


RESEARCH

Open Access

# Land use and life history constrain adaptive genetic variation and reduce the capacity for climate change adaptation in turtles



Nathan W. Byer<sup>1\*</sup> , Emily D. Fountain<sup>2</sup>, Brendan N. Reid<sup>3</sup>, Kristen Miller<sup>2</sup>, Paige J. Kulzer<sup>2</sup> and M. Zachariah Peery<sup>2</sup>

## Abstract

**Background:** Rapid anthropogenic climate change will require species to adapt to shifting environmental conditions, with successful adaptation dependent upon current patterns of genetic variation. While landscape genomic approaches allow for exploration of local adaptation in non-model systems, most landscape genomics studies of adaptive capacity are limited to exploratory identification of potentially important functional genes, often without *a priori* expectations as to the gene functions that may be most important for climate change responses. In this study, we integrated targeted sequencing of genes of known function and genotyping of single-nucleotide polymorphisms to examine spatial, environmental, and species-specific patterns of potential local adaptation in two co-occurring turtle species: the Blanding's turtle (*Emydoidea blandingii*) and the snapping turtle (*Chelydra serpentina*).

**Results:** We documented divergent patterns of spatial clustering between neutral and putatively adaptive genetic variation in both species. Environmental associations varied among gene regions and between species, with stronger environmental associations detected for genes involved in stress response and for the more specialized Blanding's turtle. Land cover appeared to be more important than climate in shaping spatial variation in functional genes, indicating that human landscape alterations may affect adaptive capacity important for climate change responses.

**Conclusions:** Our study provides evidence that responses to climate change will be contingent on species-specific adaptive capacity and past history of exposure to human land cover change.

**Keywords:** Genotype-by-sequencing, Custom amplicon, Landscape genomics, Local adaptation, Reptiles

## Background

Climate change will be among the greatest challenges facing biodiversity through the 21st century [1–3]. The intensity and speed of contemporary climate change are likely outside of the range of conditions previously experienced by most species on earth, which will require comparatively rapid adaptive responses by species unable to disperse to favorable climate conditions or with limited plasticity [3–6]. Although evolutionary rescue due

to novel mutations will be important for climate change responses [7, 8], current levels of adaptive genetic variation, defined as differences among genomes of individuals resulting from natural selection, may also be important for evolutionary responses [9]. Since the capacity for evolutionary change in the face of environmental change is predicated on sufficient adaptive genetic variation in the present day, characterizing adaptive genetic variation is important for predicting population- and species-level resiliency to climate change [10].

Given that adaptive genetic variation is a function of species-specific traits, adaptive potential is expected to vary among species [9, 11–14]. For example, short-lived

\* Correspondence: [nbyer@unr.edu](mailto:nbyer@unr.edu)

<sup>1</sup>University of Nevada–Reno, NV 89557 Reno, USA

Full list of author information is available at the end of the article



© The Author(s). 2021 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

species with relatively abbreviated generation times are expected to have a greater capacity to adapt to rapid environmental change than longer-lived species [10]. Furthermore, the likelihood of adaptive responses is also expected to vary along habitat specialist-generalist gradients, with specialists expected to respond via natural selection to environmental change and generalists expected to respond through phenotypic plasticity [15, 16]. Thus, a better understanding of how species-specific characteristics affect adaptation will be crucial for making predictions of climate change responses [10].

Since local adaptation represents an interaction between genetic polymorphisms with functional roles, spatial variation in selection pressures, and phenotypic effects on fitness, focusing on genes of known functional significance allows for tests of ecological hypotheses related to adaptation [9, 10]. Indeed, such studies have proven instrumental for identifying and understanding gene functional groups likely to be important for local adaptation to climate conditions within and among taxa. For example, the gene *GIGANTEA-GI5*, which has documented interactions with plant circadian clocks [17], has been found to be strongly associated with spatial variation in temperature, providing initial evidence that this gene is involved in local adaptation to temperature [18]. A number of additional gene functional groups – including genes related to stress response [19] and development [20] – hold promise for further study of the capacity for local adaptation to climate change.

Identifying functional genes important for climate change adaptation in non-model organisms often relies on landscape genomic approaches for investigating adaptive genetic variation [9, 21]. Tests of associations between functional genetic variation – defined in this study as genetic variants that potentially affect protein function and expression, and thus may be under selection – and environmental features can be used to explore and identify environmentally-imposed selective pressures that may have shaped adaptive genetic variation [9, 21, 22]. Moreover, characterizing spatial patterns in functionally-relevant gene regions allows for identification of populations that are expected to be less capable of responding adaptively to climate change, which can then be used to prioritize conservation efforts [18, 22]. Finally, functional genes that vary along environmental gradients represent potentially useful targets for more detailed investigations of how functional genetic variation may have consequences for fitness across the landscape [9, 22].

While substantial progress has been made in identifying and exploring local adaptation, several limitations challenge landscape genomic studies. First, most studies do not compare genes of known function to independent, unlinked neutral regions, instead relying on

screening large numbers of loci to identify either markers that are more related to environmental gradients or more divergent between populations than others; this relies on the assumption that these outlier loci hold adaptive significance, yet very few outlier loci are often found to match with genes of known function [9, 23]. Second, many landscape genomic studies of local adaptation focus on a single target species, despite evidence that patterns of local adaptation are influenced by species-specific life history traits [10, 24, 25]. The influence of species traits and life history on local adaptation means that insights into the capacity for local adaptation may not transfer among species, which will necessitate studies that delve into the species-level traits that drive local adaptation [9, 10, 25]. These factors collectively limit inferences from many landscape genomic studies of local adaptation to initial, exploratory identification of loci worth further investigation, and challenge the testing of ecological and evolutionary hypotheses.

In this study, we addressed several of the aforementioned limitations by characterizing how putatively neutral and putatively adaptive genetic variation in two turtle species is affected by species-specific life history and habitat utilization traits and environmental characteristics. We addressed questions related to adaptive genetic variation using turtles for several reasons. Turtles, with their long lifespans and delayed sexual maturity, may have a limited capacity to respond rapidly to environmental change [26–28]. However, turtles possess three characteristics directly tied to climatic conditions: (1) they are ectotherms, and many physiological processes are directly linked to thermal characteristics; (2) many species exhibit Temperature-Dependent Sex Determination (TSD) whereby hatchling sex ratios are a function of incubation temperatures [29], and (3) a number of temperature-related traits, including TSD, display documented spatial differentiation in turtles [30]. These three features suggest that increases in temperature may be more likely to lead to adaptive change for turtles than many other taxa, despite life history traits that may reduce the likelihood of adaptive responses. Therefore, landscape genomics tools may provide important insights into whether turtle populations possess sufficient adaptive genetic variation to allow evolutionary response to climate change.

We focused our analysis to two species: the Blanding's turtle (*Emydoidea blandingii*) and the snapping turtle (*Chelydra serpentina*). The Blanding's turtle (*E. blandingii*) is a medium-bodied freshwater turtle that requires a juxtaposition of open terrestrial habitat for nesting with permanent wetlands for foraging and overwintering [31]. Due to pressures posed by habitat loss and road mortality, combined with this species' long generation times (36–47 years), *E. blandingii* is currently listed on the

IUCN Red List as “Endangered”, and declines have been observed across much of the species’ range [32]. The snapping turtle (*C. serpentina*) is a large-bodied freshwater generalist that is widely distributed throughout the United States and Canada [31]. Although both species have TSD, in *E. blandingii* females are produced at incubation temperatures higher than a certain pivotal temperature, whereas for *C. serpentina* males are produced at intermediate temperatures and females are produced at extreme temperatures [31]. Importantly, the range of incubation temperatures producing male and female hatchlings exhibits heritable variation [33] and varies spatially across snapping turtle populations [34], although intraspecific variation in TSD has yet to be examined in Blanding’s turtle.

We used both genome-wide discovery and genotyping of SNPs and targeted sequencing of four target genes that hold potential functional importance for adaptation to climate change at regional and local scales to test several hypotheses. First, we hypothesized that variation in targeted gene regions may be of adaptive importance; thus, we predicted that spatial genetic variation in targeted gene regions should differ from patterns observed with putatively neutral SNP loci. Second, we hypothesized that species-level life and natural history traits should influence putatively adaptive genetic variation. From this hypothesis, we explored two contrasting predictions. One alternative is that habitat specialization leads to a predisposition towards evolutionary adaptation over plastic responses in the face of environmental change [35, 36]; under this prediction, we may expect stronger environmental associations and more putatively adaptive genetic variation in the habitat specialist, *E. blandingii*, than the generalist, *C. serpentina*. In contrast, in our system, the sensitive conservation status of the Blanding’s turtle and its lower effective population sizes may mean that it has less adaptive genetic variation overall and weaker environmental associations than the widespread *C. serpentina*. Finally, we hypothesized that environmental associations in targeted genes should reflect the putative functions fulfilled by each gene. From this we predicted stronger associations between climatic conditions than land cover in genes associated with temperature-sensing and sex determination; in contrast, we predicted associations with both land cover and climate variables in genes related to stress response. By addressing evolutionarily- and ecologically-relevant hypotheses about local adaptation with multiple lines of genetic evidence, we provide novel insights into how species-level traits and environmental features interact to constrain local adaptation.

## Results

### SNP statistics

The *E. blandingii* SNP dataset consisted of 2,811 SNPs across 37 individuals, and had a mean minor allele frequency (MAF) of 0.210, a mean  $H_O$  of 0.192, and a mean  $H_E$  of

0.294. The *E. blandingii* SNP dataset global  $F_{ST}$  was 0.0300, and 20/21 pairwise comparisons between sites were significant despite low sample sizes (95.2%; Supplementary Materials S1). For these SNPs, we detected a significant effect of isolation-by-distance based on a Mantel test of individual genetic and geographic distances ( $r = 0.306$ ,  $p = 0.015$ ). The *C. serpentina* SNP dataset consisted of 1,606 SNPs across 33 individuals, and had a mean MAF of 0.169, a mean  $H_O$  of 0.169, and a mean  $H_E$  of 0.244. Global  $F_{ST}$  was 0.068, and 9/10 pairwise comparisons between sites were significant despite low sample sizes (90%; Supplementary Materials S1). For these SNPs, we detected a significant effect of isolation-by-distance based on a Mantel test of individual genetic and geographic distances ( $r = 0.363$ ,  $p = 0.001$ ).

