Received: 25 September 2023 | Revised: 9 June 2024 | Accepted: 1 July 2024

DOI: 10.1111/mec.17464

ORIGINAL ARTICLE

MOLECULAR ECOLOGY WILEY

Spatial variation in genomic signatures of local adaptation during the cane toad invasion of Australia

Marc A. Beer[1] | Daryl R. Trumbo[2] | Rhett M. Rautsaw[1,3] | Christopher P. Kozakiewicz[4] | Brendan Epstein[5] | Paul A. Hohenlohe[6] | Ross A. Alford[7] | Lin Schwarzkopf[7] | Andrew Storfer[1]

[1]School of Biological Sciences, Washington State University, Pullman, Washington, USA

[2]Department of Biology, Colorado State University Pueblo, Pueblo, Colorado, USA

[3]Department of Integrative Biology, University of South Florida, Tampa, Florida, USA

[4]W.K. Kellogg Biological Station, Department of Integrative Biology, Michigan State University, Hickory Corners, Michigan, USA

[5]Department of Plant and Microbial Biology, University of Minnesota, St Paul, Minnesota, USA

[6]Department of Biological Sciences, University of Idaho, Moscow, Idaho, USA

[7]College of Science and Engineering, James Cook University, Townsville, Queensland, Australia

Correspondence
Andrew Storfer, School of Biological Sciences, Washington State University, Pullman, WA, USA.
Email: astorfer@wsu.edu

Funding information
National Science Foundation, Grant/Award Number: DDIG 1407335 and GRFP 1842493

Handling Editor: Sean D Schoville

Abstract

Adaptive evolution can facilitate species' range expansions across environmentally heterogeneous landscapes. However, serial founder effects can limit the efficacy of selection, and the evolution of increased dispersal during range expansions may result in gene flow swamping local adaptation. Here, we study how genetic drift, gene flow and selection interact during the cane toad's (*Rhinella marina*) invasion across the heterogeneous landscape of Australia. Following its introduction in 1935, the cane toad colonised eastern Australia and established several stable range edges. The ongoing, more rapid range expansion in north-central Australia has occurred concomitant with an evolved increase in dispersal capacity. Using reduced representation genomic data of Australian cane toads from the expansion front and from two areas of their established range, we test the hypothesis that high gene flow constrains local adaptation at the expansion front relative to established areas. Genetic analyses indicate the three study areas are genetically distinct but show similar levels of allelic richness, heterozygosity and inbreeding. Markedly higher gene flow or recency of colonisation at the expansion front have likely hindered local adaptation at the time of sampling, as indicated by reduced slopes of genetic-environment associations (GEAs) estimated using a novel application of geographically weighted regression that accounts for allele surfing; GEA slopes are significantly steeper in established parts of the range. Our work bolsters evidence supporting adaptation of invasive species post-introduction and adds novel evidence for differing strengths of evolutionary forces among geographic areas with different invasion histories.

KEYWORDS
adaptation, ecological genetics, invasive species, landscape genetics, natural selection and contemporary evolution, population genetics – empirical

1 | INTRODUCTION

Species undergoing range expansion may rapidly colonise large geographic areas and encounter environmental heterogeneity that can impede further spread in the absence of adaptive evolution (Bock et al., 2015; Gilbert et al., 2017; Szűcs et al., 2017). Indeed, rapid, contemporary evolution is increasingly recognised as necessary for understanding and predicting the successful establishment of range-expanding species in novel habitat (Colautti & Lau, 2015; Szűcs et al., 2017). Understanding factors that facilitate or constrain adaptation during range expansions may benefit management of invasive species, which are a leading global threat to biodiversity (Hogue & Breon, 2022; Leclerc et al., 2018; Molnar et al., 2008), as well as management of species imperilled by climate change whose persistence may require range shifts.

The spatial context of range expansions affects interactions among evolutionary forces. Range expansions often result in a spatial analogue of genetic drift: the sequential establishment of populations at the advancing range edge can generate a progressive decline in genetic diversity from the expansion origin to the expansion front, a phenomenon referred to as a serial founder effect (Excoffier et al., 2009; Slatkin & Excoffier, 2012). Concomitantly, the spatial spread of alleles can result in the chance establishment of allele frequency gradients across populations distributed along the direction of expansion, known as allele surfing (Edmonds et al., 2004; Fix, 1997; Klopfstein et al., 2006). Because allele surfing is a form of genetic drift (Slatkin & Excoffier, 2012), populations distributed along an environmental gradient may initially experience a strengthening or disruption to local adaptation via chance increases in the frequencies of locally adaptive or maladaptive alleles, respectively (e.g. Gralka et al., 2016; Travis et al., 2007). Following initial colonisation, genetic drift may be strong within very small expansion front populations, which can hinder local responses to selection (Polechová & Barton, 2015). Finally, reduced adaptive potential at the expansion front (due to both founder effects and enhanced drift while populations are small) may impede continued spread into areas with novel environments (Antonovics, 1976; Polechová & Barton, 2015).

Increased dispersal capacity often evolves during range expansions (e.g. via spatial sorting) (Hargreaves & Eckert, 2014; Miller et al., 2020; Phillips et al., 2008; Shine et al., 2011; Travis & Dytham, 2002). To the extent that dispersers reproduce successfully at their destinations, increased dispersal can bolster gene flow among populations near the expansion front. Gene flow can introduce genetic variation into expansion front populations, thereby ameliorating serial founder effects (Bialozyt et al., 2006; Fayard et al., 2009; Ray & Excoffier, 2010) and facilitating adaptation at the range edge (Polechová, 2018). Alternatively, substantial gene flow can swamp local adaptation among populations experiencing divergent selection (Lenormand, 2002) and thereby limit further range expansion (Fedorka et al., 2012). Owing to differences in dispersal capacity between earlier-colonised portions of a species' range and the expansion front, the magnitude of gene flow and whether it facilitates or constrains adaptive evolution are expected to vary spatially.

Although studying adaptation of range-expanding species may be critical for predicting their capacity for further expansion, range expansions pose a unique analytical challenge for genomic approaches to studying local adaptation. Allele surfing can generate spurious allele frequency – environment correlations (i.e. genetic-environment associations or GEAs) that are often interpreted as evidence that a given locus experiences spatially divergent selection and thus contributes to local adaptation (Hoban et al., 2016; Lotterhos & Whitlock, 2015; Zhao et al., 2020). In addition, typical GEA analyses assume spatial stationarity of the allele–environment relationship (Joost et al., 2013); this assumption is almost certainly violated in the context of rapid range expansions. The evolution of increased dispersal and consequent gene flow can homogenise allele frequencies among populations (Slatkin, 1985), thereby hindering the efficacy of selection at the expansion front. Geographic cline theory (Alleaume-Benharira et al., 2006; Barton & Hewitt, 1985; Endler, 1977) suggests that this phenomenon will be borne out analytically as a reduction in the slope or 'steepness' of a GEA (i.e. the allele frequency change per unit change in an environmental factor) at the expansion front relative to established portions of the range where dispersal is comparatively low. Clearly, more work is needed to disentangle the effects of drift (via allele surfing), gene flow and selection in genomic tests of local adaptation during range expansions.

The cane toad (*Rhinella marina*) colonisation of Australia represents a marquee example of a biological invasion characterised by evolution of increased dispersal at the range expansion front. Initially introduced to the northeastern (NE) coast of the Australian mainland in 1935, the cane toad became invasive and rapidly spread northwest and southeast (Urban et al., 2008) (Figure 1). Although the NE region (throughout central and northern Queensland; Figure 1) and southeastern region (S; extending from southern Queensland to New South Wales) of the cane toad's Australian range have largely remained stable since the 1970s (Macgregor et al., 2021; Urban et al., 2007), the northwestern (NW) region remains an expanding invasion front, where cane toads have spread into the upper part of the Northern Territory and into Western Australia (Doody et al., 2018). The colonisation of the NW region is characterised by an accelerating invasion speed of approximately 40–60 km per year, contrasting with an estimated invasion speed of 10–20 km per year in the earlier colonised NE and S regions (Urban et al., 2008). The increased invasion speed of the NW region is thought to have occurred as a result of selection for a dispersal-enhancing phenotype (Alford et al., 2009; Gruber et al., 2017a; Hudson et al., 2016, 2020; Phillips et al., 2008) that appears to have a genetic basis (Gruber et al., 2017b; Phillips et al., 2010). In contrast, longer-colonised areas are inhabited by toads with reduced dispersal capacity (Alford et al., 2009; Gruber et al., 2017a; Hudson et al., 2016; Phillips et al., 2006, 2008).

Prior population genomic work suggests that the cane toad invasion of Australia is an ideal testbed for advancing analytical approaches to understanding the genomic basis of local adaptation in range-expanding species. Trumbo et al. (2016) provided a test of the central-marginal hypothesis of species' range limits, which suggests

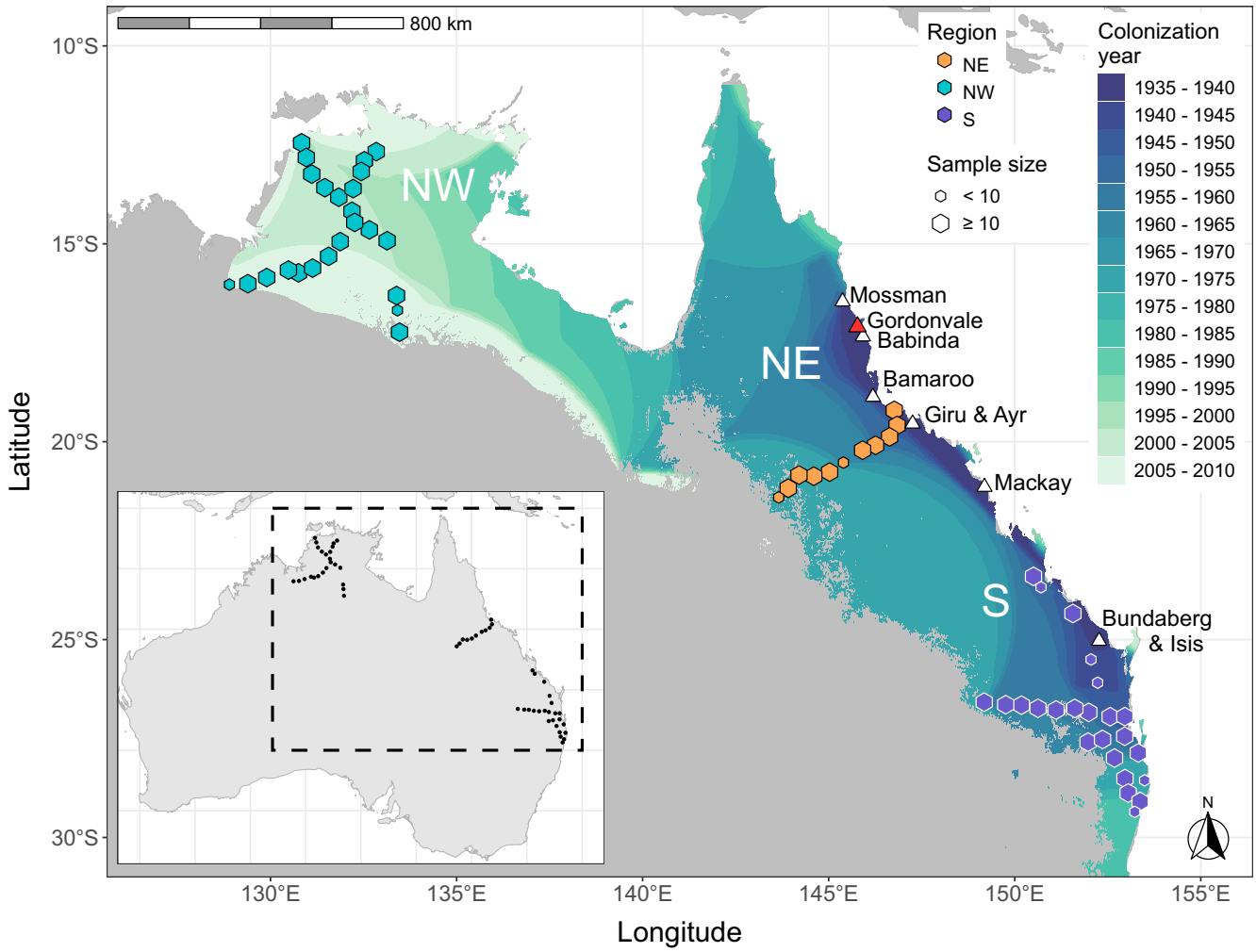

FIGURE 1 Sampling of cane toads across the species' invasive range in Australia. Hexagonal points indicate the final 59 sampling localities coloured by geographic sampling region (teal = NW, orange = NE and purple = S). Hexagonal point sizes reflect final sample sizes at localities. The red triangle (Gordonvale) identifies the initial introduction location of the cane toad in 1935; white triangles indicate subsequent introductions (1935–1937) derived from Gordonvale cane toads. Map features are superimposed on a simplified colonisation year surface showing progression of the cane toad invasion across Australia; the colonisation year surface is only shown for areas with habitat suitability ≥0.20 as inferred by Trumbo et al. (2016).

that genetic diversity and gene flow should decrease from the core of a species' geographic range to the range edges (Eckert et al., 2008). Results in Trumbo et al. (2016) were consistent with the predictions of the central-marginal hypothesis: within the NW and S regions, genetic isolation between localities increases with distance from the range core, suggesting increasing habitat fragmentation towards the inland range edge. Additionally, genetic diversity in the S region is negatively correlated with distance from the range core. Despite expectations for sequential declines in genetic diversity within the expanding NW region owing to serial founder effects, this region shows high effective population sizes (N_e) and genetic diversity (Trumbo et al., 2016). These prior results suggest that spatial variation in the relative strengths of evolutionary processes may impact the capacity for cane toad populations to become locally adapted. However, apart from an RNA sequencing study of cane toads from a limited number of localities (Selechnik et al., 2019), genomic signatures of local adaptation and geographic variation in these signatures have yet to be tested across the cane toad's invasive Australian range.

