Electronic Supplementary Materials

Study system details and population sampling. *Timema* feed and mate on the plants upon which they rest. As in past work, a 'population' of *T. cristinae* is defined as all individuals of this species collected within a homogenous patch of a single host species. Patches of the two host species used by *T. cristinae* are often distributed in adjacent patches that are in direct geographic contact with one another. Insect populations associated with such patches are referred to as 'geographically adjacent' [26]. Other host patches are separated from patches of the alternative host, usually via regions containing unsuitable hosts (termed 'geographically separated')(Table S1). The geographic distance between populations was calculated from GPS coordinates using the program Geographic Distance Matrix Generator v. 1.2.3 [44].

PCR and sequencing protocols. Restriction digestion and adaptor-ligation were carried out simultaneously on 0.5 μg of genomic DNA using the restriction endonucleases EcoRI and MseI (NEB, Inc.). Adaptor sequences and their reverse complements that allowed for ligation to the restriction sites were annealed to each other by incubating at 95˚C for five minutes and slow cooling to room temperature. We ligated adaptor sequences consisting of the Illumina adaptor, a 10 bp internal DNA barcode on the EcoRI adaptor [45], and additional bases to protect the annealed fragments from digestion (EcoRI side: 5'–CTCTTTCCCTACACGACGCTCTTCCGATCT–3' + 10 bp barcode + C; MseI side: 5'–GCAGAAGACGGCATACGAGCTCTTCCGATCT–3' + G). The Illumina/barcoded adaptor pairs were attached to digested fragments using T4 DNA ligase (NEB, Inc.), and restriction and ligation were accomplished simultaneously in 11 μL reactions subject to 18 hours of incubation at 38˚C, followed by dilution with 170 μL 0.1× TE buffer. Fragments were PCR amplified with Illumina PCR primers (1, 5' *AATGATACGGCGACCACCGAGATCTACACTCTTTCCC TACACGACGCTCTTCCGATCT–3'; 2, 5'*CAAGCAGAAGACGGCATACGAGCTCTT CCGATCT–3') (Illumina, Inc.), which amplify fragments based on the sequences of the ligated adaptors. The asterisks in the primer sequences above refer to phosphorothioation, which serves to limit exonuclease activity at the ends of the amplified fragments. These reactions contained 6 μL of the diluted restriction-ligation products, 21.7 μL 1x PCR buffer, 0.3μL Iproof high fidelity polymerase at 4 Units/μL (Bio-Rad, Inc.), and 2 μL of a 5 μM mix of forward and reverse Illumina PCR primers. PCR conditions included 20 PCR cycles (94˚C for 30 seconds, 56 ˚C for 1 minute,

72˚C for 2 minutes) and a final extension at 60˚C for 30 minutes. These reactions were run in duplicate to guard against variation across individuals in amplification of the fragment pools.

The product of these PCR reactions was then subjected to electrophoresis in 2% agarose gels at 75 volts for 210 minutes. We used 200μL pipette tips to remove uniformly sized pieces of agarose out of each electrophoresis lane at the same vertical position in the gel. DNA in the region of 300–450 bp in length was excised and purified with QiaQuick gel extraction kits (Qiagen, Inc.). We combined 10 ng of each purified DNA sample into a single pool for Illumina sequencing. The quality of libraries was assessed with quantitative electrophoresis on an Agilent Bioanalyzer and verified with qPCR. Sequencing was accomplished on two lanes of an Illumina GAIIx genetic analyzer flow cell at the National Center for Genome Resources (NCGR) in Santa Fe, NM. Sequence reads were 108 bases in length; all began with the 10 bp barcode at the EcoRI end of our amplified fragments and were of the same orientation. Consequently, aside from the 10 bp barcode and the 6 bases corresponding to the EcoRI cut site, all reads contained 92 bp of informative sequence.

Assembly. The 10 bp barcodes and the six preceding nucleotides corresponding to the EcoRI restriction site at the 5' ends of sequences were trimmed from all reads prior to further processing, and the individual IDs corresponding to each barcode were added to the info line of fastq files. The Perl script we used to parse barcodes also identified and corrected barcodes that were 1 bp away from a known barcode sequence, and thus represented errors in barcode synthesis or sequencing. For assemblies we used a gap penalty of 125, minimum match percentage of 93%, match size of 25 bp, mismatch penalty of 15, and used the repeat handling option. Further details on the parameters used in assembly are available from the authors upon request. We used custom Perl scripts along with bcftools and samtools [46] to call SNPs in the assembled contigs. We retained contigs between 96 bases and 88 bases in length and concatenated the consensus sequences from these contigs into an artificial reference template, with the separate sequences padded by Ns. Samtools processes input bam files, computes the likelihood of the data given each possible genotype and bcftools then applies a prior and executes the calling of variant sites based on a Bayesian model. We used the full prior in bcftools, only considered SNPs minimally present in 30% of the samples (this is a threshold for missing data across all samples and is not a function of allele frequencies or whether a specific allele is fixed in some populations), and required the probability of the data to be less than 0.05 under the assumption that all samples were homozygous for the reference allele. We disregarded

insertions and deletions. The data for each locus were placed in a file containing the number of reads for each SNP in each individual. We then trimmed out all genetic regions where more than two individuals appeared to have more than two haplotypes, and discarded any SNPs where counts did not fit the expected binomial distribution. We retained 46,153,271 reads averaging 92 bases in length for analysis after removing barcodes and the proceeding six bases associated with the EcoRI cut site.