### Targeted gene statistics

For *E. blandingii*, a total of two variant sites were identified for CIRBP, six for HSPA8, one for Sox9, and one for TRPV1; after excluding variant sites with the minor allele represented in fewer than two individuals for the unfiltered dataset, this left two for CIRBP, three for HSPA8, zero for Sox9, and one for TRPV1 (Table 1). Although most variant sites were synonymous or noncoding, both polymorphisms for CIRBP coded for amino acid changes (Supplementary Materials S2). Genetic diversity was generally low ( $H_O = 0.0795$ ), but was highest for one of three variants discovered in HSPA8 (0.278). Across all target genes and sample sites, global  $F_{ST}$  was lower than global  $F_{ST}$  for the SNP dataset (0.0246), and 33/120 pairwise comparisons between sample sites were significant (27.5%; Supplementary Materials S3). We failed to detect a significant effect of isolation by distance as well, based on a Mantel test of individual genetic and geographic distances ( $r = 0.107$ ,  $p = 0.097$ ).

We observed more variants in target genes overall for *C. serpentina*, with zero variant sites for CIRBP, six for HSPA8, zero for Sox9, and nine for TRPV1; however, the nine variant sites in TRPV1 characterized only two

**Table 1** Variant sites (with minor allele present in at least 2 individuals upon initial filtering) per species and target gene region. EB = Blanding’s turtle (*Emydoidea blandingii*), and CS = snapping turtle (*Chelydra serpentina*)

Gene	Species	# variants
CIRBP	EB	2
HSPA8	EB	3
Sox9	EB	0
TRPV1	EB	1
CIRBP	CS	0
HSPA8	CS	6
Sox9	CS	0
TRPV1	CS	9

distinct haplotypes, and were thus analyzed as one distinct marker (Table 1). Although most variant sites were synonymous or noncoding, three variants for TRPV1 coded for amino acid changes (Supplementary Materials S2). Mean  $H_O$  was 0.129, global  $F_{ST}$  was lower than  $F_{ST}$  for the SNP dataset for target genes (0.0590), and only 18/91 pairwise comparisons between sites were significant (18.6%; Supplementary Materials S3). We detected a significant effect of isolation by distance as well ( $r = 0.070$ ,  $p = 0.03$ ).

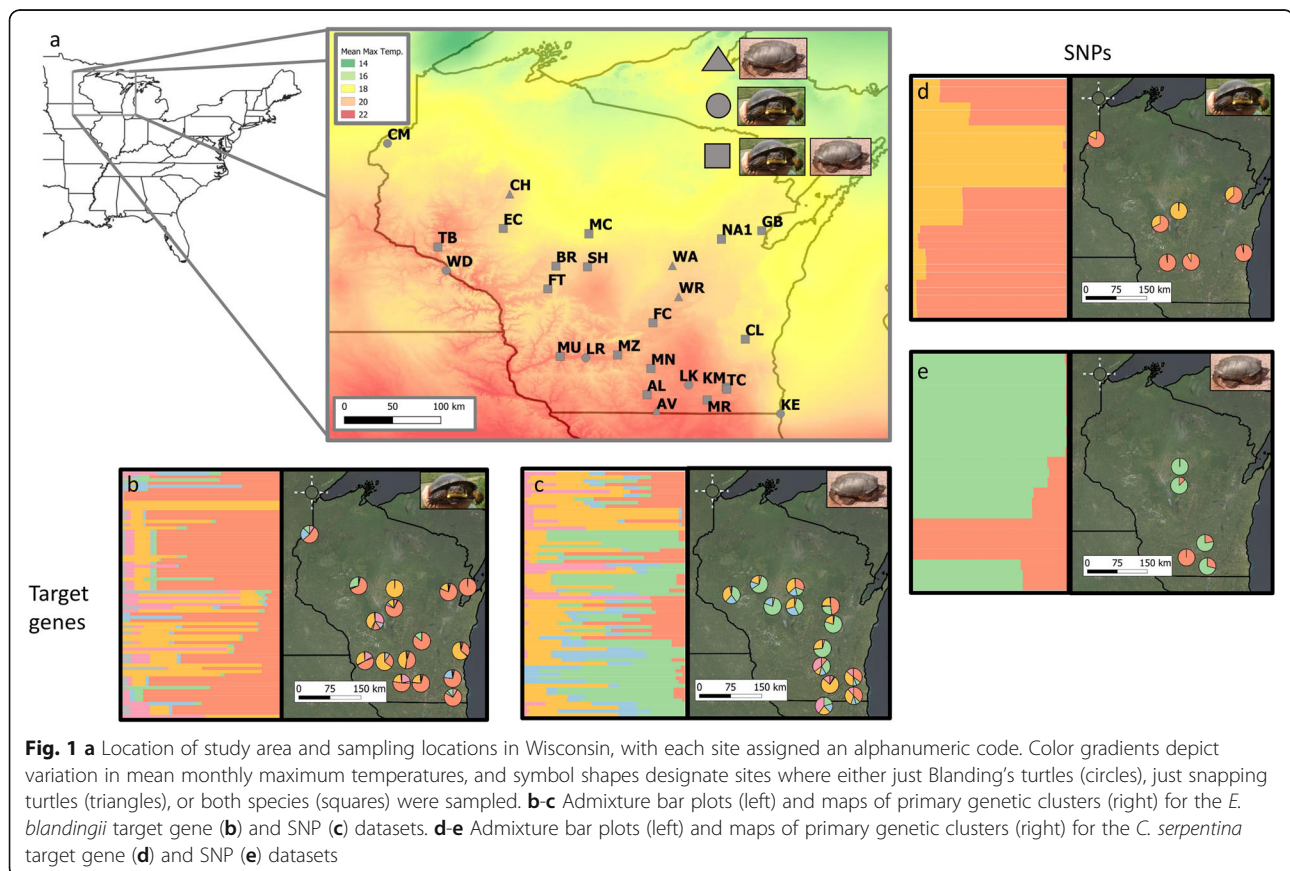
### Spatial patterns and clustering

The number of genetic clusters and their distribution on the landscape differed between clustering approaches, datasets, and species. For the *E. blandingii* target dataset, the number of inferred clusters varied between approaches, with  $K = 4$  for DAPC and  $K = 6$  for TESS3 based on observed inflection points in cross-validation scores; in contrast, the number of inferred clusters for the *E. blandingii* SNP dataset was lower, ranging from  $K = 1$  for DAPC to  $K = 2$  for TESS3, with no clear inflection points observed in cross-validation scores (Supplementary Materials S4). Although clusters for both datasets largely overlapped spatially, these clustering approaches did highlight sites that contained ancestry from relatively few clusters for target genes (FC, GB, and MC) and

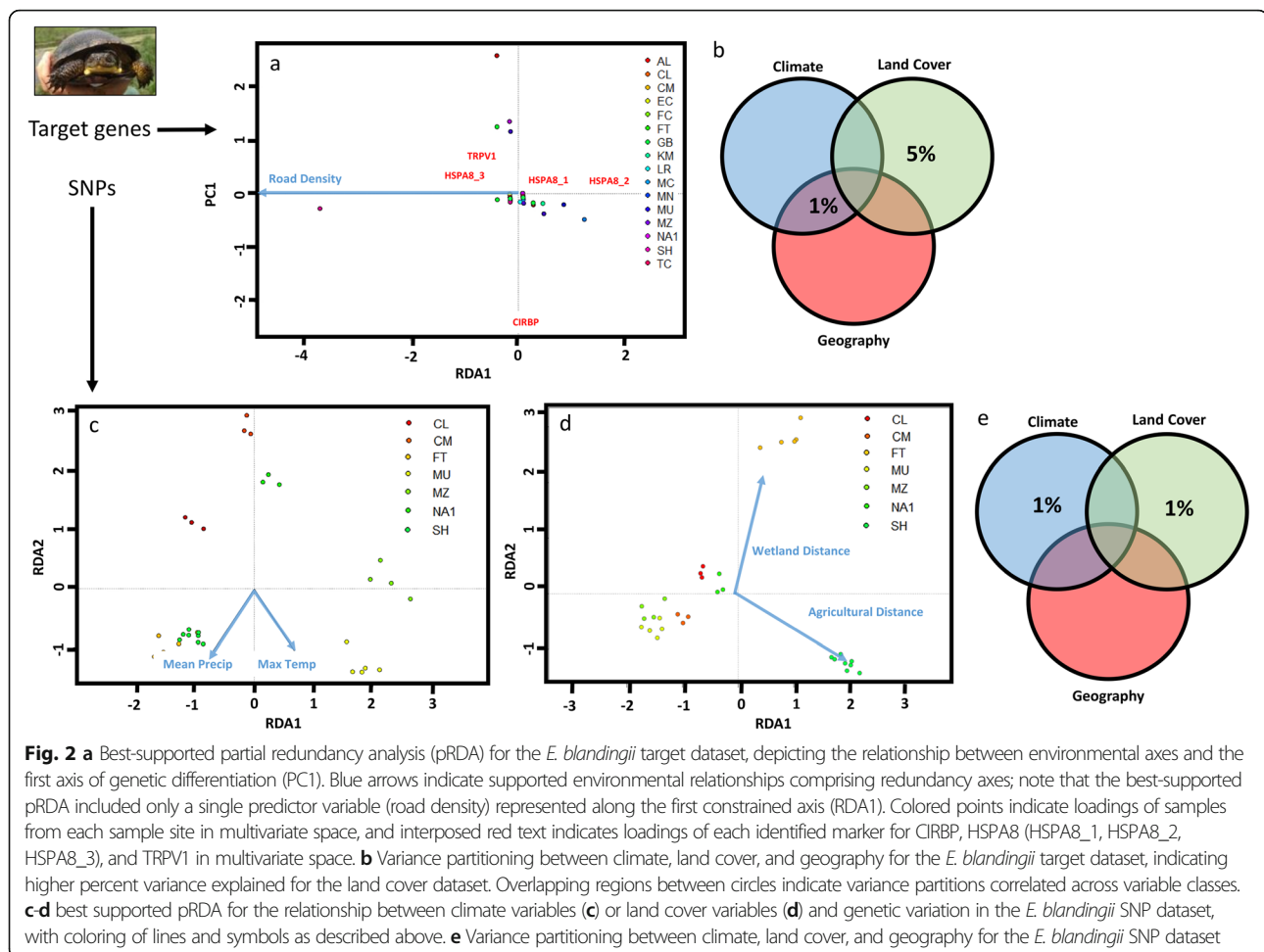
sites that represented more inferred clusters (CM, FT, MU, and SH, Fig. 1). For the *C. serpentina* target dataset, the number of inferred clusters varied from  $K = 3$  for DAPC to  $K = 5$  for TESS3; in contrast, the number of inferred clusters for the associated SNP dataset was lower, with  $K = 1$  for DAPC and  $K = 2$  for TESS3 (Fig. 1; Supplementary Materials S4). As with the *E. blandingii* SNP dataset, we observed no clear inflection points in cross-validation scores for the *C. serpentina* SNP dataset.

