major, shared alleles with alternative alleles at low frequency (0.20) in the mainland population. The nonsynonymous variant in VEGF-1 was present in both populations, but also at low frequency (0.30 in the island and 0.05 in the mainland), suggesting that these variants were not responsible for the identified adaptive divergence in venom function (table 4). We next compared F_{ST} estimates on a site-by-site basis, comparing all toxin SNPs to all neutral SNPs to identify specific outlier mutations within toxin-encoding genes; the neutral variants again provided the null distribution (fig. 4B). We identified two specific variants, both synonymous changes in NGF-1 (table 4), that exceeded the 95th percentile of the neutral distribution. Because these specific variants were synonymous, we argue that these alleles did not exhibit significant functional effects. The lack of fixed or nearly fixed nonsynonymous differences provides strong evidence that coding-region mutations were not responsible for the identified adaptation.

Table 3. Mean F_{ST} Estimates for Each Class of Loci.

Locus Type	Weir & Cockerham's Mean F_{ST}
Anchored	0.032
Long anonymous	0.033
Short anonymous	0.023
Nontoxin	0.023
Toxin	0.013

The debate as to whether mutations in regulatory or coding regions are the primary mechanism underlying adaptive phenotypic divergence has received substantial attention (e.g., Carroll 2008; Hoekstra and Coyne 2007), but the relative importance of either mutational class in polygenic traits was unclear. Coding variants are the most abundant mutations recovered in yeast laboratory experiments (Kvitek and Sherlock 2011), and some have argued that these variants are responsible for nearly all adaptive changes between closely related taxa (Hoekstra and Coyne 2007). These estimates, however, were based on traits with simple genetic bases, and genomic architecture can affect the rate and mechanism of adaptation. Fraser (2013) recently found that expression differentiation was ~10× more likely to lead to putative adaptation than changes to protein sequences in humans, although functional verification and links to specific phenotypic effects were not possible. Similarly, an evolve-and-resequence study found that the majority of adaptive changes in mammalian aerobic metabolism was linked to differential expression, but these loci evolved under artificial selection, and not all loci had an obvious association with the metabolic phenotype (Konczal et al. 2015). Here, the clear genotype–phenotype–function relationship in venom allowed us to determine that rapid adaptation in a polygenic trait proceeded predominantly, or potentially exclusively, through expression differentiation, and this may be the result of greater opportunity for expression changes relative to coding-region variation. Genetic constraints can bias adaptive trajectories, and evolution may be more likely in the direction of greater opportunity, either in the form of genetic variation (Chenoweth et al. 2010) or types of mutational mechanisms (Rokyta et al. 2015b). Disregarding the probability of generating beneficial variation, more mutational mechanisms exist for altering the amounts of proteins produced than for altering their functions through their coding-sequences. The latter can only occur through point mutations in specific regions of a protein, whereas expression evolution can occur through regulatory mutations (Manceau et al. 2011), copy-number

