Supplementary information

Climate-driven changes in the composition of New World plant communities

In the format provided by the authors and unedited

Supplementary Materials for "Climate-driven changes in the composition of New World plant communities" by K.J. Feeley, C. Bravo-Avila, B. Fadrique, T.M. Perez, D. Zuleta

Supplementary Text

<u>Acknowledgments of herbaria and data providers</u>

We acknowledge the many herbaria that contributed data to this work including: ASDM, L, AD, ACAD, AMO, FURB, HPL, IAC, INPA, IPA, MBML, MT, UBC, UFS, US, USP, TRH, BRIT, MO, NCU, NY, TEX, UNCC, A, UI, AAU, GH, ANSP, ASU, B, BAA, BAF, BC, BG, BH, BIO, SEV, BOUM, BR, CANB, CAS, CEN, CHR, CICY, CIMI, COA, CP, COL, CR, CU, CTES, DAO, DAV, E, ENCB, F, FAU, FLAS, FR, FTG, CMM, G, GB, GLM, K, GZU, HAL, HAST, HBG, HNT, HO, HU, IBUG, ICN, IEB, ILL, ABH, INEGI, GAT, UCSB, ISU, SD, USF, TALL, RELC, CATA, KMN, KSTC, LAGU, KSU, UESC, KU, LA, SUU, LAF, GMDRC, LAM, LD, LI, CNH, MACF, CATIE, MGC, IRVC, JOTR, LSU, BRY, DBG, HSC, MELU, MA, CSUSB, MBM, UCSC, UCS, OBI, MEL, MEXU, FSU, MICH, BABY, SCFS, SACT, JCT, JROH, SBBG, SJSU, MNHM, MNHN, SDSU, MOR, MSC, SFV, CNS, JEPS, CIB, MUR, VIT, MU, MUB, PGM, PASA, BOON, ND, NU, NE, NHM, NMSU, NSW, O, CHSC, CHAS, CDA, OSC, OULU, P, UPS, PFC, PH, SI, POM, QMEX, TROM, UH, RM, RSA, S, SALA, SNM, TAIF, TAMU, TU, UADY, UAM, UAS, UC, UCR, UFMA, UGDA, UJAT, ULM, UME, UNM, USON, UT, UTEP, VAL, W, WELT, WU, XAL, CUVC, PERTH.

.