### Species- and dataset-level environmental associations

Although we detected more nonsynonymous substitutions for *C. serpentina* than *E. blandingii*, stronger environmental associations for the *E. blandingii* target dataset indicated support for the prediction that habitat specialists are more likely to be locally adapted to environmental conditions. For the *E. blandingii* target dataset, we detected a significant association between genetic differentiation and road density ( $F = 3.503$ ,  $p = 0.042$ ; Fig. 2), and variance partitioning indicated that the largest proportion of variance was explained by land cover covariates (5%; Fig. 2). In contrast, for the *E. blandingii* SNP dataset, while we detected significant associations with distance to agricultural cover ( $F = 1.2187$ ,  $p = 0.001$ ),







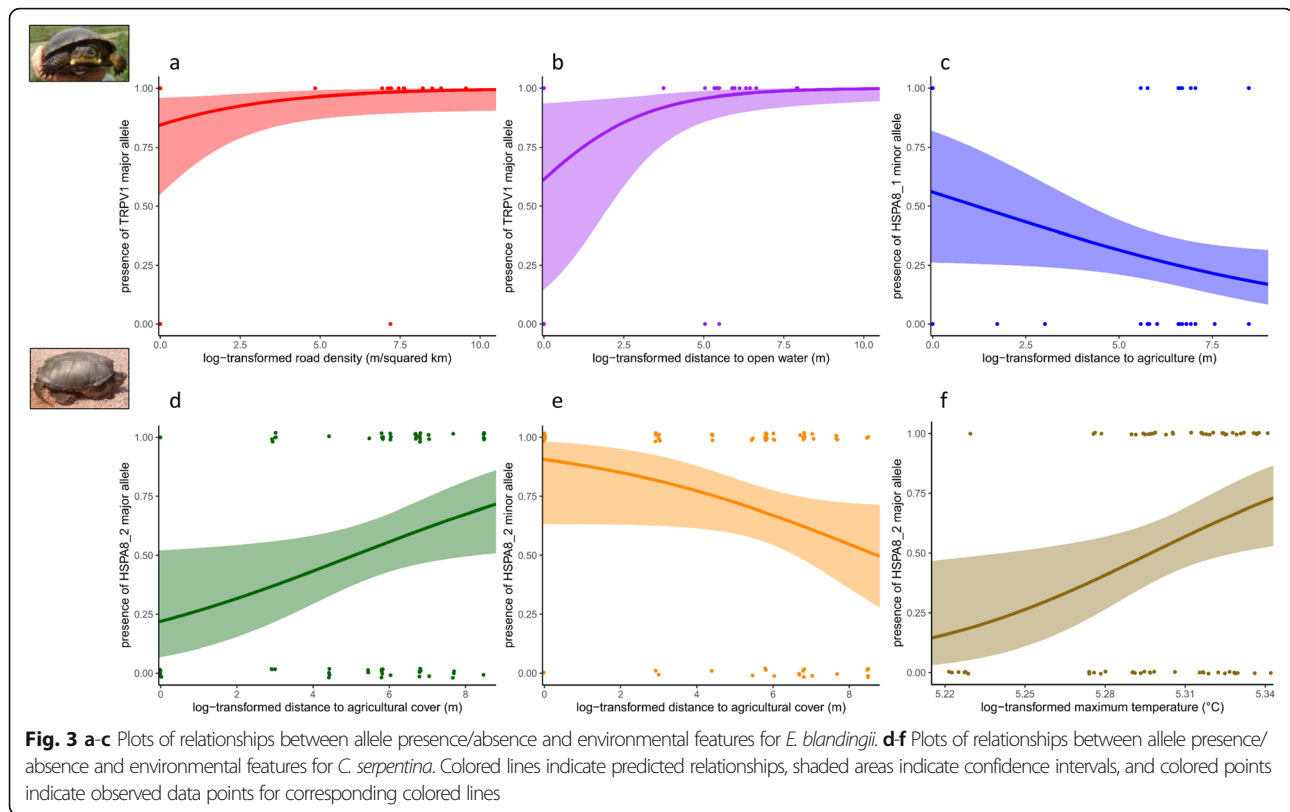
distance to wetlands ( $F = 1.1459$ ,  $p = 0.004$ ), precipitation ( $F = 1.1272$ ,  $p = 0.006$ ), and maximum temperature ( $F = 1.1920$ ,  $p = 0.001$ ), variance partitioning revealed that both climate and land cover covariates explain very little variance overall (about 1% of total variance each; Fig. 2). For the *C. serpentina* target dataset, we failed to detect any significant environmental associations, and variance partitioning indicated that less than 1% of total variance was explained by climate, land cover, and geography covariates. In addition, we failed to detect any significant associations between climate or land cover and genetic variation in the SNP dataset, although proportions of variance explained were approximately equal across climate, land cover, and geography (3–4%; Supplementary Materials S5).

#### Gene-specific environmental associations

For the *E. blandingii* target gene dataset univariate analyses, likelihood ratio tests identified three univariate logistic regression models of allele presence/absence that outperformed null expectations. First, a model of presence/absence of the TRPV1 major allele against road density was

significant ( $\chi^2 = 4.0555$ ,  $p = 0.044$ ), and indicated that presence of the major allele was associated with higher log-transformed road densities ( $\beta_{\text{roads}} = 0.3302$ , 95% CI = 0.009 to 0.7599, Fig. 3). Second, a model of presence/absence of the TRPV1 major allele against distance to open water was significant ( $\chi^2 = 4.8221$ ,  $p = 0.028$ ), and indicated that presence of the major allele was associated with greater distance to open water ( $\beta_{\text{open}} = 0.5266$ , 95% CI = 0.0627 to 1.0156, Fig. 3). Third, a model of presence/absence of the minor allele for the first HSPA8 variant against distance to agricultural cover was significant ( $\chi^2 = 4.0168$ ,  $p = 0.045$ ), and indicated that presence of the minor allele was associated with reduced distances to agricultural cover (Fig. 3). For gene-specific multivariate partial redundancy analyses conditioned on latitude and longitude, we identified only a significant association between genetic differentiation in HSPA8 and road density ( $F = 6.8965$ ,  $p = 0.011$ ; Supplementary Materials S6).

In contrast, for the *C. serpentina* target gene dataset, likelihood ratio tests identified three models of allele presence/absence that outperformed null expectations, all associated with the second documented HSPA8



variant site. First, models of presence/absence of the minor allele for the second HSPA8 variant site against maximum temperature ( $\chi^2 = 6.7824$ ,  $p = 0.009206$ ) and distance to agricultural cover ( $\chi^2 = 5.272$ ,  $p = 0.02167$ ) were both significant, and indicated positive associations with both maximum temperature ( $\beta_{\text{tmax}} = 21.746$ , 95 % CI = 5.106 to 41.554) and distance to agricultural cover ( $\beta_{\text{ag}} = 0.2494$ , 95 % CI = 0.0352 to 0.4962) and allele presence. Finally, a model of presence/absence of the major allele of the second HSPA8 variant site against distance to agricultural cover was significant, and indicated a negative association between presence of this allele and distance to agricultural cover ( $\beta_{\text{ag}} = -0.2593$ , 95 % CI = -0.5672 to -0.0107). For gene-specific multivariate partial redundancy analyses conditioned on latitude and longitude, we failed to detect any significant associations between climate or land cover variables and genetic differentiation.

## Discussion

In this study, we leveraged both genome-wide SNPs and targeted sequencing of genes potentially important for local adaptation to explore how life history affects patterns of local adaptation. By exploring potentially adaptive genetic variation in a widespread generalist with an earlier age of maturity (*C. serpentina*) and a more range-restricted specialist with substantially delayed maturity

(*E. blandingii*), we explored the genetic consequences of life history traits and habitat specialization [15, 16]. Moreover, we documented sample site-specific polymorphisms in most target genes examined. Several of these polymorphisms were nonsynonymous, and may have adaptive significance (Supplementary Materials S2). While many of these polymorphisms did not lead to changes in gene function, polymorphisms in non-coding regions may still be under positive selection and may still play a role in functional gene regulation [37]. By exploring how variation in these polymorphisms differed from variation in putatively neutral SNPs across a regional extent, we provided initial evidence that spatial heterogeneity in selective pressure associated with environmental conditions shaped among-population variation in these target genes.

## Spatial patterns in neutral and adaptive genetic variation

Neutral and adaptive genetic variation are expected to be shaped by different processes, which should produce contrasting representations of spatial clustering between putatively neutral and adaptive genetic markers [24, 38, 39]. As expected, the numbers and spatial arrangement of identified genetic clusters differed between putatively neutral SNPs and target genes (Fig. 1). In particular, although we failed to observe clear inflection points in cross-validation scores for both SNP datasets, clusters

identified for these datasets always corresponded to cohesive spatial groupings. For the *C. serpentina* SNP dataset, site AL represented a separate cluster from all other sample sites; in addition, other sites in the southern portion of the state (sites AV and MN, Fig. 1) had greater representation of this cluster as well. For the *E. blandingii* SNP dataset, site SH – one of the largest and best-studied Blanding's turtle populations in Wisconsin [40–43] – was assigned a unique cluster, highlighting the importance of protecting this site in the state. In contrast, the target gene panel was classified into a larger number of clusters that: (a) did not appear to correspond to clusters observed in the SNP datasets, and (b) overlapped spatially.

