



OPEN

Putative climate adaptation in American pikas (*Ochotona princeps*) is associated with copy number variation across environmental gradients

Bryson M. F. Sjodin¹, Danielle A. Schmidt¹, Kurt E. Galbreath² & Michael A. Russello¹✉

Improved understanding of the genetic basis of adaptation to climate change is necessary for maintaining global biodiversity moving forward. Studies to date have largely focused on sequence variation, yet there is growing evidence that suggests that changes in genome structure may be an even more significant source of adaptive potential. The American pika (*Ochotona princeps*) is an alpine specialist that shows some evidence of adaptation to climate along elevational gradients, but previous work has been limited to single nucleotide polymorphism based analyses within a fraction of the species range. Here, we investigated the role of copy number variation underlying patterns of local adaptation in the American pika using genome-wide data previously collected across the entire species range. We identified 37–193 putative copy number variants (CNVs) associated with environmental variation (temperature, precipitation, solar radiation) within each of the six major American pika lineages, with patterns of divergence largely following elevational and latitudinal gradients. Genes associated ($n = 158$) with independent annotations across lineages, variables, and/or CNVs had functions related to mitochondrial structure/function, immune response, hypoxia, olfaction, and DNA repair. Some of these genes have been previously linked to putative high elevation and/or climate adaptation in other species, suggesting they may serve as important targets in future studies.

Keywords Copy number variant, Local adaptation, Climate, RADseq, Latitudinal and elevational gradients

The complex interplay between environmental factors and local adaptation plays an essential role in the generation and maintenance of biodiversity^{1–4}. Typically, genetic studies on adaptation have focused on DNA sequences (single nucleotide polymorphisms; SNPs), given that point mutations were thought to be the predominant source of selectable variation^{5,6}; however, genomes can also vary in their physical structure across species, populations, and even individuals^{7–12}. These physical variations, collectively known as structural variants, include insertions or deletions of single or large numbers of nucleotides, duplications of genes or entire regions of the genome, inversions or changes in polarity of chromosomes, and translocations both within and among chromosomes. Structural variants can be significant factors in shaping species divergence and local adaptation^{10,12–16}. Early comparative genomics research determined that chromosomal inversions were linked to speciation in *Drosophila*⁸. More recent work is focused on the eco-evolutionary impact of structural variation within natural populations^{12,16}. For instance, Arostegui et al.¹⁷ found that a chromosomal inversion is likely responsible for ecotype differentiation in rainbow trout. Likewise, Cayuela et al.¹³ found significant genotype-environment associations (GEA) among sympatric capelin lineages in the North Atlantic Ocean linked to chromosomal rearrangements and hypothesized that they were associated with adaptation to environmental conditions at spawning sites. These examples and others demonstrate how investigations of structural variants can provide important and novel insights into local adaptation.

Copy number variants (CNVs) span different classes of structural variants (e.g., insertions/deletion, duplications, transposable elements) that vary in the number of copies among individuals¹⁸. Previous studies have

¹Department of Biology, The University of British Columbia, 3247 University Way, Kelowna, BC V1V 1V7, Canada. ²Department of Biology, Northern Michigan University, 1401 Presque Isle Ave, Marquette, MI 49855, USA. ✉email: michael.russello@ubc.ca

revealed that signal at CNVs can differ relative to SNPs, providing important insights into population differentiation and evolutionary history, including the genetic basis of adaptation (reviewed in Mérot et al.¹⁰). Importantly, recent studies have effectively detected and analyzed CNVs from reduced representation genome sequencing (e.g., restriction-site associated DNA sequencing; RADseq¹⁹), providing a cost-effective approach for investigating the role of structural variation in a range of non-model organisms, including American lobster²⁰, capelin²¹, and Columbia spotted frog²². Moreover, these analytical approaches can now be retrospectively applied to the wealth of publicly available RADseq data to investigate the potential role of CNVs underlying patterns of ecological and evolutionary relevance²³.

The American pika (*Ochotona princeps*) is a small lagomorph distributed across a large latitudinal gradient in western North America, from New Mexico (USA) to central British Columbia (Canada)^{24,25}. Pikas are alpine specialists typically found at elevations > 2000 m, although their full distribution spans an elevational gradient from 0 to 4000 m^{25–27}. These environmental gradients serve as an excellent system for investigating local adaptation, as variations in genomic architecture can be directly correlated to differences in environmental variables including temperature, precipitation, and solar radiation^{1,28}. Furthermore, American pikas can also be separated into six geographically isolated evolutionary lineages²⁹, allowing for investigations of potentially independent and parallel histories of adaptation within the same system. While previous studies have found genetic evidence for local adaptation in the American pika on regional scales^{30–32} and at the whole genome level^{33,34}, range-wide adaptation has not been explored. Moreover, these previous studies focused solely on sequence variation; investigation of structural variation may provide a novel, yet complementary, perspective on local adaptation in this species.

To complement past and on-going studies of sequence variation, we investigated the role of CNVs underlying patterns of local adaptation in the American pika using RADseq data previously collected for 36 sites across the species range²⁹. We used a combination of partial redundancy analysis and linear mixed-effect modelling to identify loci putatively associated with temperature, precipitation, and solar radiation within and across all six major American pika lineages. We subsequently examined the spatial distribution of putatively adaptive variation and population differentiation across range-wide latitudinal and elevational gradients. Finally, we annotated putatively adaptive variants and identified target genes with potential functional impacts on local adaptation.

Methods

Data and study area

We used previously generated RADseq data²⁹ collected from 348 individuals sampled from 36 localities across the entire American pika distribution spanning all six major lineages: (NRM) Northern Rocky Mountains; (CRM) Central Rocky Mountains; (SRM) Southern Rocky Mountains; (CSC) Cascades; (SN) Sierra Nevada; and (CU) Central Utah (Fig. 1; Table S1); a subset (n = 173) of these samples were used in Galbreath et al.³⁵. We removed site 18 from the NRM lineage as this site displayed significant admixture between the NRM and CRM lineages²⁹.

Initial SNP calling and filtering

Raw sequencing data from all 11 RADseq libraries were de-multiplexed using the *process_radtags* module of STACKS v2.57³⁶; during this stage, reads were trimmed to 94 bp to remove low-quality bases at the end of the read. De-multiplexed sequences were then aligned to the American pika reference genome assembly (OchPri4.0; GenBank accession ID: GCA_014633375.1³³) using BWA-mem v2.2.1³⁷ under default parameters. The aligned sequence data was then processed using the *gstacks* module of STACKS v2.57³⁶ to generate a catalog of RADtags for downstream analyses.

Due to low range-wide heterozygosity^{38,39} that interfered with downstream copy number variant (CNV) detection (see Discussion for further explanation), we called SNPs independently within each lineage from the aligned sequence data using STACKS v2.57³⁶ retaining polymorphic, autosomal SNPs present in $\geq 70\%$ of all individuals within that lineage. This design also allowed us to assess patterns of parallel evolution by considering each lineage as a separate pseudoreplicate. We further filtered loci to only retain genotypes with a genotyping quality ≥ 20 using VCFtools v0.1.15⁴⁰ and only retained loci present in ≥ 2 individuals within each sample site. Following Dorant et al.²⁰, we performed a final filtering step to retain loci which genotyped in $\geq 70\%$ of individuals within a sample site (allowing for 10% of populations to fail this threshold; maximum of 2 populations per lineage) and had a minimum minor allele sample (i.e., the minimum number of samples possessing the minor/rare allele) of 5% of individuals within a lineage (rounded up to the nearest individual) across all individuals within a lineage using the *05_filter_vcf_fast.py* script downloaded from https://github.com/enormandeu/stacks_workflow (see Table S1 for a summary of filtering values). Also during this step, individual genotypes at a locus with a total read depth < 4 (i.e., minimum allele depth) were reclassified as missing data following Dorant et al.²⁰.

Putative CNV identification

We used the *HDplot* method to detect duplicated SNPs representing putative CNVs from the above dataset following methods initially described by McKinney et al.⁴¹ and modified by Dorant et al.²⁰. Putatively duplicated loci were identified using a combination of four parameters calculated for each locus: proportion of heterozygotes; inbreeding coefficient; median allele ratio for heterozygotes; and proportion of rare homozygotes. Each parameter was calculated using the *08_extract_snp_duplication_info.py* script downloaded from https://github.com/enormandeu/stacks_workflow. We plotted the four parameters against each other and used graphical cut-offs to categorize loci based on their position in the distribution using a modified *09_classify_snps.R* script downloaded from https://github.com/enormandeu/stacks_workflow. Putative CNVs were those loci categorized as duplicated, highly diverged, or had a high depth of coverage following Dorant et al.²⁰ and Cayuela et al.²¹ (see Table S2 for summary of graphical cut-offs). We then extracted read depth for each putative CNV to use as a proxy for copy number^{20,41} using VCFtools v0.1.15⁴⁰. To account for differences in sequencing effort across individuals,