Supplementary Fig. S1.
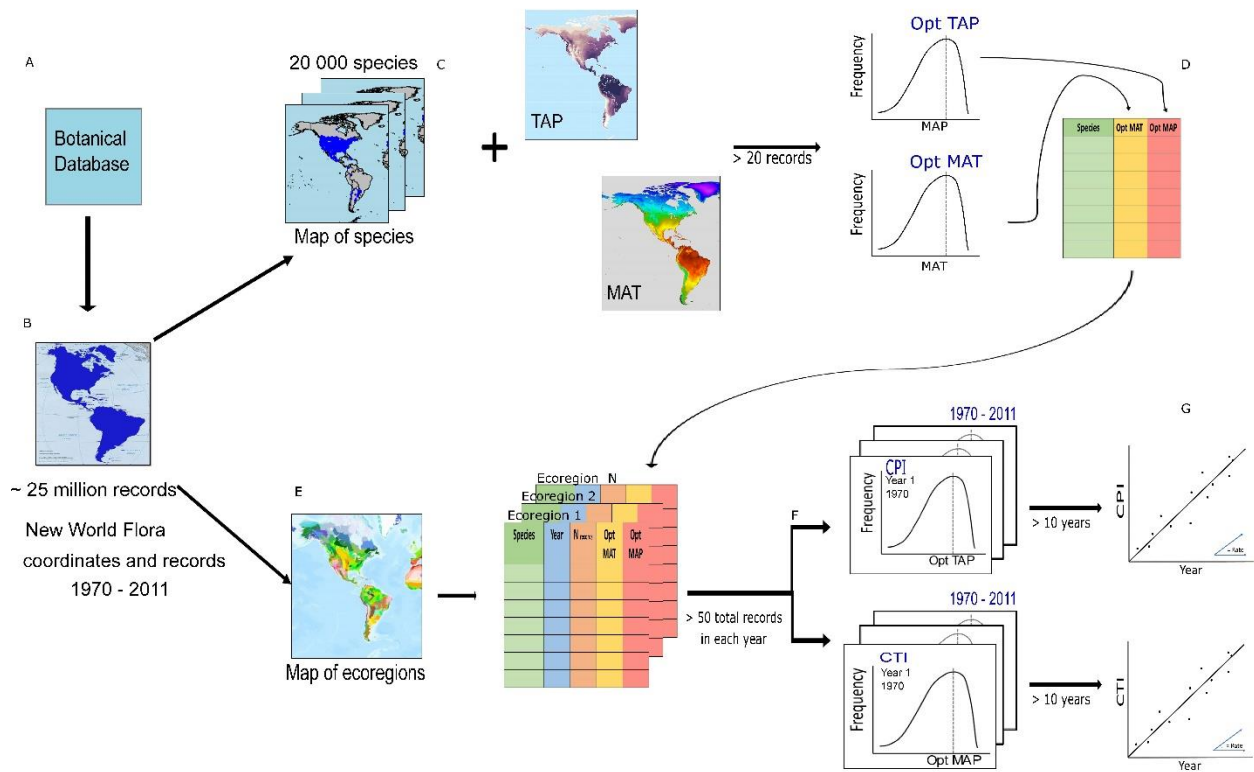
Supplementary Fig. S1 Legend. A conceptual diagram of the workflow for calculating the community climatic indices (CTI and CPI) and annualized change rates (TR and MR) of ecoregions based on the climatic distribution of observations and collections: A) All observation and collection records from the New World were downloaded from BIEN database version 4.1.1 using the BIEN package in R (http://bien.nceas.ucsb.edu/bien/). B) Collection records were filtered to include only those with georeferenced coordinates and with a year of collection between 1970 to 2011. C) Based on record coordinates, we extracted the Mean Annual Temperature (MAT, $^{\circ}$C) and Total Annual Precipitation (TAP, mm) estimates from the CHELSA extrapolated climate rasters (http://chelsa-climate.org/). D) Species' optimal Mean Annual Temperature (MAT$_{opt}$, $^{\circ}$C) and optimal Total Annual Precipitation (TAP$_{opt}$, mm) were calculated using the mean MAT and TAP, respectively, of all collections from 1970 to 1985. We only calculated MAT$_{opt}$ and TAP$_{opt}$ for species with more than 20 records in this initial 15-year period. E) The Community Temperature Index (CTI, $^{\circ}$C) and Community Precipitation Index (CPI, mm) were calculated for each New World ecoregion in each calendar year as the mean MAT$_{opt}$ and TAP$_{opt}$, respectively, of the incumbent plant species weighted by the number of records in that year within the ecoregion. CTI and CPI were only calculated for ecoregions with more than 50 total collections in the year a correction procedure was applied to account for potential sampling biases (see Supplementary Fig. S5). F) For all ecoregions with ten or more CTI/CPI estimates, Thermophilization Rates (TR, $^{\circ}$C yr^{-1}) and Mesophilization Rates (MR, mm yr^{-1}) were calculated as the slope of the linear least square regression between CTI or CPI, respectively, and year.