Although this appears to support our hypothesis – that spatial patterns in genetic differentiation in target gene and SNP datasets should reflect contrasting processes governing adaptive and neutral genetic variation, respectively – there are several caveats to these findings. First, although our intention for sequencing these particular gene regions was to target regions of potential adaptive importance, the limited number of genomic regions targeted by this approach relative to genotype-by-sequencing (four versus several thousand, respectively) may have produced differences in spatial genetic structure purely due to sampling effects, rather than effects of ecological and evolutionary processes. Second, any clustering in target gene datasets incorporated variant sites identified across all genes, as low numbers of variant sites identified overall made gene-specific clustering intractable. Given that we observed differences in environmental associations in both gene and allele-level analyses (see below), the spatial overlap in target gene clusters likely indicates that different processes are producing patterns in each gene region. Third, low sample sizes per site likely led to incomplete depictions of spatial patterns in allele frequencies. Although we attempted to use approaches throughout that do not have documented biases at low site-level sample sizes, clustering algorithms are affected by choices in sampling strategy [44]. Despite these caveats, these findings do suggest that highly divergent spatial genetic patterns may be observed between gene-specific and genome-wide datasets, and provide justification for further investigations into environmental associations.

#### Species-specific environmental associations

In addition to documenting divergent patterns of spatial clustering across datasets, we also documented significant environmental associations with adaptive genetic variation, suggesting that environmental heterogeneity may exert selective pressure on these genetic regions [9, 14]. Furthermore, our comparison of two species at opposite ends of a habitat specialist-generalist gradient allowed for

consideration of how habitat specialization affects local adaptation. Although habitat specialists are expected to respond adaptively to climate change [15, 16], habitat specialists often persist in smaller populations than generalists, and small populations are generally expected to have reduced capacity to respond adaptively due to inbreeding effects on fitness and lower expected rates of mutations per unit time [36, 45, 46]. Our habitat specialist species, *E. blandingii*, requires co-occurrence of undisturbed wetlands habitat for residency and upland sandy bluffs for nesting [26, 31, 47], and many populations throughout Wisconsin are expected to be declining [42, 48, 49]; furthermore, the capacity for local adaptation in this species has been uncertain until now.

We considered three criteria when evaluating the potential effects of habitat specialization on local adaptation: (a) the number of nonsynonymous substitutions in each species-specific dataset, (b) the relative explanatory power of environmental variables for target vs. SNP datasets for each species, and (c) the relative explanatory power of environmental variables for target gene datasets across species. Below, we discuss each of these criteria in turn. First, we documented similar numbers of nonsynonymous substitutions for each species (three in *C. serpentina* and two in *E. blandingii*); furthermore, these nonsynonymous substitutions were observed in different genes across both species, with all three in *C. serpentina* observed in TRPV1 and both in *E. blandingii* in CIRBP. In contrast, climate and land cover covariates explained more variance in the *E. blandingii* target gene dataset than the SNP dataset (5% vs. <1%, Fig. 2) and the *C. serpentina* target gene dataset (5% vs. 0%; Supplementary Materials S5).

While diversity in these gene regions may facilitate adaptive responses to climate change, the predominance of land cover associations in these target gene datasets may indicate that spatial patterns in diversity for these gene regions has largely been shaped by anthropogenic habitat alterations – which have likely reduced effective population sizes, connectivity, and genetic diversity for both of these species [48, 50, 51] (Fig. 3). Although we expected to document climate associations for both species, we failed to detect any significant climate associations in *C. serpentina* for any genes, and land cover covariates (particularly road density) explained more genetic variation in target genes than climate covariates for *E. blandingii* (Fig. 2). Areas of increased road density have previously been found to cause reduced genetic diversity and biased sex ratios in *E. blandingii*, likely due to increased road mortality risk in areas with more roads [48]. Assuming that these target gene regions represent a reasonable proxy for adaptive genetic variation, our results thus indicate that these same landscape features that reduce neutral genetic diversity and negatively

influence population demographics may also affect adaptive potential for this species. Given how many *E. blandingii* populations are located within a matrix of anthropogenic habitat alterations, efforts to protect habitats that are relatively isolated from human activities and reduce or mitigate the effects of roads near critical habitat areas may represent effective strategies for increasing population sizes and maximizing adaptive potential.

#### Gene-specific environmental associations

Environmental associations also varied among gene functional categories. First, since both species exhibit temperature-dependent sex determination and prior studies have suggested that plastic nesting behaviors may be insufficient to demographically buffer turtle populations against climate change [52, 53], we predicted that variation in two genes associated with sex determination (CIRBP and Sox9) would be related to environmental conditions. Although both polymorphisms in CIRBP for *E. blandingii* were nonsynonymous (Supplementary Materials S2), neither of these polymorphisms were associated with any environmental gradients, potentially indicating no adaptive divergence in CIRBP due to environmental factors. This is consistent with prior studies that indicate that TSD is a polygenic trait with relatively little heritable variability in wild populations [33, 54, 55]. Furthermore, within-population variation in several traits associated with TSD (sex ratio and transitional range of incubation temperatures, or TRTs) is often much higher than among-population variation [30]. When interpreted with relatively low levels of standing genetic variation documented in our study, this indicates that substantial physiological plasticity in TSD will be more important than standing genetic variation in gene regions associated with TSD for adequate climate responses.

Since such plasticity would depend upon adequate sensing and response to increasing temperatures and TRPV1 has functional associations with temperature-sensing [46], we expected variation in this gene to fall along similar environmental gradients to CIRBP and Sox9. As before, we failed to detect significant associations with climate for TRPV1 in either *E. blandingii* or *C. serpentina*. However, for *E. blandingii*, we did document significant associations between allele presence and land cover covariates (road density and distance to open water). As with the species-specific environmental association analyses, this may indicate that genetic variation in genes critical for adequately adapting to climate change is in some ways decoupled from contemporary climate conditions, and is instead shaped by the demographic and genetic impacts of anthropogenic habitat alterations [48].

Unlike sex-determining genes, we documented more polymorphisms and stronger environmental associations in our stress response gene (HSPA8) than expected. While several previous studies have suggested the importance of heat shock proteins and other genes associated with stress response for screens of local adaptation, little is known about the environmental factors that might select for particular polymorphisms in heat shock proteins [19, 56–59]. Since HSPA8 polymorphisms were associated with road density in *E. blandingii*, anthropogenic landscape alterations may have exerted selective pressures upon this gene [57, 60, 61]. Although physiological stress responses are often not considered in climate change vulnerability assessments, temperature increases through the 21st century will likely exceed the thermal tolerance limits of many species [56]. Given how many heat shock proteins exist further study is needed to determine which heat shock proteins may be implicated in local adaptation, and whether these genes may play a role in responses to environmental change [56, 57, 60, 61]. Although further investigation of the gene region targeted by our HSPA8 primer set indicated that these polymorphisms may be in a noncoding region (Supplementary Materials S2), numerous studies have suggested that adaptation may still be facilitated by variation in noncoding regions [62, 63]. Although it is often unclear if intronic polymorphisms are themselves functional or if they are simply within linkage disequilibrium with unidentified functional polymorphisms potentially in exonic regions, it appears that exploration of noncoding polymorphisms within gene regions may still provide insights into adaptation [64].

#### Conclusions

In summary, we leveraged one of the many strengths of landscape genomic approaches – the testing of evolutionarily-relevant hypotheses in field systems– to explore the impacts of habitat specialization and gene function on patterns of putatively adaptive genetic variation [23]. While our conclusions were limited by the lack of additional genomic resources for either species, our study points to several gene regions – in particular HSPA8 and TRPV1 – that may represent promising targets for further, targeted investigations of spatial patterns of local adaptation in these and other non-model organisms [11–13, 23, 65]. As annotated genomes in non-model organisms become increasingly available, the genomic context of functional genes can then be considered in more detail, allowing studies of the genomics of local adaptation to move beyond initial exploratory analyses and towards causal, evolutionarily-grounded investigations of the likelihood of climate change adaptation [23].

Given the scope of conservation challenges posed by climate change, landscape genomic studies will increasingly



be leveraged to provide insights into effective conservation and wildlife management [66–68]. In particular, such may facilitate more effective spatial area prioritization and climate change vulnerability assessments for sensitive taxa. In our study, observed differences in the distribution of adaptive and neutral genetic variation mean that spatial prioritization based solely on neutral markers is unlikely to capture the full spectrum of adaptive genetic variation in either species [69, 70]. As with many other landscape genomic studies, however, we were not able to connect genetic variation in these functional genes to phenotypic differences or fitness consequences [38, 71]. As a result, although we were able to identify populations with unique adaptive potential, we do not know which variants in any of these genes confer increased resiliency to environmental change. Although the actual functional significance of spatial variation in our chosen target genes is currently unknown for these two species, prioritizing sites that cover the full range of potentially adaptive genetic variation may maintain the potential for evolutionary adaptation [69, 72]. In the absence of information on how fitness is associated with spatial genetic variation in any of these genes, ecological studies linking phenotypic and genotypic variation in these functional traits to fitness outcomes may provide additional evidence for local adaptation [13, 65, 73].

Studies of local adaptation provide additional advantages beyond spatial prioritization as well, particularly if significant environmental associations with adaptive genetic variation are detected, as these environmental associations may provide evidence of the selective forces underpinning the distribution of adaptive genetic variation [9, 13]. For our specialist species (*E. blandingii*), potentially adaptive genetic variation was more strongly associated with land cover than climate, with particularly strong effects of road density. Given previous findings that roads have demographic and genetic impacts on turtle populations, this may indicate an additional threat posed to turtles by roads, whereby roads reduce the potential for local adaptation [48, 50, 51]. However, we caution that adaptation in each of these species may take many generations; thus, associations between potentially adaptive genetic variability and contemporary land cover, climate, and geographical characteristics may not represent conditions that shaped historical adaptive responses [9, 14]. Nonetheless, associations between contemporary patterns of potentially adaptive genetic variation and environmental heterogeneity provide insights into whether adaptive genetic variation is currently distributed along environmental gradients expected to drive selection in the future [9, 11–13]. Information on these potentially important environmental gradients – when combined with effective spatial prioritization –

will likely prove to be critical for predicting the role of local adaptation in species persistence in a changing world.