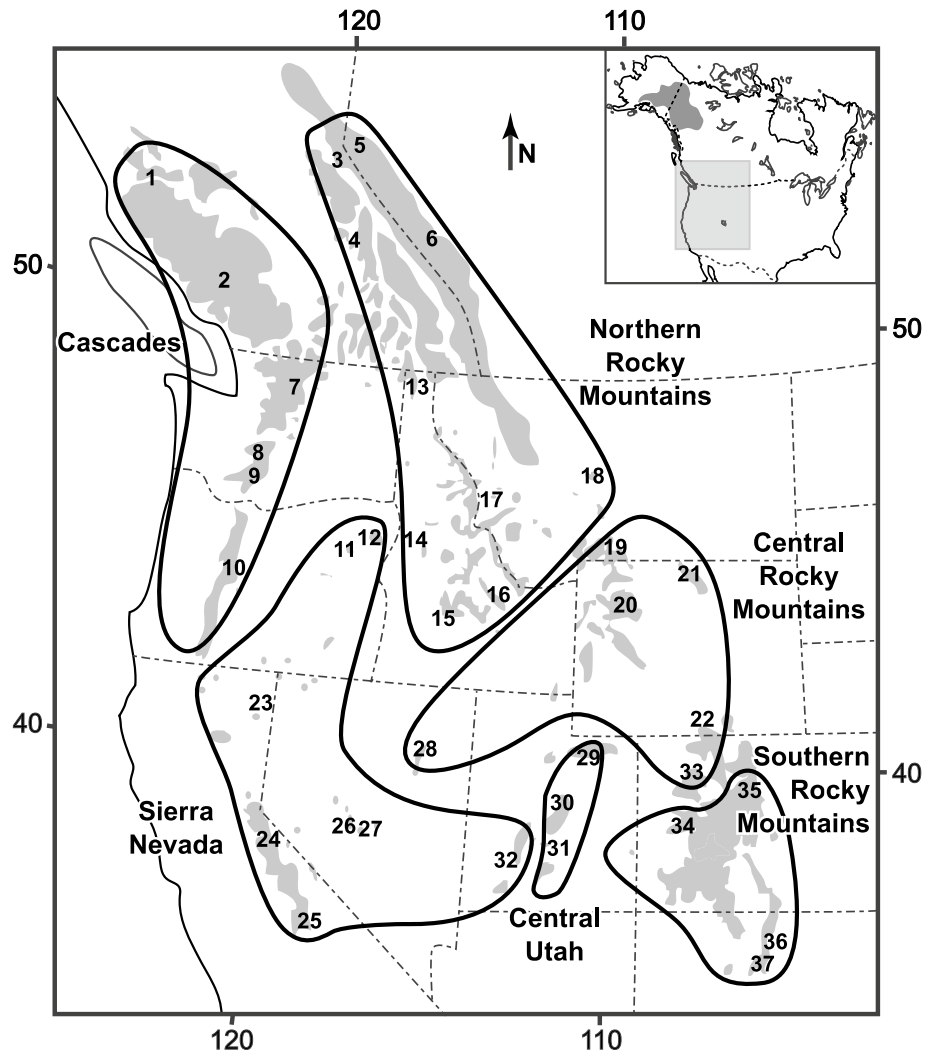


Figure 1. Sampling sites and lineage delineations (thick black lines) for the American pika (*Ochotona princeps*). Shaded regions indicate the approximate American pika distribution as modified from Galbreath et al.³⁵.

we normalized read counts using a trimmed mean of M -values method using the *edgeR* R-package⁴² following Dorant et al.²⁰. Missing genotypes for an individual were imputed as the mean read depth for that locus across individuals within the same sample site.

Genotype-environment association analyses

We performed several genotype-environment association (GEA) analyses to identify CNVs with putative links to climate adaptation. For this, we downloaded climate data for 27 variables from ClimateNA⁴³ (Table S3) and separated them into three climate variable categories: temperature ($n = 16$); precipitation ($n = 6$); and solar radiation ($n = 5$). We separated variables into categories to assess the relative impact of temperature, precipitation, and solar radiation independently and to minimize redundancy across categories due to multicollinearity. To further reduce multicollinearity, we calculated correlation coefficients between each pair of climate variables within each variable category; for each pair with $|r| > 0.70$, we removed the variable with largest mean absolute correlation using the *findCorrelation* function as part of the *caret* R-package⁴⁴.

We first used partial redundancy analysis (pRDA) to detect CNVs with climate associations separately for each climate variable category using the *vegan* R-package⁴⁵. Climate variables were assigned as predictors with normalized read depth as the response. Due to sequencing batch effects detected in the normalized read depth matrices (Figure S1), we included the sequencing library ID for each individual as a covariate. To validate the absence of multicollinearity, we calculated the variance inflation factor (VIF) for each predictor variable and removed those with $VIF > 10$ ⁴⁶. We assessed model significance using global and marginal analyses of variance (ANOVAs) with 1000 permutations and retained all significant ($p \leq 0.05$) axes for downstream analysis. In the event of no significant axes, only the first RDA axis was retained for outlier detection. Climate-associated CNVs (i.e., outlier loci) were then classified as those with a loading > 2.25 standard deviations from the mean loading along each retained RDA axis ($p \leq 0.01$)⁴⁷.

We also used linear mixed-effect models (LMM) to detect climate-CNV associations for each climate variable category using the *lme4* R-package⁴⁸. We used log-transformed normalized read depth as the response variable and the same climate variables as used for the pRDA as fixed effects; we included the sequencing library ID as a random intercept to account for batch effects. We assessed significance of each climate variable using the likelihood ratio test (LRT) as implemented by the *drop1* function in the *lmerTest* R-package⁴⁹. To control for false positives, we corrected *p*-values using the Benjamini–Hochberg false discovery rate method with a corrected significance threshold of $\alpha = 0.05$. As a final step to reduce false positives, we retained only those loci detected as outliers in both the LMM and pRDA and classified these as “robust outliers”. Finally, we re-ran the pRDAs using the methods above with only the robust outliers to estimate the amount of variation explained by the climate variables for these loci.

Genetic differentiation and adaptive divergence

We estimated pairwise genetic differentiation using the variant fixation index V_{ST} ⁵⁰ via a custom R function, which is analogous to the estimator of population differentiation θ^{51} , and is commonly used to identify differentiated CNVs between populations^{20,50}. Using the robust outliers for each climate category, V_{ST} estimates were calculated on a per locus basis and averaged across each pair of sites to obtain mean pairwise estimates. We assessed significance using a bootstrap resampling procedure implemented in the *boot* R-package⁵² for mean pairwise V_{ST} estimates across 10,000 replicates.

We also examined patterns of adaptive divergence within lineages using a hierarchical clustering approach. Using the robust outliers for each climate category, we generated a matrix of pairwise genetic distance by calculating Bray–Curtis distances for each pair of individuals using the *ecodist* R-package⁵³ then calculated the mean distance for each pair of sites. Using this distance matrix as an input, we performed a hierarchical clustering analysis employing the Ward’s minimum variance method⁵⁴ using the *hclust* R-function. The resulting dendrograms were bootstrapped with 10,000 replicates to assess robustness using *boot.phylo* function in the *ape* R-package⁵⁵.

To assess the impact of geography on adaptive divergence, we performed linear regressions on each robust locus with either elevation or latitude as the explanatory variable and normalized read depth as the response. Linear regressions were performed in R using the base packages⁵⁶.

Annotation of putatively adaptive variants

To assess the putative functional implications of climate-associated CNVs, we annotated all robust outliers using the Ensembl Variant Effect Predictor v103.1⁵⁷ and identified CNVs found within protein-coding genes. We then performed a literature search to explore the function of genes that had linked CNVs from across multiple lineages, were linked to multiple CNVs, or were linked to CNVs with significant associations to multiple climate variable categories.

Results

SNP identification and CNV detection

Initial SNP calling resulted in 359,569–1,170,051 loci genotyped per lineage (mean = 784,092 loci) with 60,383–150,978 loci remaining post-filtering (mean = 94,428 loci; Table S1). From the filtered datasets, we identified between 2208 and 9585 putative CNVs per lineage with a mean sequencing depth of $11.0x \pm 3.62$ SD (ranging from 4.3x to 118.7x) for downstream analysis (Table S1). Mean missing data were $4\% \pm 4.2$ SD per locus (range: 0–30%).

GEA analyses and robust outliers

After removing colinear variables, we retained two to three variables within each climate variable category within each lineage (Table S4). For the pRDAs, all models were significant ($p < 0.05$) for all climate variable categories and all lineages except for the solar radiation model in the CU lineage (Table S4). Each RDA had at least one significant axis ($p < 0.05$) except for the CRM precipitation model ($p_{RDA1} = 0.072$), the SRM solar radiation model ($p_{RDA1} = 0.120$) and the CU solar radiation model ($p_{RDA1} = 0.233$); for these models, we still retained the first pRDA axis for outlier detection. The temperature model explained the most variation within each lineage (adjusted $r^2 = 0.018$ – 0.064) except for SRM (adjusted $r^2 = 0.015$), which had the precipitation model explaining the most variation (adjusted $r^2 = 0.017$). We detected between 35 and 468 climate-associated CNVs for each of the pRDAs after removing duplicate loci (Table S4; Figs. S2–S7).

For the LMMs, we detected between 46 and 216 unique loci associated with at least one climate variable within each lineage (Tables S5–S10; Figs. S2–S7). We found the most significant associations with temperature for the NRM (79 unique loci), CSC (139 unique loci), and CU (116 unique loci) lineages, while precipitation had the most for the SRM (29 unique loci) and SN (112 unique loci) lineages. Solar radiation had the largest number of outlier loci for the CRM lineage (28 unique loci; Tables S5–S10; Fig. S2–S7).

We found from 37 to 193 total robust outliers detected by both methods within each lineage (Figs. S2–S7). Repeating the pRDAs using only the robust outliers for each climate variable category resulted in substantially greater variation explained by climate (between 19.8 and 59.1%; Table 1; Figs. 2, 3, 4). All models were highly significant across all climate variable categories and lineages ($p < 0.001$; Table 1; Figs. 2, 3, 4).

Genetic differentiation and adaptive divergence

During preliminary analyses, we saw minimal to no evidence of population structure (Figure S1) and only weak population differentiation (all pairwise $V_{ST} < 0.07$) within each lineage using all putative CNVs. Using the robust outliers, we found consistent elevational and latitudinal patterns of population structure and population differentiation across analyses, though with slight differences between lineages and climate variable categories

| Lineage | Category | N_{robust} | r^2 | r^2_{adjusted} | df | F | p |
|---------|-----------------|---------------------|-------|-------------------------|----|--------|-------|
| NRM | Temperature | 53 | 0.256 | 0.229 | 3 | 8.621 | 0.001 |
| | Precipitation | 26 | 0.228 | 0.210 | 2 | 11.175 | 0.001 |
| | Solar radiation | 39 | 0.224 | 0.206 | 2 | 10.777 | 0.001 |
| CRM | Temperature | 13 | 0.271 | 0.245 | 3 | 9.206 | 0.001 |
| | Precipitation | 9 | 0.267 | 0.252 | 2 | 14.163 | 0.001 |
| | Solar radiation | 15 | 0.205 | 0.176 | 3 | 6.506 | 0.001 |
| SRM | Temperature | 17 | 0.321 | 0.299 | 2 | 11.244 | 0.001 |
| | Precipitation | 23 | 0.370 | 0.351 | 2 | 14.044 | 0.001 |
| | Solar radiation | 8 | 0.337 | 0.314 | 2 | 11.489 | 0.001 |
| CSC | Temperature | 64 | 0.269 | 0.230 | 4 | 6.255 | 0.001 |
| | Precipitation | 7 | 0.217 | 0.198 | 2 | 9.548 | 0.001 |
| | Solar radiation | 31 | 0.198 | 0.178 | 2 | 8.732 | 0.001 |
| SN | Temperature | 62 | 0.383 | 0.351 | 3 | 9.834 | 0.001 |
| | Precipitation | 90 | 0.280 | 0.256 | 2 | 9.417 | 0.001 |
| | Solar radiation | 41 | 0.376 | 0.357 | 2 | 14.372 | 0.001 |
| CU | Temperature | 13 | 0.591 | 0.562 | 3 | 11.929 | 0.001 |
| | Precipitation | 33 | 0.507 | 0.486 | 2 | 12.807 | 0.001 |
| | Solar radiation | 90 | 0.515 | 0.496 | 2 | 13.845 | 0.001 |

Table 1. Summary values for redundancy analysis (RDA) and ANOVA performed on robust outlier loci for six American pika (*Ochotona princeps*) lineages. Robust outliers (N_{robust}) were those loci detected via RDA and linear mixed modelling.