Genotype estimation. The genotype probabilities that we use for PCA and LD were estimated assuming independent conditional priors for the genotypes in each population. These priors were the expected genotype frequencies in each population based on the allele frequencies in that population. We estimated genotypes and allele frequencies simultaneously. We also estimate genotypes for the F_{ST} analysis and these could vary slightly for individual pairwise comparisons. See main text for details.

Population differentiation. We used a hierarchical Bayesian implementation of the F-model to quantify genetic differentiation among populations. The F-model treats F_{ST} as an evolutionary parameter rather than a simple summary of the population allele frequencies. Our implementation of the F-model follows [37] and allows information sharing among loci. Related implementations of the F-model have been developed [47-50]. The model assumes Hardy-Weinberg and linkage equilibrium within populations, but incorporates uncertainty in genotypic state as described for the allele frequency model. F_{ST} is allowed to vary across the genome to reflect potential genome-wide variation in selection or other key evolutionary parameters. The model assumes that F_{ST} follows a normal distribution across the genome, where $F_{ST\,i}$ denotes F_{ST} for the *i*th locus. The mean (μ) and precision (τ) of the genome-wide F_{ST} distribution are treated as model parameters and estimated from the data. This normal distribution is the conditional prior for the locus-specific F_{ST}. This prior specification assumes conditional independence of F_{ST} among loci (i.e., the F_{STi} are independent draws from a common distribution. Tightly linked loci are expected to violate the assumption of conditional independence. Thus, we examined the degree of statistical independence among loci using linkage disequilibrium analyses, as described below. The hierarchical Bayesian F-model is described fully by [37]. We used MCMC to estimate the posterior probability distributions of each F_{STi} and the genome-wide mean and variance for each pair of populations. We ran a single chain for each analysis for 25,000 iterations and retained every fifth sampled parameter value for parameter

estimation. Although this analytical method accounts for genotype uncertainty, it does not directly account for sequencing error or PRC duplicates (a potential issue when many cycles of PCR are used). Future work examining these issues is warranted.

Linkage disequilibrium analyses. We estimated Burrow's composite measure of Hardy Weinberg and linkage disequilibrium (Δ) for each pair of variable sites [38]. This measure does not assume Hardy-Weinberg equilibrium or require phased data, but instead provides a joint metric of intralocus and interlocus disequilibria based solely on genotype frequencies [38]. We used a Monte Carlo algorithm to incorporate uncertainty in genotypic state into our estimates of Δ and estimated Δ separately for each population. Specifically, we first sampled genotypes for each individual based on the posterior probability of genotypic state for that individual (these posterior probabilities were obtained from the allele frequency estimation model described previously). We then calculated Δ{ii'} for each locus pair based on the sampled genotypes. We iterated this procedure 100 times for each locus pair and used the mean value of Δ{ii'} as an estimate of Δ{ii'}. We summarized estimates of Δ{ii'} for all locus pairs (approximately 1.9 billion locus pairs) and pairs of outlier loci (approximately 3.98 million pairs) for each population. We also estimated linkage disequilibrium for SNPs in the same versus different contigs. We used C to implement this Monte Carlo estimation procedure. We tested whether levels of linkage disequilibrium within populations differed for outlier loci versus all loci using a paired t-test.

Matrix correlation analyses. We used simple and partial Mantel tests to compare various matrices to one another (45). A simple Mantel test compares the association between two distance matrices. The significance of this relationship is tested by comparing the linear correlation between the two matrices to a null distribution generated by randomizing rows and columns of one matrix while holding the other constant and recalculating the linear association. A partial Mantel test allows one to compare three distance matrices. In this case, the partial Mantel statistic estimates the correlation between two matrices while controlling for the effect of a third and is computed in the same way as a partial correlation coefficient. The test of significance of the partial Mantel compares this partial correlation coefficient to a null distribution generated by randomizing the rows and columns of one of the two focal matrices, while holding the other two constant, and recalculating the partial correlation coefficients. In this way we evaluated simple, ‘uncorrected’ associations between two matrices and partial Mantel tests to evaluate association between two matrices while controlling for

a third one. Analyses were implemented using 1000 randomizations in the program Isolation by Distance v. 1.52 (46), which reports 1-tailed probabilities. Because divergence in host use (i.e., same-host pair or different-host pair) was unrelated to the number of outliers observed (e.g., simple Mantel tests, all $p > 0.40$) further analyses with host divergence were not conducted.

Associations with bioclimatic variables. We extracted bioclimatic data from the WorldClim website (http://www.worldclim.org/). Specifically, we used 18 of the 19 provided bioclimatic variables at the resolution of 30 arc-seconds to derive two principle components (PC) axes representing climatic variation (variable 14 was excluded due to lack of variation among sites). These two axes explained 75 and 24 percent of the variance, respectively. PC axis loadings from the climatic data were as follows (PC1 score, PC2 score, respectively): BIO1 = Annual Mean Temperature, -0.98, 0.18; BIO2 = Mean Diurnal Range (Mean of monthly (max temp - min temp)), -0.30, 0.93; BIO3 = Isothermality (BIO2/BIO7) (* 100), -0.93, -0.36; BIO4 = Temperature Seasonality (standard deviation *100), 0.94, 0.34; BIO5 = Max Temperature of Warmest Month, 0.16, 0.98; BIO6 = Min Temperature of Coldest Month, 1.0, -0.03; BIO7 = Temperature Annual Range (BIO5-BIO6), 0.85, 0.54; BIO8 = Mean Temperature of Wettest Quarter, -1.0, 0.02; BIO9 = Mean Temperature of Driest Quarter, -0.66, 0.74; BIO10 = Mean Temperature of Warmest Quarter, -0.65, 0.76; BIO11 = Mean Temperature of Coldest Quarter, -1.0,-0.02; BIO12 = Annual Precipitation, 0.98, 0.20; BIO13 = Precipitation of Wettest Month, 0.94, 0.33; BIO14 = Precipitation of Driest Month, excluded; BIO15 = Precipitation Seasonality (Coefficient of Variation), -0.71, 0.69; BIO16 = Precipitation of Wettest Quarter, 0.96, 0.29; BIO17 = Precipitation of Driest Quarter, 0.94, -0.29; BIO18 = Precipitation of Warmest Quarter, 0.98, -0.17; BIO19 = Precipitation of Coldest Quarter, 0.95, 0.30.