We set out to test for evidence of local adaptation using GEA analyses (Hoban et al., 2016; Rellstab et al., 2015). However, because range-expanding species can confound signals of selection via genetic drift, unlike Trumbo et al. (2016), we also tested the relative strengths of gene flow, selection and drift among populations in established range regions (NE and S) and the expanding NW region (Figure 1). Furthermore, we re-genotyped the sequencing data from Trumbo et al. (2016) for 932 cane toads collected in 2010–2011 and used the reference genome published in the interim (Edwards et al., 2018) to assist genotyping and annotate loci with nearby genes. Environmental heterogeneity is replicated across the three sampling regions (Figure S1), providing a rare opportunity to evaluate the correspondence between the strength of genomic signatures of local adaptation and the magnitudes of genetic drift and gene flow among

populations. Additionally, the availability of 76 years of georeferenced Australian cane toad occurrence records provides the potential to explicitly model allele surfing and control for its impact on GEA analyses. We hypothesised that high gene flow within the expanding NW region swamps local adaptation, as measured by a reduction in the slope of GEAs in the NW relative to the comparatively dispersal-limited NE and S regions. To test this hypothesis, we (1) reconstructed the cane toad's spatiotemporal spread across Australia; (2) tested for differences in the magnitudes of genetic drift and gene flow among regions by estimating genome-wide genetic diversity and structure and (3) used a novel application of geographically weighted regression (GWR) to assess spatial variation in the slopes of GEAs and thereby test the strength of local adaptation to abiotic environmental heterogeneity while also accounting for allele surfing.

2 | MATERIALS AND METHODS

2.1 | Sequence data

We re-processed raw sequencing data generated by Trumbo et al. (2016). These data were from tissue samples collected from 1123 individual cane toads at 62 Australian mainland localities in 2010 and 2011 between January and April; sampled cane toads were typically adults (Trumbo et al., 2016). The localities spanned the Australian range of the cane toad at the time of sampling, and localities corresponded to three broad geographic regions with varying invasion histories. Specifically, the NW region remains a range expansion front, while the NE and southern (S) regions are stabilised portions of the cane toad's Australian range (Figure 1).

Trumbo et al. (2016) generated sequencing libraries using a double-digest restriction site-associated sequencing (ddRADseq) protocol (Peterson et al., 2012). The restriction enzymes PstI and EcoRI, which each recognise different 6-bp restriction sites, were used in restriction digestion. Further details regarding tissue sample processing in preparation for sequencing can be found in Trumbo et al. (2016). In total, Trumbo et al. (2016) sequenced eight libraries separately on Illumina HiSeq 2000 lanes at the University of Oregon Genomics Core Facility using 100-bp single-end reads. We downloaded these sequence data from the NCBI BioProject Accession PRJNA328156.

As this study's aims differed from Trumbo et al. (2016), and a draft genome assembly has since been made available for the cane toad (Edwards et al., 2018), we reprocessed the data from raw sequence reads. First, we removed seven individuals represented by <10,000 reads. Next, we employed an integrated alignment procedure (Paris et al., 2017) by assembling RAD loci de novo in Stacks 2.52 (Rochette et al., 2019) and subsequently aligning the catalogue of assembled RAD loci to the draft genome generated from an Australian cane toad (Edwards et al., 2018). We created the de novo RAD locus catalogue using the Stacks script *denovo_map.pl*, allowing a minimum of three reads to form a stack ($m=3$), a maximum of two mismatches between read stacks within individuals ($M=2$) and a maximum of three mismatches between read stacks across individuals ($n=3$). We aligned the de novo RAD locus catalogue to the cane toad draft genome using GSNAP (Wu & Nacu, 2010); we specified a maximum of five mismatches, minimum coverage of 0.95, and terminal alignments disabled. We did not align sequence reads directly to the reference genome because the integrated alignment procedure can be more effective at discovering RAD loci (Paris et al., 2017). Finally, we retained only the RAD loci with unique alignments to the reference genome, which should increase single nucleotide polymorphism (SNP) quality by removing loci that may represent collapsed paralogous sequences (O'Leary et al., 2018).

We then used an iterative data filtering scheme (O'Leary et al., 2018) to improve genotyping quality and reduce missing data (Table S1). The filters included progressively strengthened individual-level missing data thresholds and SNP-level missing data thresholds whereby sampling localities were grouped into 12 spatial clusters (Figure S2) before removing SNPs showing high missing data in at least one spatial cluster from the entire dataset (e.g. Beer et al., 2022, 2024). This process leads to similar levels of missing data at the same SNPs across the entire study area. The spatial clusters of localities were generated using hierarchical clustering of geographic coordinates and truncating the resulting tree to generate 12 groups; higher numbers of spatial clusters led to excessive loss of SNPs during filtering. Note that this procedure differs from the filtering used by Trumbo et al. (2016), who filtered SNPs separately by transect and region; this methodology was appropriate for the study question in Trumbo et al. (2016) but is not appropriate for our research, as SNPs must be consistently genotyped across the entire study area to make inter-region comparisons of signatures of local adaptation. SNP filtering was conducted in VCFtools version 0.1.16 (Danecek et al., 2011).

The final dataset consisted of 5723 SNPs and 932 individuals distributed across 59 sampling localities (mean = 15.80 samples/locality; SD = 4.52; range = 5–20). Individual sequencing depth averaged across SNPs was generally high globally (mean = 19.10 reads/SNP; SD = 11.23; range = 6.30–88.43). Individual sequencing depth was also generally high when averaging across individuals grouped by sampling locality (mean = 18.83 reads/SNP; SD = 6.58; range = 8.36–31.48). SNP sequencing depth averaged across individuals was also generally high globally (mean = 19.56 reads/SNP/individual; SD = 4.69; range = 10.51–105.20). Global individual-level missing data was generally low (mean = 9.35%; SD = 11.56%; range = 0.63%–51.46%), as was global SNP-level missing data (mean = 9.35%; SD = 5.03%; range = 0.32%–34.33%). Per-locality individual missing data was also low (mean = 9.76%; SD = 7.01%; range = 1.59%–32.72%).

2.2 | Local colonisation year estimation

We reconstructed the cane toad's spread across Australia in order to obtain an estimate of the year each of our sampling localities was first colonised; we subsequently used these colonisation year estimates to test for serial founder effects and account for allele surfing in GEA analyses. Previous reconstructions of the cane toad's spread across Australia (Phillips & Shine, 2004; Urban et al., 2008) did not

extend through the timeframe of our sample collection; we generated a novel reconstruction up to 2010. We obtained 19,762 observations of cane toads across mainland Australia from the Global Biodiversity Information Facility (GBIF; Edwards et al., 2000). We removed observations lacking precise geographic coordinates and dates; we also removed observations with implausible locations based on prior knowledge of the cane toad's spread (e.g. we removed observations in the inland Australian deserts). We then added our own observations for the localities we sampled in 2010–2011, such that our final dataset totalled 19,804 unique observations (Figure S3). We then separated observations for each year from 1936 to 2010. Broadly following Urban et al. (2008), we generated an alpha shape describing the spatial extent of cane toad observations each year. 1936 was the earliest year used because ≥3 observations are necessary to construct alpha shapes. Alpha shapes were generated using the function *getDynamicAlphaHull* in the R package rangeBuilder (Davis Rabosky et al., 2016). The yearly alpha shapes were then collapsed into a single surface describing the putative year that cane toads colonised different locations across Australia.

2.3 | Environmental data processing

We obtained environmental data from the WorldClim v2 database (Fick & Hijmans, 2017). We initially collected data for mean annual temperature (BIO1), mean temperature diurnal range (BIO2), isothermality (BIO3), temperature seasonality (BIO4), temperature annual range (BIO7), annual precipitation (BIO12), precipitation seasonality (BIO15) and elevation. Temperature- and precipitation-related variables relate to amphibian stressors (Daszak et al., 2005) and may broadly act as selective pressures (Snyder & Weathers, 1975). Dissolved oxygen concentration affects embryonic development in a different toad species (Dmitrieva, 2015), and elevation may be a proxy for this factor. These variables also contribute to cane toad habitat suitability across Australia (Trumbo et al., 2016).

We first estimated Pearson's correlation coefficient r between each environmental factor and colonisation year at sampling localities; we removed environmental factors correlated with colonisation year at $|r| > .70$ to reduce the potential for mistaking SNPs showing allele surfing along environmental gradients as being subject to spatially divergent selection. For the remaining environmental factors, we removed one environmental factor when a pair had $|r| > .70$ (Figure S4). Our final set of four environmental factors consisted of mean diurnal temperature range (BIO2), temperature seasonality (BIO4), annual precipitation (BIO12) and elevation.

2.4 | Population genomic structure

In order to identify consensus genetic groupings of samples on which to run downstream population genetic analyses, we employed four methods to characterise the number of genetic clusters represented in our dataset: TESS3, ConStruct, principal components analysis (PCA) and Treemix. TESS3 analyses were performed using the R package tess3r. TESS3 uses geographically constrained non-negative matrix factorisation to characterise population genomic structure (Caye et al., 2018), including the evaluation of statistical support for the number of ancestral genetic clusters (K). We tested $K = 1–40$ (10 replicates each) and assessed statistical support using root mean squared error and cross entropy. $K = 4–9$ showed similar support when summarising across replicates; for each of these values of K, we retained the individual replicate that showed the strongest statistical support (i.e. minimum RMSE).

Genetic clustering algorithms can overestimate the number of ancestral genetic clusters when there is continuous genetic differentiation across space (Bradburd et al., 2018). Therefore, we used the R package conStruct as another method to characterise the number of genetic clusters in our dataset. ConStruct partitions genetic variation across K geo-genetic 'layers' (genetic clusters). Genetic similarity is allowed to decay with geographic distance within each layer, thereby allowing allele frequencies to vary spatially within each genetic cluster. We tested $K = 1–10$ and evaluated statistical support using 20 cross-validation replicates with 75/25% training/testing data partitions and 10,000 iterations. We chose as optimal the value of K at which predictive accuracy reached an asymptote and at which each genetic cluster explained ≥5% of total genetic covariance; larger values of K added genetic clusters that explained little covariance. We re-ran conStruct on the full dataset using the optimal value of K to generate a final model.

We also implemented PCA of individual genotypes to visualise genomic variation using the R package adegenet (Jombart, 2008). PCA does not allow missing data, so we imputed missing genotypes using the R package LEA (Frichot & François, 2015). Using the function *snmf*, we ran three replicates of each value of $K = 1–40$ and identified the replicate with the best genotype prediction accuracy. Using this value of K, we used the function *impute* to replace missing genotypes. We also conducted PCA of locality allele frequencies using function *rda* in the vegan R package (Oksanen et al., 2020); 20 missing locality allele frequencies were replaced using median allele frequencies at the appropriate SNPs. We centred and scaled data prior to both individual- and locality-based PCAs.

We used Treemix to infer population trees based on locality allele frequency data (Pickrell & Pritchard, 2012). Treemix estimates a drift parameter describing the magnitude of drift experienced along branches in the population tree, as indicated by branch lengths. We also tested directional migration edges connecting branches in the population tree using 100 replicate runs with $m = 0$ (i.e. no migration edges modelled) and 10 replicate runs each for $m = 1–20$. We evaluated relative evidence for values of m using the delta-m method implemented in the R package OptM (Fitak, 2021); we identified the value of m at which (1) the composite log-likelihood reached an asymptote and (2) the second-order rate of change in the composite log-likelihood was maximised; this value of m was considered the optimal number of migration edges.

To estimate gene flow, we calculated Weir and Cockerham's estimator of F_{ST} (Weir & Cockerham, 1984) using the R package

diveRsity (Keenan et al., 2013). Pairwise F_{ST} values and their 95% confidence intervals (CIs) were estimated using the *diffCalc* function; CIs were estimated by bootstrapping across SNPs for 5000 iterations. Pairwise F_{ST} was estimated among sampling localities as well as among TESS3 genetic clusters. We tested for isolation-by-distance (IBD) by fitting linear models using generalised least squares (GLS); we specified a correlation structure following the method of maximum likelihood population effects (MLPE) to account for non-independence of pairwise observations (Clarke et al., 2002). Specifically, we regressed pairwise linearised F_{ST} (calculated as $F_{ST}/[1-F_{ST}]$) against the natural logarithm of geographic distance (Rousset, 1997). We also fitted an intercept-only null model. All models were fitted using the *gls* function in the R package nlme, and the correlation structure was specified using the function *corMLPE* in the R package MLPE. The geographic distance model was preferred over the intercept-only null model when the geographic distance model had the lowest AICc, and the null model had $\Delta AICc > 2$ relative to the geographic distance model. Models were fitted based on pairwise F_{ST} between sampling localities within regions as well as F_{ST} between the K genetic clusters inferred by TESS3 for $K = 4–9$. Given that multiple sampling localities are subsumed into each of K genetic clusters, we calculated geographic distances between genetic clusters using the mean of the geographic coordinates of their constituent sampling localities.

2.5 | Genetic diversity

We calculated Weir and Cockerham's estimator of F_{IS} using the *basicStats* function in diveRsity, with 95% CIs calculated from 5000 bootstraps (Keenan et al., 2013). Observed and expected SNP heterozygosity, as well as allelic richness (Ar), were similarly estimated using the *basicStats* function. For Ar, we enabled rarefaction to account for sample size differences across localities (Keenan et al., 2013). We estimated effective population size (N_e) for each sampling locality using the linkage disequilibrium (LD) method in NeEstimator v2.1 (Do et al., 2014); we also estimated overall N_e for each region by jointly analysing all samples within each region. F_{IS} and N_e were estimated using putatively neutral SNPs (i.e. after removing SNPs showing significant GEAs; see below). Observed heterozygosity, expected heterozygosity and Ar were estimated separately for putatively neutral SNPs and SNPs with significant GEAs. Locality-specific N_e, observed heterozygosity, expected heterozygosity and Ar were regressed against locality colonisation year using the *lm* function in R v4.1.2.