(Figs. 2, 3, 4, 5). Within the NRM lineage, we detected two main population clusters that largely followed latitudinal gradients (Figures S8–S10), though there were also weak clinal trends between normalized read depth of the top correlated outlier loci and elevation (Figures S11–S13). Population structure was less apparent within the CRM lineage, especially for the temperature (Fig. 2) and solar radiation (Fig. 4) outliers, though there were noticeable elevational patterns among precipitation outliers (Fig. S12). For the SRM lineage, we saw consistent patterns in population structure between temperature (Fig. 2) and precipitation (Fig. 3), with slightly different clustering for solar radiation (Fig. 4); however, we did see a distinct pattern between elevation (Figs. S11–S13) and latitude (Figures S8–S10) and normalized read depth. Temperature outliers seemed to be strongly correlated with elevation and latitude in the CSC lineage, while precipitation outliers seemed to be primarily associated with elevation, and radiation outliers followed latitudinal gradients (similar to temperature; Fig. 5). Structure within the SN lineage followed both elevational and latitudinal gradients for temperature and solar radiation outliers (Figs. 2, 4, 5); there was still distinct clustering by sample site using the precipitation outliers (Fig. 3), though there was no consistent pattern between either latitude or elevation and read depth (Figs. S9, S12). Population structure and differentiation most clearly followed elevational and latitudinal gradients in the CU lineage for all variable categories (Figs. 2, 3, 4, 5).

Annotation of adaptive variants

We found that 207 of 508 unique robust outliers were located within introns, exons, or untranslated regions of protein-coding genes, with hits to 158 unique genes (Table S11). Temperature outliers had the most gene hits ($n=85$ unique), followed by precipitation ($n=65$ unique) and solar radiation ($n=55$). Of these genes, 31 had associations across climate variables, 25 had associations across multiple CNVs, and five had associations across multiple lineages (Table S11). Additionally, one gene (*CHCHD3*) that had two associated CNVs was identified independently in both the NRM and CU lineages (Table 2, Table S11). Following our literature review, we identified 12 genes that had putative implications for local adaptation including those with functions regarding: mitochondrial structure and function (i.e., *CHCHD3*, *FBXL3*, *MCU*); immune response (i.e., *DOCK1*, *HPSE2*, *ONECUT1*); transcription (i.e., *MED12L*); hemoglobin structure and function (i.e., *HPSE2*); response to hypoxia (i.e., *LOC101529014*, *FBXL3*); olfaction (i.e., *LOC101517538*, *LOC101532007*); and DNA repair (i.e., *LOC101529014*, *HUS1*, *MACROD1*; Table 2).

Discussion

Structural variation and local adaptation

Various environmental factors can serve as drivers of local adaptation. For example, Muir et al.⁵⁸ found that temperature significantly impacted larval period and growth rate in the common frog (*Rana temporaria*) distributed over an elevational gradient. Precipitation has been shown to be one of the most significant drivers of adaptation and natural selection on both continental and global scales^{59,60}. High levels of solar (UV) radiation have led to rapid, convergent evolution of genes related to DNA repair in species residing on the Qinghai-Tibetan Plateau, the highest plateau on the planet²⁸. Studying both the genetic and phenotypic impacts of environmental factors

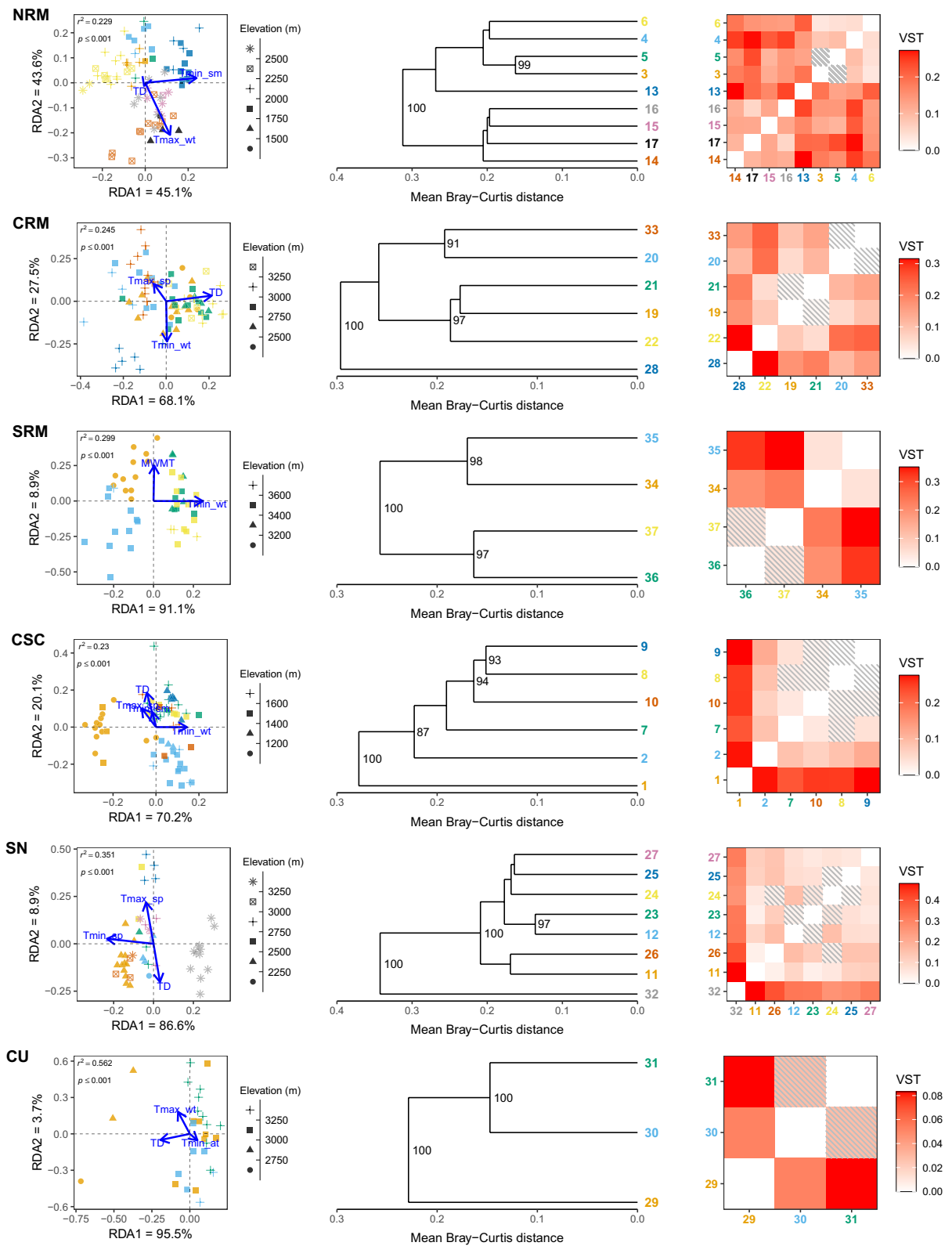


Figure 2. Adaptive divergence of robust temperature outliers in the American pika (*Ochotona princeps*). Partial redundancy analysis (left) was run using the *vegan* R-package⁴⁵; sample sites are indicated by colour and correspond to the sample site numbers on the dendrograms (middle) and heatmaps (right). Sample site numbers correspond with those found on Fig. 1. Dendrograms (middle) were created by hierarchical clustering of Bray-Curtis distances with bootstrap values shown on the nodes; for clarity, only values > 80 are shown. Heatmaps (right) display pairwise population differentiation (V_{ST}); shaded tiles indicate non-significant results from a bootstrapping resampling procedure with 10,000 replicates ($p > 0.05$). *NRM* Northern Rocky Mountains, *CRM* Central Rocky Mountains, *SRM* Southern Rocky Mountains, *CSC* Cascades, *SN* Sierra Nevada, *CU* Central Utah.