PC1 loaded strongly for most variables representative of temperature (negative loadings) and precipitation (positive loadings). PC1 is thus was a general temperature and precipitation PC, with high scores being indicative of cold and wet climates. PC2 loaded strongly for variables associated with climatic variability (e.g., mean diurnal range) and is thus indicative of climatic stability, with high scores indicating more variable climates. We examined climatic divergence between sites (i.e., the difference between two sites in their PC scores) and found that once the geographic distance between populations was controlled for, divergence in climatic PCs was not significantly positively correlated with number of outliers observed (all relationships negative or $p > 0.15$, partial Mantel tests). We then use ANCOVA analyses to test for associations between allele

158 frequencies within populations at different classes of loci (Table S6 for details) and four factors:
159 climate PC1, climate PC2, longitude (these three variables were covariates) and host (factor). For
160 these analyses, we examined 16 outlier loci that were most highly replicated across population pairs
161 (i.e., those which appeared in three or more different host comparisons, three or more same host
162 comparisons, and the one locus which appeared in more than one comparison between adjacent
163 pairs, Table S4 for details) and 16 randomly chosen loci.
164
165 **Approximate Bayesian Computation to test for gene flow.** We used approximate Bayesian
166 computation (ABC)[51, 52] to model historical divergence and gene flow among the 28 *Timema*
167 population pairs. Specifically, we compared two alternative models for each pair of populations: (1)
168 divergence without gene flow, and (2) divergence with gene flow. The divergence without gene
169 flow model assumes that population pairs split from an ancestral population at time t in the past, and
170 have diverged in the absence of gene flow. N_eP_1 and N_eP_2 denote the effective sizes of populations
171 one and two and t denotes the divergence time in generations. This model further assumes constant
172 mutation-scaled effective population sizes for the ancestral (Θ_A) and extant populations (Θ_1 and
173 Θ_2). The divergence with gene flow model is specified similarly, but assumes a constant rate of
174 gene flow between the diverging populations. Forward in time, gives the number of migrant
175 individuals from population one into population two (mp1) and gives the number of migrant
176 individuals from population two into population one (mp2). We used the average of these two
177 parameters in the analyses reported in the main text.
178 We specified a prior probability for each model and specified uninformative priors for the
179 model parameters: $N_eP_1 \sim$ loguniform(min = 100, max = 10^5), $N_eP_2 \sim$ loguniform(min = 100, max =
180 10^5), $N_eP_A \sim$ loguniform(min = 100, max = 10^5), $m \sim$ uniform(min = 5 x 10^{-5}, max = 0.1) , $t \sim$
181 uniform(min = 10, max = 10^6), and the mean per locus mutation rate $u \sim$ uniform(min = 10^{-8}, max =
182 10^{-6}). We allowed mutation rate to vary among loci and sampled locus-specific mutation rates from
183 a gamma distribution with a shape parameter of 2 and a rate parameter of $u/2$ [51].
184 We simulated one million sets of DNA sequences for each pair of populations using
185 coalescent methods. For each set we first sampled a model ($m = 0$ or $m > 0$) and then model
186 parameters from the appropriate priors. We used the software ms [53] to simulate sequence data
187 according to the sampled model and parameters. We simulated sequence data to match the sampling
188 of the observed sequence data. Specifically, we simulated $2n$ (i.e., twice the number of sampled
189 individuals) gene copies per population and locus and subsampled the simulated sequences with

replacement to match the number of observed sequences per individual and locus. This procedure appropriately incorporates genotype uncertainty and missing data associated with high-throughput DNA sequencing.

We calculated 27 summary statistics to describe genetic variation in the simulated and observed sequence data. These statistics are the mean, variance and skew across the sequence loci of the following metrics: expected heterozygosity in population one, expected heterozygosity in population two, Nei's G_{ST} [54], the number of segregating sites in the sample, the mean number of differences between pairs of sequences in population one (π) [55], the mean number of differences between pairs of sequences in population two, the net nucleotide difference between populations one and two , the number of haplotypes unique to population one, and the number of haplotypes unique to population two (π^{net}) [56]. These metrics are informative about genetic diversity, divergence time, and gene flow. We performed an orthogonal transformation of the summary statistics to reduce their dimensionality [57]. We retained the seven scaled and rotated summary statistics with the highest eigenvalues for inference. These new summary statistics explained 89.7% of the variation in the original 27 summary statistics and were correlated with the model parameters in simulated data sets.