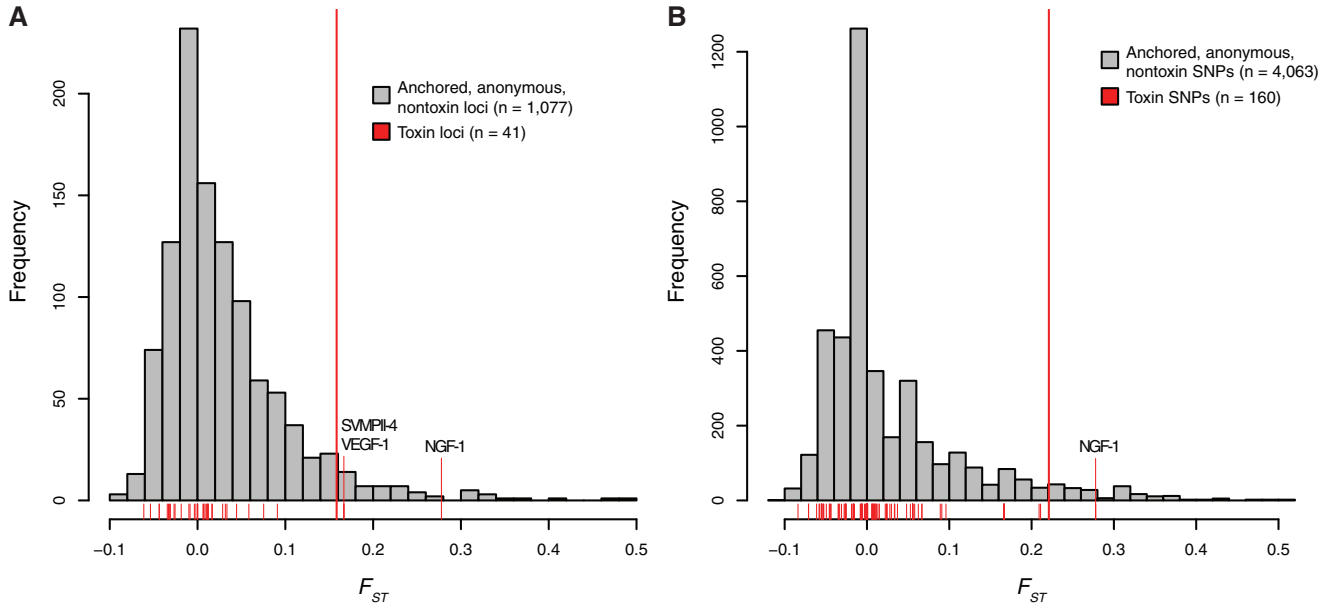

Fig. 4. The paucity of sequence divergence in toxin-encoding genes indicated that rapid adaptation proceeded through expression differentiation, not point mutations within coding regions. (*A*) We compared F_{ST} estimates for each toxin locus to that of all the neutral loci to identify specific toxin outliers; the neutral loci provided the null distribution. We identified SNPs in 41 of the 59 toxin loci, and three toxin loci exceeded the 95th percentile of the neutral distribution (vertical red line): NGF-1, SVMPII-4, and VEGF-1. Only VEGF-1 possessed a nonsynonymous variant found in the island population, albeit at low frequency (0.30). (*B*) We compared F_{ST} estimates on a site-by-site basis, comparing all toxin SNPs to that of all neutral SNPs to identify specific outlier mutations within toxin-encoding genes; the neutral variants provided the null distribution. We identified two synonymous variants, both in NGF-1, that exceeded the 95th percentile of the neutral distribution. Abbreviations: NGF, nerve-growth factor; SVMPII, snake venom metalloproteinase type II; VEGF, vascular endothelial growth factor.

Table 4. The Three Significantly Differentiated Toxin Loci across Island and Mainland Populations.

Outlier Locus (no. of SNPs)	SNP F_{ST}	Position	Mutation	Syn./Nonsyn.	Substitution Codon (AA)
NGF-1 (2)	0.278	210	G → C	Syn.	–
	0.278	393	A → G	Syn.	–
SVMPII-4 (4)	0.167	919	C → G	Nonsyn.	CAT → GAT (H → D)
	0.167	932	G → C	Nonsyn.	CGG → CCG (R → P)
	0.167	1,114	G → A	Nonsyn.	GAT → AAT (D → N)
	0.167	1,378	A → G	Nonsyn.	ATA → GTA (I → V)
VEGF-1 (1)	0.167	157	G → A	Nonsyn.	GAT → AAT (D → N)

Note.—All four nonsynonymous SNPs in SVMPII-4 were absent from the island population, and the nonsynonymous change in VEGF-1 was at low frequency (0.30) in the island population. Toxin abbreviations: NGF, nerve-growth factor; SVMP, snake venom metalloproteinase; VEGF, vascular endothelial growth factor.

variation (Margres et al. 2017), microRNA regulation (Huntzinger and Izaurralde 2011), and translational efficiency among other mechanisms. We therefore might expect that, without any innate, trait-specific bias, traits may preferentially evolve through mutations affecting expression. Our data clearly support this expectation, but additional systematic comparisons of coding-region divergence and expression variation are needed to characterize the relative frequencies of different molecular mechanisms underlying adaptive polygenic evolution.

The Relationship between Local Adaptation and Gene Flow

Traditional theory described gene flow as a homogenizing force that opposes directional selection and prevents local adaptation (Slatkin 1987), but more recent work has recognized that the interaction between gene flow and selection can be complex (Garant et al. 2007). High levels of gene flow may not only disrupt local adaptation, but also stimulate adaptive evolution in certain scenarios, such as in inbred populations, by increasing genetic and phenotypic variation on which selection can act (Fitzpatrick et al. 2015).

To determine whether the identified adaptive expression differentiation between island and mainland populations occurred in the presence or absence of gene flow, we designed four migration models using Migrate-n v. 4.2.8 (Beerli 2008) and used marginal likelihoods to compare models that differed in the number of migration parameters to test the directionality of gene flow (fig. 5). For all three data sets (anchored, long anonymous, and short anonymous loci; see

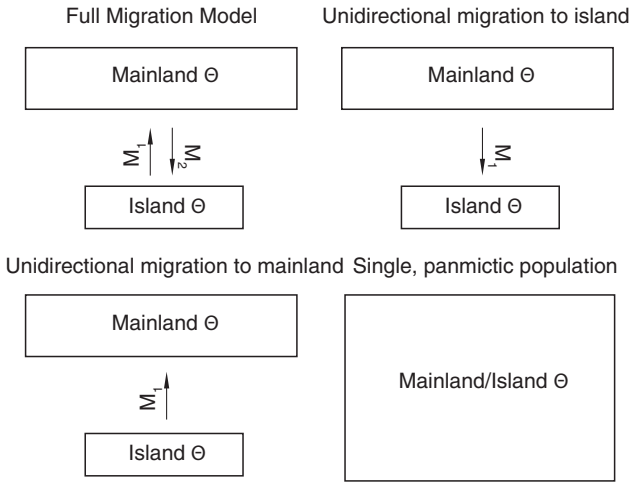

Fig. 5. Migrate-n models. We used Migrate-n, a coalescent-based approach, to estimate migration rates (M) and effective population sizes (θ) across four different models: 1) a full migration model with two population sizes and bidirectional migration, 2) two population sizes and unidirectional migration from mainland to island, 3) two population sizes and unidirectional migration from island to mainland, and 4) a panmictic model where we estimate a single population size with no migration. Bayes factors were used to select the best model.

Materials and Methods), the best-supported model was the panmixia model, where a single effective population size was estimated, indicating that island and mainland individuals belong to the same randomly mating population. This model had very strong support (model probability ~1.0) across all data sets.