## Methods

### Study species and sample selection

Blanding's turtles and snapping turtles were sampled across more than 20 sites in Wisconsin between 2010 and 2015 [41, 48, 74] (Fig. 1). Turtles were trapped using a variety of approaches, including hoop traps, crayfish traps, minnow traps, and hand captures. For each captured turtle, we obtained blood samples from the dorsal coccygeal vein, and animals were released following individual marking. We also opportunistically collected tissue samples from museum specimens, road-killed animals, and captive-reared juveniles from egg clutches with known localities [41]. Genomic DNA was extracted from samples using a Qiagen DNeasy Blood and Tissue Kit, as described elsewhere [41, 48]. We subsampled sites and samples from this larger sample set for two separate sequencing efforts: (a) SNP discovery and genotyping and (b) investigation of variation in four functional genes. We selected samples to maximize: (a) variation in climatic conditions across sample sites, as indicated by Worldclim climate data layers [75], and (b) individuals sampled per site. We then randomly selected 3 or more individuals for each species from each site across these sample sites; if a selected site has fewer than 3 individuals for a given species, we used all samples instead (see Table 2 for sample sizes for each species and dataset). For SNP discovery and genotyping, we also only used samples with DNA concentrations suitable for sequencing (>20 ng/uL). While we initially intended to use the same samples for both functional gene and SNP analyses, stringent concentration requirements for the SNP discovery component of the project and low sample volumes meant that some samples were used solely for SNP discovery. Final sample sets were: 50 *C. serpentina* across 16 sites for SNP genotyping, 86 *C. serpentina* across 21 sites for target gene sequencing, 60 *E. blandingii* across 13 sites for SNP genotyping, and 95 *E. blandingii* across 22 sites for target gene sequencing.

### SNP discovery and genotyping

We used genotype-by-sequencing (GBS) to discover and genotype SNPs for each species (Elshire et al. 2011). Samples were submitted for library preparation and sequencing to the University of Wisconsin-Madison Biotechnology Center. Each sample was digested using *apeKI*, indexed and then sequenced on an Illumina HiSeq2500. Two libraries of the same 87 samples (51 *E. blandingii*, 36 *C. serpentina*) were single-end 100 bp sequenced in their own lanes; given poor overlap in loci

**Table 2** Final sample sizes for each species, dataset, and site

Population	# EB samples		# CS samples	
	Target	SNP	Target	SNP
AL	4	0	0	4
AV	0	0	3	3
BR	0	0	4	0
CH	0	0	0	0
CL	3	3	0	0
CM	3	3	0	0
EC	3	0	4	0
FC	3	0	4	0
FT	7	5	0	0
GB	3	0	0	0
KM	3	0	6	0
LR	3	0	0	0
MC	3	0	7	10
MN	4	0	3	3
MR	0	0	6	0
MU	6	5	4	0
MZ	4	4	0	0
NA	5	3	0	0
SH	20	8	7	3
TB	0	0	6	0
TC	3	0	4	0
WA	0	0	4	0
WR	0	0	5	0

among individuals for these libraries, an additional library of 87 samples (containing 37/51 previously sequenced *E. blandingii* and 27/36 previously sequenced *C. serpentina* samples, with an additional 9 *E. blandingii* and 14 *C. serpentina* samples unique to this library) was also pair-end 100 bp sequenced on a RAPID run in two lanes. Since genomes were not yet available for either species at time of analysis, we then used *de novo* assembly in iPyrad [73] to call SNPs and genotype individuals for these SNPs. Before assembly, we merged raw sequencing output per individual across runs. We outline the iPyrad assembly steps in brief below, which were applied to each species separately. Step 1 involved demultiplexing input fastq files; for this step, we set max\_barcode\_mismatch to 0 but maintained most other default settings. Step 2 involved filtering and editing the demultiplexed data; for this step, we used the following settings: max\_low\_qual\_bases = 5, filter\_adapters = 2, filter\_min\_trim\_len = 35, trim\_reads = 0, and Phred\_Qscore\_offset = 33. Step 3 involved clustering and mapping demultiplexed and filtered reads; we set a clust\_threshold of 0.85 for this step. Step 4 involved joint estimation of heterozygosity and sequencing error rates; we set max\_

alleles\_consens to 2 for this step. Given that choices of parameter values for minimum cluster depth (mind\_eplth\_statistical and mindepth\_majrule) has downstream influences on the number and quality of final SNPs called and genotyped, we tested several parameter sets that maximized both the number of individuals and loci retained and the quality of the final SNP datasets. During this parameter optimization we discovered that most clusters (>70 %) were at coverage levels of 2–4; thus, we ultimately set mindepth\_majrule to 5 and mindepth\_stat to 10. We also set maxdepth to 10,000 for this step. Step 5 involved consensus base calling and filtering; we set max\_Ns\_consens to 5 for this step. Step 6 involved clustering and mapping reads across samples and alignment of clustered sequences; we used default settings for this step. Finally, step 7 involved final filtering and creation of output files. For this step, we tried several parameter values for min\_samples\_locus initially ranging from 5 to 60, but discovered high levels of individual-level missing data as a result of this optimization procedure; thus, we opted to use a low value for min\_samples\_locus of 5 to facilitate downstream filtering [76]. Other parameter values were as follows: max\_Indels\_locus = 8, max\_shared\_Hs\_locus = 8, and trim\_loci = 0.

Final SNP datasets for each species were exported as vcf files for further filtering using package “dartR” [77]. First, we filtered out monomorphs using “gl.filter.monomorphs”. We then excluded any SNPs with MAF less than 0.05. Next, in keeping with [76], we iteratively and alternately filtered call rates at individual and locus-level using “gl.filter.callrate”, starting with a threshold tolerance for missing data of 0.5 for loci and 0.1 for individuals and ending with a threshold of 0.7 for loci and 0.5 for individuals; this progression was chosen to mimic FS 5 and 6 from [76]. We then used “gl.filter.secondaries” to select one SNP per fragment with the least missing data. We then used “gl.report.hwe” to identify SNPs that deviated from Hardy-Weinberg Equilibrium across all sites, and removed these SNPs using “gl.drop.loc”. Finally, we filtered for heterozygote excess by first calculated observed heterozygosity ( $H_O$ ) for each locus using “gl.basic.stats”, and then used “gl.drop.loc” to drop any loci with  $H_O > 0.5$ . The *E. blandingii* dataset comprised 557,829 SNPs across 64 individuals and 12 sites, represented at a mean coverage of 28.58x, each of which was present in at least four individuals. After filtering for MAF, missing data by individual and locus, Hardy-Weinberg Equilibrium, heterozygote excess, and for a single SNP per RAD locus, a total of 2,811 SNPs across 37 individuals and 11 sites were available for *E. blandingii*, with a total missing data percentage of 18.9%. Excluding sites with fewer than three samples produced a final dataset of 2,810 SNPs across 31 individuals in seven sites. The *C. serpentina* dataset comprised 795,296 SNPs

across 52 individuals and 15 sites, represented at mean coverage of 25.25X, each of which was present in at least four individuals. After filtering for MAF, missing data by individual and locus, Hardy-Weinberg Equilibrium, heterozygote excess, and for a single SNP per RAD locus, a total of 1,606 SNPs across 33 individuals and 13 sites were available for *C. serpentina*, with a total missing data percentage of 17.23%. Excluding sites with fewer than three samples produced a final dataset of 1,600 SNPs across 23 individuals in five sites. Hereafter, we refer to these final genome-wide datasets as our SNP datasets. Given low sample sizes and limited spatial representation of sample sites, we primarily used these datasets as a point of comparison with the targeted sequencing efforts described below.

#### **Amplification of targeted functional genes**

We searched NCBI GenBank [78] for literature-supported functional genes that: (1) play a role in thermoregulatory behaviors, such as genes associated with heat- or cold-sensing; (2) code for heat shock proteins, as these have been shown to correlate with stress related to climatic extremes [60, 61]; and (3) play some role in temperature-dependent sex determination. We then selected genes that have a high degree of literature support for putative functional roles. Given (a) the lack of functional genomic information for either species and (b) the need to develop and optimize novel primers for each selected gene, we limited our inquiry to four genes total: CIRBP, a gene that plays a role with temperature-dependent sex determination [79]; Sox9, a frequently studied gene region implicated in sex determination [80, 81]; HSPA8, a heat-shock protein associated with stress response [82]; and one Transient Receptor Potential Cation Channel (TRPV1) which has been implicated in temperature-sensing [83]. We downloaded available coding sequences for turtles for each gene from GenBank (Supplementary Materials S1), and imported these sequences into MEGA7 [84]. We then used the alignment algorithm MUSCLE [85] to align sequences across taxa using default settings. We then exported these alignments to Primacle to identify potential primer sets to use for each gene, and selected primers that maximized the number of variant sites within each amplicon of a target size of 250 bp [86]; given the relative scarcity of functional genomic information for these species and rather than subdividing effort between multiple poorly known genomic regions per gene, we opted to investigate only the 250 bp region within each gene with the most variant sites. Furthermore, although we attempted to target only coding regions for selected genes, the lack of available genomic information made delineation of intron and exon boundaries difficult *a priori*; as a result, some primers inadvertently targeted non-coding

sequences (Supplementary Materials S7). However, these non-coding regions may still be under positive selection or play a role in functional gene regulation, and likely still reflect patterns in adjacent coding regions to some degree [37].

We tagged the primers for each of these five primer sets (CIRBP, HSPA8, Sox9, and TRPV1) with Illumina TruSeq Universal adapters for downstream Illumina MiSeq sequencing. We then amplified each tagged primer set for each sample using Qiagen Taq PCR core kits and Phusion High-Fidelity DNA Polymerase kits. See supplementary materials S1 for PCR reactions and parameters. We cleaned amplified product using Ampure XP beads, and quantified DNA concentrations for cleaned product using a Qubit fluorometer. Amplified product was then combined in equimolar proportions for each sample, and custom i5 and i7 indices with stem adapters attached were attached using KAPA HiFi Hot-Start ReadyMix. The indexed samples were then cleaned with Ampure XP beads to remove unincorporated adapters, quantified and combined into two libraries of 96 samples. Libraries were then sequenced on two Illumina MiSeq 2 × 250 nano runs by the University of Wisconsin-Madison Biotechnology Center. All unique gene sequences will be available online (see supplementary materials S2).