We used data from 100 loci (here locus = contig and can include more than one SNP) to make the ABC analysis computationally feasible. Data from a 100 loci should be sufficient for our goal of estimating average genome-wide rates of gene flow, and is a greater number of loci than used in many population genetic studies of gene flow. Moreover, summary statistics calculated based on a haphazardly selected subset of 100 loci were highly correlated (r = 0.87) with those calculated on the full data set, and thus using 100 loci, rather than more, likely had little effect on our results. Simulation for ABC analysis using 100 loci required approximately 30,000 CPU hours on a Linux computer cluster.

We estimated model posterior probabilities by treating the evolutionary model as a categorical variable and using a multinomial logit model to estimate P ($\gamma = j \mid S^* = s^*$) [58]. Model posterior probabilities are based on the 1% of simulations with summary statistics closest to the summary statistics for the observed data (s*). We computed posterior parameter estimates using local, weighted multivariate regression based on the same 1% of simulations [51]. We used the Bayesian Model Averaging technique to estimate model parameters while incorporating uncertainty in the evolutionary model (i.e., m = 0 or m > 0). We log transformed all model parameters prior to analysis and summarized posterior densities based on the median and 95% equal-tail probability

intervals. The ABC analyses were conducted using ms, custom Perl scripts, the GNU Scientific Library, standard R functions, and the *postpr* and *abc* functions from the R package 'abc' [52, 53, 59]. In the analyses reported in the text we used the mean of mp1 (median gene flow parameter $4N_em$ from population 1 into population 2) and mp2 (median gene flow parameter $4N_em$ from population 2 into population 1) as our estimate of overall gene flow. Highly congruent results were obtained using other estimates of gene flow.

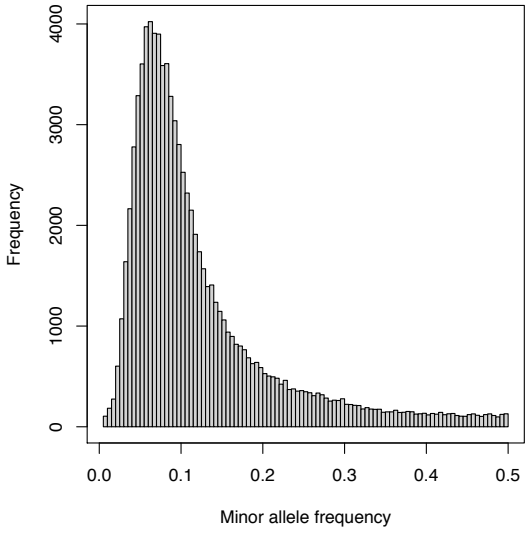
Transplantation data. Insects were transplanted to one general area (approximately an 1 km^2 area surrounding N34 30.958 W119 48.050), but from two populations that varied in their distance from the transplant site. The first transplant involved the population R12C, which was also considered in the genomic data, and is roughly 25km from the transplantation site. This is similar to the maximum distance in our genomic sampling, and thus represents a 'distant' transplant. The second transplant involved the population FHA (N34 31.089 W119 48.166), which is <1km from the transplant site, thus representing a 'near' transplant. To increase the similarity of the age distributions of individuals between the two transplants, each transplant was conducted at a slightly different time period. Because survival was monitored soon after release (eight days later, and this species typically lives for months in the lab), differences between transplants likely reflect differences in general survival ability, although further experiments are required to confirm this. Individuals for the distant transplant were collected on March 27th, 2011. Individuals were transplanted on March 29th onto both *Ceanothus* and *Adenostoma.* A paired blocks design was used (n = 5 blocks), with insects released onto one plant individual of each host species per block. In all instances, 50 *T. cristinae* were released per plant individual. Survivors were recaptured on April 6th using previously published protocols, which have been shown to be efficient at recapturing all survivors, or very near so [29, 30, 32]. Individuals for the near transplant were collected on April 14th, released on April 16th, and recaptured on April 24th. All other procedures were as for the distant transplant. For all individuals from both transplants a tissue sample (from middle leg, left dorsal side of the body) was taken as part of another study.

ESM references

44. Ersts, P.J., *Geographic Distance Matrix Generator (version 1.2.3)* [Internet], American Museum of Natural History, Center for Biodiversity and Conservation.