2.6 | Genetic-environment association analyses

Our final sample size of 932 individuals from 59 localities surpasses minimum statistical recommendations (e.g. a total sample size of ≥200 individuals) for GEA analyses (Selmoni et al., 2020). Environmental conditions were also replicated across the three sampling regions (Figure S1). Thus, we were able to make high-quality inter-region comparisons of GEAs.

We implemented GEA tests using Bayenv2 and GWR. Bayenv2 tests for GEAs at individual SNPs while accounting for genome-wide neutral genetic structure using a population covariance matrix (Günther & Coop, 2013). We generated the covariance matrix of our sampling localities by calculating the median matrix based on five replicates of 500,000 iterations each. For each SNP – environmental factor combination, we then estimated Bayes Factors describing the relative support for a model containing an environmental factor compared to a null model excluding the environmental factor. We also estimated Spearman's Rho, a correlation coefficient that is robust to outlier populations. We ran five replicates of 750,000 iterations each and calculated the median Bayes factor and median absolute value of Spearman's Rho for each SNP – environmental factor combination. SNPs that were in the top 0.05 quantile of Spearman's Rho and had Bayes Factors >10 were considered as showing significant GEAs; a Bayes Factor >10 alone is considered strong evidence in favour of the model including the environmental factor (Kass & Raftery, 1995); thus, our use of thresholds for both Spearman's Rho and Bayes Factors should increase confidence in our results (Günther & Coop, 2013).

We also performed one of the first applications of GWR to investigate variation in slopes of GEAs across the three sampling regions. In short, GWR estimates geographically local response – covariate relationships. Specifically, GWR estimates a local response – covariate beta coefficient (i.e. slope) for a focal observation (here, locality) based on nearby observations, with more geographically distant observations contributing less to the estimation of the local beta coefficient; this process is repeated for every observation in the dataset (Brunsdon et al., 1996). Thus, GWR accommodates the possibility that the steepness and/or direction of the relationship between two variables (here, allele frequency at a given SNP and an environmental factor) may vary across space (Brunsdon et al., 1996). GWR has previously been suggested and used for studying barriers to gene flow (Diniz-Filho et al., 2016; Storfer et al., 2007); GWR has also been suggested for detecting signatures of spatially divergent selection at individual loci (Manel et al., 2010), but to our knowledge, it has not previously seen practical implementation as a GEA test.

We implemented GWR using the R package GWmodel (Gollini et al., 2015). The extent to which geographically distant observations contribute to estimation of local beta coefficients at a focal observation depends partly on a parameter referred to as bandwidth. Using a bisquare kernel function, bandwidth governs the spatial rate of decay in the influence of distant observations and defines the maximum geographic distance at which observations influence the estimation of local beta coefficients at a focal observation. We used the function *bw.gwr* to optimise model bandwidth using cross-validation (i.e. the bandwidth that minimises out-of-sample prediction error was selected as optimal). Subsequently, we used the function *gwr.basic* to fit GWR models and obtain local estimates of beta coefficients. Further details of model fitting and model comparison are provided below.

For each SNP, we specified a null model where locality allele frequency was regressed against locality values for genome-wide PC1 and PC2 scores and colonisation year. Colonisation year is a proxy for range expansion direction, and alleles correlated with colonisation year thus carry a signature of allele surfing. Thus, this null model explicitly tests for the effects of neutral population structure and allele surfing on allele frequencies. We compared models that additionally include one of the four environmental factors (alongside genome-wide PC scores and colonisation year) against this null model by calculating an approximate Bayes Factor based on the models' Bayesian Information Criterion scores (Wagenmakers, 2007). As both the null and alternative models include colonisation year, a Bayes Factor >1 indicates support for an effect of the environment on allele frequencies that is not satisfactorily explained by allele surfing alone. We considered SNPs in the top 0.05 quantile of Bayes Factors as showing significant GEAs; in practice, these SNPs had Bayes Factors >10.

Because our work represents a novel application of GWR as a GEA test, we sought to be conservative in our use of the method. Accordingly, we intersected significant SNPs identified by GWR with the SNPs identified by Bayenv2; subsequent analyses used only the overlapping SNPs. We also tested whether the observed overlap between Bayenv2 and GWR was significantly greater than expected by random chance by using a permutation test with 500,000 replicates. For each replicate, we completed two random draws of SNPs without replacement: we randomly drew 287 SNPS to represent the results from GWR, and we randomly drew 130 SNPs to represent the results from Bayenv2. These numbers of SNPs are identical to the numbers of SNPs identified by GWR and Bayenv2 in our real results. Next, we computed the number of SNPs that overlapped between the two random draws and used this as our test statistic. 500,000 replicates of the above procedure were used to construct a distribution of overlap, which we subsequently used to calculate the *p*-value of the overlap we observed in our real analyses.

To evaluate differences in genomic signatures of local adaptation across our sampling regions, we focused on spatial variation in the environmental beta coefficient (i.e. allele frequency–environment slope) estimated by GWR models for the SNPs that overlapped between GWR and Bayenv2. The use of the environmental beta coefficient as a metric to compare genomic evidence for local adaptation in different geographic regions follows from the fact that local adaptation depends on allele frequency differentiation among populations experiencing different selective environments (Hoban et al., 2016); small environmental beta coefficients (i.e. shallow slopes) in our regression models indicate little allele frequency differentiation at putatively adaptive loci (and vice versa) among populations inhabiting different environments. Additionally, geographic cline theory posits that the slope of a cline (conceptually analogous to the environmental beta coefficient estimated here) reflects the efficacy of selection at a locus given the relative strengths of selection and other forces affecting allele frequencies, such as gene flow (Barton & Hewitt, 1985; Endler, 1977).

For each SNP that overlapped between GWR and Bayenv2, we re-fit GWR models using centred and scaled environmental data so that the environmental beta coefficient represents the change in the SNP's allele frequency per standard deviation change in the environment; this procedure makes beta coefficients obtained from models including different environmental factors comparable. Within each region, we then averaged the absolute value of each SNP's locality-specific environmental beta coefficients such that each SNP had one mean absolute beta coefficient per region. We statistically compared environmental beta coefficients across regions using multiple pairwise Wilcoxon signed-rank tests with a family-wise alpha of .05 after Bonferroni correction. The Wilcoxon signed-rank test leverages paired data; in our case, the beta coefficient of a given SNP was paired across regions. We used the nonparametric Wilcoxon signed-rank test because paired differences in beta coefficients between the NW and NE regions showed a strong but non-significant deviation from normality (Shapiro–Wilk test $W=0.86$; $p=.06$). The Wilcoxon signed-rank test assumes that observations (here, SNPs) are independent, so we pruned SNPs by LD. We used the VCFtools commands *geno-r2* and *interchrom-geno-r2* to calculate the LD estimator r^2 for the significant SNPs that overlapped between GWR and Bayenv2. We calculated r^2 between SNPs separately for the NW, NE and S sampling regions. We removed one SNP from a pair of SNPs that had $r^2>.1$ within at least one sampling region.

2.7 | Candidate gene identification

To gain insight into the genomic basis of local adaptation and generate hypotheses for future work, we recorded genes nearest each SNP with a significant GEA using the BEDTools *closest* command (Quinlan & Hall, 2010). We obtained gene ontology (GO) terms for the candidate genes from Xenbase (Xenbase.org), which is a biological database that focuses on *Xenopus* frog species. Next, we used the PANTHER-powered system (Mi et al., 2019) on the GO Consortium webpage to employ GO enrichment analysis (The Gene Ontology Consortium, 2019). We specified genes nearest all 5723 SNPs in our dataset as the reference gene set. We specified the reference species as the western clawed frog (*Xenopus tropicalis*), as it is the most closely related reference species available. We tested for significant over- or under-representation of biological processes, molecular functions and cellular components GO terms among the genes nearest the SNPs with significant GEAs; we used Fisher's exact test and a false discovery rate of 0.05 (Benjamini & Hochberg, 1995).

3 | RESULTS

3.1 | Local colonisation year estimation

Using 76 years of georeferenced Australian cane toad occurrence records, we reconstructed the cane toad's spatiotemporal spread across Australia. Our estimation of cane toad colonisation

years across Australia shows strong agreement with prior studies (Figure 1; Figure S3) (Phillips & Shine, 2004; Urban et al., 2008) but extends estimates through 2010. NW localities were colonised in approximately 1996–2010, NE localities 1937–1974 and S localities 1941–1980. As cane toads have a generation time of approximately 1–2 years (Kosmala et al., 2018), some NW localities may have been sampled herein within one generation of colonisation. Note that downstream uses of estimated colonisation years involve correlative analyses, so minor deviations from true colonisation years are unlikely to obscure general patterns.

3.2 | Population genomic structure

Genetic data filtering (Table S1) generated a final dataset of 5723 SNPs and 932 cane toads from 59 sampling localities in three broad geographic sampling regions (Figure 1). TESS3, ConStruct and Treemix suggest that the NW, NE and S sampling regions correspond to biologically meaningful and genetically distinct groups of samples. PCA indicated broad clustering of individuals and localities based on sampling region (Figure 2a; Figure S5). TESS3 identified similar support across replicates for between four and nine genetic clusters (i.e. $K=4$–9; Figure S6). $K=4$ grouped the sampling localities such that the NW and NE regions each comprised a single genetic cluster and the S region comprised two genetic clusters (Figure S7a), and this partitioning of sampling localities into genetic clusters is broadly consistent with PCA (Figure 2a; Figure S5). $K>4$ primarily further subdivided the S region and split the NE region into two genetic clusters occupying the coast and more inland geographic areas (Figure S7b–f). ConStruct indicated weaker support for a non-spatial model of population structure compared to a spatial model that explicitly models IBD within genetic clusters (Figure S8a,b) (Bradburd et al., 2018). Comparisons among different values of K for the spatial model suggested that there are three meaningful spatial genetic clusters represented in the dataset (i.e. each cluster contributes ≥5% to total covariance; Figure S8c,d). These three genetic clusters largely correspond to the three sampling regions (Figure S9).

The maximum pairwise F_{ST} value ($F_{ST}=0.218$) occurred between a locality in the NW region and a locality in the S region. Pairwise F_{ST} values between localities within the NW sampling region were low (mean = 0.008; range = 0–0.038) compared to pairwise F_{ST} values in the NE (mean = 0.051; range = 0.009–0.121) and S regions (mean = 0.123; range = 0.014–0.192); this pattern holds for pairs of localities separated by similar geographic distances (Figure 2b). For each region, the MLPE model regressing linearised F_{ST} (i.e. $F_{ST}/[1-F_{ST}]$) against the natural logarithm of geographic distance outperformed an intercept-only null model with ΔAICc > 2, evidencing IBD (Table S2); the coefficient estimated for geographic distance was smaller in the NW than in the NE and S, reflecting weaker IBD among localities in the NW region (Figure 2b; Table S2). With respect to TESS3, $K=5$–9 showed significant signatures of IBD; $K=4$ did not show significant IBD despite a general positive correlation between genetic and geographic distances, possibly due to the small number of data points (Figure S10, Table S3).

Treemix indicated that sampling localities clustered together phylogeographically in a way that reflected sampling region and geographic proximity within regions. The NW and S regions share a common ancestor that in turn shares a common ancestor with inland NE localities (Figure 2c). Treemix indicated that localities in the S region have experienced substantially greater independent genetic drift than localities in the NE and NW regions, as represented by relative branch lengths in Figure 2c. Localities in the NW region have experienced relatively little independent drift. The addition of one directional migration edge had high support, with the next best models suggesting the presence of four and 11 migration edges respectively (Figures S11 and S12). Migration edges typically occurred between groups of localities in the southernmost S region (Figure S12). Taken together, these analyses consistently support the NE, NW and S as biologically meaningful and genetically distinct groups of samples. Thus, we focus on population genomic patterns within and across these three sampling areas for the remainder of our analyses.

3.3 | Genetic diversity

Point estimates of locality effective population sizes (N_e) were generally highest in the most recently invaded NW region (median $N_e=180.7$), lower in the NE region (median $N_e=63.7$) closer to the site of original introduction and the lowest in the S region (median $N_e=40.8$; Figure 3a). We report median locality N_e here because the earliest-colonised locality had an extremely high outlier estimate ($N_e=1091.4$; Figure 3b), which skewed the mean. However, we note that locality N_e estimates were typically imprecise, with upper 95% confidence limits often returned as infinite. N_e was significantly negatively correlated with colonisation year in the NE region (Pearson's $r=-.76$, $p<.0167$), although this correlation was non-significant after removing the outlier N_e value ($r=-.6$, $p=.08$). A negative but non-significant correlation between N_e and colonisation year was identified in the S region ($r=-.43$, $p=.05$). The NW region showed no trend between N_e and colonisation year ($r=.03$, $p=.91$; Figure 3b). Estimates at the region scale revealed relatively high N_e in the NW region ($N_e=178.2$), intermediate N_e in the NE region ($N_e=90.8$) and relatively low N_e in the S region ($N_e=44.2$).

Allelic richness was comparable among regions (Figure 3c). The S region showed a significant negative correlation between Ar and colonisation year ($r=-.67$, $p<.01$; Figure 3d); we found no association between Ar and colonisation year for the NE and NW regions. Observed and expected SNP heterozygosity showed similar patterns to Ar (Figure 3e–h). F_{IS} was negative in all localities except one locality in the S region (Figure 3i), and we found no association between F_{IS} and colonisation year in any region (Figure 3j). Note that the trends in genetic diversity characterised here are consistent with a previous study based on the same raw sequencing data (Trumbo et al., 2016), although here we explicitly model the relationship between genetic variation and colonisation year.