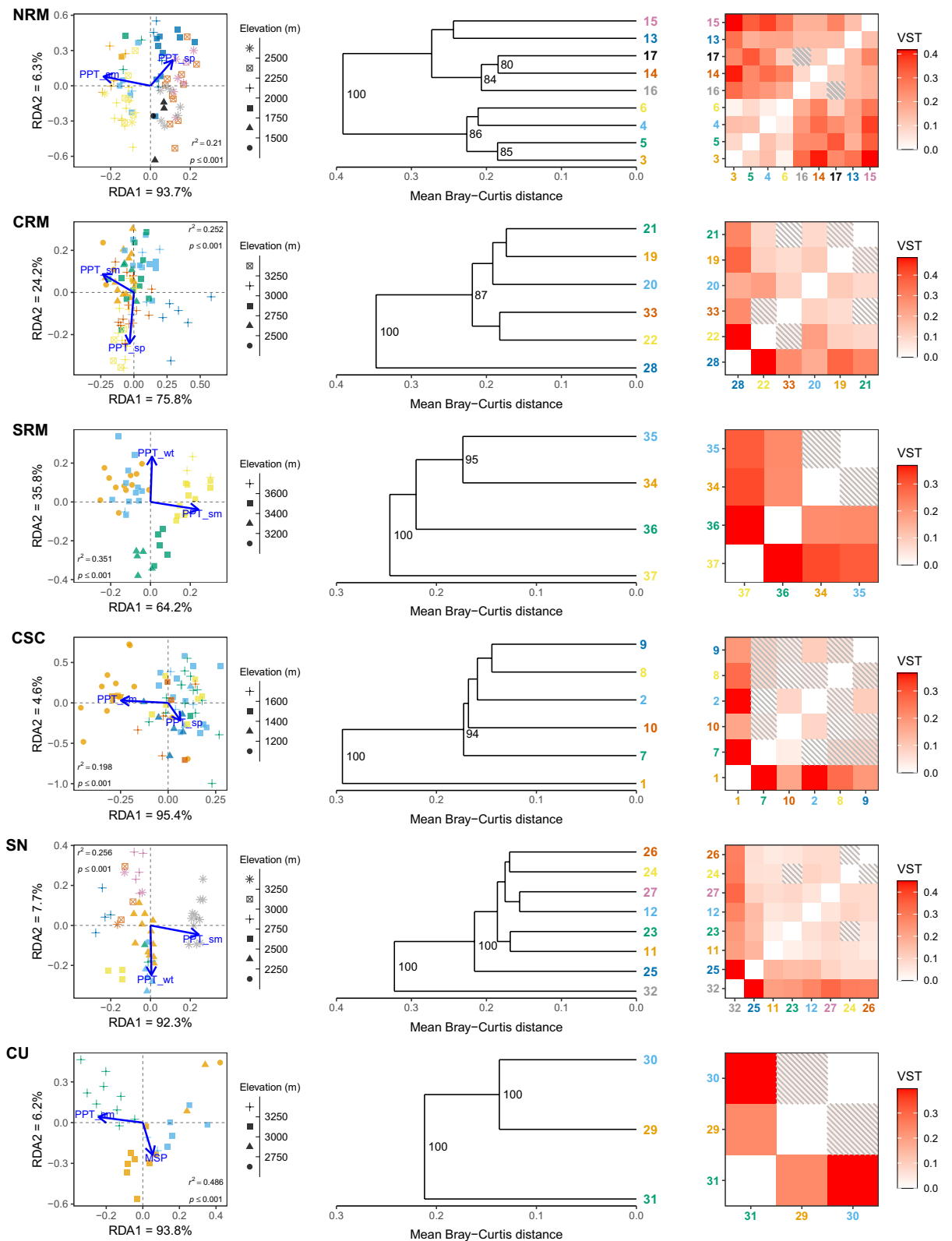


Figure 3. Adaptive divergence of robust precipitation outliers in the American pika (*Ochotona princeps*). Partial redundancy analysis (left) was run using the *vegan* R-package⁴⁵; sample sites are indicated by colour and correspond to the sample site numbers on the dendrograms (middle) and heatmaps (right). Sample site numbers correspond with those found on Fig. 1. Dendrograms (middle) were created by hierarchical clustering of Bray-Curtis distances with bootstrap values shown on the nodes; for clarity, only values > 80 are shown. Heatmaps (right) display pairwise population differentiation (V_{ST}); shaded tiles indicate non-significant results from a bootstrapping resampling procedure with 10,000 replicates ($p > 0.05$). *NRM* Northern Rocky Mountains, *CRM* Central Rocky Mountains, *SRM* Southern Rocky Mountains, *CSC* Cascades, *SN* Sierra Nevada, *CU* Central Utah.

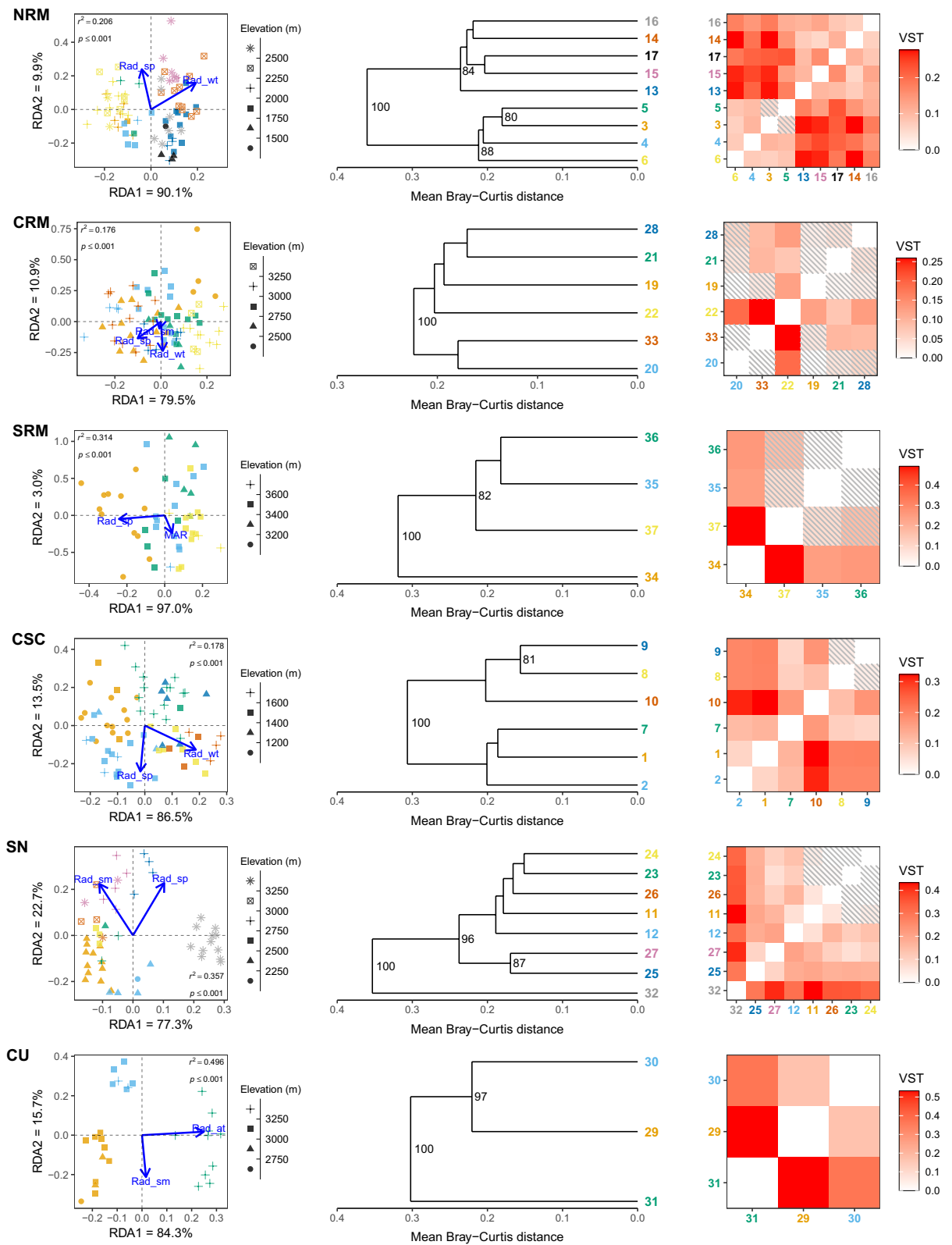


Figure 4. Adaptive divergence of robust solar radiation outliers in the American pika (*Ochotona princeps*). Partial redundancy analysis (left) was run using the *vegan* R-package⁴⁵; sample sites are indicated by colour and correspond to the sample site numbers on the dendrograms (middle) and heatmaps (right). Sample site numbers correspond with those found on Fig. 1. Dendrograms (middle) were created by hierarchical clustering of Bray–Curtis distances with bootstrap values shown on the nodes; for clarity, only values > 80 are shown. Heatmaps (right) display pairwise population differentiation (V_{ST}); shaded tiles indicate non-significant results from a bootstrapping resampling procedure with 10,000 replicates ($p > 0.05$). *NRM* Northern Rocky Mountains, *CRM* Central Rocky Mountains, *SRM* Southern Rocky Mountains, *CSC* Cascades, *SN* Sierra Nevada, *CU* Central Utah.

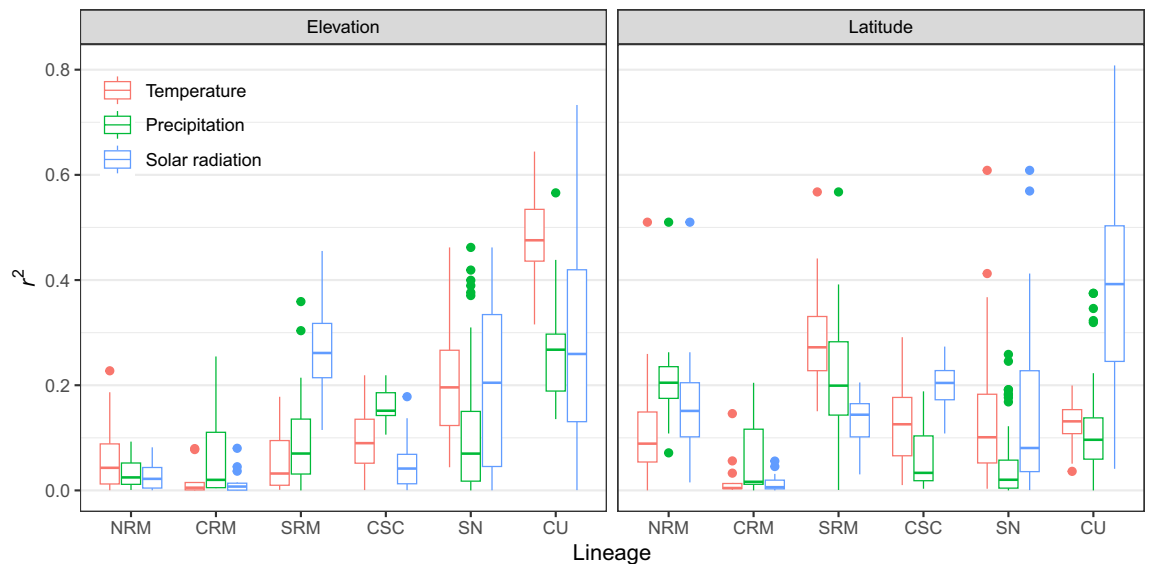


Figure 5. Boxplots showing the distribution of r^2 values from linear regressions among robust climate-associated outlier loci in the American pika (*Ochotona princeps*). We performed independent linear regressions on each locus with either elevation or latitude as the explanatory variable and normalized read depth as the response. *NRM* Northern Rocky Mountains, *CRM* Central Rocky Mountains, *SRM* Southern Rocky Mountains, *CSC* Cascades, *SN* Sierra Nevada, *CU* Central Utah.

within and among species can improve understanding of biodiversity, speciation, and adaptation to heterogeneous and changing landscapes.