Supplementary Fig. S2.
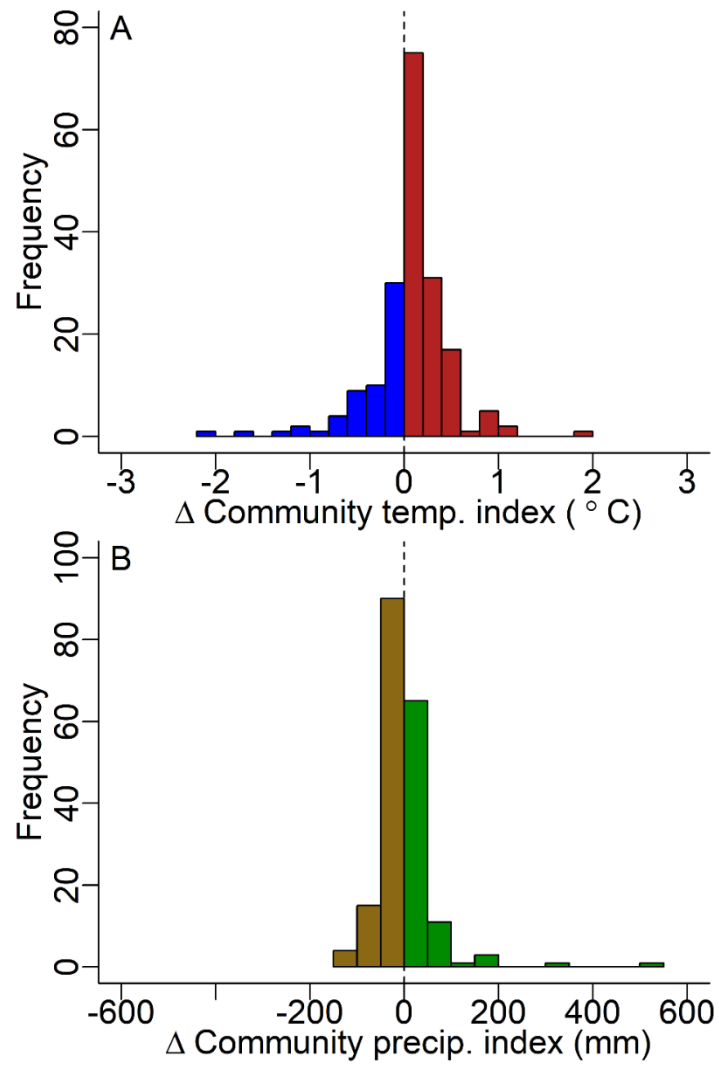
Supplementary Fig. S2 Legend. The distribution of differences in A) CTI and B) CTI between the initial (1970-1985) and final (1996-2011) communities in ecoregions.