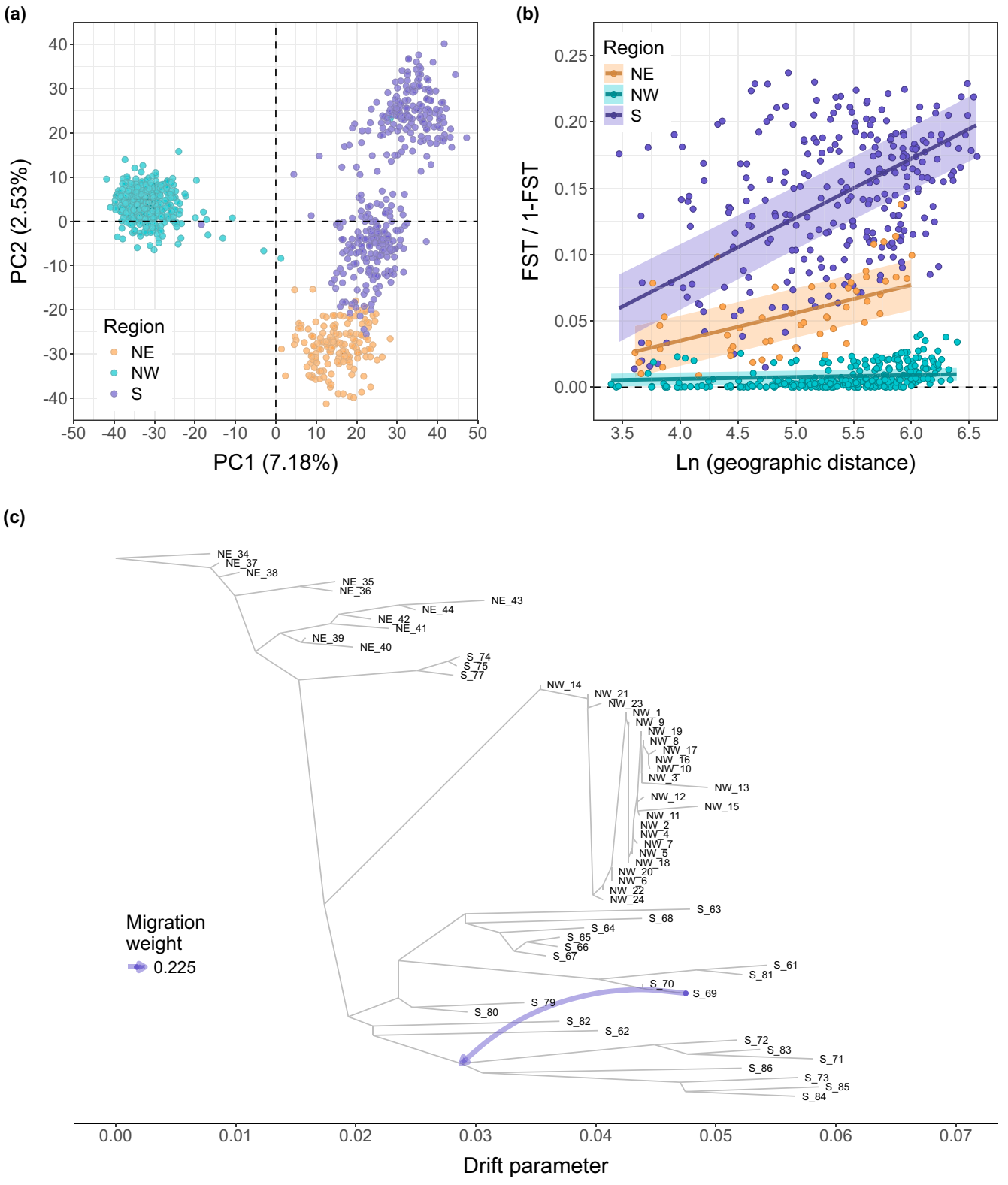

FIGURE 2 Population genomic structure among cane toads in Australia. (a) Principal components analysis of individual genotypes, with individuals coloured by sampling region. (b) Isolation-by-distance among localities within regions. Points represent estimates of genetic differentiation and geographic distance between pairs of localities, which are coloured by sampling region. Lines and 95% confidence intervals visualise region-specific linear regression models fitted with maximum likelihood population effects. (c) Treemix phylogeny with one migration edge. Branch lengths reflect the magnitude of genetic drift. The replicate with the highest log-likelihood is plotted. Localities are labelled by region and a locality number corresponding to those mapped in Figure S2.

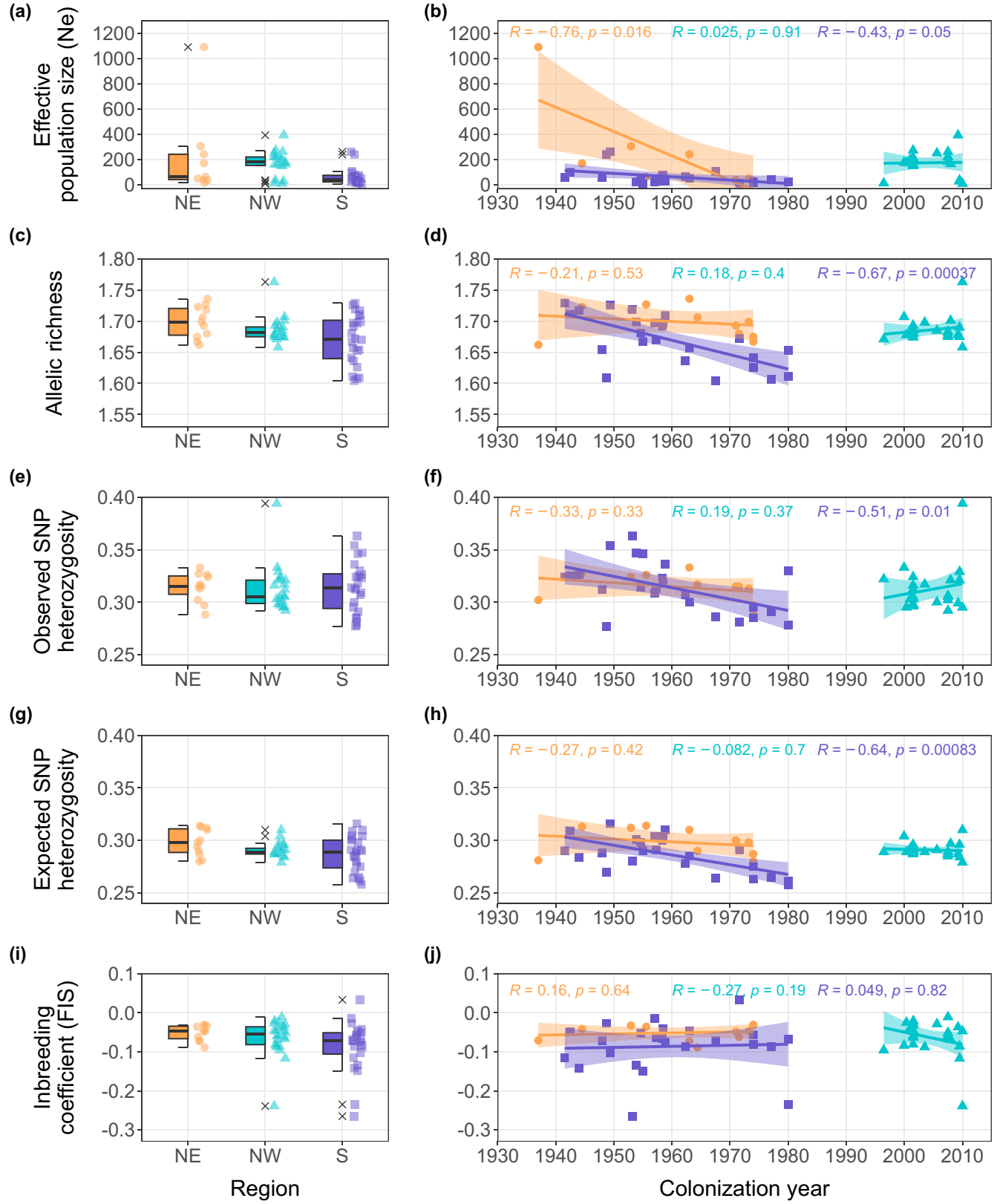

FIGURE 3 Inter-region comparisons of within-locality genetic variation at putatively neutral SNPs. (a) Boxplots of effective population sizes at sampling localities within regions. (b) Correlations between effective population sizes and local colonisation year estimates. Subsequent pairs of plots (rows) represent the same analyses as panels (a) and (b) for other genetic variation statistics (indicated along the leftmost *y*-axes). (c, d) Allelic richness. (e, f) Observed single nucleotide polymorphism (SNP) heterozygosity. (g, h) Expected SNP heterozygosity. (i, j) Inbreeding coefficient, F_{IS}. Outliers in the boxplots (marked with the symbol X) were defined as observations beyond 1.5 times the interquartile range below or above the first and third quartiles respectively.

3.4 | Genetic-environment association analyses

We detected 287 SNPs with significant GEAs using GWR. Our implementation of GWR accounted for allele surfing by including locality principal components scores (representing neutral population structure) and local colonisation year in a null model; a locus was identified as having a significant GEA indicative of spatially divergent selection only when the addition of an environmental factor explained additional spatial variation in allele frequencies compared to the null allele surfing model. We also used Bayenv2 to identify 130 SNPs with significant GEAs (hereafter, 'significant SNPs'). Collectively, GWR and Bayenv2 identified 399 significant SNPs, 18 of which overlapped between the two analyses. The observed overlap is significantly greater than expected by chance (permutation test, $p < .001$; Figure S13).

Among the 18 overlapping SNPs, one was associated with annual precipitation, four were associated with elevation and 13 were associated with temperature seasonality. Locality-specific allele frequencies and locality-specific GWR environmental beta coefficients for each of the 18 overlapping significant SNPs are mapped in Figures S14 and S15 respectively. All 18 SNPs had GWR environmental beta coefficients that shared the same sign in the NE and S regions (i.e. the same allele at a SNP increased or decreased in frequency with an increase in the environmental factor in both the NE and S regions). Twelve of the 18 SNPs had GWR environmental beta coefficients that shared the same sign across all three regions. Notably, six SNPs showed a reversal in the sign of the environmental beta coefficient only in the NW region (Figure S15).

After removing SNPs showing strong LD (i.e. removing one SNP from a pair with $r^2 > .1$), we retained 11 of the 18 overlapping significant SNPs to test for differences in the absolute values of GWR environmental beta coefficients (i.e. absolute regression slopes representing the magnitude of change in allele frequency per standard deviation change in an environmental factor; see Section 2). Absolute environmental beta coefficients in the recently invaded NW region were significantly smaller than those of the NE and S regions (NW-NE Wilcoxon signed-rank test $T = 1$, $p = 1.9E\text{-}03$; NW-S $T = 0$, $p = 9.8E\text{-}04$; Figure 4). Absolute environmental beta coefficients in the NE region were also significantly smaller than in the S region (Wilcoxon signed-rank test $T = 18$, $p = 4.9E\text{-}03$; Figure 4).

Small GWR environmental beta coefficients in the NW are unlikely to be statistical artefacts of regression analysis. Environmental variance among localities was comparable across regions (Figure S1a–d). Per-locality genetic diversity (i.e. Ar, observed heterozygosity and expected heterozygosity) at the 18 overlapping significant SNPs was also comparable among regions (Figure S16); this was also true for the collective set of 399 significant SNPs identified by either GWR or Bayenv2 (Figure S16). Thus, regional differences in allele frequency variance are likely not driving variation in the slopes of GEAs (i.e. near-fixation of alleles at significant SNPs in the NW could otherwise lead to near-zero environmental beta coefficients). Smaller SNP-environment beta coefficients in the NW were also not

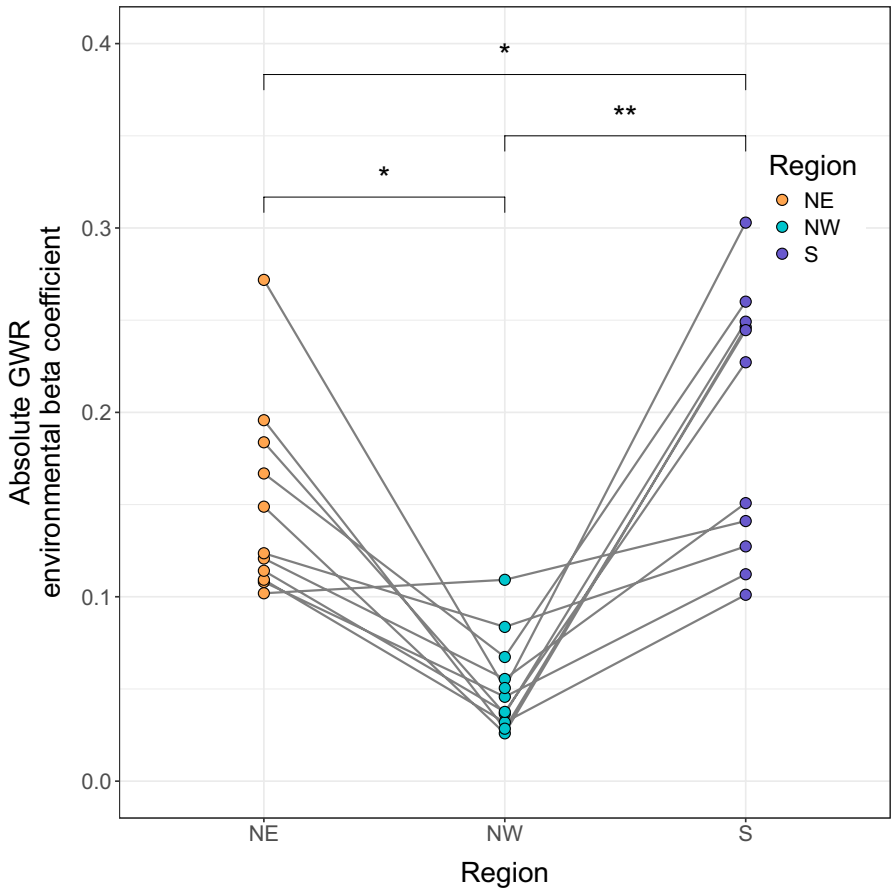

FIGURE 4 Regional variation in the strength of genetic-environment associations detected by geographically weighted regression (GWR) for the 11 unlinked significant SNPs identified by both GWR and Bayenv2. Environmental factors were centred and scaled such that the beta coefficient represents the change in allele frequency per standard deviation change in the environment. Points represent individual SNPs, with lines pairing the same SNP across regions. Significance was evaluated using multiple pairwise Wilcoxon-signed rank tests. All comparisons were significant after Bonferroni correction at familywise $\alpha = .05$. Significance symbols correspond as follows: * for a significant test at $p < .0167$ and ** for $p < .01$.

driven by greater confounding between the environment and colonisation year within the NW; correlations between the four environmental factors and colonisation year were of similar magnitude for the NW and at least one of the other two regions (Figure S4b–d). Indeed, most of the 18 overlapping significant SNPs were associated with temperature seasonality in GWR models, and this environmental factor was most greatly confounded with colonisation year in the NE region, not the NW region (Figure S4b–d).