There is a growing body of evidence that changes in genome structure may be an even more significant source of evolutionary potential than other, more well-studied markers such as SNPs^{10,12,61–64}. Here, we found that structural variation in the form of copy number variation may be associated with local adaptation in the American pika. Specifically, we observed that putatively adaptive variation in this system largely followed elevational and latitudinal gradients. In the southwestern lineages (SN, CU), putatively adaptive variation was most strongly associated with elevational gradients, particularly with temperature (Figs. 2, 3, 4, 5). Populations of pikas in these regions are often limited to higher elevations, likely due to higher temperatures^{25,26}; in fact, both recent and historical population declines have been documented in southwestern portions of the American pika range, primarily at lower elevations, and have been linked to warmer climates^{65–68}. In the northern lineages (NRM, CSC), putatively adaptive variation closely followed latitudinal gradients, with elevational patterns also evident in the CSC lineage (Figs. 2, 3, 4, 5). Due to cooler temperatures at northern latitudes, pikas in these lineages can occupy a greater range of elevations and can be found as low as sea-level^{69,70}. Additionally, the strength of latitudinal temperature gradients tends to increase moving northward from the tropics⁷¹ and could explain the difference in the effect of latitude between the northern and southern lineages. Furthermore, the northern lineages span a larger latitudinal gradient than the southern lineages, suggesting latitude has a greater potential to correlate with genetic variation in these regions. On the other hand, several studies have found that elevation plays a significant role in shaping putatively adaptive variation in northern populations of the American pika within the CSC lineage^{31,32,72}, indicating that both spatial factors likely influence local adaptation in this system.

In contrast to the other southern lineages, we saw strong, latitudinal trends among outlier loci in the SRM lineage, though solar radiation in this lineage also strongly correlated with elevation (Fig. 5, Figs. S8–S10). Samples in this lineage were collected at high elevations (all > 3150 m) over a relatively small elevational gradient (~ 500 m difference between lowest and highest sites), which could explain the minimal impact of elevation in this lineage. The CRM lineage did not display any clear patterns and had significantly less structure when compared to the other lineages (Figs. 2, 3, 4, 5), possibly due to a narrow and intermediate latitudinal distribution that could limit differences in climate among sample sites. Alternatively, the genetic variation resulting from the relatively disjunct distribution between sample sites in this lineage, particularly in the case of site 28, may be masking signals of climate adaptation. This lineage is also the most recently diverged of the six phylogroups²⁹ and may have had insufficient time post-divergence for selection to leave a significant signature of adaptive variation. Nevertheless, these patterns highlight the importance in sampling over appropriate environmental gradients to detect local adaptation⁷³.

Potential relationships between CNVs, gene expression, and phenotype

Copy number variation can have significant phenotypic and adaptive consequences^{22,74–76}. CNVs are the most abundant form of structural variation, accounting for up to 10% of the total length of the human genome. Moreover, they even occur more frequently than SNPs^{50,77–79} and can serve as a significant source of selectable material, particularly when genes are located within a CNV^{10,12,14,15}. CNVs can directly alter gene expression through increases/decreases in copy number, cause the inactivation of genes via duplication, and lead to amino acid changes and/or reading frame shifts when genes are only partially covered by a variant. In fact, CNV-linked

| Gene ID | CNV | Variable associations | Lineages |
|-----------------------------|-------------------------------|---------------------------------------|----------|
| CHCHD3 | NC_050563.1_30967720_2611485 | Radiation | CSC |
| | NC_050563.1_31126318_2611814 | Temperature | NRM |
| | NC_050563.1_31126337_2611814 | Temperature, precipitation, radiation | NRM, CU |
| | NC_050563.1_31126355_2611814 | Temperature, precipitation, radiation | NRM, CU |
| | NC_050563.1_31126371_2611814 | Precipitation, radiation | CU |
| | NC_050563.1_31126379_2611814 | Precipitation, radiation | CU |
| DOCK1 | NC_050553.1_3509915_140166 | Precipitation | SRM |
| | NC_050553.1_3651078_140637 | Precipitation, radiation | SN |
| FBXL7 | NC_050546.1_30893897_950917 | Temperature | CSC |
| | NC_050546.1_31094969_951406 | Temperature | SN |
| | NC_050568.1_2668793_3906688 | Temperature | CU |
| | NC_050568.1_2668814_3906688 | Temperature | CU |
| HPSE2 | NC_050553.1_25946746_201125 | Precipitation | SN |
| | NC_050553.1_26022614_201248 | Temperature | NRM |
| HUS1 | NC_050545.1_42800208_3781296 | Temperature, precipitation | SRM |
| LOC101517538 (VMN2R116) | NC_050553.1_69408172_305279 | Temperature | NRM |
| | NC_050553.1_69408231_305279 | Temperature | NRM |
| LOC101529014 (TRRAP) | NC_050548.1_133443909_2930290 | Radiation | CU |
| | NC_050548.1_133443965_2930289 | Radiation | CU |
| | NC_050548.1_133596957_2931039 | Precipitation | CU |
| | NC_050548.1_133596959_2931039 | Precipitation | CU |
| | NC_050548.1_37778753_2726048 | Radiation | CU |
| | NC_050548.1_37779077_2726047 | Radiation | CU |
| | NC_050548.1_37779097_2726047 | Radiation | CU |
| | NC_050536.1_13032794_1943468 | Temperature | CRM |
| | NC_050536.1_13033154_1943468 | Temperature, radiation | CRM |
| | NC_050541.1_35683184_4832659 | Precipitation | SN |
| | NC_050541.1_35703827_4832693 | Radiation | CU |
| | NC_050541.1_35703843_4832693 | Radiation | CU |
| | NC_050541.1_35703860_4832693 | Radiation | CU |
| | NC_050541.1_35704145_4832693 | Radiation | CU |
| | NC_050541.1_35704147_4832693 | Radiation | CU |
| | NC_050541.1_35704162_4832693 | Radiation | CU |
| | NC_050541.1_35704225_4832693 | Radiation | CU |
| | NC_050542.1_9477885_1402332 | Precipitation | NRM |
| NC_050542.1_9477910_1402332 | Precipitation | NRM | |
| LOC101532007 (OLFR147) | NC_050536.1_13032794_1943468 | Temperature | CRM |
| | NC_050536.1_13033154_1943468 | Temperature, radiation | CRM |
| MACROD1 | NC_050536.1_108586515_2151246 | Temperature, precipitation | CSC |
| MCU | NC_050553.1_37066530_231578 | Temperature, precipitation, radiation | SN |
| MED12L | NC_050549.1_14187010_4510924 | Temperature | SN |
| | NC_050549.1_14187037_4510924 | Temperature | SN |
| | NC_050549.1_14206971_4511008 | Temperature | NRM |
| | NC_050539.1_2811551_2405314 | Temperature | CSC |
| | NC_050539.1_2811556_2405314 | Temperature | CSC |
| ONECUT1 | NC_050552.1_65357252_2319361 | Precipitation, radiation | CU |

Table 2. Subset of gene annotations for robust climate outliers detected using GEA analyses on six American pika (*Ochotona princeps*) lineages. NRM Northern Rocky Mountains, SRM Southern Rocky Mountains, CRM Central Rocky Mountains, CSC Cascades, SN Sierra Nevada, CU Central Utah

genes more often have functions related to environmental response compared to basic cellular processes⁷⁴. Copy number variation can also lead to high elevation adaptation. A study involving five species of domestic livestock found that mtDNA copy number decreased among high elevation populations compared to low elevation populations, which was hypothesized to be the result of chronic hypoxia⁸⁰. Copy number variation has also been linked to high elevation adaptation in a number of other species, including yak⁸¹, ground tit⁸², and humans⁸³.

Further investigations of CNVs and high elevation environments may lead to a more thorough understanding of adaptations to hypoxia, cold stress, and UV radiation.

We found numerous CNVs within genes with putative links to climate and high elevation adaptation related to mitochondrial function, response to hypoxia, and DNA repair. *CHCHD3* had the greatest support for being under selection of all the annotated genes, as it was detected across all variable categories, several lineages (NRM, CSC, and CU), and multiple CNVs ($n = 6$; Table 2). This gene is critical in the formation of mitochondrial cristae. In fact, in vitro knockdown of *CHCHD3* results in significantly reduced oxygen consumption and glycolytic rates^{84,85}, indicating that this gene could have potential consequences for cold tolerance and adaptation to hypoxia. *FBXL7*, another gene identified in our study as a potential target of selection, is also an important regulator of mitochondrial function. Expression levels of *FBXL7* were down-regulated under hypoxic conditions in the marine medaka⁸⁶, and constitute an important predictor of the severity of asthma symptoms in humans⁸⁷. We also found one CNV annotated to the mitochondrial calcium uniporter (*MCU*), which is an integral component of the mitochondrial inner membrane⁸⁸. These results are consistent with previous work that identified functional enrichment and positively selected genes associated with mitochondrial structure and function in the American pika reference genome^{33,34}, providing further evidence for adaptation to high elevation environments in this species.

In addition, we found several genes associated with hypoxia response to be potentially under selection. The gene *HPSE2* encodes for the enzyme heparanase-2, which plays a role in extracellular matrix remodelling as well as embryo implantation. *HPSE2* has also been linked to hemoglobin-related traits, including fetal hemoglobin in North African human populations, which could have potential implications for adaptation to hypoxia^{89,90}. We found that a predicted *TRRAP* ortholog (LOC101529014) also may be under selection. This gene is part of the *INO80* family of chromatin remodelers, which appear to have putative links to the response to hypoxia by interacting with hypoxia inducible factor-1^{91,92}.