45. Corp., L.S., *Using multiplex identifier (MID) adaptors for the GS FLX Titanium chemistry-extended MID set. Tech. rep., Technical Bulletin: Genome Sequencer FLX System.* . 2009.
46. Li, J.H., et al., *An algorithmic model for constructing a linkage and linkage disequilibrium map in outcrossing plant populations.* Genetics Research, 2009. **91**(1): p. 9-21.
47. Foll, M. and O. Gaggiotti, *A genome-scan method to identify selected loci appropriate for both dominant and codominant Markers: a Bayesian perspective.* Genetics, 2008. **180**(2): p. 977-993.
48. Nicholson, G., et al., *Assessing population differentiation and isolation from single-nucleotide polymorphism data.* Journal of the Royal Statistical Society Series B-Statistical Methodology, 2002. **64**: p. 695-715.
49. Falush, D., M. Stephens, and J.K. Pritchard, *Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies.* Genetics, 2003. **164**(4): p. 1567-1587.
50. Guo, F., D.K. Dey, and K.E. Holsinger, *A Bayesian Hierarchical Model for Analysis of Single-Nucleotide Polymorphisms Diversity in Multilocus, Multipopulation Samples.* Journal of the American Statistical Association, 2009. **104**(485): p. 142-154.
51. Beaumont, M.A., W.Y. Zhang, and D.J. Balding, *Approximate Bayesian computation in population genetics.* Genetics, 2002. **162**(4): p. 2025-2035.
52. Csillery, K., et al., *Approximate Bayesian Computation (ABC) in practice.* Trends in Ecology & Evolution, 2010. **25**(7): p. 410-418.
53. Hudson, R.R., *Generating samples under a Wright-Fisher neutral model of genetic variation.* Bioinformatics, 2002. **18**(2): p. 337-338.
54. Nei, M., *Analysis of gene diversity in subdivided populations.* Proceedings of the National Academy of Sciences of the United States of America, 1973. **70**(12): p. 3321-3323.
55. Tajima, F., *Evolutionary relationships of DNA-sequences in finite populations.* Genetics, 1983. **105**(2): p. 437-460.
56. Nei, M. and W.H. Li, *Mathematical model for studying genetic variation in terms of restriction endonucleases.* Proceedings of the National Academy of Sciences of the United States of America, 1979. **76**(10): p. 5269-5273.
57. Bazin, E., K.J. Dawson, and M.A. Beaumont, *Likelihood-Free Inference of Population Structure and Local Adaptation in a Bayesian Hierarchical Model.* Genetics, 2010. **185**(2): p. 587-602.
58. Beaumont, M.A., in *Simulations, genetics and human prehistory*, S. Matsumura, P. Forster, and C. Renfrew, Editors. 2008, McDonald Institute for Archaeological Research: Joint determination of topology, divergence time and immigration in population trees. p. 135-154.
59. Galassi, M., et al., *{GNU} Scientific Library: Reference Manual*, 2009, Network Theory Limited.

60. Nosil, P., S.P. Egan, and D.J. Funk, *Heterogeneous genomic differentiation between walking-stick ecotypes: "Isolation by adaptation" and multiple roles for divergent selection.* Evolution, 2008. **62**(2): p. 316-336.



290

291 **Figure S1.** The minor allele frequency spectrum.

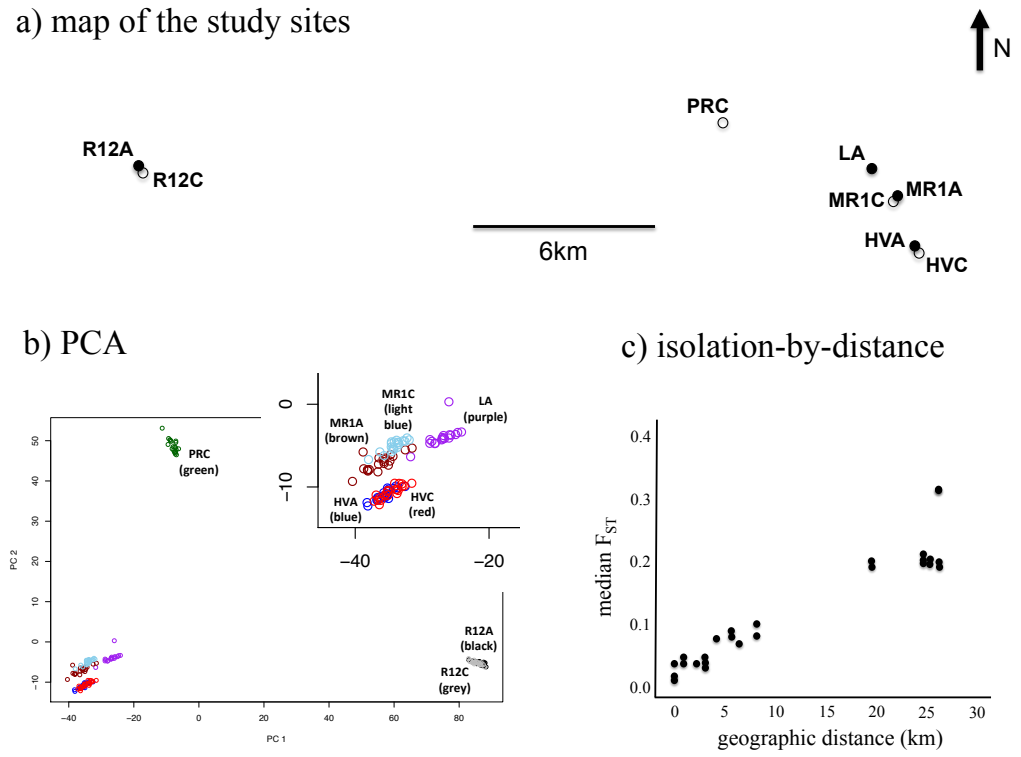


Figure S2. Location of the study sites and genetic divergence between them. a) Map of the eight study populations. Table S1 for GPS coordinates. b) Principal components analysis of the eight populations based upon 86,130 SNPs. The first two principal components explained 18% and 6% of the variation in the data. c) The relationship between genetic and geographic distance (i.e., isolation-by-distance, r = 0.96 and 0.89 for Mantel tests on raw and log-transformed distances respectively, both p = 0.001).

Table S1. Description of the natural study populations. C = *Ceanothus spinosus*. A = *Adenostoma fasciculatum*. Pop. = population (numbers used in figures are given in parentheses). 'n' refers to the number of individual specimens sequenced. Mean coverage in the genomic data is also provided for each population.