3.5 | Candidate gene identification

We considered genes nearest to the significant SNPs as being candidate genes contributing to local adaptation. Of the 18 significant SNPs identified by both GWR and Bayenv2, 15 SNPs were near a total of 17 genes (two significant SNPs were located within two genes each). 12 of these 17 genes were orthologues of genes described in other species (Table S4). Among the broader set of 399 significant SNPs identified by either GWR or Bayenv2, 357 SNPs were near a total of 375 genes. 254 of these 375 genes were orthologues of genes described in other species (Table S5). All genes were <170 kb away from their corresponding significant SNPs (<75 kb for the 18 SNPs identified by both GWR and Bayenv2) (Tables S4 and S5). GO terms extracted from Xenbase are provided in Tables S4 and S5. No GO terms pertaining to biological processes, molecular functions or cellular components were significantly over- or under-represented among candidate genes near the narrower set of 18 significant SNPs nor the broader set of 399 significant SNPs. We comment on the plausibility of candidate genes and their potential connections to ecological observations in Section 4.

4 | DISCUSSION

Understanding how interactions between evolutionary forces influence species' adaptive responses to environmental heterogeneity is critical for predicting the geographic extent of range expansions. Serial founder effects can make genetic drift a powerful evolutionary force that can constrain adaptive evolution of species undergoing rapid range expansions or range shifts (Polechová, 2018; Polechová & Barton, 2015). Allele surfing can generate spurious relationships between allele frequencies and environmental variables (i.e. GEAs), which are often interpreted as evidence of divergent selection (Hoban et al., 2016; Lotterhos & Whitlock, 2015; Zhao et al., 2020). Additionally, the evolution of increased dispersal during range expansions can obscure genomic signatures of local adaptation by driving the spatial non-stationarity of GEAs. We overcame these challenges with a novel application of GWR for detecting candidate loci under selection while implementing an appropriate null hypothesis that accounts for allele surfing.

Prior knowledge of colonisation times and dispersal dynamics of the cane toad in Australia allowed us to characterise genetic drift, gene flow and selection among populations occupying three geographic regions with different invasion histories. Despite each geographic region harbouring similar distributions of abiotic environmental conditions, rapidly dispersing toads in the expanding NW region show little evidence for local adaptation to the abiotic environment. That is, GEA slopes are shallower in the NW compared to the more established NE and S regions. Despite higher effective population sizes in the NW and the expectation that high genetic variation may facilitate local adaptation, apparent poor local adaptation in the NW may be explained by a combination of (1) selection having had little time to generate GEAs due to the recency of colonisation and (2) substantially higher gene flow among NW localities constraining local adaptation.

4.1 | Genetic diversity and structure

Recently colonised areas are expected to show more extreme signatures of serial founder effects than more established portions of a species' range due to ongoing gene flow in the latter (Jangjoo et al., 2016). Contrary to expectations, the cane toad exhibits the inverse pattern across its invasive range throughout Australia. The NE and S regions, both of which were colonised >20 generations prior to sampling, show the strongest negative correlations between colonisation year and genetic diversity. In contrast, localities in the NW were colonised <15 generations prior to sampling and show little evidence for sequential declines in genetic diversity. A previous study similarly found evidence for serial loss of genetic variation at microsatellite loci in the S region but not in the NW region (Estoup et al., 2004). Indeed, median per-locality effective population sizes in the NW are higher than in the NE and S regions, although these estimates are imprecise (likely due to poor power to differentiate a signal of genetic drift from sampling effects; Do et al., 2014). Heterozygosity and allelic richness in the NW are also comparable to estimates in the NE and S regions. Negative F_{IS} values in most localities may reflect small local effective population sizes (Balloux, 2004; Kardos et al., 2016; Luikart & Cornuet, 1999).

Relatively low genetic diversity in the S region may reflect the repeated range expansion further southward and subsequent retreat northward driven by cold temperatures (Macgregor et al., 2021), which can increase the effects of genetic drift (Davies et al., 2016). This is consistent with the population tree inferred by Treemix, which shows strong genetic drift and evidence for directional migration events among S localities. Indeed, the southernmost localities near the area of repeated range expansion and retreat were often the sources and recipients of directional migration events. The expected loss of diversity in the recently colonised NW region may have been buffered by high gene flow among localities, as evidenced by low pairwise F_{ST} values and little independent drift inferred by Treemix. Indeed, simulations suggest that substantial gene flow via long-distance dispersal can attenuate loss of genetic diversity during range expansions (Bialozyt et al., 2006; Fayard et al., 2009; Ray & Excoffier, 2010). Additionally, the NW region has high habitat suitability (Trumbo et al., 2016), which may bolster population sizes

and genetic diversity. Although repeated introductions of invasive species can increase genetic diversity (Dlugosch & Parker, 2008), repeated introductions of the cane toad in 1935–1937 all occurred along the Queensland coast (near the NE and S sampling regions; Figure 1) and used offspring of the single Australian founding population (i.e. the first introduction near Gordonvale) (Sabath et al., 1981). Thus, repeated introductions of the cane toad are unlikely to have bolstered genetic diversity in the NW region, which is geographically distant and was invaded approximately 63 years after the final introduction.

Geographic clines in local genetic diversity may also exist in contexts other than range expansions. For example, stable geographic ranges may show such clines when less suitable edge habitats support smaller effective population sizes than core habitats, a prediction posited by the central-marginal hypothesis for evolutionary determinants of species' range limits (Eckert et al., 2008). Therefore, putative serial founder effects identified in the stabilised NE and S regions may instead reflect core–edge patterns characteristic of many species (Eckert et al., 2008; Trumbo et al., 2016). However, the cane toad was initially introduced in an area of high habitat suitability (Trumbo et al., 2016) such that its inland spread into less suitable habitat has generated a negative correlation between colonisation year and habitat suitability. Thus, it is difficult to discriminate whether spatial variation in genetic diversity reflects a persistent signature of a serial founder effect or a stable range core–edge pattern driven by habitat suitability.

4.2 | Spatial variation in genetic-environment associations

Although the propensity for gene flow to constrain local adaptation has long been considered in theoretical work in population genetics (Lenormand, 2002; Slatkin, 1987), few studies have provided empirical evidence of this phenomenon (Fedorka et al., 2012; Kottler et al., 2021; Storfer, 1999; Storfer et al., 1999; Storfer & Sih, 1998). In one of the first empirical uses of GWR as a GEA analysis, we identified geographic variation in the slopes of GEAs at loci putatively contributing to local adaptation. Cane toads in the expanding NW region show significantly shallower GEA slopes than in the stabilised NE and S regions. Shallower GEA slopes in the NW coincide with substantially higher gene flow among localities. Indeed, the magnitude of IBD aligns with the magnitude of GEA slopes across all three regions (i.e. IBD is strongest and GEAs steepest in the S region; IBD is weakest and GEAs shallowest in the NW region; NE region intermediate). These patterns suggest that gene flow may constrain local adaptation to abiotic environmental variation to varying extents across the cane toad's Australian range. Although strong genetic drift can also impede local adaptation (Polechová & Barton, 2015), genetic variation is relatively high in the NW (see above) and thus does not explain the shallower GEAs.

It is also possible that the recency of NW colonisation has contributed to shallow GEA slopes: <15 generations had elapsed between colonisation of NW localities and our 2010–2011 sampling efforts. Accordingly, selection has had little time to generate GEAs and genetic differentiation at putatively adaptive loci in the NW. Local adaptation may occur later despite high gene flow, depending on the relative strengths of each evolutionary force. The reversal in the sign of GEA slopes in the NW relative to the NE and S for some SNPs may also reflect transient local maladaptation driven by recent colonisation dynamics, which selection has not had time to reverse. However, this explanation is speculative as geographic variation in GEAs is typically not evaluated and causes for this 'slope switching' phenomenon have not previously been investigated.

To the extent that high gene flow in the NW region is driven by the dispersal-enhancing phenotype inhabiting that region, evidence of poor local adaptation to the abiotic environment in the NW suggests that this phenotype may have a mean population fitness cost (i.e. increased immigration from populations inhabiting alternative environmental conditions may increase the frequencies of locally maladaptive alleles, resulting in heightened migration load) (Bolnick & Nosil, 2007). Thus, the evolution of increased dispersal capacity resulting from spatial sorting ('spatial selection' sensu Phillips et al., 2008) can interfere with spatially divergent 'classical' natural selection driven by the abiotic environment. Additional studies suggest that resource investment in the dispersal-enhancing phenotype of expansion front toads trades off with reproductive investment (Friesen & Shine, 2019; Hudson et al., 2016). Thus, the dispersal-enhancing phenotype may be costly outside of the expansion front, and these trade-offs may explain why longer-established cane toad populations have reduced dispersal capacity (Alford et al., 2009; Gruber et al., 2017a; Hudson et al., 2016; Phillips et al., 2006). Consequently, fitness advantages of reduced dispersal capacity may indirectly facilitate local adaptation to the abiotic environment in established portions of the cane toad's invasive range.

4.3 | Management implications and future directions

Genomic evidence for local adaptation in some parts of the cane toad's Australian range suggests that genetic variation has been maintained at ecologically functional loci. Indeed, we found comparable genetic variation at SNPs with significant GEAs and putatively neutral SNPs. The maintenance of genetic variation at environmentally associated loci has also been supported by an RNA sequencing study of cane toads across their native range, the Hawaiian source population, and a smaller portion of their Australian range (Selechnik et al., 2019). There are currently vast areas of putatively suitable but uncolonised habitat across Australia (Urban et al., 2007), and our results suggest that genetic variation is unlikely to be a limiting factor in the cane toad's capacity for further spread.

Given that poor local adaptation or local maladaptation can reduce the rate of range expansion (Andrade-Restrepo et al., 2019; Gilbert et al., 2017; Szűcs et al., 2017), we expected that further

ecological monitoring of the cane toad since the collection of our samples in 2010 and 2011 would reveal a reduction in invasion speed. Contrary to expectations, as of 2015, the cane toad has continued to progress across NW Australia at approximately 50 km/year (Doody et al., 2018). This invasion speed is similar to the rate estimated for the initial invasion of the lower NW region circa 2000 (Urban et al., 2008). Although the expected reduction in invasion speed has not been observed, it is possible that factors such as putatively reduced landscape resistance to movement in the NW (associated with increased availability of suitable tropical habitat; Urban et al., 2008) may have attenuated this phenomenon. Further genetic monitoring of cane toads in the NW may capture the steepening of GEA slopes over time as divergent selection progressively generates allele frequency differentiation among populations and as the dispersal-enhancing phenotype of the cane toad is replaced by weaker dispersers following initial colonisation of habitat (Phillips et al., 2006).

Knowledge of the genomic basis of local adaptation among cane toad populations may yield important insights into evolutionary changes that may facilitate population establishment and geographic spread (Chown et al., 2015). Among the candidate genes near the 18 significant SNPs identified by both GWR and Bayenv2, we note that the gene *GDPD1* (near a SNP associated with temperature seasonality and annual precipitation) has previously been implicated in local adaptation of chickens to geographic variation in precipitation as well as transcriptional responses to heat and drought stress in several ant and oak species respectively (Kotrade et al., 2019; Perez et al., 2021). In addition to vulnerability to desiccation like many amphibians, cane toad metamorphs experience a trade-off between risk of desiccation and risk of competition or cannibalisation when selecting habitat farther away or closer to ponds occupied by conspecifics respectively (Child, Phillips, Brown, & Shine, 2008; Child, Phillips, & Shine, 2008); in precipitation-limited locations, alleles that improve drought tolerance may improve fitness by enabling metamorphs to spend less time exposed to competitors or cannibals at ponds. Several other candidate genes (i.e. *ZCCHC3*, *OSBPL1A* and *USP45*) are implicated in immune responses (Lian, Wei, et al., 2018; Lian, Zang, et al., 2018; Machuka et al., 2022; Taye et al., 2017), suggesting that signatures of selection related to the abiotic environment identified herein may instead reflect biotic selective pressures; indeed, there is geographic variation in rates of parasitism of Australian cane toads by several pathogen taxa (Freeland et al., 1986; Russo et al., 2021).