#### **Functional gene amplicon processing**

We used QIIME2 [87], a program designed for filtering and processing custom amplicon sequencing output, to filter and separate pooled reads by individual and gene region. First, we used “dada2” to trim and filter demultiplexed sequences and identify unique genetic variants present in each sample [87]. We then trained a naïve Bayes classifier on sequences downloaded from GenBank for each gene to classify genetic variants identified previously into the five amplified genes [87]. Whenever possible, we attempted to use sequences from our two study species; if gene-specific sequences for our species were unavailable, we used other available sequences from turtles instead. We re-ran this procedure with multiple filtering, trimming, and classification confidence settings to identify optimal settings for each gene that: (1) maximize the number of identified genetic variants for each gene, and (2) minimize spurious or incorrectly-classified sequences for each gene. Although this optimization procedure reduced the number of spurious and incorrectly classified sequences, some chimeras, spurious classified sequences, and potential incidences of barcode-hopping were still observed. Thus, we visually inspected each representative sequence and used NCBI BLAST to filter out genetic variants with <60% query coverage or <80% identity for the targeted genes. We then excluded individual-level variants with less than certain sequence count thresholds; for

CIRBP, Sox9, and TRPV1, we excluded variants for individuals with less than five reads (due to lower representation of these genes in overall libraries), but used a higher threshold of 10 for HSPA8.

The first Illumina MiSeq nano run generated a total of 726,112 sequences of 250 bp each, whereas the second generated a total of 621,932 sequences of 250 bp each. For most genes, aligning forward and reverse reads resulted in relatively high data loss post-filtering; thus, only forward reads were retained for downstream analyses, most of which were either truncated to 200 bp or left unaltered (depending on the gene and species; see supplementary materials S7). Furthermore, HSPA8 was typically over-represented relative to other gene products; however, sufficient reads were retained for each gene for downstream analysis. After excluding samples that failed to sequence reliably, this left 88 individuals across 22 for *E. blandingii* and 79 individuals across 22 sites for *C. serpentina*. After further removing sites with fewer than 3 individuals, this left 77 individuals across 16 sites for *E. blandingii* and 67 individuals across 14 sites for *C. serpentina*.

Genetic variants within each target gene were summarized for each sample and exported to MEGA7 for alignment. The MUSCLE algorithm (using default settings) was used for multiple sequence alignment of all genetic variants across samples, and variant sites present in greater than one individual were identified using MEGA7 and exported to comma-separated value (csv) format. Genotypes were numerically coded based on minor allele frequencies of each polymorphic site for each individual, and converted to a “genind” object using package “adegenet” in R [88]. For all analyses described below, we refer to target gene amplicons as our amplicon dataset, used primarily to represent putatively adaptive genetic variation. In contrast, our final SNP datasets are primarily used to represent neutral genetic variation given low numbers of SNPs per species. Given low sample sizes per site, we have largely chosen to use approaches that lack strict requirements for site-level sample sizes.

### Spatial genetic divergence

In order to address our first hypothesis – that there are differences in spatial genetic structure between target gene and SNP datasets – we quantified population genetic summary statistics and inferred numbers of genetic clusters. First, we quantified expected heterozygosity ( $H_e$ ) and observed heterozygosity ( $H_o$ ) for the SNP and target gene datasets for each sample site and species. We also quantified pairwise hierarchical  $F$ -statistics between sample sites for each species and dataset using packages “hierfstat”, “pegas”, and “adegenet” based on the Weir and Cockerham (1984) estimator [89–91]; despite low

per-site sample sizes, we believe this choice of estimator is appropriate given recent findings that this estimator is relatively robust at small sample sizes, particularly when using more than 1,000 SNP markers [92]. While low numbers of markers for the target gene datasets have likely reduced power for finding statistically-significant population structure, we have retained such analyses to provide a point of comparison with the SNP  $F_{ST}$  calculations. In all cases, we assessed significance using permutation tests with 100 bootstrapped replicates. We tested for isolation by distance for each dataset and species using Mantel tests by relating a matrix of individual-level genetic distances to a matrix of Euclidian geographic distances between samples [93]. While we initially intended to use the SNP datasets to calculate effective population size for each sample site and species, low numbers of genotyped SNPs made such analyses intractable.

In order to identify whether genetic variation was geographically structured among sample sites, we used several approaches to identify genetic clusters in each dataset. First, we used successive  $K$ -means clustering and discriminant analysis of principal components (DAPC) in package “adegenet” in R to capture spatial clustering in the SNP and amplicon datasets. In order to do so, we used the “find.clusters” function to detect the optimal number of clusters for each species and dataset based on Bayesian Information Criterion (BIC), which generally coincided with a sharp elbow in the curve of BIC values as a function of  $K$  [37, 86]. We then used DAPC to depict the clusters in multivariate space for each species and dataset. Second, we identified clusters using “tess3r”, a spatially-explicit clustering approach [94]. This approach estimates ancestry coefficients while accounting for information on spatial sample coordinates to control for population structure. As above, we visualized cross-validation scores for each value of  $K$ , and inferred the optimal  $K$  based on when cross-validation scores plateau or start increasing [94]. For each algorithm, we tested all values from  $K=1$  to 12.

### Environmental associations

We sampled environmental covariates from a variety of sources, with a similar methodology to that described in [42]. Rather than representing individual sampling locations within each site when sampling environmental conditions, we used a single site-level coordinate for each individual; thus, each individual sampled at a particular sample site is assumed to experience the same environmental conditions. We then extracted climate, land cover, and environmental covariates from a variety of data sources for each sample site. Given that each chosen gene region had functional associations with climate conditions [79, 80, 82, 83], we used WorldClim 1.0



[75] to calculate three climate and weather variables throughout the active season (April to November): mean monthly maximum temperature, mean monthly minimum temperature, and mean monthly precipitation. We calculated distances to several core land cover types critical for population persistence for these species (open water and wetland cover; [31]) or known to reduce gene flow between turtle populations (urban and agricultural cover; [48]) using the WISCLAND dataset (<https://dnr.wi.gov/maps/WISCLAND.html>) by first converting raster representations of land cover to polygons, and then calculating Euclidian distance from sample sites to each land cover type. Finally, since road density has documented impacts on genetic diversity for *E. blandingii* [48] we extracted road densities (meters/km<sup>2</sup> pixel) from the TIGER dataset (<https://www.census.gov/geo/maps-data/data/tiger.html>). We grouped these covariates into three categories: climate/weather (mean maximum and minimum monthly temperatures, mean monthly precipitation), land cover (distance to urban cover, agricultural cover, open water, and wetlands, and road density), and geographic (latitude and longitude).

In order to address our hypothesis that species-level traits will lead to divergent environmental trends in adaptive differentiation, we used similar genotype-environment association tests as described above. First, for each overall dataset (SNP and target gene) and species (*C. serpentina* and *E. blandingii*), we used the variance partitioning function “varpart” in package “vegan” to assess the variance in allele counts explained by climate, land cover, and geographic variables for each genetic variant in each gene. Next, we ran similar pRDA models to those described above with forward and backward stepwise variable selection to select climate and land cover covariates associated with overall genetic variation in each dataset and species, and used permutation tests to assess variable significance. We used output from these tests to assess relative support for alternative interpretations of our hypothesis. If cumulative variance explained by land cover and climate covariates for the *E. blandingii* target dataset (a) has more nonsynonymous substitutions than the *C. serpentina* dataset, (b) exceeds the proportion for the putatively neutral *E. blandingii* SNP dataset and (c) exceeds the proportion explained for the *C. serpentina* target dataset, our prediction that habitat specialists will show stronger adaptive responses to environmental gradients would be considered supported. Alternatively, if the opposite of all three of these conditions is true, then we will interpret that as evidence that the sensitive conservation status and lower effective population sizes of the Blanding’s turtle may lead to a lower potential of adaptive responses to environmental gradients than the widespread *C. serpentina*. Between these two alternatives, mixed evidence in support of

either or both predictions may be possible as well, and may illustrate the relative roles of drift and selection in driving adaptive divergence for these species.

In order to test our hypothesis that environmental associations in targeted genes should reflect the putative functions fulfilled by each gene, we used a combination of variance partitioning approaches and multivariate environmental association tests. First, we classified each polymorphism as noncoding or coding based on alignment of amplicon sequences to previously-described GenBank records; if a polymorphism was determined to be in a coding sequence, we then classified polymorphisms as synonymous or non-synonymous using MEGA7. Next, we used a similar variance partitioning approach to the one described above to assess overall gene-specific differences in the proportion of the genetic variation explained by climate, land cover, and geography. Next, we used both univariate and multivariate tests of environmental associations to assess relationships between variants and environmental gradients potentially driving selection. For univariate tests, we used generalized linear models with allele presence/absence as the response and environmental variables as predictor variables. We limited inquiry to one-predictor models to facilitate downstream analysis, and evaluated significance using likelihood ratio tests. Finally, we used partial redundancy analysis (pRDAs) in the “vegan” library to conduct multivariate tests of genotype-environment associations [14, 95]. Before doing so, we used the function “findCorrelation” in package “caret” to identify and drop environmental covariates in each target dataset that were highly correlated ( $r > 0.6$ ). We then used a forward and backward stepwise variable selection procedure in package “vegan” to select covariates associated with genetic variation in target genes while conditioning for geographic coordinates (latitude + longitude), and used permutation tests to assess variable significance [95].

### Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12864-021-08151-7>.

#### Additional file 1.

### Acknowledgements

We would like to thank the Wisconsin Department of Natural Resources for providing critical support, housing, and approval for this project. We also thank J. Millonzi, J. Summers, B. Witek, J. Wogsland, P. Riddle, A. Swichtenberg, S. Tomke, and R. Klausch for assisting with field research for this project.

### Authors’ contributions

NWB, BNR, EDF, and MZP designed the study and collaborated on initial planning; NWB and BNR conducted field work; NWB, BNR, EDF, KM, and PJK performed laboratory work; NWB and EDF led computational analyses; NWB wrote the manuscript; and all authors revised and approved the final submission.

### Funding

Funding for field and laboratory work was provided by a Hatch Act formula grant (to MZP; WIS01865), an Environmental Protection Agency (EPA)-STAR fellowship (to BNR), an Experiment.com crowdfunding effort (to NWB and BNR), and an American Turtle Observatory Grant (to NWB, BNR, EDF, and MZP).