Pair	Scale	Pop.	n	Host	Latitude (N)	Longitude (W)	Mean coverage
1.	Adjacent patches	MR1C (3C)	20	C	34 30.859	119 47.986	24.07x
		MR1A (3A)	20	A	34 30.872	119 47.988	34.46x
2.	Adjacent patches	HVC (2C)	21	C	34 29.309	119 47.180	33.02x
		HVA (2A)	20	A	34 29.305	119 47.191	30.36x
3.	Adjacent patches	R12C (1C)	21	C	34 30.902	120 04.267	31.53x
		R12A (1A)	20	A	34 30.899	120 04.275	35.50x
4.	Separated patches	PRC (5C)	20	C	34 32.000	119 51.458	15.88x
		LA (4A)	19	A	34 30.464	119 47.694	14.93x

Table S2. Characteristics of the 28 pairwise population comparisons of *T. cristinae*. 'SH' = 'same host' and refers to use of the same (1) versus different (0) host species. 'Gdis' is the geographical distance between a population pair (the three directly adjacent pairs are denoted in bold). Median F_{ST} estimates (and 95% credible intervals in parentheses) are shown from a Bayesian F_{ST} model using the following parameters: 25,000 MCMC steps, thin = 10, burn-in = 2,000. Variance (Var.), kurtosis and skew were calculated using the distribution of F_{ST} values among individual loci. # outliers = "# outs" is the number of outlier loci, defined as those whose probability exceeded the 95th quantile of the genome distribution. Also given are the differences between population pairs in PC scores generated from climatic variables. PC1 is an index of temperature and precipitation and PC2 of climatic variability (see methods for details and PC axis loadings). Also shown are the results of ABC analyses testing for gene flow. Shown are the posterior probabilities for a model with zero gene flow (PP(m=0)) and the median estimates of migration parameters mp1 and mp2 (methods for details). Of the comparison below, six were also included in a past study using AFLP genotypes (F_{ST} at AFLPs was as follows: PRC x HVA = 0.048, PRC x HVC = 0.119, HVA x LA = 0.022, PRC x LA = 0.057, HVC x HVA = 0.063, HVC x LA = 0.024)[60]. The correlation between F_{ST} values calculated from SNPs versus AFLPs was moderate (r = 0.66).

Pop. pair	SH	Gdis (km)	F_{ST}	Var.	kurtosis	skew	# outs	PC1	PC2	PP (m=0)	mp1 median	mp2 median
1. hva x hvc	0	0.02	0.013 (0.012 – 0.013)	0.0000	73.96	7.17	140	0.00	0.00	0.00	36.03	185.90
2. hva x la	1	2.28	0.031 (0.031 – 0.032)	0.0000	38.59	4.62	34	1.63	1.39	0.00	86.84	213.51
3. hva x mr1a	1	3.15	0.040 (0.039 – 0.041)	0.0001	23.93	3.94	67	1.76	1.53	0.00	146.17	56.92

4. hva x mr1c	0	3.13	0.028 (0.027 – 0.029)	0.0000	32.33	4.52	84	1.76	1.53	0.00	122.99	125.71
5. hva x prc	0	8.22	0.073 (0.072 – 0.074)	0.0003	26.74	3.98	247	0.51	0.00	0.00	41.59	49.00
6. hva x r12a	1	26.29	0.306 (0.303 – 0.310)	0.0047	6.34	2.29	1329	0.53	2.31	0.99	0.32	0.49
7. hva x r12c	0	26.28	0.186 (0.184 – 0.188)	0.0049	6.30	2.29	1364	0.53	2.31	0.99	0.23	0.39
8. hvc x la	0	2.28	0.031 (0.030 – 0.031)	0.0000	39.27	4.67	34	1.63	1.39	0.00	140.17	281.59
9. hvc x mr1a	0	3.15	0.039 (0.038 – 0.040)	0.0001	20.92	3.82	68	1.76	1.53	0.00	186.50	38.78
10. hvc x mr1c	1	3.13	0.032 (0.031 – 0.033)	0.0000	39.58	4.79	90	1.76	1.53	0.00	205.45	145.78
11. hvc x prc	1	8.23	0.094 (0.093 – 0.097)	0.0003	25.41	3.88	229	0.51	0.00	0.00	98.12	38.12
12.	0	26.30	0.189	0.0046	6.37	2.30	1294	0.53	2.31	0.99	0.67	0.43

hvc x r12a			(0.187 – 0.190)									
13. hvc x r12c	1	26.29	0.185 (0.183 – 0.187)	0.0048	6.32	2.29	1278	0.53	2.31	0.98	0.28	0.31
14. la x mr1a	1	0.88	0.041 (0.040 – 0.042)	0.0000	25.22	4.03	42	0.13	0.14	0.00	185.95	28.64
15. la x mr1c	0	0.86	0.033 (0.032 – 0.033)	0.0000	39.27	4.86	60	0.13	0.14	0.00	171.10	66.51
16. la x prc	0	6.42	0.063 (0.063 – 0.065)	0.0004	30.64	4.43	265	1.12	1.40	0.01	67.64	20.50
17. la x r12a	1	25.36	0.195 (0.913 – 0.197)	0.0038	6.35	2.28	943	2.16	0.91	0.96	0.34	0.09
18. la x r12c	0	25.35	0.191 (0.189 – 0.193)	0.0040	6.31	2.28	954	2.16	0.91	0.93	0.41	0.15
19. mr1a x mr1c	0	0.02	0.030 (0.029 – 0.031)	0.0000	30.51	4.62	62	0.00	0.00	0.00	82.87	213.86
20. mr1a	0	5.70	0.083 (0.082	0.0002	23.24	3.67	173	1.25	1.54	0.00	16.19	54.11