Given these results, whether the identified adaptive divergence occurred in the presence or absence of gene flow was unclear given the timescale; the lack of divergence across all neutral loci (table 3) could have been the result of high migration rates or a lack of time for neutral differentiation to occur. If the best model accurately captured historical migration rates, the identified adaptation would have occurred under relatively strong directional selection given the timeframe and high migration rates. Under this scenario, we would expect few adaptive loci of large effect because loci under the strongest divergent selection would most readily resist gene flow (Savolainen et al. 2013). If the best model did not accurately capture historical migration rates, and the lack of neutral divergence simply reflected the extremely short timescale over which divergence occurred, then adaptation occurred in the absence, or with low levels, of gene flow. Under this scenario, we would expect many loci of small effect; given the time constraint, however, many adaptive loci of any effect size seemed unlikely. Therefore, it was not possible to distinguish between these two alternatives.

Conclusion

We constructed a genotype–phenotype–fitness map for a polygenic trait in natural populations and found that expression differentiation was exclusively the genetic basis of polygenic adaptation in venoms. The identified expression differentiation was biased toward low-expression proteins, suggesting that high-expression toxins may perform generic killing functions and low-expression toxins may exhibit prey-specific effects (Margres et al. 2016a); even though differentially expressed proteins were at low-abundance in the venom proteome, these proteins appear to exhibit large fitness effects. The lack of sequence divergence indicated that each population possessed the same toxin arsenal, demonstrating that the relative amounts of specific proteins can significantly alter venom toxicity and phenotypic function, perhaps through synergistic activities.

Although our fitness assays demonstrated that even the most susceptible *S. hispidus* populations exhibited significantly greater resistance to *C. adamanteus* venom than lab mice (LD_{50} ~2–4 mg/kg; Rokyta et al. 2017), indicative of past selection on *S. hispidus*, venom, not venom resistance, exhibited a pattern of local adaptation. Predator local adaptation is surprising given theoretical predictions based on population size, generation time, and selective pressures, all of which should lead to a more rapid evolutionary response in the prey (Dawkins and Krebs, 1979). Predator local adaptation however, at least in rattlesnakes, may be a more general trend than previously thought (Holding et al. 2016). We propose three noncompeting explanations for this nonintuitive pattern.

First, differences in genomic architecture can directly influence the rate and mechanism of adaptive evolution (Rosenblum et al. 2010; Lourenco et al. 2013), and perhaps the most important aspect of coevolutionary adaptation for predator–prey systems is the relative rates at which they coevolve. Most identified venom resistance factors are monogenic (e.g., Biardi et al. 2011; Jansa and Voss 2011), and the polygenic architecture of the venom system may increase the evolvability of venom to where it exceeds the evolvability of the resistance phenotype in prey. Second, differences in migration rates between species can contribute to the identified pattern of predator adaptation, with the more mobile species able to adapt more rapidly than the more sedentary species (Thompson et al. 2002; Garant et al. 2007). Therefore, we may expect greater migration rates in snake predators than in their prey, although our migration analyses were equivocal in regards to the effective number of *C. adamanteus* migrants per generation. Lastly, the presumed strong selection on *S. hispidus* may not be strong enough to elicit a coevolutionary response due to diffuse selection; if many species consume *S. hispidus* (e.g., mesopredators, avian predators, other snake species), the importance of the *C. adamanteus*–*S. hispidus* interaction may have been unidirectional where *S. hispidus* was an important prey source for *C. adamanteus*, but *C. adamanteus* was not an important source of mortality for *S. hispidus*.

Overall, our results provide a clear example of the importance of expression differentiation in rapid, polygenic adaptation, and more integrative studies are needed to systematically compare the relative frequencies of different molecular mechanisms underlying rapid, adaptive evolution in polygenic traits.

Materials and Methods

Sampling

We collected venom and/or tissue samples from 15 mainland and 19 island adult *C. adamanteus* and used Sherman live-traps to collect 45 mainland and 50 island *S. hispidus* (fig. 1). Island individuals were collected from Little St. George Island, a Holocene formation (< 5,000 years-old) located in the Gulf of Mexico ~7 km from the mouth of the Apalachicola River Delta (Lopez and Rink 2007). Mainland individuals were collected from the adjacent Florida panhandle east of the Apalachicola River and west of the Econfina River. Samples were collected under Florida Fish and Wildlife Conservation Commission permits LSSC-13-00004, LSSC-09-0399, and LSSC-12-00071A and Institutional Animal Care and Use Committee (IACUC) protocols #0924 and #1333.

Fitness Assays

Sigmodon hispidus were brought to the lab for LD_{50} assays; these assays determined the amount of venom required to kill half of an experimental group over a specified amount of time. We pooled the venoms of 10 mainland and 9 island adult *C. adamanteus* in equal quantities. Individual *S. hispidus* were allocated to a standardized concentration of venom to be injected, and appropriate amounts of pooled venom were dissolved in 100 µl of 0.9% sterile saline solution; venom doses were calculated based on individual *S. hispidus* mass. Because specific venoms were expected to possess different toxicities in different prey populations, we used an up-and-down procedure to determine the next concentration of venom to be injected, and concentrations ranged from 2 to 500 mg/kg. Individuals were injected pseudointraperitoneally in the lower right quadrant at a depth of 2–4 mm as this mimicked the natural route of delivery during envenomation. To minimize the number of animals used, we used 3–6 specimens per concentration including the control, and 6–8 concentrations per assay. We conducted four assays: island venom injected into island prey ($n = 26$), island venom injected into mainland prey ($n = 18$), mainland venom injected into island prey ($n = 24$), and mainland venom injected into mainland prey ($n = 27$). The number of individuals that did not survive the assay over a 48 h period were used to calculate toxicity. We calculated toxicity as both the LD_{50} and median effective dose (ED_{50}) for each assay. To calculate the LD_{50}, we used the Trimmed–Spearman Karber method (Hamilton et al. 1977) via the *tsk* function in R (Stone 2015) and log doses. To provide a second, independent confirmation of our toxicity calculations, we calculated the ED_{50}; we used the *mselect* function from the *drc* R package (Ritz et al. 2015) to select a model for the dose~response function. The best model was selected based on Akaike's information criterion (AIC), and we tested log-logistic (2–5 parameters) and Weibull (1–4 parameters) functions for each of the four estimates. We then used the *drm* function to fit a dose~response function (given the best model from *mselect*) to calculate the ED_{50}. The *comped* function (i.e., ratio test; Wheeler et al. 2006) was then used to determine if ED_{50} estimates were significantly different; here, the ratio of ED_{50}s was calculated, and no significant difference existed between the estimates if the confidence intervals for the ratio contained 1. These procedures were approved under IACUC protocol #1333.