Several caveats are worth noting. Experimental study is necessary for functional validation of putatively adaptive loci in the cane toad; as with the majority of landscape genomics studies, the environmental associations we identified should be interpreted with caution because it is unclear whether the environmental factors included herein truly act as selective pressures or simply correlate with unmeasured causal factors (Hoban et al., 2016). Additionally, it is possible that putatively adaptive loci are affected by selection differently in the NW than in the NE and S. For example, genotypic redundancy may result in regional differences in the genomic architecture underlying local adaptation (Hoban et al., 2016; Lotterhos, 2023), or putatively adaptive loci identified herein may experience conditional neutrality in the NW (Anderson et al., 2013; Mee & Yeaman, 2019). Interrogating a larger fraction of the genome in the future will (1) enable tests for selection leveraging information from linked loci (e.g. Abondio et al., 2022; Booker et al., 2024); (2) enable better characterisation of the genomic architecture underlying local adaptation of cane toad populations (Lowry et al., 2017) and (3) help clarify which environmental factors act as selective pressures.

4.4 | Broader implications for evolutionary study of range expansions

Our work adds to the growing documentation of rapid adaptive evolution of invasive species post-introduction (Butin et al., 2005; Colautti & Lau, 2015; Gong et al., 2022; Maron et al., 2004; Prentis et al., 2008; Woods & Sultan, 2022). We additionally provide novel evidence of spatial variation in the outcomes of interactions between drift, gene flow and spatially divergent selection during a biological invasion. Although we found genomic evidence for local adaptation across the established portion of the cane toad's invasive range, it is unclear whether local adaptation facilitated subsequent invasion or simply evolved concurrently. Thus, our work does not explicitly support the 'adaptation hypothesis' of biological invasions, which posits that success of an invasion depends on adaptive evolutionary responses to environmental conditions in the introduced range (Enders et al., 2020). Indeed, the rapid expansion of cane toads in the NW with little signature of local adaptation to the environment suggests that adaptive evolution may have a substantial temporal lag following establishment of populations rather than being a driver of invasion, although counterexamples exist (Szűcs et al., 2017). Continuing to study spatial heterogeneity in the interactions of evolutionary forces may help clarify the extent to which local adaptation facilitates biological invasions. Accordingly, our work suggests that efforts aimed at understanding potential drivers of invasive species should at a minimum survey geographic regions known to differ in invasion dynamics.

Our results may also reflect eco-evolutionary dynamics operating during range expansions in general, outside the context of biological invasions. For example, numerous species are undergoing distributional shifts owing to climate change (Chen et al., 2011; Pacifici et al., 2020; Parmesan & Yohe, 2003). Low genetic diversity is commonly suggested as a constraint to adaptation during range expansions and perhaps a constraint to range expansion itself (Polechová, 2018), but our work suggests that interactions between evolutionary forces may constrain adaptation at a range expansion front even when genetic diversity is not limiting. Indeed, the relative magnitudes of spatially divergent selection and gene flow may strongly impact the extent to which range-expanding species adapt to novel environmental conditions. Accounting for this phenomenon may improve genomics-based predictions of species' capacity to shift ranges, which often consider evolutionary forces in isolation. However, evolutionary dynamics and resulting population genetic patterns of non-invasive species

undergoing climate change-driven distributional shifts may differ from dynamics of invasive species because the former may have greater initial population sizes (Bialozyt et al., 2006), new habitat may become available more incrementally (Li & Park, 2020; Platts et al., 2019), and biotic interactions such as competition may be limiting (Legault et al., 2020). In contrast, invasive species may be introduced to vast areas of resource-rich, suitable habitat and may be comparatively unhindered by biotic factors (Enders et al., 2020). Thus, more work is necessary to understand whether genetic patterns among cane toad populations are typical of invasive species and range-expanding species more generally.

AUTHOR CONTRIBUTIONS

M.A.B. conducted all statistical analyses, prepared figures, and wrote the manuscript with contributions from all authors. D.R.T. collected field data and conducted all laboratory work. C.P.K. and B.E. performed bioinformatic work. P.A.H. directed laboratory work. R.A.A. and L.S. directed fieldwork. A.S. directed the project.

ACKNOWLEDGEMENTS

We thank Savannah L. Bartel, Dale Clement and Baily McCulloch for constructive feedback during project development. Funding was provided by the National Science Foundation (NSF) Doctoral Dissertation Improvement Grant (DDIG) Award Number 1407335 to A.S. and D.R.T. and the National Science Foundation Graduate Research Fellowship Program under Award 1842493 to M.A.B.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

Raw sequencing data underlying this article are publicly available at the National Center for Biotechnology Information (NCBI) under BioProject PRJNA328156. Data analysis scripts and files associated with this article are publicly available in a GitHub repository (https://github.com/marcabeer/canetoad_landscape_genomics).

ORCID

Marc A. Beer https://orcid.org/0000-0003-1374-0705
Daryl R. Trumbo https://orcid.org/0000-0002-8438-2856
Christopher P. Kozakiewicz https://orcid.org/0000-0002-4868-9252
Brendan Epstein https://orcid.org/0000-0001-7083-1588
Paul A. Hohenlohe https://orcid.org/0000-0002-7616-0161

REFERENCES

Abondio, P., Cilli, E., & Luiselli, D. (2022). Inferring signatures of positive selection in whole-genome sequencing data: An overview of haplotype-based methods. *Genes*, *13*(5), 926. https://doi.org/10.3390/genes13050926

Alford, R. A., Brown, G. P., Schwarzkopf, L., Phillips, B. L., & Shine, R. (2009). Comparisons through time and space suggest rapid evolution of dispersal behaviour in an invasive species. *Wildlife Research*, *36*(1), 23. https://doi.org/10.1071/WR08021

Alleaume-Benharira, M., Pen, I. R., & Ronce, O. (2006). Geographical patterns of adaptation within a species' range: Interactions between drift and gene flow. *Journal of Evolutionary Biology*, *19*(1), 203–215. https://doi.org/10.1111/j.1420-9101.2005.00976.x

Anderson, J. T., Lee, C.-R., Rushworth, C. A., Colautti, R. I., & Mitchell-Olds, T. (2013). Genetic trade-offs and conditional neutrality contribute to local adaptation. *Molecular Ecology*, *22*, 699–708.

Andrade-Restrepo, M., Champagnat, N., & Ferrière, R. (2019). Local adaptation, dispersal evolution, and the spatial eco-evolutionary dynamics of invasion. *Ecology Letters*, *22*(5), 767–777. https://doi.org/10.1111/ele.13234

Antonovics, J. (1976). The nature of limits to natural selection. *Annals of the Missouri Botanical Garden*, *63*(2), 224. https://doi.org/10.2307/2395303

Balloux, F. (2004). Heterozygote excess in small populations and the heterozygote-excess effective population size. *Evolution*, *58*(9), 1891–1900. https://doi.org/10.1111/j.0014-3820.2004.tb00477.x

Barton, N. H., & Hewitt, G. M. (1985). Analysis of hybrid zones. *Annual Review of Ecology and Systematics*, *16*, 113–148.

Beer, M. A., Kane, R. A., Micheletti, S. J., Kozakiewicz, C. P., & Storfer, A. (2022). Landscape genomics of the streamside salamander: Implications for species management in the face of environmental change. *Evolutionary Applications*, *15*(2), 220–236. https://doi.org/10.1111/eva.13321

Beer, M. A., Proft, K. M., Veillet, A., Kozakiewicz, C. P., Hamilton, D. G., Hamede, R., McCallum, H., Hohenlohe, P. A., Burridge, C. P., Margres, M. J., Jones, M. E., & Storfer, A. (2024). Disease-driven top predator decline affects mesopredator population genomic structure. *Nature Ecology & Evolution*, *8*(2), 293–303. https://doi.org/10.1038/s41559-023-02265-9

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B: Methodological*, *57*(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Bialozyt, R., Ziegenhagen, B., & Petit, R. J. (2006). Contrasting effects of long distance seed dispersal on genetic diversity during range expansion. *Journal of Evolutionary Biology*, *19*(1), 12–20. https://doi.org/10.1111/j.1420-9101.2005.00995.x

Bock, D. G., Caseys, C., Cousens, R. D., Hahn, M. A., Heredia, S. M., Hübner, S., Turner, K. G., Whitney, K. D., & Rieseberg, L. H. (2015). What we still don't know about invasion genetics. *Molecular Ecology*, *24*(9), 2277–2297. https://doi.org/10.1111/mec.13032

Bolnick, D. I., & Nosil, P. (2007). Natural selection in populations subject to a migration load. *Evolution*, *61*(9), 2229–2243. https://doi.org/10.1111/j.1558-5646.2007.00179.x

Booker, T. R., Yeaman, S., Whiting, J. R., & Whitlock, M. C. (2024). The WZA: A window-based method for characterizing genotype–environment associations. *Molecular Ecology Resources*, *24*(2), e13768. https://doi.org/10.1111/1755-0998.13768

Bradburd, G. S., Coop, G. M., & Ralph, P. L. (2018). Inferring continuous and discrete population genetic structure across space. *Genetics*, *210*, 33–52.

Brunsdon, C., Fotheringham, A. S., & Charlton, M. E. (1996). Geographically weighted regression: A method for exploring spatial nonstationarity. *Geographical Analysis*, *28*(4), 281–298. https://doi.org/10.1111/j.1538-4632.1996.tb00936.x

Butin, E., Porter, A. H., & Elkinton, J. (2005). Adaptation during biological invasions and the case of *Adelges tsugae*. *Evolutionary Ecology Research*, *7*, 887–900.

Caye, K., Jay, F., Michel, O., & François, O. (2018). Fast inference of individual admixture coefficients using geographic data. *The Annals of Applied Statistics*, *12*(1), 586–608. https://doi.org/10.1214/17-AOAS1106

Chen, I.-C., Hill, J. K., Ohlemüller, R., Roy, D. B., & Thomas, C. D. (2011). Rapid range shifts of species associated with high levels of climate

warming. *Science*, *333*(6045), 1024–1026. https://doi.org/10.1126/science.1206432

Child, T., Phillips, B. L., Brown, G. P., & Shine, R. (2008). The spatial ecology of cane toads (*Bufo marinus*) in tropical Australia: Why do metamorph toads stay near the water? *Austral Ecology*, *33*(5), 630–640. https://doi.org/10.1111/j.1442-9993.2007.01829.x

Child, T., Phillips, B. L., & Shine, R. (2008). Abiotic and biotic influences on the dispersal behavior of metamorph cane toads (*Bufo marinus*) in tropical Australia. *Journal of Experimental Zoology Part A: Ecological Genetics and Physiology*, *309A*(4), 215–224. https://doi.org/10.1002/jez.450

Chown, S. L., Hodgins, K. A., Griffin, P. C., Oakeshott, J. G., Byrne, M., & Hoffmann, A. A. (2015). Biological invasions, climate change and genomics. *Evolutionary Applications*, *8*(1), 23–46. https://doi.org/10.1111/eva.12234

Clarke, R. T., Rothery, P., & Raybould, A. F. (2002). Confidence limits for regression relationships between distance matrices: Estimating gene flow with distance. *Journal of Agricultural, Biological, and Environmental Statistics*, *7*(3), 361–372. https://doi.org/10.1198/108571102320

Colautti, R. I., & Lau, J. A. (2015). Contemporary evolution during invasion: Evidence for differentiation, natural selection, and local adaptation. *Molecular Ecology*, *24*(9), 1999–2017. https://doi.org/10.1111/mec.13162

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., Durbin, R., & 1000 Genomes Project Analysis Group. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*(15), 2156–2158. https://doi.org/10.1093/bioinformatics/btr330

Daszak, P., Scott, D. E., Kilpatrick, A. M., Faggioni, C., Gibbons, J. W., & Porter, D. (2005). Amphibian population declines at Savannah River site are linked to climate, not chytridiomycosis. *Ecology*, *86*(12), 3232–3237. https://doi.org/10.1890/05-0598

Davies, I. D., Cary, G. J., Landguth, E. L., Lindenmayer, D. B., & Banks, S. C. (2016). Implications of recurrent disturbance for genetic diversity. *Ecology and Evolution*, *6*(4), 1181–1196. https://doi.org/10.1002/ece3.1948

Davis Rabosky, A. R., Cox, C. L., Rabosky, D. L., Title, P. O., Holmes, I. A., Feldman, A., & McGuire, J. A. (2016). Coral snakes predict the evolution of mimicry across New World snakes. *Nature Communications*, *7*(1), 11484. https://doi.org/10.1038/ncomms11484

Diniz-Filho, J. A. F., Soares, T. N., & de Campos Telles, M. P. (2016). Geographically weighted regression as a generalized wombling to detect barriers to gene flow. *Genetica*, *144*(4), 425–433. https://doi.org/10.1007/s10709-016-9911-4

Dlugosch, K. M., & Parker, I. M. (2008). Founding events in species invasions: Genetic variation, adaptive evolution, and the role of multiple introductions. *Molecular Ecology*, *17*(1), 431–449. https://doi.org/10.1111/j.1365-294X.2007.03538.x

Dmitrieva, E. V. (2015). Influence of the concentration of dissolved oxygen on embryonic development of the common toad (*Bufo bufo*). *Russian Journal of Developmental Biology*, *46*, 368–380.