We further found that several genes related to DNA repair may be under selection. For example, *TRRAP* (discussed above) is also involved in DNA repair by binding with the *MRN*-complex to detect and repair double-strand breaks (DSBs)^{93,94}. Knockout and knockdown of *TRRAP* results in the reduced efficiency and precision of end-joining following DSBs in mice and HeLa cells, suggesting this gene plays an important role in DSB signaling and repair⁹⁴. The gene *HUS1* is part of the Rad9-Hus1-Rad1 complex, an important component of the DNA repair pathway. This complex loads onto damaged chromatin (for example, from UV exposure), promoting DSB repair⁹⁵. Again, these results are consistent with previous studies showing putative adaptation to increased UV radiation at high elevations in American pika^{33,34}.

Other putatively adaptive genes

We found several other genes with putative links to local adaptation in the American pika. For example, the gene *MED12L*—a transcriptional coactivator of RNA poly II-dependent genes—was significantly associated with temperature in the NRM, CSC, and SN lineages (Table 2), and has been linked to elevational gradients in North American deer mice⁹⁶ as well as mean annual temperature in Mediterranean cattle breeds⁹⁷. We also found two genes, LOC101532007 (olfactory receptor 147-like) and LOC101517538 (vomeronasal type-2 receptor 116-like), which encode for olfactory receptors that could have an impact on foraging. American pikas do not hibernate; rather, they remain active throughout the winter and cache food into “hay piles” to ensure adequate food supplies⁹⁸. These hay piles often consist of the highest quality vegetation available, likely detected and assessed via olfaction⁹⁹; additionally, many of these cached foods contain high levels of secondary compounds that can help preserve biomass and nutrient availability into the winter months¹⁰⁰. Differences in vegetation quality could also be linked to variation in precipitation, though the CNVs detected here were significantly associated with temperature. Evidence for putative adaptations related to olfaction has been previously shown in the American pika genome³⁴.

We found further evidence for adaptations related to immune response. *DOCK1* is a gene required for phagocytosis of apoptotic cells and has been linked to immune response and climate adaptation in Middle Eastern sheep¹⁰¹ and disease resistance in dolphins¹⁰². *HPSE2* (discussed above) is also involved in the immune response, and expression levels of this gene were significantly associated with white blood cell count in pigs¹⁰³. Similarly, *ONECUT1* is involved in B cell differentiation and has been linked to local adaptation in Ethiopian cattle¹⁰⁴ and three-spined stickleback¹⁰⁵. American pikas experience relatively high levels of parasitism and may be experiencing a spillover of parasites from other small mammals, particularly rodents, which could result in immune response adaptations^{106–108}. Additionally, some populations of American pikas may be physiologically stressed due to harsh environmental conditions^{109–113}; enhanced immune response may confer a greater ability for this species to survive and thrive.

Limitations and future directions

Although we identified putatively adaptive variation in CNVs within the American pika, there were limitations to the approach employed in this study. We used climate data downloaded from ClimateNA which comes with several considerations⁴³. First, this dataset uses climate data collected from 4891 weather stations distributed throughout North America to interpolate climate variables for any given set of coordinates, adjusting for elevation⁴³. While this method allows for the estimation of climate data anywhere on the continent without requiring direct sampling, it also introduces potential error to climate variables. Wang et al.⁴³ found that temperature estimates using this method generally correlated with direct measurements, whereas precipitation varied considerably more; additionally, elevation affected the accuracy of climate variables, with mountainous regions having lower accuracy in point estimates compared to flat regions. Second, this dataset measures total solar radiation of which harmful UV radiation only accounts for a small percentage (~5%)^{114,115}. The relative

contribution of UV radiation to total solar radiation also varies with elevation. However, both total solar radiation and UV radiation increase with elevation; as such, total solar radiation could operate as a crude proxy for UV radiation in lieu of a better estimate. These limitations could explain why both precipitation and solar radiation were relatively less correlated with copy number variation than temperature in this study. Lastly, these data only include ambient temperatures and do not account for the microclimates found underneath the talus that American pikas use for behavioural thermoregulation to prevent over-heating¹¹⁶. Recent findings suggest that subsurface microclimates important for pika thermoregulation may have changed at a faster rate than ambient temperatures over the past few decades in the Southern Rocky Mountains¹¹⁷. Consequently, the sole reliance on ambient temperature estimates may mask potential associations between CNVs and microclimate variation, which represents a potentially interesting avenue for future inquiry.

Our initial study design included plans to detect CNVs using a range-wide dataset. To take advantage of existing sequencing data, we used the *HDplot* method to detect CNVs^{13,20,41}. This method visually detects putative paralogous loci by plotting several heterozygosity and genetic diversity metrics, and qualitatively identifies those that deviate from a central distribution. However, American pikas generally have low heterozygosity as a species at a range-wide level^{38,39}, meaning standard SNP filtering procedures remove many variants. We found that the variants that remained showed very low levels of polymorphism with minimal deviations from the central distribution, greatly inhibiting our ability to confidently and accurately call putative CNVs using the *HDplot* method. To improve upon the ability to detect CNVs on a range-wide rather than within lineage scale, future work could employ whole genome resequencing that would provide a greater breadth of coverage and offer the added capability of detecting additional classes of structural variants, such as inversions and translocations¹¹⁸ that may be more directly associated with climate adaptation in the American pika.

Conclusions

Here, we present a novel analysis of local adaptation in the American pika based on copy number variation. We found that CNVs were significantly associated with temperature, precipitation, and solar radiation within each lineage, and trends in putatively adaptive variation largely followed elevational and latitudinal gradients. Additionally, we identified several genes related to putative high elevation and climate adaptation that could serve as important targets in future studies, including those explicitly involving gene expression. Overall, our work adds to a growing body of literature revealing the novel insights that may be obtained by explicitly examining structural variation in the genome for investigating species-level responses to changing environments^{20,119}. With climate change significantly altering habitats worldwide, a fuller understanding of how organisms may respond, including sentinel species such as the American pika^{65,112}, will be critical for maintaining global biodiversity moving forward.

Data availability

Previously archived sequencing data are available from the NCBI sequence read archive (BioProject ID: PRJNA1075342). Normalized read depth for identified CNVs and SNP genotypic data are available in DRYAD (<https://doi.org/10.5061/dryad.2bvq83bzf>). Climate data are available in the Supplementary Tables document (Table S3). Scripts used for CNV detection are publicly available at https://github.com/enormandeu/stacks_workflow. Benefits from this research accrue from the sharing of our data and results within public databases as described above.

Received: 17 November 2023; Accepted: 8 April 2024

Published online: 13 April 2024

References

1. Blanquart, F., Kaltz, O., Nuismer, S. L. & Gandon, S. A practical guide to measuring local adaptation. *Ecol. Lett.* **16**, 1195–1205 (2013).
2. Darwin, C. *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life.* (John Murray, 1859).
3. Darwin, C. & Wallace, A. R. On the tendency of species to form varieties; and on the perpetuation of varieties and species by natural means of selection. *Proc. Linn. Soc.* **3**, 46–62 (1858).
4. Gregory, T. R. Understanding natural selection: Essential concepts and common misconceptions. *Evol. Educ. Outreach* **2**, 156–175 (2009).
5. Morin, P. A., Luikart, G., Wayne, R. K., & the SNP workshop group. SNPs in ecology, evolution and conservation. *Trends Ecol. Evolut.* **19**, 208–216 (2004).
6. Sachidanandam, R. *et al.* A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature* **409**, 928–933 (2001).
7. Bridges, C. B. The Bar 'gene' a duplication. *Science* **83**, 210–211 (1936).
8. Dobzhansky, Th. & Sturtevant, A. H. Inversions in the chromosomes of *Drosophila pseudoobscura*. *Genetics* **23**, 28–64 (1938).
9. McClintock, B. The origin and behavior of mutable loci in maize. *Proc. Natl. Acad. Sci.* **36**, 344–355 (1950).
10. Mérot, C., Oomen, R. A., Tigano, A. & Wellenreuther, M. A roadmap for understanding the evolutionary significance of structural genomic variation. *Trends Ecol. Evolut.* **35**, 561–572 (2020).
11. Sturtevant, A. H. The linear arrangement of six sex-linked factors in *Drosophila*, as shown by their mode of association. *J. Exp. Zool.* **14**, 43–59 (1913).
12. Wellenreuther, M., Mérot, C., Berdan, E. & Bernatchez, L. Going beyond SNPs: The role of structural genomic variants in adaptive evolution and species diversification. *Mol. Ecol.* **28**, 1203–1209 (2019).
13. Cayuela, H. *et al.* Shared ancestral polymorphisms and chromosomal rearrangements as potential drivers of local adaptation in a marine fish. *Mol. Ecol.* **29**, 2379–2398 (2020).
14. Chain, F. J. J. & Feulner, P. G. D. Ecological and evolutionary implications of genomic structural variations. *Front. Genet.* **5**, 13 (2014).
15. Spielmann, M., Lupiáñez, D. G. & Mundlos, S. Structural variation in the 3D genome. *Nat. Rev. Genet.* **19**, 453–467 (2018).