x prc			– 0.084)									
21. mr1a x r12a	1	24.90	0.200 (0.198 – 0.202)	0.0042	6.53	2.32	1296	2.29	0.77	0.99	0.24	0.37
22. mr1a x r12c	0	24.89	0.197 (0.195 – 0.198)	0.0044	6.48	2.31	1328	2.29	0.77	0.99	0.24	0.37
23. mr1c x prc	1	5.71	0.075 (0.074 – 0.076)	0.0002	24.31	3.81	181	1.25	1.54	0.99	0.24	0.37
24. mr1c x r12a	0	24.90	0.197 (0.195 – 0.198)	0.0040	6.54	2.31	1209	2.29	0.77	0.99	0.24	0.37
25. mr1c x r12c	1	24.89	0.193 (0.191 – 0.195)	0.0042	6.50	2.31	1222	2.29	0.77	1.00	0.24	0.22
26. prc x r12a	0	19.70	0.193 (0.191 – 0.195)	0.0017	8.18	2.44	414	1.04	2.31	0.28	7.44	2.83
27. prc x r12c	1	19.69	0.189 (0.187 – 0.191)	0.0018	8.01	2.43	447	1.04	2.31	0.21	6.17	4.32
28. r12a x 12c	0	0.01	0.007 (0.007 –	0.0000	1022.62	20.13	353	0.00	0.00	0.00	102.92	642.00

			0.007)									

Table S3. Results of linkage disequilibrium analyses. Shown is the mean and s.d. of Burrow's composite measure of Hardy Weinberg and linkage disequilibrium (Δ) within each study population. Results are shown for all loci and for the 15,207 loci that were categorized as statistical outliers, as well as for SNPs within the same contig versus those from different contigs.

population	mean all loci	mean outlier loci	mean for SNPs in same contig	mean for SNPs in different contigs
HVA	0.00377	0.00341	0.00766	0.00370
HVC	0.00354	0.00314	0.00721	0.00346
LA	0.00302	0.00288	0.00547	0.00300
MR1A	0.00401	0.00339	0.00812	0.00398
MR1C	0.00341	0.00303	0.00714	0.00339
PRC	0.00264	0.00446	0.00526	0.00263
R12A	0.00479	0.00153	0.01370	0.00479
R12C	0.00425	0.00123	0.01254	0.00432

Table S4. Simple and partial Mantel tests for associations between number of outlier loci (or various other aspects of the F_{ST} distribution) and geographic distance / the geographic arrangement of populations (geographically separated versus geographically adjacent; separated/adjacent – coded as 0 and 1, respectively). See Table S2 for values of F_{ST}, kurtosis and skewness for each pairwise population comparison. Significant results are in bold.

Test	Raw data		Log transformed	
	r	p	r	p
Number of outlier loci				
Separated/Adjacent (simple)	-0.24	0.900	-0.17	0.82
Geographic Distance (simple)	0.94	**0.002**	0.59	**0.004**
Separated/Adjacent (partial)	0.95	**0.001**	0.88	**0.001**
Geographic Distance (partial)	0.45	**0.001**	0.92	**0.001**
Mean F_{ST}				
Separated/Adjacent (simple)	-0.39	**<0.001**	-0.61	**<0.001**
Geographic Distance (simple)	0.96	**0.001**	0.89	**0.001**
Separated/Adjacent (partial)	-0.06	0.63	0.77	**0.006**
Geographic Distance (partial)	0.95	**0.001**	0.93	**0.006**
Variance in F_{ST}				
Separated/Adjacent (simple)	-0.30	**0.023**	-0.35	**<0.001**
Geographic Distance (simple)	0.97	**0.001**	0.73	**0.001**
Separated/Adjacent (partial)	0.37	**0.003**	0.88	**0.001**
Geographic Distance (partial)	0.97	**0.001**	0.94	**0.001**
Kurtosis				
Separated/Adjacent (simple)	0.59	**0.005**	0.62	**0.010**
Geographic Distance (simple)	-0.30	**0.000**	-0.85	**0.002**
Separated/Adjacent (partial)	0.54	**0.025**	-0.47	**0.013**

Geographic Distance (partial)	-0.07	0.452	-0.80	**0.004**
Skewness				
Separated/Adjacent (simple)	0.68	**0.003**	0.66	**0.003**
Geographic Distance (simple)	-0.52	**0.000**	-0.87	**0.001**
Separated/Adjacent (partial)	0.61	**0.013**	-0.42	**0.020**
Geographic Distance (partial)	-0.38	**0.012**	-0.80	**0.002**

14

15

16

17

Table S5. Number of outliers as a function of the number of pairwise comparisons the locus was found to be an outlier in, for different classes of loci. 'All outliers' refers to each and any locus that was an outlier. 'Adjacent-pair outliers' refers to outliers found in geographically adjacent population pairs. 'Adjacent-pair specific outliers' refers to outliers only between directly-adjacent population pairs (i.e., those which were observed in geographically adjacent pairs, but were never outliers in comparisons between geographically separated pairs). 'Different-host outliers' refers to outliers found in different-host population pairs. 'Different-host specific outliers' refers to outliers only in different-host population pairs. 'Same-host outliers' refers to outliers found in same-host population pairs. 'Same-host specific outliers' refers to outliers only in same-host population pairs. Values of '0' are used to denote cases where zero outliers were found in a possible number of comparisons whereas dashes ('-') indicate cases beyond the maximum number of pairwise comparisons of a particular type (e.g., there were only three possible comparisons between adjacent pairs and only 12 possible same-host comparisons).