Proteomic Analyses

Twenty-five RP-HPLC protein peaks were quantified per venom sample and analyzed using a nonparametric MANOVA as previously described (Margres et al. 2016a). Raw RP-HPLC data are in supplementary file 1, Supplementary Material online.

Nanospray LC/MS^E was performed by the Florida State University College of Medicine Translational Science Laboratory using the Synapt G2 HD Mass Spectrometer with a nanoAcquity UPLC (Waters Corp.) as previously described (Rokyta et al. 2015a). Five individual venom samples from each population were run in triplicate. Proteins were identified using the ProteinLynx Global SERVER IdentityE algorithm to search a database containing entries specific to the venom-gland transcriptome (Rokyta et al. 2017) and the standard. All proteins retained in the final analyses had 0% false positive rates (FPR), and any protein not detected in all three replicates was excluded; the all-peptide quantification method used here and described elsewhere (Rokyta et al. 2015a) is based on the same principles underlying iBAQ (Schwanhausser et al. 2011). Because these data were compositional, we used the centered logratio (clr) transformation (Egozcue et al. 2003) to transform the data using the robCompositions package (Templ et al. 2011) in R prior to downstream analyses.

Probe Design

Toxin probes were designed as previously described (Margres et al. 2017). Probes for 200 nontoxin exons identified in the venom-gland transcriptome (Margres et al. 2015) were designed as previously described (Margres et al. 2017). Anchor target loci were obtained from Ruane et al. (2015), which were derived from Lemmon et al. (2012). To increase enrichment efficiency, low coverage genomic reads from four pit-viper species (five individuals; described in Margres et al. 2017) were mapped to the Indian python (*Python molurus*) and the brown reed snake (*Calamaria pavimentata*) reference anchor sequences using the spaced-kmer approach previously described (Ruane et al. 2015). The resulting homologous consensus sequences were then aligned using MAFFT (v7.023b; Katoh and Standley 2013) and trimmed in Geneious (v.7; Kearse et al. 2012). A total of 348 anchor targets were retained in this process. Anonymous target loci were identified by Grazziotin et al. (unpublished); the jararaca snake (*Bothrops jararaca*) and the lancehead (*Bothrops atrox*) genomes were each sequenced at 15× coverage on an Illumina 2500 sequencer with 150 PE reads. Overlapping reads were merged (Rokyta et al. 2012) and evaluated for coverage and length. For each species, 10,000 reads with ~15× coverage and 175–225 bp were selected as preliminary targets. Merged reads were then reciprocally mapped to identify loci sufficiently similar across species; well-aligned regions were selected as targets (1,500 regions at 200 bp and 500 regions at 2000 bp). To develop enrichment probes suitable

across pit vipers, genomic reads from the five individuals used for the anchor loci (described above) were mapped to the *Bothrops* anonymous targets using a spaced k-mer approach (Ruane et al. 2015). Consensus sequences were then aligned for each locus using MAFFT. After trimming to remove poorly-aligned regions, loci with repetitive elements were identified and removed (Hamilton et al. 2016). A total of 240 long anonymous (~2,000 bp) and 829 short anonymous (~250 bp) targets were retained. The final probe kit contained 57,292 probes (Margres et al. 2017).

Library Preparation and Sequencing

Samples were sequenced as previously described (Margres et al. 2017). Briefly, genomic DNA was extracted for 140 *C. adamanteus* and hybrid enrichment data were collected through the Center for Anchored Phylogenomics at Florida State University (www.anchoredphylogeny.com). Sequencing was performed in the Translational Science Laboratory in the College of Medicine at Florida State University, and all samples were sequenced on either (1) one PE150 Illumina HiSeq2500 lane, or (2) one PE150 Illumina HiSeq2500 lane and again on one PE200 Illumina HiSeq2500 as previously described (Margres et al. 2017). Ten mainland and ten island individuals were extracted from the full data set for this study.

Processing and Alignments

To generate reference sequences for all neutral loci to be used in alignments for variant calling, reads were merged with PEAR (Zhang et al. 2014). Merged and unmerged reads for all 140 samples were assembled with NGen version 12.3.1 (Rokyta et al. 2015b); contigs comprised of $\geq$ 10 reads were retained. Assembled contigs for all samples with lengths $\geq$ 300 nt were clustered using cd-hit-est version 4.6 (Li and Godzik 2006) with a global match proportion of 0.98, and only clusters including sequences from at least 20% of all 140 samples were retained. To restrict our set of reference contigs to only those loci we targeted, the clustered contigs were clustered again with each set of reference sequences (used in probe design) from each type of locus (e.g., anchored) with cd-hit-est with a global match proportion of 0.95. To avoid inclusion of spurious or repeat sequences in our references, we truncated each identified contig 150 nucleotides from each end of the target sequence.