### Availability of data and materials

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

### Declarations

#### Ethics approval and consent to participate

All methods were carried out in accordance with relevant guidelines and regulations. Furthermore, this study was carried out in compliance with Animal Research: Reporting of In Vivo Experiments (ARRIVE) guidelines. This work was conducted under Endangered and Threatened Resources (#681) permit, Scientific Collector's (#135) permit, and Scientific Research License (SRL-SOD-001-2016) issued by the Wisconsin DNR and approved by UW-Madison's Animal Care and Use Committee (#A3368-01 and #A5358).

#### Consent for publication

Not applicable.

#### Competing interests

The authors declare that they have no competing interests.

#### Author details

<sup>1</sup>University of Nevada–Reno, NV 89557 Reno, USA. <sup>2</sup>University of Wisconsin–Madison, 53706 Madison, WI, USA. <sup>3</sup>W.K. Kellogg Biological Station, Michigan State University, MI 49060 Hickory Corners, USA.

Received: 10 December 2020 Accepted: 1 November 2021

Published online: 18 November 2021

### References

- Walther G-R, Post E, Convey P, Menzel A, Parmesan C, Beebee TJC, Fromentin J-M, Hoegh-Guldberg O, Bairlein F. Ecological responses to recent climate change. *Nature*. 2002;416:389–95.
- Thomas CD, Cameron A, Green RE, Bakkenes M, Beaumont LJ, Collingham YC, Erasmus BFN, de Siqueira MF, Grainger A, Hannah L, et al. Extinction risk from climate change. *Nature*. 2004;427:145–8.
- Chen IC, Hill JK, Ohlemüller R, Roy DB, Thomas CD. Rapid range shifts of species associated with high levels of climate warming. *Science*. 2011;333:1024–6.
- Marlon JR, Bartlein PJ, Walsh MK, Harrison SP, Brown KJ, Edwards ME, Higuera PE, Power MJ, Anderson RS, Briles C, et al. Wildfire responses to abrupt climate change in North America. *Proc Natl Acad Sci USA*. 2009;106:2519–24.
- Beever Ea, Ray C, Mote PW, Wilkening JL. Testing alternative models of climate-mediated extirpation. *Ecol Appl*. 2010;20:164–78.
- Lawler JJ, Ackerly DD, Albano CM, Anderson MG, Dobrowski SZ, Gill JL, Heller NE, Pressey RL, Sanderson EW, Weiss SB. The theory behind, and the challenges of, conserving nature's stage in a time of rapid change. *Conserv Biol*. 2015;29(3):618–29.
- González EJ, Rees M, Martorell C. Identifying the demographic processes relevant for species conservation in human-impacted areas: does the model matter? *Oecologia*. 2013;171:347–56.
- Bell G. Evolutionary rescue. *Annu Rev Ecol Syst*. 2017;48:605–27.
- Schoville SD, Bonin A, François O, Lobreaux S, Melodelima C, Manel S. Adaptive genetic variation on the landscape: methods and cases. *Annu Rev Ecol Syst*. 2011;43:120830113150004.
- Hoffmann AA, Sgro CM. Climate change and evolutionary adaptation. *Nature*. 2011;470:479–85.
- Manel S, Joost S, Epperson BK, Holderegger R, Storfer A, Rosenberg MS, Scribner KT, Bonin A, Fortin MJ. Perspectives on the use of landscape genetics to detect genetic adaptive variation in the field. *Mol Ecol*. 2010;19(17):3760–72.
- Manel S, Poncet BN, Legendre P, Gugerli F, Holderegger R. Common factors drive adaptive genetic variation at different spatial scales in *Arabis alpina*. *Mol Ecol*. 2010;19(17):3824–35.
- Sork VL, Aitken SN, Dyer RJ, Eckert AJ, Legendre P, Neale DB. Putting the landscape into the genomics of trees: approaches for understanding local adaptation and population responses to changing climate. *Tree Genet Genomes*. 2013;9(4):901–11.
- Sork VL, Squire K, Gugger PF, Steele SE, Levy ED, Eckert AJ. Landscape genomic analysis of candidate genes for climate adaptation in a California endemic oak, *Quercus lobata*. *Am J Bot*. 2016;103(1):33–46.
- Van Tienderen PH. Generalists, specialists, and the evolution of phenotypic plasticity in sympatric populations of distinct species. *Evolution*. 1997;51(5):1372–80.
- Kassen R. The experimental evolution of specialists, generalists, and the maintenance of diversity. *J Evol Biol*. 2002;15(2):173–90.
- Cao S, Ye M, Jiang S. Involvement of GIGANTEA gene in the regulation of the cold stress response in *Arabidopsis*. *Plant Cell Rep*. 2005;24(11):683–90.
- Fitzpatrick MC, Keller SR. Ecological genomics meets community-level modelling of biodiversity: mapping the genomic landscape of current and future environmental adaptation. *Ecol Lett*. 2015;18(1):1–16.
- Tangwancharoen S, Moy GW, Burton RS. Multiple modes of adaptation: regulatory and structural evolution in a small heat shock protein gene. *Mol Biol Evol*. 2018;35(9):2110–9.
- Andrews RM. Changing perspectives on reptile eggs: one biologist's journey from demography to development. *J Herpetol*. 2018;52:243–51.
- Storfer A, Antolin MF, Manel S, Epperson BK, Scribner KT. Genomic approaches in landscape genetics. In: Balkenhol N, Cushman SA, Storfer AT, Waits LP, editors. *Landscape genetics: concepts, methods, applications*. Chichester: John Wiley & Sons; 2016. p. 149–64.
- Razgour O, Taggart JB, Manel S, Juste J, Ibanez C, Rebelo H, Alberdi A, Jones G, Park K. An integrated framework to identify wildlife populations under threat from climate change. *Mol Ecol Resour*. 2018;18(1):18–31.
- Hoban S, Kelley JL, Lotterhos KE, Antolin MF, Bradburd G, Lowry DB, Poss ML, Reed LK, Storfer A, Whitlock MC. Finding the genomic basis of local adaptation: pitfalls, practical solutions, and future directions. *Am Nat*. 2016;188(4):000–000.
- Hoffmann AA, Willi Y. Detecting genetic responses to environmental change. *Nat Rev Genet*. 2008;9(6):421–32.
- Forester BR, Jones MR, Joost S, Landguth EL, Lasky JR. Detecting spatial genetic signatures of local adaptation in heterogeneous landscapes. *Mol Ecol*. 2016;25(1):104–20.
- Congdon JD, Dunham AE, Sels RCVL. Delayed sexual maturity and demographics of blanding's turtles (*Emydoidea blandingii*): implications for conservation and management of long-lived organisms. *Conserv Biol*. 1993;7:826–33.
- Congdon JD, Dunham AE, Sels RCVL. Demographics of common snapping turtles (*Chelydra serpentina*): implications for conservation and management of long-lived organisms. *Am Zool*. 1994;34:397–408.
- Gibbons JW, Scott DE, Ryan TJ, Buhmann Ka, Tuberville TD, Metts BS, Greene JL, Mills T, Leiden Y, Poppy S, et al. The global decline of reptiles, Déjà Vu amphibians. *BioScience*. 2000;50:653–66.
- Ewert M, Nelson C. Sex determination in turtles: diverse patterns and some possible adaptive values. *Copeia*. 1991;1991:50–69.
- Carter AL, Bodensteiner BL, Iverson JB, Milne-Zelman CL, Mitchell TS, Refsnider JM, Warner DA, Janzen FJ. Breadth of the thermal response captures individual and geographic variation in temperature-dependent sex determination. *Funct Ecol*. 2019;33(10):1928–39.
- Ernst CH, Lovich JE. *Turtles of the United States and Canada*. Baltimore: JHU Press; 2009.
- Van Dijk PP, Rhodin AGJ. *Emydoidea blandingii*. The IUCN Red List of Threatened Species 2011: e.T7709A155088836; 2011.
- Janzen FJ. Heritable variation for sex ratio under environmental sex determination in the common snapping turtle (*Chelydra serpentina*). *Genetics*. 1992;131(1):155–61.
- Ewert MA, Lang JW, Nelson CE. Geographic variation in the pattern of temperature-dependent sex determination in the American snapping turtle (*Chelydra serpentina*). *J Zool*. 2005;265(1):81–95.
- Packer L, Zayed A, Grixti JC, Ruz L, Owen RE, Vivallo F, Toro H. Conservation genetics of potentially endangered mutualisms: reduced levels of genetic variation in specialist versus generalist bees. *Conserv Biol*. 2005;19(1):195–202.