16. Wellenreuther, M. & Bernatchez, L. Eco-evolutionary genomics of chromosomal inversions. *Trends Ecol. Evolut.* **33**, 427–440 (2018).
17. Arostegui, M. C., Quinn, T. P., Seeb, L. W., Seeb, J. E. & McKinney, G. J. Retention of a chromosomal inversion from an anadromous ancestor provides the genetic basis for alternative freshwater ecotypes in rainbow trout. *Mol. Ecol.* **28**, 1412–1427 (2019).
18. Clop, A., Vidal, O. & Amills, M. Copy number variation in the genomes of domestic animals. *Anim. Genet.* **43**, 503–517 (2012).
19. Baird, N. A. *et al.* Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLOS ONE* **3**, e3376 (2008).
20. Dorant, Y. *et al.* Copy number variants outperform SNPs to reveal genotype–temperature association in a marine species. *Mol. Ecol.* **29**, 4765–4782 (2020).
21. Cayuela, H. *et al.* Thermal adaptation rather than demographic history drives genetic structure inferred by copy number variants in a marine fish. *Mol. Ecol.* **30**, 1624–1641 (2021).
22. Cayuela, H. *et al.* Genomic signatures of thermal adaptation are associated with clinal shifts of life history in a broadly distributed frog. *J. Anim. Ecol.* **91**, 1222–1238 (2022).
23. Tigano, A. A population genomics approach to uncover the CNVs, and their evolutionary significance, hidden in reduced-representation sequencing data sets. *Mol. Ecol.* **29**, 4749–4753 (2020).
24. Hafner, D. J. & Smith, A. T. Revision of the subspecies of the American pika, *Ochotona princeps* (Lagomorpha: Ochotonidae). *J. Mammal.* **91**, 401–417 (2010).
25. Smith, A. T. & Weston, M. L. *Ochotona princeps*. *Mammalian Species*. Vol. 1. <https://doi.org/10.2307/3504319> (1990).
26. Millar, C. I. & Westfall, R. D. Distribution and climatic relationships of the American pika (*Ochotona princeps*) in the Sierra Nevada and Western Great Basin, U.S.A.; periglacial landforms as refugia in warming climates. *Arctic Antarctic Alpine Res.* **42**, 76–88 (2010).
27. Smith, A. T. & Beaver, E. *Ochotona princeps* (2016).
28. Sun, Y.-B. *et al.* Species groups distributed across elevational gradients reveal convergent and continuous genetic adaptation to high elevations. *Proc. Natl. Acad. Sci. USA* **115**, E10634–E10641 (2018).
29. Schmidt, D. A., Galbreath, K. E. & Russello, M. A. Phylogenomics of American pika (*Ochotona princeps*) lineage diversification. *Mol. Phylogenet. Evolut.* **193**, 108030 (2024).
30. Lemay, M. A., Henry, P., Lamb, C. T., Robson, K. M. & Russello, M. A. Novel genomic resources for a climate change sensitive mammal: Characterization of the American pika transcriptome. *BMC Genomics* **14**, 311 (2013).
31. Schmidt, D. A., Waterhouse, M. D., Sjodin, B. M. F. & Russello, M. A. Genome-wide analysis reveals associations between climate and regional patterns of adaptive divergence and dispersal in American pikas. *Heredity* **127**, 443–454 (2021).
32. Waterhouse, M. D., Erb, L. P., Beaver, E. A. & Russello, M. A. Adaptive population divergence and directional gene flow across steep elevational gradients in a climate-sensitive mammal. *Mol. Ecol.* **27**, 2512–2528 (2018).
33. Sjodin, B. M. F., Galbreath, K. E., Lanier, H. C. & Russello, M. A. Chromosome-level reference genome assembly for the American pika (*Ochotona princeps*). *J. Hered.* **112**, 549–557 (2021).
34. Sjodin, B. M. F. & Russello, M. A. Comparative genomics reveals putative evidence for high-elevation adaptation in the American pika (*Ochotona princeps*). *G3 Genes|Genomes|Genet.* **12**, 241 (2022).
35. Galbreath, K. E., Hafner, D. J. & Zamudio, K. R. When cold is better: Climate-driven elevation shifts yield complex patterns of diversification and demography in an alpine specialist (American pika, *Ochotona princeps*). *Evolution* **63**, 2848–2863 (2009).
36. Rochette, N. C., Rivera-Colón, A. G. & Catchen, J. M. Stacks 2: Analytical methods for paired-end sequencing improve RADseq-based population genomics. *Mol. Ecol.* **28**, 4737–4754 (2019).
37. Li, H. *Aligning Sequence Reads, Clone Sequences and Assembly Contigs with BWA-MEM*. [arXiv:1303.3997](https://arxiv.org/abs/1303.3997) [q-bio] (2013).
38. Beaver, E. A. *et al.* Geographic and taxonomic variation in adaptive capacity among mountain-dwelling small mammals: Implications for conservation status and actions. *Biol. Conserv.* **282**, 109942 (2023).
39. Klingler, K. B., Jahner, J. P., Parchman, T. L., Ray, C. & Peacock, M. M. Genomic variation in the American pika: Signatures of geographic isolation and implications for conservation. *BMC Ecol. Evol.* **21**, 2 (2021).
40. Danecek, P. *et al.* The variant call format and VCFtools. *Bioinformatics* **27**, 2156–2158 (2011).
41. McKinney, G. J., Waples, R. K., Seeb, L. W. & Seeb, J. E. Paralogs are revealed by proportion of heterozygotes and deviations in read ratios in genotyping-by-sequencing data from natural populations. *Mol. Ecol. Resour.* **17**, 656–669 (2017).
42. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: A bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
43. Wang, T., Hamann, A., Spittlehouse, D. & Carroll, C. Locally downscaled and spatially customizable climate data for historical and future periods for North America. *PLOS ONE* **11**, e0156720 (2016).
44. Kuhn, M. Building predictive models in R using the caret package. *J. Stat. Softw.* **28**, 1–26 (2008).
45. Oksanen, J. *et al.* *vegan: Community Ecology Package*. (2022).
46. Zuur, A. F., Ieno, E. N. & Elphick, C. S. A protocol for data exploration to avoid common statistical problems. *Methods Ecol. Evolut.* **1**, 3–14 (2010).
47. Forester, B. R., Lasky, J. R., Wagner, H. H. & Urban, D. L. Comparing methods for detecting multilocus adaptation with multivariate genotype–environment associations. *Mol. Ecol.* **27**, 2215–2233 (2018).
48. Bates, D., Mächler, M., Bolker, B. & Walker, S. Fitting linear mixed-effects models using lme4. *J. Stat. Soft.* **67**, 1–48 (2015).
49. Kuznetsova, A., Brockhoff, P. B. & Christensen, R. H. B. lmerTest package: Tests in linear mixed effects models. *J. Stat. Soft.* **82**, 1–26 (2017).
50. Redon, R. *et al.* Global variation in copy number in the human genome. *Nature* **444**, 444–454 (2006).
51. Weir, B. S. & Cockerham, C. C. Estimating F-statistics for the analysis of population structure. *Evolution* **38**, 1358–1370 (1984).
52. Canty, A. & Ripley, B. D. *boot: Bootstrap R (S-Plus) Functions*. (2022).
53. Goslee, S. C. & Urban, D. L. The ecodist package for dissimilarity-based analysis of ecological data. *J. Stat. Soft.* **22**, 1–19 (2007).
54. Murtagh, F. & Legendre, P. Ward's hierarchical agglomerative clustering method: Which algorithms implement Ward's criterion?. *J. Classif.* **31**, 274–295 (2014).
55. Paradis, E. & Schliep, K. ape 5.0: An environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* **35**, 526–528 (2019).
56. R Core Team. *R: A Language and Environment for Statistical Computing*. (R Foundation for Statistical Computing, 2022).
57. McLaren, W. *et al.* The Ensembl variant effect predictor. *Genome Biol.* **17**, 122 (2016).
58. Muir, A. P., Biek, R., Thomas, R. & Mable, B. K. Local adaptation with high gene flow: Temperature parameters drive adaptation to altitude in the common frog (*Rana temporaria*). *Mol. Ecol.* **23**, 561–574 (2014).
59. Moreira, L. R. & Smith, B. T. Convergent genomic signatures of local adaptation across a continental-scale environmental gradient. *Sci. Adv.* **9**, eadd0560 (2023).
60. Siepielski, A. M. *et al.* Precipitation drives global variation in natural selection. *Science* **355**, 959–962 (2017).
61. Barth, J. M. I. *et al.* Disentangling structural genomic and behavioural barriers in a sea of connectivity. *Mol. Ecol.* **28**, 1394–1411 (2019).
62. Fuller, Z. L., Koury, S. A., Phadnis, N. & Schaeffer, S. W. How chromosomal rearrangements shape adaptation and speciation: Case studies in *Drosophila pseudoobscura* and its sibling species *Drosophila persimilis*. *Mol. Ecol.* **28**, 1283–1301 (2019).
63. Hooper, D. M., Griffith, S. C. & Price, T. D. Sex chromosome inversions enforce reproductive isolation across an avian hybrid zone. *Mol. Ecol.* **28**, 1246–1262 (2019).