Reads for ten mainland and ten island individuals were merged with PEAR (Zhang et al. 2014). We used the BWA-MEM algorithm (Li and Durbin 2009) to generate alignments using the neutral references and the 59 toxin transcripts (Rokyta et al. 2017) as the index. We used merged and unmerged reads for assembly, and reads with $>$ two gaps/mismatches relative to the reference were removed. We then conducted local realignments in GATK (McKenna et al. 2010; DePristo et al. 2011).

Variant Calling

Variant calling was performed in GATK (McKenna et al. 2010; DePristo et al. 2011) for 10 mainland and 10 island individuals across two data sets: a neutral data set containing alignments to the anchored, long anonymous, short anonymous, and nontoxin loci, and a toxin data set containing alignments to the 59 toxin transcripts from Rokyta et al. (2017). Variants matching any of the following criteria were removed: quality by depth (QD) $<$ 2.0, Phred-scaled *P*-value (FS) $>$ 60.0, root mean square of the mapping quality (MQ) $<$ 40.0, MQ rank sum test approximation (MQRankSum) of 12.5, and a read position rank sum test approximation (ReadPosRankSum) of 8. Individual samples with a genotype quality $<$ 99 for 80% of the variants and/or coverage $<$ 10 for 50% of the variants were removed; individual SNPs where $\geq$ 9 individuals had depth of coverage $<$ 10 and/or a genotype quality $<$ 99 were removed. Because most toxins belong to large gene families (Casewell et al. 2009; Margres et al. 2017), the toxin data set required additional filtering to eliminate confounded paralogs and/or pseudogenes in the alignments. All SNPs with a genotype quality $<$ 99 and depth of coverage $<$ 10 were removed, and all alignments containing putative toxin SNPs were verified manually.

Population Genetic Estimates

We estimated Weir and Cockerham's mean F_{ST} using vcftools (Danecek et al. 2011) for the neutral and toxin SNP data sets.

To generate sequence alignments for migration rate estimation, we performed the following assembly approach for the anchored, short anonymous, and long anonymous loci for ten individuals from each population. Following sequencing, paired reads were merged prior to assembly following (Rokyta et al., 2012). In short, for each degree of overlap each read pair was evaluated with respect to the probability of obtaining the observed number of matches by chance. The overlap with the lowest probability was chosen if the *P*-value was $<10^{-10}$. Read pairs with a *P*-value below the threshold were merged and quality scores were recomputed for overlapping bases. Read pairs failing to merge were utilized but left unmerged during the assembly.

Divergent reference assembly was used to map reads to the probe regions and extend the assembly into the flanking regions; *Calamaria pavimentata* was chosen as the reference for the anchored assembly (Prum et al. 2015; Ruane et al. 2015) and *B. jararaca* was chosen as the reference for the anonymous assemblies. Matches were called if 17 bases matched a library of spaced 20-mers derived from the conserved reference regions. Preliminary reads were then considered mapped if the 55 matches were found over 100 consecutive bases in the reference sequences (all possible gap-free alignments between the read and the reference were considered). The approximate alignment position of mapped reads were estimated using the position of the spaced 20-mer, and all 60-mers existing in the read were stored in a hash table used by the de novo assembler. The de novo assembler identifies exact matches between a read and one of the 60-mers found in the hash table. Simultaneously, using the two levels of assembly described above, the three read files were traversed repeatedly until a pass through the reads produced no additional mapped reads. For each locus, a list of all 60-mers found in the mapped reads was compiled and 60-mers were clustered if found together in at least two reads. Contigs were estimated from

60-mer clusters. In the absence of contamination, low coverage, or gene duplication, each locus should produce one assembly cluster. Consensus bases were called from assembly clusters as unambiguous base calls if polymorphisms could be explained as sequencing error (assuming a binomial probability model with the probability of error = 0.1 and $\alpha = 0.05$). Otherwise ambiguous bases were called (e.g., "R" was used if "A" and "G" were observed). Called bases were soft-masked for sites with coverage < 5. Assembled contigs derived from < 25 reads were removed in order to reduce the effects of cross contamination and rare sequencing errors in index reads.

After grouping homologous sequences obtained by enrichment, orthology was determined for each locus following Prum et al. (2015). For each locus, pairwise distances among homologs were computed using an alignment-free approach based on percent overlap of continuous and spaced 20-mers. Using the distance matrix, sequences were clustered using a Neighbor-Joining algorithm, but allowing at most one sequence per individual to be in a given cluster. Note that flanks recovered through extension assembly contain more variable regions and allow gene copies to be sorted efficiently. In order to reduce the effects of missing data, clusters containing fewer than 50% of the individuals were removed from downstream processing.