Do, C., Waples, R. S., Peel, D., Macbeth, G. M., Tillett, B. J., & Ovenden, J. R. (2014). NeEstimator v2: Re-implementation of software for the estimation of contemporary effective population size (N_e) from genetic data. *Molecular Ecology Resources*, *14*(1), 209–214. https://doi.org/10.1111/1755-0998.12157

Doody, J. S., McHenry, C., Letnic, M., Everitt, C., Sawyer, G., & Clulow, S. (2018). Forecasting the spatiotemporal pattern of the cane toad invasion into north-western Australia. *Wildlife Research*, *45*(8), 718. https://doi.org/10.1071/WR18091

Eckert, C. G., Samis, K. E., & Lougheed, S. C. (2008). Genetic variation across species' geographical ranges: The central–marginal hypothesis and beyond. *Molecular Ecology*, *17*(5), 1170–1188. https://doi.org/10.1111/j.1365-294X.2007.03659.x

Edmonds, C. A., Lillie, A. S., & Cavalli-Sforza, L. L. (2004). Mutations arising in the wave front of an expanding population. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(4), 975–979. https://doi.org/10.1073/pnas.0308064100

Edwards, J. L., Lane, M. A., & Nielsen, E. S. (2000). Interoperability of biodiversity databases: Biodiversity information on every desktop. *Science*, *289*(5488), 2312–2314. https://doi.org/10.1126/science.289.5488.2312

Edwards, R. J., Tuipulotu, D. E., Amos, T. G., O'Meally, D., Richardson, M. F., Russell, T. L., Vallinoto, M., Carneiro, M., Ferrand, N., Wilkins, M. R., Sequeira, F., Rollins, L. A., Holmes, E. C., Shine, R., & White, P. A. (2018). Draft genome assembly of the invasive cane toad, *Rhinella marina*. *GigaScience*, *7*(9), 1–13. https://doi.org/10.1093/gigascience/giy095

Enders, M., Havemann, F., Ruland, F., Bernard-Verdier, M., Catford, J. A., Gómez-Aparicio, L., Haider, S., Heger, T., Kueffer, C., Kühn, I., Meyerson, L. A., Musseau, C., Novoa, A., Ricciardi, A., Sagouis, A., Schittko, C., Strayer, D. L., Vilà, M., Essl, F., … Jeschke, J. M. (2020). A conceptual map of invasion biology: Integrating hypotheses into a consensus network. *Global Ecology and Biogeography*, *29*(6), 978–991. https://doi.org/10.1111/geb.13082

Endler, J. A. (1977). *Geographic variation, speciation, and clines* (Vol. *10*). Princeton University Press.

Estoup, A., Beaumont, M., Sennedot, F., Moritz, C., & Cornuet, J.-M. (2004). Genetic analysis of complex demographic scenarios: Spatially expanding populations of the cane toad, *Bufo marinus*. *Evolution*, *58*(9), 2021–2036. https://doi.org/10.1111/j.0014-3820.2004.tb00487.x

Excoffier, L., Foll, M., & Petit, R. J. (2009). Genetic consequences of range expansions. *Annual Review of Ecology, Evolution, and Systematics*, *40*(1), 481–501. https://doi.org/10.1146/annurev.ecolsys.39.110707.173414

Fayard, J., Klein, E. K., & Lefèvre, F. (2009). Long distance dispersal and the fate of a gene from the colonization front. *Journal of Evolutionary Biology*, *22*(11), 2171–2182. https://doi.org/10.1111/j.1420-9101.2009.01832.x

Fedorka, K. M., Winterhalter, W. E., Shaw, K. L., Brogan, W. R., & Mousseau, T. A. (2012). The role of gene flow asymmetry along an environmental gradient in constraining local adaptation and range expansion: Gene flow symmetry and range limits. *Journal of Evolutionary Biology*, *25*(8), 1676–1685. https://doi.org/10.1111/j.1420-9101.2012.02552.x

Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: New 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, *37*(12), 4302–4315. https://doi.org/10.1002/joc.5086

Fitak, R. R. (2021). *OptM*: Estimating the optimal number of migration edges on population trees using *Treemix*. *Biology Methods & Protocols*, *6*(1), bpab017. https://doi.org/10.1093/biomethods/bpab017

Fix, A. G. (1997). Gene frequency clines produced by kin-structured founder effects. *Human Biology*, *69*, 663–673.

Freeland, W., Delvinqueir, B., & Bonnin, B. (1986). Food and parasitism of the cane toad, *Bufo marinus*, in relation to time since colonization. *Wildlife Research*, *13*(3), 489. https://doi.org/10.1071/WR9860489

Frichot, E., & François, O. (2015). LEA: An R package for landscape and ecological association studies. *Methods in Ecology and Evolution*, *6*(8), 925–929. https://doi.org/10.1111/2041-210X.12382

Friesen, C. R., & Shine, R. (2019). At the invasion front, male cane toads (*Rhinella marina*) have smaller testes. *Biology Letters*, *15*(7), 20190339. https://doi.org/10.1098/rsbl.2019.0339

Gilbert, K. J., Sharp, N. P., Angert, A. L., Conte, G. L., Draghi, J. A., Guillaume, F., Hargreaves, A. L., Matthey-Doret, R., & Whitlock, M. C. (2017). Local adaptation interacts with expansion load during range expansion: Maladaptation reduces expansion load. *The American Naturalist*, *189*(4), 368–380. https://doi.org/10.1086/690673

Gollini, I., Lu, B., Charlton, M., Brunsdon, C., & Harris, P. (2015). GWmodel: An *R* package for exploring spatial heterogeneity using geographically weighted models. *Journal of Statistical Software*, *63*(17), 1–50. https://doi.org/10.18637/jss.v063.i17

Gong, W., Wang, Y., Chen, C., Xiong, Y., Zhou, Y., Xiao, F., Li, B., & Wang, Y. (2022). The rapid evolution of an invasive plant due to increased selection pressures throughout its invasive history. *Ecotoxicology and Environmental Safety*, *233*, 113322. https://doi.org/10.1016/j.ecoenv.2022.113322

Gralka, M., Stiewe, F., Farrell, F., Möbius, W., Waclaw, B., & Hallatschek, O. (2016). Allele surfing promotes microbial adaptation from standing variation. *Ecology Letters*, *19*(8), 889–898. https://doi.org/10.1111/ele.12625

Gruber, J., Brown, G., Whiting, M. J., & Shine, R. (2017a). Geographic divergence in dispersal-related behaviour in cane toads from range-front versus range-core populations in Australia. *Behavioral Ecology and Sociobiology*, *71*(2), 38. https://doi.org/10.1007/s00265-017-2266-8

Gruber, J., Brown, G., Whiting, M. J., & Shine, R. (2017b). Is the behavioural divergence between range-core and range-edge populations of cane toads (*Rhinella marina*) due to evolutionary change or developmental plasticity? *Royal Society Open Science*, *4*(10), 170789. https://doi.org/10.1098/rsos.170789

Günther, T., & Coop, G. (2013). Robust identification of local adaptation from allele frequencies. *Genetics*, *195*(1), 205–220. https://doi.org/10.1534/genetics.113.152462

Hargreaves, A. L., & Eckert, C. G. (2014). Evolution of dispersal and mating systems along geographic gradients: Implications for shifting ranges. *Functional Ecology*, *28*(1), 5–21. https://doi.org/10.1111/1365-2435.12170

Hoban, S., Kelley, J. L., Lotterhos, K. E., Antolin, M. F., Bradburd, G., Lowry, D. B., Poss, M. L., Reed, L. K., Storfer, A., & Whitlock, M. C. (2016). Finding the genomic basis of local adaptation: Pitfalls, practical solutions, and future directions. *The American Naturalist*, *188*(4), 379–397. https://doi.org/10.1086/688018

Hogue, A. S., & Breon, K. (2022). The greatest threats to species. *Conservation Science and Practice*, *4*(5), 1–9. https://doi.org/10.1111/csp2.12670

Hudson, C. M., McCurry, M. R., Lundgren, P., McHenry, C. R., & Shine, R. (2016). Constructing an invasion machine: The rapid evolution of a dispersal-enhancing phenotype during the cane toad invasion of Australia. *PLoS ONE*, *11*(9), e0156950. https://doi.org/10.1371/journal.pone.0156950

Hudson, C. M., Vidal-García, M., Murray, T. G., & Shine, R. (2020). The accelerating anuran: Evolution of locomotor performance in cane toads (*Rhinella marina*, Bufonidae) at an invasion front. *Proceedings of the Royal Society B: Biological Sciences*, *287*(1938), 20201964. https://doi.org/10.1098/rspb.2020.1964

Jangjoo, M., Matter, S. F., Roland, J., & Keyghobadi, N. (2016). Connectivity rescues genetic diversity after a demographic bottleneck in a butterfly population network. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(39), 10914–10919. https://doi.org/10.1073/pnas.1600865113

Jombart, T. (2008). Adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics*, *24*(11), 1403–1405. https://doi.org/10.1093/bioinformatics/btn129

Joost, S., Vuilleumier, S., Jensen, J. D., Schoville, S., Leempoel, K., Stucki, S., Widmer, I., Melodelima, C., Rolland, J., & Manel, S. (2013). Uncovering the genetic basis of adaptive change: On the intersection of landscape genomics and theoretical population genetics. *Molecular Ecology*, *22*(14), 3659–3665. https://doi.org/10.1111/mec.12352

Kardos, M., Taylor, H. R., Ellegren, H., Luikart, G., & Allendorf, F. W. (2016). Genomics advances the study of inbreeding depression in the wild. *Evolutionary Applications*, *9*(10), 1205–1218. https://doi.org/10.1111/eva.12414

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, *90*(430), 773–795.

Keenan, K., McGinnity, P., Cross, T. F., Crozier, W. W., & Prodöhl, P. A. (2013). diveRsity: An R package for the estimation and exploration of population genetics parameters and their associated errors. *Methods in Ecology and Evolution*, *4*(8), 782–788. https://doi.org/10.1111/2041-210X.12067

Klopfstein, S., Currat, M., & Excoffier, L. (2006). The fate of mutations surfing on the wave of a range expansion. *Molecular Biology and Evolution*, *23*(3), 482–490. https://doi.org/10.1093/molbev/msj057

Kosmala, G. K., Brown, G. P., Christian, K. A., Hudson, C. M., & Shine, R. (2018). The thermal dependency of locomotor performance evolves rapidly within an invasive species. *Ecology and Evolution*, *8*(9), 4403–4408. https://doi.org/10.1002/ece3.3996

Kotrade, P., Sehr, E. M., Wischnitzki, E., & Brüggemann, W. (2019). Comparative transcriptomics-based selection of suitable reference genes for normalization of RT-qPCR experiments in drought-stressed leaves of three European Quercus species. *Tree Genetics & Genomes*, *15*(3), 38. https://doi.org/10.1007/s11295-019-1347-4

Kottler, E. J., Dickman, E. E., Sexton, J. P., Emery, N. C., & Franks, S. J. (2021). Draining the swamping hypothesis: Little evidence that gene flow reduces fitness at range edges. *Trends in Ecology & Evolution*, *36*(6), 533–544. https://doi.org/10.1016/j.tree.2021.02.004

Leclerc, C., Courchamp, F., & Bellard, C. (2018). Insular threat associations within taxa worldwide. *Scientific Reports*, *8*(1), 6393. https://doi.org/10.1038/s41598-018-24733-0

Legault, G., Bitters, M. E., Hastings, A., & Melbourne, B. A. (2020). Interspecific competition slows range expansion and shapes range boundaries. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(43), 26854–26860. https://doi.org/10.1073/pnas.2009701117

Lenormand, T. (2002). Gene flow and the limits to natural selection. *Trends in Ecology & Evolution*, *17*(4), 183–189. https://doi.org/10.1016/S0169-5347(02)02497-7

Li, F., & Park, Y. (2020). Habitat availability and environmental preference drive species range shifts in concordance with climate change. *Diversity and Distributions*, *26*(10), 1343–1356. https://doi.org/10.1111/ddi.13126

Lian, H., Wei, J., Zang, R., Ye, W., Yang, Q., Zhang, X.-N., Chen, Y.-D., Fu, Y.-Z., Hu, M.-M., Lei, C.-Q., Luo, W.-W., Li, S., & Shu, H.-B. (2018). ZCCHC3 is a co-sensor of cGAS for dsDNA recognition in innate immune response. *Nature Communications*, *9*(1), 3349. https://doi.org/10.1038/s41467-018-05559-w

Lian, H., Zang, R., Wei, J., Ye, W., Hu, M.-M., Chen, Y.-D., Zhang, X.-N., Guo, Y., Lei, C.-Q., Yang, Q., Luo, W.-W., Li, S., & Shu, H.-B. (2018). The zinc-finger protein ZCCHC3 binds RNA and facilitates viral RNA sensing and activation of the RIG-I-like receptors. *Immunity*, *49*(3), 438–448.e5. https://doi.org/10.1016/j.immuni.2018.08.014

Lotterhos, K. E. (2023). The paradox of adaptive trait clines with nonclinal patterns in the underlying genes. *Proceedings of the National Academy of Sciences of the United States of America*, *120*(12), e2220313120. https://doi.org/10.1073/pnas.2220313120

Lotterhos, K. E., & Whitlock, M. C. (2015). The relative power of genome scans to detect local adaptation depends on sampling design and statistical method. *Molecular Ecology*, *24*(5), 1031–1046. https://doi.org/10.1111/mec.13100

Lowry, D. B., Hoban, S., Kelley, J. L., Lotterhos, K. E., Reed, L. K., Antolin, M. F., & Storfer, A. (2017). Breaking RAD: An evaluation of the utility of restriction site-associated DNA sequencing for genome scans of adaptation. *Molecular Ecology Resources*, *17*(2), 142–152. https://doi.org/10.1111/1755-0998.12635

Luikart, G., & Cornuet, J.-M. (1999). Estimating the effective number of breeders from heterozygote excess in progeny. *Genetics*, *151*(3), 1211–1216. https://doi.org/10.1093/genetics/151.3.1211

Macgregor, L. F., Greenlees, M., de Bruyn, M., & Shine, R. (2021). An invasion in slow motion: The spread of invasive cane toads (*Rhinella*

marina) into cooler climates in southern Australia. *Biological Invasions*, *23*(11), 3565–3581. https://doi.org/10.1007/s10530-021-02597-2

Machuka, E. M., Juma, J., Muigai, A. W. T., Amimo, J. O., Pelle, R., & Abworo, E. O. (2022). Transcriptome profile of spleen tissues from locally-adapted Kenyan pigs (*Sus scrofa*) experimentally infected with three varying doses of a highly virulent African swine fever virus genotype IX isolate: Ken12/busia.1 (ken-1033). *BMC Genomics*, *23*(1), 522. https://doi.org/10.1186/s12864-022-08754-8

Manel, S., Joost, S., Epperson, B. K., Holderegger, R., Storfer, A., Rosenberg, M. S., Scribner, K. T., Bonin, A., & Fortin, M.-J. (2010). Perspectives on the use of landscape genetics to detect genetic adaptive variation in the field. *Molecular Ecology*, *19*(17), 3760–3772. https://doi.org/10.1111/j.1365-294X.2010.04717.x

Maron, J. L., Vilà, M., Bommarco, R., Elmendorf, S., & Beardsley, P. (2004). Rapid evolution of an invasive plant. *Ecological Monographs*, *74*(2), 261–280. https://doi.org/10.1890/03-4027

Mee, J. A., & Yeaman, S. (2019). Unpacking conditional neutrality: Genomic signatures of selection on conditionally beneficial and conditionally deleterious mutations. *The American Naturalist*, *194*(4), 529–540. https://doi.org/10.1086/702314

Mi, H., Muruganujan, A., Ebert, D., Huang, X., & Thomas, P. D. (2019). PANTHER version 14: More genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. *Nucleic Acids Research*, *47*(D1), D419–D426. https://doi.org/10.1093/nar/gky1038

Miller, T. E. X., Angert, A. L., Brown, C. D., Lee-Yaw, J. A., Lewis, M., Lutscher, F., Marculis, N. G., Melbourne, B. A., Shaw, A. K., Szűcs, M., Tabares, O., Usui, T., Weiss-Lehman, C., & Williams, J. L. (2020). Eco-evolutionary dynamics of range expansion. *Ecology*, *101*(10), e03139. https://doi.org/10.1002/ecy.3139

Molnar, J. L., Gamboa, R. L., Revenga, C., & Spalding, M. D. (2008). Assessing the global threat of invasive species to marine biodiversity. *Frontiers in Ecology and the Environment*, *6*(9), 485–492. https://doi.org/10.1890/070064

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P. R., O'Hara, R. B., Simpson, G. L., Solymos, P., Stevens, M. H. H., Szoecs, E., & Wagner, H. (2020). *Vegan: Community ecology package*. R Package Version 2.5-7.