64. Yoshida, K. *et al.* Functional divergence of a heterochromatin-binding protein during stickleback speciation. *Mol. Ecol.* **28**, 1563–1578 (2019).
65. Beever, E. A., Brussard, P. F. & Berger, J. Patterns of apparent extirpation among isolated populations of pikas (*Ochotona princeps*) in the Great Basin. *J. Mammal.* **84**, 37–54 (2003).
66. Beever, E. A., Ray, C., Mote, P. W. & Wilkening, J. L. Testing alternative models of climate-mediated extirpations. *Ecol. Appl.* **20**, 164–178 (2010).
67. Beever, E. A., Ray, C., Wilkening, J. L., Brussard, P. F. & Mote, P. W. Contemporary climate change alters the pace and drivers of extinction. *Glob. Change Biol.* **17**, 2054–2070 (2011).
68. Wilkening, J. L., Ray, C., Beever, E. A. & Brussard, P. F. Modeling contemporary range retraction in Great Basin pikas (*Ochotona princeps*) using data on microclimate and microhabitat. *Quat. Int.* **235**, 77–88 (2011).
69. Henry, P., Sim, Z. & Russello, M. A. Genetic evidence for restricted dispersal along continuous altitudinal gradients in a climate change-sensitive mammal: the American pika. *PLOS ONE* **7**, e39077 (2012).
70. Robson, K. M., Lamb, C. T. & Russello, M. A. Low genetic diversity, restricted dispersal, and elevation-specific patterns of population decline in American pikas in an atypical environment. *J. Mammal.* **97**, 464–472 (2016).
71. Sorte, F. A. L., Butchart, S. H. M., Jetz, W. & Böhning-Gaese, K. Range-wide latitudinal and elevational temperature gradients for the world's terrestrial birds: Implications under global climate change. *PLOS ONE* **9**, e98361 (2014).
72. Henry, P. & Russello, M. A. Adaptive divergence along environmental gradients in a climate-change-sensitive mammal. *Ecol. Evol.* **3**, 3906–3917 (2013).
73. Forester, B. R., Jones, M. R., Joost, S., Landguth, E. L. & Lasky, J. R. Detecting spatial genetic signatures of local adaptation in heterogeneous landscapes. *Mol. Ecol.* **25**, 104–120 (2016).
74. Korbel, J. O. *et al.* The current excitement about copy-number variation: How it relates to gene duplication and protein families. *Curr. Opin. Struct. Biol.* **18**, 366–374 (2008).
75. Schrider, D. R. & Hahn, M. W. Gene copy-number polymorphism in nature. *Proc. R. Soc. B Biol. Sci.* **277**, 3213–3221 (2010).
76. Stranger, B. E. *et al.* Relative impact of nucleotide and copy number variation on gene expression phenotypes. *Science* **315**, 848–853 (2007).
77. Katju, V. & Bergthorsson, U. Copy-number changes in evolution: Rates, fitness effects and adaptive significance. *Front. Genet.* **4**, 273 (2013).
78. Levy, S. *et al.* The diploid genome sequence of an individual human. *PLOS Biol.* **5**, e254 (2007).
79. Zarrei, M., MacDonald, J. R., Merico, D. & Scherer, S. W. A copy number variation map of the human genome. *Nat. Rev. Genet.* **16**, 172–183 (2015).
80. Liu, R. *et al.* Analysis of mitochondrial DNA sequence and copy number variation across five high-altitude species and their low-altitude relatives. *Mitochond. DNA Part B* **3**, 847–851 (2018).
81. Guang-Xin, E. *et al.* Genome-wide selective sweep analysis of the high-altitude adaptability of yaks by using the copy number variant. *3 Biotech* **10**, 259 (2020).
82. Cai, Q. *et al.* Genome sequence of ground tit *Pseudopodoces humilis* and its adaptation to high altitude. *Genome Biol.* **14**, R29 (2013).
83. Xu, S. *et al.* A genome-wide search for signals of high-altitude adaptation in Tibetans. *Mol. Biol. Evol.* **28**, 1003–1011 (2011).
84. Darshi, M. *et al.* ChChd3, an inner mitochondrial membrane protein, is essential for maintaining crista integrity and mitochondrial function. *J. Biol. Chem.* **286**, 2918–2932 (2011).
85. Wang, G., He, Y. & Luo, Y. Expression of OPA1 and Mic60 genes and their association with mitochondrial cristae morphology in Tibetan sheep. *Cell Tissue Res.* **376**, 273–279 (2019).
86. Lai, K. P. *et al.* Hypoxia causes sex-specific hepatic toxicity at the transcriptome level in marine medaka (*Oryzias melastigma*). *Aquatic Toxicol.* **224**, 105520 (2020).
87. Park, H.-W. *et al.* Genetic predictors associated with improvement of asthma symptoms in response to inhaled corticosteroids. *J. Allergy Clin. Immunol.* **133**, 664–669.e5 (2014).
88. Baughman, J. M. *et al.* Integrative genomics identifies MCU as an essential component of the mitochondrial calcium uniporter. *Nature* **476**, 341–345 (2011).
89. Caro-Consuegra, R., Lucas-Sánchez, M., Comas, D. & Bosch, E. Identifying signatures of positive selection in human populations from North Africa. *Sci. Rep.* **13**, 8166 (2023).
90. Menzel, S. & Thein, S. L. Genetic modifiers of fetal haemoglobin in sickle cell disease. *Mol. Diagn. Ther.* **23**, 235–244 (2019).
91. Collier, H., Albanese, A., Kwok, C.-S., Kou, J. & Rocha, S. Functional crosstalk between chromatin and hypoxia signalling. *Cell. Signal.* **106**, 110660 (2023).
92. Melvin, A. & Rocha, S. Chromatin as an oxygen sensor and active player in the hypoxia response. *Cell. Signal.* **24**, 35–43 (2012).
93. Murr, R. *et al.* Histone acetylation by Trrap–Tip60 modulates loading of repair proteins and repair of DNA double-strand breaks. *Nat. Cell Biol.* **8**, 91–99 (2006).
94. Robert, F. *et al.* The transcriptional histone acetyltransferase cofactor TRRAP associates with the MRN repair complex and plays a role in DNA double-strand break repair. *Mol. Cell. Biol.* **26**, 402–412 (2006).
95. Ohashi, E., Takeishi, Y., Ueda, S. & Tsurimoto, T. Interaction between Rad9–Hus1–Rad1 and TopBP1 activates ATR–ATRIP and promotes TopBP1 recruitment to sites of UV-damage. *DNA Repair* **21**, 1–11 (2014).
96. Schweizer, R. M. *et al.* Broad concordance in the spatial distribution of adaptive and neutral genetic variation across an elevational gradient in deer mice. *Mol. Biol. Evol.* **38**, 4286–4300 (2021).
97. Flori, L. *et al.* A genomic map of climate adaptation in Mediterranean cattle breeds. *Mol. Ecol.* **28**, 1009–1029 (2019).
98. Huntly, N. J., Smith, A. T. & Ivins, B. L. Foraging behavior of the pika (*Ochotona princeps*), with comparisons of grazing versus haying. *J. Mammal.* **67**, 139–148 (1986).
99. Smith, J. A. & Erb, L. P. Patterns of selective caching behavior of a generalist herbivore, the American pika (*Ochotona princeps*). *Arctic Antarctic Alpine Res.* **45**, 396–403 (2013).
100. Dearing, M. D. The manipulation of plant toxins by a food-hoarding herbivore, *Ochotona princeps*. *Ecology* **78**, 774–781 (1997).
101. Eydivandi, S., Roudbar, M. A., Karimi, M. O. & Sahana, G. Genomic scans for selective sweeps through haplotype homozygosity and allelic fixation in 14 indigenous sheep breeds from Middle East and South Asia. *Sci. Rep.* **11**, 2834 (2021).
102. Batley, K. C. *et al.* Whole genomes reveal multiple candidate genes and pathways involved in the immune response of dolphins to a highly infectious virus. *Mol. Ecol.* **30**, 6434–6448 (2021).
103. Tao, C. *et al.* Molecular characterization, expression profiles, and association analysis with hematologic parameters of the porcine HPSE and HPSE2 genes. *J. Appl. Genet.* **54**, 71–78 (2013).
104. Saravanan, K. A. *et al.* Genomic scans for selection signatures revealed candidate genes for adaptation and production traits in a variety of cattle breeds. *Genomics* **113**, 955–963 (2021).
105. Pritchard, V. L. *et al.* Regulatory architecture of gene expression variation in the threespine stickleback *Gasterosteus aculeatus*. *G3 Genes|Genomes|Genet.* **7**, 165–178 (2017).
106. Brinkerhoff, R. J., Rinsland, H. S., Sato, S., Maruyama, S. & Ray, C. Vector-borne pathogens in ectoparasites collected from high-elevation pika populations. *EcoHealth* **17**, 333–344 (2020).
107. Galbreath, K. E. & Hoberg, E. P. Return to Beringia: Parasites reveal cryptic biogeographic history of North American pikas. *Proc. R. Soc. B* **279**, 371–378 (2012).

108. Foley, P., Roth, T., Foley, J. & Ray, C. Rodent–pika parasite spillover in western North America. *J. Med. Entomol.* **54**, 1251–1257 (2017).
109. Jeffress, M. R., Rodhouse, T. J., Ray, C., Wolff, S. & Epps, C. W. The idiosyncrasies of place: Geographic variation in the climate-distribution relationships of the American pika. *Ecol. Appl.* **23**, 864–878 (2013).
110. Ray, C., Beever, E. & Loarie, S. Retreat of the American pika: Up the mountain or into the void? In *Wildlife Conservation in a Changing Climate* (eds. Brodie, J.F., Post, E.S. & Doak, D.F.) 245–270 (University of Chicago Press, 2012).
111. Waterhouse, M. D. *et al.* Individual-based analysis of hair corticosterone reveals factors influencing chronic stress in the American pika. *Ecol. Evol.* **7**, 4099–4108 (2017).
112. Wilkening, J. L. & Ray, C. Characterizing predictors of survival in the American pika (*Ochotona princeps*). *J. Mammal.* **97**, 1366–1375 (2016).
113. Whipple, A. L. *et al.* Temporal vs. spatial variation in stress-associated metabolites within a population of climate-sensitive small mammals. *Conserv Physiol.* **9**, coab024 (2021).
114. *A Review of Human Carcinogens*. (International Agency for Research on Cancer, 2012).
115. Wang, Q.-W., Hidema, J. & Hikosaka, K. Is UV-induced DNA damage greater at higher elevation?. *Am. J. Bot.* **101**, 796–802 (2014).
116. MacArthur, R. A. & Wang, L. C. H. Behavioral thermoregulation in the pika *Ochotona princeps*: A field study using radiotelemetry. *Can. J. Zool.* **52**, 353–358 (1974).
117. Monk, E. M. & Ray, C. Revisiting talus and free-air temperatures after 50 years of change at an American pika (*Ochotona princeps*) study site in the southern Rockies. *PLOS Clim.* **1**, e0000049 (2022).
118. Mahmoud, M. *et al.* Structural variant calling: The long and the short of it. *Genome Biol.* **20**, 246 (2019).
119. Layton, K. K. S. & Bradbury, I. R. Harnessing the power of multi-omics data for predicting climate change response. *J. Anim. Ecol.* **91**, 1064–1072 (2022).

Acknowledgements

We would like to thank Anna Tigano and Yann Dorant for advice on data analysis. This work was funded by a Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery grant (RGPIN-2019-04621) to MAR. Specimens used in this study were originally collected in Galbreath *et al.* (2009) with support from the US National Science Foundation (NSF grant DEB 0506042 to KEG). Computational resources were made available by the Digital Research Alliance of Canada through the Resources for Research Group (MAR).

Author contributions

BMFS contributed to study design, carried out all data analyses, and drafted the manuscript. DAS contributed the initial data and provided feedback on the manuscript. KEG contributed samples and provided feedback on the manuscript. MAR conceived and coordinated the study, helped draft the manuscript, and obtained funding.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-024-59157-6>.

Correspondence and requests for materials should be addressed to M.A.R.

Reprints and permissions information is available at www.nature.com/reprints.

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



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

© The Author(s) 2024