Orthologous sequences were processed using a combination of automated and manual steps in order to generate high quality alignments. Sequences in each orthologous cluster were aligned using MAFFT (v7.023 b; Katoh and Standley 2013), with "–genafpair" and "–maxiterate 1000" flags utilized. The alignment for each locus was then trimmed/masked using the steps from Prum et al. (2015). First, each alignment site was identified as conserved if the most commonly observed character was present in >40% of the sequences. Second, using a 20 bp window, each sequence was scanned for regions that did not contain at least 10 characters matching to the common base at the corresponding conserved site. Characters from regions not meeting this requirement were masked. Third, sites with fewer than 10 unmasked bases were removed from the alignment. A visual inspection of each masked alignment was carried out in Geneious (v7; Kearse et al. 2012). Regions of sequences identified as obviously misaligned or paralogous were removed. Individual loci were removed if any individual was absent from the alignment and/or if the shortest sequence was <50% of the length of the longest sequence in the alignment. Final alignments were generated by eliminating gaps and trimming each locus to the shortest locus in the alignment to eliminate missing data. The final alignments for analyses were as follows:

(1) Anchored loci ($n = 329$; 535,894 sites)
(2) Long anonymous loci ($n = 266$; 421,196 sites)
(3) Short anonymous loci ($n = 721$; 453,235 sites)

We designed migration models using Migrate-n v. 4.2.8 (Beerli 2008) to estimate effective population sizes ($\theta = 4N_e\mu$) and historical migration rates ($M = m/\mu$) across island and mainland populations for *C. adamanteus*. Migrate-n uses a coalescent framework and assumes an n-island model in which genetic material is exchanged only through migration and not population divergence. As previously described by Beerli and others (Beerli and Felsenstein 1999, 2001; Beerli and Palczewski 2010; Barrow et al. 2015), marginal likelihoods can be used to compare models that differ in the number of migration parameters to test the directionality of gene flow and connectivity of populations, and models that combine populations can be compared with those that divide populations in order to determine whether individuals come from the same randomly mating population. We used three neutral loci data sets described above to compare the following four models (fig. 5):

(1) Two effective population sizes and bidirectional migration (mainland ↔ island; four parameters)
(2) Two effective population sizes and unidirectional migration (mainland → island; three parameters)
(3) Two effective population sizes and unidirectional migrational (island → mainland; three parameters)
(4) A single panmictic population; one effective population size and no migration (single parameter)

All analyses were run on the High Performance Computing Cluster at the Research Computing Center at Florida State University, and each locus type (e.g., anchored) was analyzed independently. We varied the prior settings with a full model (i.e., all parameters estimated) for all three data sets before choosing uniform priors for θ (Minimum = 0, Maximum = 0.10, Delta = 0.01) and M (Minimum = 0, Maximum = 10,000.00, Delta = 1,000.00). We conducted 20 replicates for each model with a static heating scheme using four chains, discarded 40,000 trees as burn-in, and recorded 40,000 steps with an increment of 100 (resulting in 80 million samples) as described by Barrow et al. (2015). We used the same prior settings and search parameters when testing all models. Parameter posterior distributions were used to assess convergence (i.e., single peaks with smooth curves). Bayes Factors were calculated as the ratio of the marginal likelihoods (approximated by thermodynamic integration with Bezier approximation; Beerli and Palczewski 2010) to calculate model probabilities; these probabilities were used to select the most appropriate model.

Supplementary Material

Supplementary data are available at *Molecular Biology and Evolution* online.

Acknowledgments

This work was supported by the National Science Foundation (DEB 1145987 to D.R.R., E.M.L., and A.R.L.), the Gopher Tortoise Council (to M.J.M.), and Florida State University (to D.R.R. and M.J.M.). The authors thank Nathanael Herrera, Pierson Hill, Kristin Klimley, Mark S. Margres, Flavio Morrissiey, Joe Pfaller, and Sarah Smiley for their assistance in the field. The authors thank Megan Lamb, Danielle Jones, Ethan Bourque, Jennifer Wanat, and Rebecca Bernard with the Florida DEP and Apalachicola River NERR for access to

field sites. The authors thank Peter Beerli for his assistance with the Migrate-n analyses, Margaret Seavy and Rakesh Singh for their assistance with proteomic analyses, and the Florida State University Coastal and Marine Laboratory for logistical support. Reads were deposited in the NCBI Sequence Read Archive under accessions SRR5485689–SRR5485719 (BioSample SAMN06845833–SAMN06845852). Division of Environmental Biology, 1145987.

References

Barrow L, Bigelow A, Phillips C, Lemmon E. 2015. Phylogeographic inference using Bayesian model comparison across a fragmented chorus frog species complex. *Mol Ecol*. 24(18):4739–4758.

Beerli P. 2008. Migrate version 3.0: a maximum likelihood and Bayesian estimator of gene flow using the coalescent. Available from: http://popgen.sc.fsu.edu/Migrate/Migrate-n.html.

Beerli P, Felsenstein J. 1999. Maximum-likelihood estimation of migration rates and effective population numbers in two populations using a coalescent approach. *Genetics* 152:763–773.

Beerli P, Felsenstein J. 2001. Maximum likelihood estimation of a migration matrix and effective population sizes in *n* subpopulations by using a coalescent approach. *Proc Natl Acad Sci U S A*. 98(8):4563–4568.

Beerli P, Palczewski M. 2010. Unified framework to evaluate panmixia and migration direction among multiple sampling locations. *Genetics* 185(1):313–326.

Biardi JE, Ho CYL, Marcinczyk J, Nambiar KP. 2011. Isolation and identification of a snake venom metalloproteinase inhibitor from California ground squirrel (*Spermophilus beecheyi*) blood sera. *Toxicon* 58(6–7):486–493.

Carroll S. 2008. Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. *Cell* 134(1):25–36.

Casewell NR, Harrison RA, Wüster W, Wagstaff SC. 2009. Comparative venom gland transcriptome surveys of the saw-scaled vipers (Viperidae: *Echis*) reveal substantial intra-family gene diversity and novel venom transcripts. *BMC Genomics*. 10(1):564.

Casewell NR, Wagstaff SC, Harrison RA, Renjifo C, Wüster W. 2011. Domain loss facilitates accelerated evolution and neofunctionalization of duplicate snake venom metalloproteinase toxin genes. *Mol Biol Evol*. 28(9):2637–2649.