Number of comparisons	All outliers	Adjacent-pair outliers	Adjacent-pair specific outliers	Different-host outliers	Different-host specific outliers	Same-host outliers	Same-host specific outliers
1	3883	553	492	1109	470	1001	320
2	2621	1	1	620	47	650	43
3	2029	0	0	523	11	446	2
4	1738	-	-	361	1	366	1
5	1365	-	-	341	0	336	0
6	1111	-	-	64	0	59	0
7	815	-	-	5	0	3	0
8	711	-	-	1	0	0	0
9	454	-	-	0	0	0	0
10	347	-	-	0	0	0	0
11	76	-	-	0	0	0	0
12	50	-	-	0	0	0	0
13	3	-	-	0	0	-	-

14	3	-	-	0	0	-	-
15	1	-	-	0	0	-	-
16	0	-	-	0	0	-	-

Table S6. Results of ANCOVA analyses testing for associations between allele frequencies within populations at different classes of loci and four factors: climate PC1, climate PC2, longitude (all covariates) and host (factor). See methods for details. DHS = 'different-host specific' outliers. SHS = 'same-host specific' outliers. APS = 'adjacent-pair' specific outlier. Random = randomly chosen loci. Parentheses denote the number of pairwise comparisons that a locus was an outlier in. Note that more significant relationships (denoted in bold text) were found for outlier than for randomly chosen loci. Principle components analysis on allele frequencies for the 12 DHS outliers listed below resulted in two PC axes that explained 69 and 26% of allele frequency variance, respectively. Both PCs revealed clines overlaid on host effects. For example, PC1 was related to both longitude and host (both $p < 0.05$, ANCOVA). PC2 was related to longitude, both climate PCs, and host (all $p < 0.05$, ANCOVA).

locus		climate PC1		climate PC2		longitude		host	
Type	number	F	p	F	p	F	p	F	p
DHS(4x)	14004	0.02	0.900	0.89	0.415	0.60	0.495	3.21	0.171
DHS(3x)	11734	62.03	**0.004**	88.62	**0.003**	182.97	**0.001**	0.29	0.629
DHS(3x)	20240	0.03	0.883	1.52	0.306	126.56	**0.002**	19.49	**0.022**
DHS(3x)	34286	1.54	0.302	2.13	0.240	0.53	0.521	1.68	0.286
DHS(3x)	38276	24.68	**0.016**	26.87	**0.014**	27.29	**0.014**	3.09	0.177
DHS(3x)	43125	37.67	**0.009**	66.13	**0.004**	131.97	**0.001**	5.24	0.106
DHS(3x)	44357	1.96	0.256	2.79	0.194	4.81	0.116	5.82	0.095
DHS(3x)	46553	33.75	**0.010**	36.72	**0.009**	41.37	**0.008**	0.54	0.514
DHS(3x)	47063	12.49	**0.039**	15.61	**0.029**	3.10	0.177	1.27	0.342
DHS(3x)	47457	325.81	**0.000**	429.87	**0.000**	349.19	**0.000**	1.05	0.382
DHS(3x)	63047	123.79	**0.002**	144.35	**0.001**	75.75	**0.003**	20.97	**0.020**
DHS(3x)	84640	22.65	**0.018**	28.13	**0.013**	36.79	**0.009**	7.05	0.077
SHS(4x)	76305	58.02	**0.005**	85.50	**0.003**	18.54	**0.023**	0.92	0.409
SHS(3x)	30642	4.44	0.126	7.88	0.067	0.20	0.684	0.75	0.451
SHS(3x)	45373	11.43	**0.043**	11.83	**0.041**	47.94	**0.006**	0.26	0.645
APS(2x)	69502	0.01	0.948	0.02	0.908	0.04	0.853	1.20	0.354

Random	56	0.64	0.483	0.23	0.663	1.46	0.313	0.90	0.413
Random	1025	0.08	0.799	0.17	0.707	2.41	0.218	1.11	0.370
Random	2433	0.06	0.825	0.12	0.750	0.58	0.502	1.36	0.328
Random	10346	0.34	0.602	0.01	0.928	1.99	0.254	1.62	0.292
Random	12245	0.28	0.635	0.62	0.489	3.17	0.173	1.36	0.328
Random	26452	0.00	0.994	0.00	0.996	0.83	0.430	1.33	0.332
Random	27854	1.82	0.270	2.74	0.197	3.20	0.171	2.02	0.251
Random	35652	4.29	0.130	3.72	0.149	3.13	0.175	1.00	0.392
Random	45122	1.00	0.390	1.53	0.304	2.51	0.212	1.57	0.299
Random	46258	0.75	0.449	4.02	0.139	2.23	0.232	1.20	0.354
Random	48997	0.00	0.984	0.05	0.831	0.52	0.524	0.86	0.421
Random	58412	5.21	0.107	8.21	0.064	26.28	**0.014**	1.07	0.378
Random	59485	0.70	0.464	0.14	0.732	0.17	0.708	0.28	0.634
Random	60111	0.27	0.638	1.08	0.374	1.78	0.274	0.00	0.964
Random	66259	103.13	**0.002**	132.80	**0.001**	98.48	**0.002**	0.09	0.781
Random	74581	0.43	0.560	0.37	0.584	0.21	0.682	0.72	0.459