O'Leary, S. J., Puritz, J. B., Willis, S. C., Hollenbeck, C. M., & Portnoy, D. S. (2018). These aren't the loci you'e looking for: Principles of effective SNP filtering for molecular ecologists. *Molecular Ecology*, *27*(16), 3193–3206. https://doi.org/10.1111/mec.14792

Pacifici, M., Rondinini, C., Rhodes, J. R., Burbidge, A. A., Cristiano, A., Watson, J. E. M., Woinarski, J. C. Z., & Di Marco, M. (2020). Global correlates of range contractions and expansions in terrestrial mammals. *Nature Communications*, *11*(1), 2840. https://doi.org/10.1038/s41467-020-16684-w

Paris, J. R., Stevens, J. R., & Catchen, J. M. (2017). Lost in parameter space: A road map for STACKS. *Methods in Ecology and Evolution*, *8*(10), 1360–1373. https://doi.org/10.1111/2041-210X.12775

Parmesan, C., & Yohe, G. (2003). A globally coherent fingerprint of climate change impacts across natural systems. *Nature*, *421*(6918), 37–42. https://doi.org/10.1038/nature01286

Perez, R., Souza Araujo, N., Defrance, M., & Aron, S. (2021). Molecular adaptations to heat stress in the thermophilic ant genus *Cataglyphis*. *Molecular Ecology*, *30*(21), 5503–5516. https://doi.org/10.1111/mec.16134

Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S., & Hoekstra, H. E. (2012). Double digest RADseq: An inexpensive method for De novo SNP discovery and genotyping in model and non-model species. *PLoS ONE*, *7*(5), e37135. https://doi.org/10.1371/journal.pone.0037135

Phillips, B. L., Brown, G. P., & Shine, R. (2010). Evolutionarily accelerated invasions: The rate of dispersal evolves upwards during the range advance of cane toads: Dispersal evolution during range advance. *Journal of Evolutionary Biology*, *23*(12), 2595–2601. https://doi.org/10.1111/j.1420-9101.2010.02118.x

Phillips, B. L., Brown, G. P., Travis, J. M. J., & Shine, R. (2008). Reid's paradox revisited: The evolution of dispersal kernels during range expansion. *The American Naturalist*, *172*(S1), S34–S48. https://doi.org/10.1086/588255

Phillips, B. L., Brown, G. P., Webb, J. K., & Shine, R. (2006). Invasion and the evolution of speed in toads. *Nature*, *439*(7078), 803. https://doi.org/10.1038/439803a

Phillips, B. L., & Shine, R. (2004). Adapting to an invasive species: Toxic cane toads induce morphological change in Australian snakes. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(49), 17150–17155. https://doi.org/10.1073/pnas.0406440101

Pickrell, J. K., & Pritchard, J. K. (2012). Inference of population splits and mixtures from genome-wide allele frequency data. *PLoS Genetics*, *8*(11), e1002967. https://doi.org/10.1371/journal.pgen.1002967

Platts, P. J., Mason, S. C., Palmer, G., Hill, J. K., Oliver, T. H., Powney, G. D., Fox, R., & Thomas, C. D. (2019). Habitat availability explains variation in climate-driven range shifts across multiple taxonomic groups. *Scientific Reports*, *9*(1), 15039. https://doi.org/10.1038/s41598-019-51582-2

Polechová, J. (2018). Is the sky the limit? On the expansion threshold of a species' range. *PLoS Biology*, *16*(6), e2005372. https://doi.org/10.1371/journal.pbio.2005372

Polechová, J., & Barton, N. H. (2015). Limits to adaptation along environmental gradients. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(20), 6401–6406. https://doi.org/10.1073/pnas.1421515112

Prentis, P. J., Wilson, J. R. U., Dormontt, E. E., Richardson, D. M., & Lowe, A. J. (2008). Adaptive evolution in invasive species. *Trends in Plant Science*, *13*(6), 288–294. https://doi.org/10.1016/j.tplants.2008.03.004

Quinlan, A. R., & Hall, I. M. (2010). BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics*, *26*(6), 841–842. https://doi.org/10.1093/bioinformatics/btq033

Ray, N., & Excoffier, L. (2010). A first step towards inferring levels of long-distance dispersal during past expansions. *Molecular Ecology Resources*, *10*(5), 902–914. https://doi.org/10.1111/j.1755-0998.2010.02881.x

Rellstab, C., Gugerli, F., Eckert, A. J., Hancock, A. M., & Holderegger, R. (2015). A practical guide to environmental association analysis in landscape genomics. *Molecular Ecology*, *24*(17), 4348–4370. https://doi.org/10.1111/mec.13322

Rochette, N. C., Rivera-Colón, A. G., & Catchen, J. M. (2019). Stacks 2: Analytical methods for paired-end sequencing improve RADseq-based population genomics. *Molecular Ecology*, *28*(21), 4737–4754. https://doi.org/10.1111/mec.15253

Rousset, F. (1997). Genetic differentiation and estimation of gene flow from *F*-statistics under isolation by distance. *Genetics*, *145*(4), 1219–1228. https://doi.org/10.1093/genetics/145.4.1219

Russo, A. G., Harding, E. F., Yan, G. J. H., Selechnik, D., Ducatez, S., DeVore, J. L., Zhou, J., Sarma, R. R., Lee, Y. P., Richardson, M. F., Shine, R., Rollins, L. A., & White, P. A. (2021). Discovery of novel viruses associated with the invasive cane toad (*Rhinella marina*) in its native and introduced ranges. *Frontiers in Microbiology*, *12*, 733631. https://doi.org/10.3389/fmicb.2021.733631

Sabath, M. D., Boughton, W. C., & Easteal, S. (1981). Expansion of the range of the introduced toad *Bufo marinus* in Australia from 1935 to 1974. *Copeia*, *1981*(3), 676. https://doi.org/10.2307/1444573

Selechnik, D., Richardson, M. F., Shine, R., DeVore, J. L., Ducatez, S., & Rollins, L. A. (2019). Increased adaptive variation despite reduced overall genetic diversity in a rapidly adapting invader. *Frontiers in Genetics*, *10*, 1221. https://doi.org/10.3389/fgene.2019.01221

Selmoni, O., Vajana, E., Guillaume, A., Rochat, E., & Joost, S. (2020). Sampling strategy optimization to increase statistical power in

landscape genomics: A simulation-based approach. *Molecular Ecology Resources*, *20*(1), 154–169. https://doi.org/10.1111/1755-0998.13095

Shine, R., Brown, G. P., & Phillips, B. L. (2011). An evolutionary process that assembles phenotypes through space rather than through time. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(14), 5708–5711. https://doi.org/10.1073/pnas.1018989108

Slatkin, M. (1985). Gene flow in natural populations. *Annual Review of Ecology and Systematics*, *16*, 393–430.

Slatkin, M. (1987). Gene flow and the geographic structure of natural populations. *Science*, *236*, 787–792. https://doi.org/10.1126/science.3576198

Slatkin, M., & Excoffier, L. (2012). Serial founder effects during range expansion: A spatial analog of genetic drift. *Genetics*, *191*(1), 171–181. https://doi.org/10.1534/genetics.112.139022

Snyder, G. K., & Weathers, W. W. (1975). Temperature adaptations in amphibians. *The American Naturalist*, *109*(965), 93–101. https://doi.org/10.1086/282976

Storfer, A. (1999). Gene flow and endangered species translocations: A topic revisited. *Biological Conservation*, *8*, 173–180.

Storfer, A., Cross, J., Rush, V., & Caruso, J. (1999). Adaptive coloration and gene flow as a constraint to local adaptation in the streamside salamander, *Ambystoma barbouri*. *Evolution*, *53*(3), 889–898. https://doi.org/10.1111/j.1558-5646.1999.tb05383.x

Storfer, A., Murphy, M. A., Evans, J. S., Goldberg, C. S., Robinson, S., Spear, S. F., Dezzani, R., Delmelle, E., Vierling, L., & Waits, L. P. (2007). Putting the 'landscape' in landscape genetics. *Heredity*, *98*(3), 128–142. https://doi.org/10.1038/sj.hdy.6800917

Storfer, A., & Sih, A. (1998). Gene flow and ineffective antipredator behavior in a stream-breeding salamander. *Evolution*, *52*(2), 558–565. https://doi.org/10.1111/j.1558-5646.1998.tb01654.x

Szűcs, M., Vahsen, M. L., Melbourne, B. A., Hoover, C., Weiss-Lehman, C., & Hufbauer, R. A. (2017). Rapid adaptive evolution in novel environments acts as an architect of population range expansion. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(51), 13501–13506. https://doi.org/10.1073/pnas.1712934114

Taye, B., Yeo, D., Lee, R., Tan, B., Sugrue, R., & Maurer-Stroh, S. (2017). Inter-species host gene expression differences in response to human and avian influenza a virus strains. *International Journal of Molecular Sciences*, *18*(11), 2295. https://doi.org/10.3390/ijms18112295

The Gene Ontology Consortium. (2019). The gene ontology resource: 20 years and still GOing strong. *Nucleic Acids Research*, *47*(D1), D330–D338. https://doi.org/10.1093/nar/gky1055

Travis, J. M. J., & Dytham, C. (2002). Dispersal evolution during invasions. *Evolutionary Ecology Research*, *4*, 1119–1129.

Travis, J. M. J., Munkemuller, T., Burton, O. J., Best, A., Dytham, C., & Johst, K. (2007). Deleterious mutations can surf to high densities on the wave front of an expanding population. *Molecular Biology and Evolution*, *24*(10), 2334–2343. https://doi.org/10.1093/molbev/msm167

Trumbo, D. R., Epstein, B., Hohenlohe, P. A., Alford, R. A., Schwarzkopf, L., & Storfer, A. (2016). Mixed population genomics support for the central marginal hypothesis across the invasive range of the cane toad (*Rhinella marina*) in Australia. *Molecular Ecology*, *25*(17), 4161–4176. https://doi.org/10.1111/mec.13754

Urban, M. C., Phillips, B. L., Skelly, D. K., & Shine, R. (2007). The cane toad's (*Chaunus* [*Bufo*] *marinus*) increasing ability to invade Australia is revealed by a dynamically updated range model. *Proceedings of the Royal Society B: Biological Sciences*, *274*(1616), 1413–1419. https://doi.org/10.1098/rspb.2007.0114

Urban, M. C., Phillips, B. L., Skelly, D. K., & Shine, R. (2008). A toad more traveled: The heterogeneous invasion dynamics of cane toads in Australia. *The American Naturalist*, *171*(3), E134–E148. https://doi.org/10.1086/527494

Wagenmakers, E.-J. (2007). A practical solution to the pervasive problems of p values. *Psychonomic Bulletin & Review*, *14*(5), 779–804. https://doi.org/10.3758/BF03194105

Weir, B. S., & Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution*, *38*(6), 1358–1370.

Woods, E. C., & Sultan, S. E. (2022). Post-introduction evolution of a rapid life-history strategy in a newly invasive plant. *Ecology*, *103*(11), e3803. https://doi.org/10.1002/ecy.3803

Wu, T. D., & Nacu, S. (2010). Fast and SNP-tolerant detection of complex variants and splicing in short reads. *Bioinformatics*, *26*(7), 873–881. https://doi.org/10.1093/bioinformatics/btq057

Zhao, W., Sun, Y., Pan, J., Sullivan, A. R., Arnold, M. L., Mao, J., & Wang, X. (2020). Effects of landscapes and range expansion on population structure and local adaptation. *New Phytologist*, *228*(1), 330–343. https://doi.org/10.1111/nph.16619

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Beer, M. A., Trumbo, D. R., Rautsaw, R. M., Kozakiewicz, C. P., Epstein, B., Hohenlohe, P. A., Alford, R. A., Schwarzkopf, L., & Storfer, A. (2024). Spatial variation in genomic signatures of local adaptation during the cane toad invasion of Australia. *Molecular Ecology*, *00*, e17464. https://doi.org/10.1111/mec.17464