Casewell NR, Huttley GA, Wüster W. 2012. Dynamic evolution of venom proteins in squamate reptiles. *Nat Commun*. 3:1066.

Chenoweth S, Rundle H, Blows M. 2010. The contribution of selection and genetic constraints to phenotypic divergence. *Am Nat*. 175(2):186–196.

Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, McVean G, Durbin R. 2011. The variant call format and VCFtools. *Bioinformatics* 27(15):2156–2158.

Daub J, Hofer T, Cutivet E, Dupanloup I, Quintana-Murci L, Ribonson-Rechavi M, Excoffier L. 2013. Evidence for polygenic adaptation to pathogens in the human genome. *Mol Biol Evol*. 30(7):1544–1558.

Dawkins R, Krebs J. 1979. Arms races between and within species. *Proc R Soc B*. 205(1161):489–511.

DePristo M, Banks E, Poplin R, Garimella K, Maguire J, Hartl C, Philippakis A, del Angel G, Rivas M, Hanna M, et al. 2011. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet*. 43(5):491–498.

Egozcue JJ, Pawlowsky-Glahn V, Mateu-Figueras G, Barceló-Vidal C. 2003. Isometric logratio transformations for compositional data analysis. *Math Geol*. 35(3):279–300.

Fitzpatrick S, Gerberich J, Kronenberger J, Angeloni L, Funk W. 2015. Locally adapted traits maintained in the face of high gene flow. *Ecol Lett*. 18(1):37–47.

Fraser H. 2013. Gene expression drives local adaptation in humans. *Genome Res*. 23(7):1089–1096.

Garant D, Forde SE, Hendry AP. 2007. The multifarious effects of dispersal and gene flow on contemporary adaptation. *Funct Ecol*. 21(3):434–443.

Geffeney SL, Fujimoto E, Brodie ED, Brodie ED, Ruben PC. 2005. Evolutionary diversification of TTX-resistant sodium channels in a predator–prey interaction. *Nature* 434(7034):759–763.

Hamilton C, Lemmon A, Lemmon E, Bond J. 2016. Expanding anchored hybrid enrichment to resolve both deep and shallow relationships within the spider tree of life. *BMC Evol Biol*. 16(1):212.

Hamilton M, Russo R, Thurston R. 1977. Trimmed Spearman–Karber method for estimating median lethal concentrations in toxicity bioassays. *Environ Sci Technol*. 11(7):714–719.

Hill W, Zhang X-S. 2012. Assessing pleiotropy and its evolutionary consequences: pleiotropy is not necessarily limited, nor need it hinder the evolution of complexity. *Nat Rev Genet*. 13(4):296.

Hoekstra H, Hirschmann R, Bundey R, Insel P, Crossland J. 2006. A single amino acid mutation contributes to adaptive beach mouse color pattern. *Science* 313(5783):101–104.

Hoekstra HE, Coyne JA. 2007. The locus of evolution: evo devo and the genetics of adaptation. *Evolution* 61(5):995–1016.

Holding M, Biardi J, Gibbs H. 2016. Coevolution of venom function and venom resistance in a rattlesnake predator and its squirrel prey. *Proc R Soc B*. 283(1829):20152841.

Huntzinger E, Izaurralde E. 2011. Gene silencing by microRNAs: contributions of translational repression and mRNA decay. *Nat Rev Genet*. 12(2):99–110.

Innan H, Kim Y. 2008. Detecting local adaptation using the joint sampling of polymorphism data in the parental and derived populations. *Genetics* 179(3):1713–1720.

Jansa SA, Voss RS. 2011. Adaptive evolution of the venom-targeted vWF protein in opossums that eat pitvipers. *PLoS ONE*. 6(6):e20997.

Katoh K, Standley D. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol*. 30(4):772–780.

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, et al. 2012. Genious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28(12):1647–1649.

Konczal M, Babik W, Radwan J, Sadowska E, Koteja P. 2015. Initial molecular-level response to artificial selection for increased aerobic metabolism occurs primarily through changes in gene expression. *Mol Biol Evol*. 32(6):1461–1473.

Kvitek DJ, Sherlock G. 2011. Reciprocal sign epistasis between frequently experimentally evolved adaptive mutations causes a rugged fitness landscape. *PLoS Genet*. 7(4):e1002056.

Lamichhaney S, Han F, Berglund J, Wang C, Almen M, Webster M, Grant B, Grant P, Andersson L. 2016. A beak size locus in Darwin's finches facilitated character displacement during a drought. *Science* 353(6284):470–474.

Lemmon A, Emme S, Lemmon E. 2012. Anchored hybrid enrichment for massively high-throughput phylogenomics. *Syst Biol*. 61(5):727–744.

Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows–Wheeler Transform. *Bioinformatics* 25(14):1754–1760.

Lopez G, Rink W. 2007. Characteristics of the burial environment related to quartz SAR-OSL dating at St. Vincent Island, NW Florida, USA. *Quart Geochronol*. 2(1–4):65–70.

Lourenco J, Glemin S, Galtier N. 2013. The rate of molecular adaptation in a changing environment. *Mol Biol Evol*. mst026.

Manceau M, Domingues V, Mallarino R, Hoekstra H. 2011. The developmental role of Agouti in color pattern evolution. *Science* 331(6020):1062–1065.

Margres M, McGivern J, Wray K, Seavy M, Calvin K, Rokyta D. 2014. Linking the transcriptome and proteome to characterize the venom of the eastern diamondback rattlesnake (*Crotalus adamanteus*). *J Proteomics*. 96:145–158.