36. Jensen JD, Bachtrog D. Characterizing the influence of effective population size on the rate of adaptation: Gillespie's Darwin domain. *Genome Biol Evol.* 2011;3:687–701.
37. Kornienko AE, Guenzl PM, Barlow DP, Pauler FM. Gene regulation by the act of long non-coding RNA transcription. *BMC Biol.* 2013;11(1):59.
38. Holderegger R, Kamm U, Gugerli F. Adaptive vs. neutral genetic diversity: implications for landscape genetics. *Landscape Ecol.* 2006;21(6):797–807.
39. Meyer-Lucht Y, Mulder KP, James MC, McMahon BJ, Buckley K, Piertney SB, Höglund J. Adaptive and neutral genetic differentiation among Scottish and endangered Irish red grouse (*Lagopus lagopus scotica*). *Conserv Genet.* 2016;17(3):615–30.
40. Reid BN, Thiel RP, Peery MZ. Population dynamics of endangered blanding's turtles in a restored area. *J Wildl Manage.* 2016;80(3):553–62.
41. Reid BN, Mladenoff DJ, Peery MZ. Genetic effects of landscape, habitat preference and demography on three co-occurring turtle species. *Mol Ecol.* 2017;26(3):781–98.
42. Byer NW, Reid BN, Peery M. Genetically-informed population models improve climate change vulnerability assessments. *Landscape Ecol.* 2020;35:1215–28.
43. Byer NW, Reid BN, Thiel RP, Peery MZ. Strong climate associations but no temporal trends in nesting phenology of blanding's turtles (*Emydoidea blandingii*). *Herpetologica.* 2020;76:396.
44. Schwartz MK, McKelvey KS. Why sampling scheme matters: the effect of sampling scheme on landscape genetic results. *Conserv Genet.* 2009;10(2):441.
45. Willi Y, Van Buskirk J, Hoffmann AA. Limits to the adaptive potential of small populations. *Annu Rev Ecol Syst.* 2006;37:433–58.
46. Raghwani J, Bhatt S, Pybus OG. Faster adaptation in smaller populations: counterintuitive evolution of HIV during childhood infection. *PLoS Comput Biol.* 2016;12(1):e1004694.
47. Beaudry F, deMaynadier PG, Hunter ML. Seasonally dynamic habitat use by spotted (*Clemmys guttata*) and blanding's turtles (*Emydoidea blandingii*) in Maine. *J Herpetol.* 2009;43:636–45.
48. Reid BN, Peery MZ. Land use patterns skew sex ratios, decrease genetic diversity and trump the effects of recent climate change in an endangered turtle. *Divers Distrib.* 2014;20:1425–37.
49. Hamilton CM, Bateman BL, Gorzo JM, Reid B, Thogmartin WE, Peery MZ, Heglund PJ, Radeloff VC, Pidgeon AM. Slow and steady wins the race? Future climate and land use change leaves the imperiled Blanding's turtle (*Emydoidea blandingii*) behind. *Biol Conserv.* 2018;222:75–85.
50. Steen Da, Aresco MJ, Beilke SG, Compton BW, Condon EP, Kenneth Dodd C, Forrester H, Gibbons JW, Greene JL, Johnson G, et al. Relative vulnerability of female turtles to road mortality. *Anim Conserv.* 2006;9:269–73.
51. Beaudry F, deMaynadier PG, Hunter ML. Identifying road mortality threat at multiple spatial scales for semi-aquatic turtles. *Biol Conserv.* 2008;141:2550–63.
52. Refsnider JM, Janzen FJ. Temperature-dependent sex determination under rapid anthropogenic environmental change: evolution at a turtle's pace? *J Hered.* 2015;107(1):61–70.
53. Schwanz LE, Janzen FJ. Climate change and temperature-dependent sex determination: can individual plasticity in nesting phenology prevent extreme sex ratios? *Physiol Biochem Zool.* 2008;81(6):826–34.
54. Bull JJ, Vogt RC, McCoy CJ. Sex determining temperatures in turtles: a geographic comparison. *Evolution.* 1982;36:326–32.
55. McGaugh S, Janzen F. Effective heritability of targets of sex-ratio selection under environmental sex determination. *J Evol Biol.* 2011;24(4):784–94.
56. Somero G. The physiology of climate change: how potentials for acclimatization and genetic adaptation will determine 'winners' and 'losers.' *J Exp Biol.* 2010;213(6):912–20.
57. Tedeschi JN, Kennington WJ, Tomkins JL, Berry O, Whiting S, Meekan MG, Mitchell NJ. Heritable variation in heat shock gene expression: a potential mechanism for adaptation to thermal stress in embryos of sea turtles. *Proc Biol Sci.* 2016;283(1822):20152320.
58. Urquhart KR, Zhao Y, Baker JA, Lu Y, Yan L, Cook MN, Jones BC, Hamre KM, Lu L. A novel heat shock protein alpha 8 (*Hspa8*) molecular network mediating responses to stress and ethanol-related behaviors. *Neurogenetics.* 2016;17(2):91–105.
59. Mkiye LS, Zishiri OT. Novel single nucleotide polymorphisms in the heat shock protein 70.1 gene in South African Nguni crossbred cattle. *Trop Anim Health Prod.* 2020;52:893–901.
60. Tomanek L. The importance of physiological limits in determining biogeographical range shifts due to global climate change: the heat-shock response. *Physiol Biochem Zool.* 2008;81(6):709–17.
61. Tomanek L. Variation in the heat shock response and its implication for predicting the effect of global climate change on species' biogeographical distribution ranges and metabolic costs. *J Exp Biol.* 2010, 213(6):971–979.
62. Haygood R, Babbitt CC, Fedrigo O, Wray GA. Contrasts between adaptive coding and noncoding changes during human evolution. *Proc Natl Acad Sci USA.* 2010;107(17):7853–7.
63. Lai Y-T, Yeung CK, Omland KE, Pang E-L, Hao Y, Liao B-Y, Cao H-F, Zhang B-W, Yeh C-F, Hung C-M. Standing genetic variation as the predominant source for adaptation of a songbird. *Proc Natl Acad Sci.* 2019;116(6):2152–7.
64. Cooper DN. Functional intronic polymorphisms: buried treasure awaiting discovery within our genes. *Hum Genomics.* 2010;4:284.
65. Savolainen O, Lascoux M, Merilä J. Ecological genomics of local adaptation. *Nat Rev Genet.* 2013;14(11):807.
66. Hand BK, Lowe WH, Kovach RP, Muhlfeld CC, Luikart G. Landscape community genomics: understanding eco-evolutionary processes in complex environments. *Trends Ecol Evol.* 2015;30(3):161–8.
67. Shaffer HB, Gidiş M, McCartney-Melstad E, Neal KM, Oyamauchi HM, Tellez M, Toffelmier EM. Conservation genetics and genomics of amphibians and reptiles. *Annu Rev Anim Biosci.* 2015;3(ch\_1):113–38.
68. Meek MH, Larson WA. The future is now: amplicon sequencing and sequence capture usher in the conservation genomics era. *Mol Ecol Resour.* 2019;19:795–803.
69. Gebremedhin B, Ficetola G, Naderi S, Rezaei HR, Maudet C, Rioux D, Luikart G, Flagstad Ø, Thuiller W, Taberlet P. Frontiers in identifying conservation units: from neutral markers to adaptive genetic variation. *Anim Conserv.* 2009;12(2):107–9.
70. Xuereb A, D'Aloia CC, Andreello M, Bernatchez L, Fortin MJ. Incorporating putatively neutral and adaptive genomic data into marine conservation planning. *Conserv Biol.* 2020;35:909–20.
71. Kawecki TJ, Ebert D. Conceptual issues in local adaptation. *Ecol Lett.* 2004; 7(12):1225–41.
72. Walsworth TE, Schindler DE, Colton MA, Webster MS, Palumbi SR, Mumby PJ, Essington TE, Pinsky ML. Management for network diversity speeds evolutionary adaptation to climate change. *Nat Clim.* 2019;9(8):632–6.
73. Bernatchez L. On the maintenance of genetic variation and adaptation to environmental change: considerations from population genomics in fishes. *J Fish Biol.* 2016;89(6):2519–56.
74. Reid BN, Thiel RP, Palsbøll PJ, Peery MZ. Linking genetic kinship and demographic analyses to characterize dispersal: methods and application to Blanding's turtle. *J Hered.* 2016;107:603–14.
75. Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. Very high resolution interpolated climate surfaces for global land areas. *Int J Climatol.* 2005; 25(15):1965–78.
76. O'Leary SJ, Puritz JB, Willis SC, Hollenbeck CM, Portnoy DS. These aren't the loci you're looking for: principles of effective SNP filtering for molecular ecologists. *Mol Ecol.* 2018;27:1–14.
77. Gruber B, Unmack PJ, Berry OF, Georges A. dartr: an R package to facilitate analysis of SNP data generated from reduced representation genome sequencing. *Mol Ecol Resour.* 2018;18(3):691–9.
78. Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. GenBank. *Nucleic Acids Res.* 2016;44(D1):D67–D72.
79. Schroeder AL, Metzger KJ, Miller A, Rhen T. A novel candidate gene for temperature-dependent sex determination in the common snapping turtle. *Genetics.* 2016;203(1):557–71.
80. Spotila LD, Spotila JR, Hall SE. Sequence and expression analysis of *WT1* and *Sox9* in the red-eared slider turtle, *Trachemys scripta*. *J Exp Zool.* 1998; 281(5):417–27.
81. Schoch CL, Sung G-H, López-Giráldez F, Townsend JP, Miadlikowska J, Hofstetter V, Robbertse B, Matheny PB, Kauff F, Wang Z, et al. The Ascomycota tree of life: a phylum-wide phylogeny clarifies the origin and evolution of fundamental reproductive and ecological traits. *Syst Biol.* 2009; 58:224–39.
82. Spinks PQ, Thomson RC, McCartney-Melstad E, Shaffer HB. Phylogeny and temporal diversification of the New World pond turtles (*Emyidae*). *Mol Phylogenet Evol.* 2016;103:85–97.
83. Seebacher F, Murray SA. Transient receptor potential ion channels control thermoregulatory behaviour in reptiles. *PLoS One.* 2007;2(3):e281.
84. Kumar S, Stecher G, Tamura K. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol.* 2016;33(7):1870–4.
85. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 2004;32(5):1792–7.

86. Gadberry MD, Malcomber ST, Doust AN, Kellogg EA. Primaclade—a flexible tool to find conserved PCR primers across multiple species. *Bioinformatics*. 2005;21(7):1263–4.
87. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet C, Al-Ghalith GA, Alexander H, Alm EJ, Arumugam M, Asnicar F. QIIME 2: reproducible, interactive, scalable, and extensible microbiome data science. *PeerJ*. 2018. Preprints.
88. Jombart T, Ahmed I. adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics*. 2011;27(21):3070–1.
89. Goudet J. Hierfstat, a package for R to compute and test hierarchical F-statistics. *Mol Ecol Notes*. 2005;5(1):184–6.
90. Paradis E. pegas: an R package for population genetics with an integrated–modular approach. *Bioinformatics*. 2010;26(3):419–20.
91. Weir BS, Cockerham CC. Estimating F-statistics for the analysis of population structure. *Evolution*. 1984;38:1358–70.
92. Willing E-M, Dreyer C, Van Oosterhout C. Estimates of genetic differentiation measured by FST do not necessarily require large sample sizes when using many SNP markers. *PLOS ONE*. 2012;7(8):e426492012.
93. Rousset F. Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics*. 1997;145(4):1219–28.
94. Caye K, Deist TM, Martins H, Michel O, François O. TESS3: fast inference of spatial population structure and genome scans for selection. *Mol Ecol Resour*. 2016;16(2):540–8.
95. Dixon P. VEGAN, a package of R functions for community ecology. *J Veg Sci*. 2003;14(6):927–30.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