Margres M, McGivern J, Seavy M, Wray K, Facente J, Rokyta D. 2015. Contrasting modes and tempos of venom expression evolution in two snake species. *Genetics* 199(1):165–176.

Margres M, Wray K, Seavy M, McGivern J, Herrera N, Rokyta D. 2016a. Expression differentiation is constrained to low-expression proteins over ecological timescales. *Genetics* 202(1):273–283.

Margres M, Walls R, Suntravat M, Lucena S, Sanchez E, Rokyta D. 2016b. Functional characterizations of venom phenotypes in the eastern diamondback rattlesnake (*Crotalus adamanteus*) and evidence for expression-driven divergence in toxic activities among populations. *Toxicon* 119:28–38.

Margres M, Bigelow A, Lemmon E, Lemmon A, Rokyta D. 2017. Selection to increase expression, not sequence diversity, precedes gene family origin and expansion in rattlesnake venom. *Genetics* 206(3):1569–1580.

McCue MD. 2006. Cost of producing venom in three North American pitviper species. *Copeia* 2006(4):818–825.

McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, DePristo M. 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 20(9):1297–1303.

Means D. (2017). Diamonds in the rough, natural history of the eastern diamondback rattlesnake. Tallahassee (FL): Tall Timbers Press.

Prum R, Berv J, Dornburg A, Field D, Townsend J, Lemmon E, Lemmon A. 2015. A comprehensive phylogeny of birds (Aves) using targeted next-generation DNA sequencing. *Nature* 526(7574):569–573.

Ritz C, Baty F, Streibig J, Gerhard D. 2015. Dose–response analysis using R. *PLoS ONE*. 10(12):e0146021.

Rokyta D, Margres M, Calvin K. 2015a. Post-transcriptional mechanisms contribute little to phenotypic variation in snake venoms. *G3: Genes Genomes Genetics*. g3:115.

Rokyta D, Wray K, McGivern J, Margres M. 2015b. The transcriptomic and proteomic basis for the evolution of a novel venom phenotype within the Timber Rattlesnake (*Crotalus horridus*). *Toxicon* 98:34–48.

Rokyta D, Margres M, Ward M, Sanchez E. 2017. The genetics of venom ontogeny in the eastern diamondback rattlesnake (*Crotalus adamanteus*). *PeerJ* 5:e3249.

Rokyta DR, Joyce P, Caudle SB, Miller C, Beisel CJ, Wichman HA, Malik HS. 2011. Epistasis between beneficial mutations and the phenotype-to-fitness map for a ssDNA virus. *PLoS Genet.* 7(6):e1002075.

Rokyta DR, Lemmon AR, Margres MJ, Aronow K. 2012. The venom-gland transcriptome of the eastern diamondback rattlesnake (*Crotalus adamanteus*). *BMC Genomics*. 13(1):312.

Rosenblum E, Rompler H, Schoneberg T, Hoekstra H. 2010. Molecular and functional basis of phenotypic convergence in white lizards at White Sands. *Proc Natl Acad Sci U S A*. 107(5):2113–2117.

Ruane S, Raxworthy C, Lemmon A, Lemmon E, Burbrink F. 2015. Comparing species tree estimation with large anchored phylogenomic and small Sanger-sequenced molecular datasets: an empirical study on *Malagasy pseudoxyrhophiine* snakes. *BMC Evol Biol*. 15(1):1.

Savolainen O, Lascoux M, Merila J. 2013. Ecological genomics of local adaptation. *Nat Rev Genet*. 14(11):807–820.

Schwanhausser B, Busse D, Li N, Dittmar G, Schuchhardt J, Wolf J, Chen W, Selbach M. 2011. Global quantification of mammalian gene expression control. *Nature* 473(7347):337–342.

Slatkin M. 1987. Gene flow and the geographic structure of natural populations. *Science* 236(4803):787–792.

Stone B. (2015). tsk: Trimmed Spearman–Karber Method. R package version 1.2.

Storz J. 2016. Causes of molecular convergence and parallelism in protein evolution. *Nat Rev Genet*. 17(4):239–250.

Templ M, Hron K, Filzmoser P. (2011). *Compositional Data Analysis. Theory and Applications*. Chapter robCompositions: an R-package for robust statistical analysis of compositional data. Chichester (United Kingdom): John Wiley & Sons, p. 341–355.

Thompson J, Nuismer S, Gomulkiewicz R. 2002. Coevolution and maladaptation. *Integr Comp Biol*. 42(2):381–387.

Li W, Godzik A. 2006. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* 22(13):1658–1659.

Wellenreuther M, Hansson B. 2016. Detecting polygenic evolution: problems, pitfalls, and promises. *Trends Genet*. 32(3):156–164.

Wheeler M, Park R, Bailer A. 2006. Comparing median lethal concentration values using confidence interval overlap or ratio tests. *Environ Toxicol Chem*. 25(5):1441–1444.

Yokoyama S, Tada T, Zhang H, Britt L. 2008. Elucidation of phenotypic adaptations: molecular analyses of dim-light vision in proteins in vertebrates. *Proc Natl Acad Sci U S A*. 105(36):13480–13485.

Zhang J, Kobert K, Flouri T, Stamatakis A. 2014. PEAR: a fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics* 30(5):614–620.

Zhang L, Reed R. 2016. Genome editing in butterflies reveals that *spalt* promotes and *Distal-less* represses eyespot colour patterns. *Nat Commun*. 7:11769.