Supplementary Fig. S3.
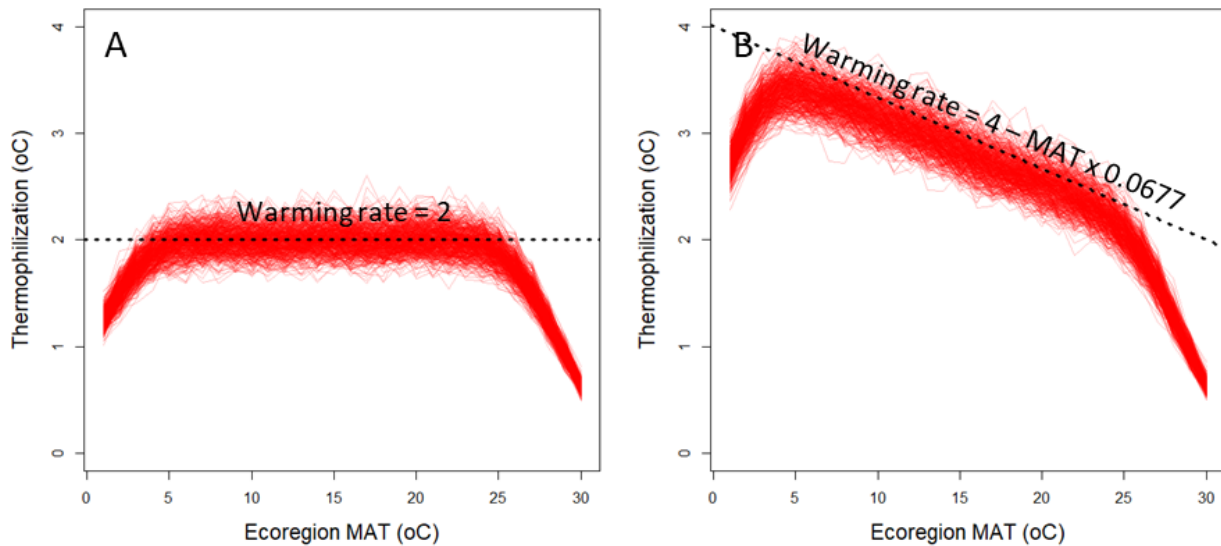
Supplementary Fig. S3 Legend. Results of a simulation illustrating the effects of niche truncation on thermophilization in ecoregions at different MATs. In the simulation, 1000 species are assigned true optimal temperatures (MATopt) that are uniformly distributed between -5 and 35°C. The species are observed at temperatures normally distributed around their MATopt (st. dev. = 1.5) but only between the thermal limits of 0-30°C (i.e., colder and hotter observations are missing due to niche truncation). The observed MATopt (MATopt.obs) is then calculated for each species on the basis of these observations and the Community Temperature Index (CTI.current) is calculated for 1°C temperature bands (i.e., ecoregions) from the MATopt.obs weighted by the number of collections per band. CTI is then recalculated under set warming rates and perfect species migrations - perfect migration is achieved in the simulation by decreasing the true MATopt of each species by the warming rates. Warming rates are indicated by the dashed lines in the figures and are simulated to be either A) constant (+2) or B) fastest in the coldest ecoregions (i.e., +4 at high/cold latitudes and +2 at low/hot latitudes). The thermophilization of each band/ecoregion is calculated as the difference between the original and final CTI. The process is reiterated 500 times. Plotting the simulation results (one line per iteration) shows that even when all species are migrating perfectly, the observed magnitude of thermophilization will be decreased at hot and cold thermal extremes due to niche truncation coupled with a lack of immigrants into the hottest areas.

```
#####R code for simulation
CTI.current = CTI.future = TR = list()

for(z in 1:1000){
 MATopt = runif(1000, -5, 35)
 collections = collections.future = list()
 MATopt.obs = c()
 for(i in 1:1000){
 collections[[i]] = rnorm(n=100, mean=MATopt[i], sd=1.5)
 MATopt.obs[i] = mean(collections[[i]][collections[[i]]<=30 & collections[[i]]>=0])
 warm.rate = 2 # for A
 warm.rate = 4 - 0.06667*MATopt[i] # for B
 collections.future[[i]] = rnorm(n=100, mean=MATopt[i]-warm.rate, sd=1.5)
 }
```

4

```
 CTI1 = CTI2 = c()
 for(i in 1:30){
 n.col = n.col.future = c()
 for(j in 1:1000){
 n.col[j] = sum(collections[[j]]<i & collections[[j]]>=(i-1) & collections[[j]]<=30 &
collections[[j]]>=0)
 n.col.future[j] = sum(collections.future[[j]]<i & collections.future[[j]]>=(i-1) &
collections.future[[j]]<=30 & collections.future[[j]]>=0)
}

 CTI1[i] = weighted.mean(MATopt.obs, w=n.col, na.rm=T)
 CTI2[i] = weighted.mean(MATopt.obs, w=n.col.future, na.rm=T)
 }

 CTI.current[[z]] = CTI1
 CTI.future[[z]] = CTI2
 TR[[z]] = CTI2-CTI1
}

#plot results
require(scales)
plot(TR[[1]], type='n', xlab="Ecoregion MAT (oC)", ylab="Thermophilization (oC)",
ylim=c(0,4), cex.lab=1.4)
for(i in 1:500){
points(TR[[i]], col=alpha('red', .25), type='l')}

abline(h=2, lty=3, lwd=3)    # for A
abline(4, -0.06667, lty=3, lwd=3)  # for B
```
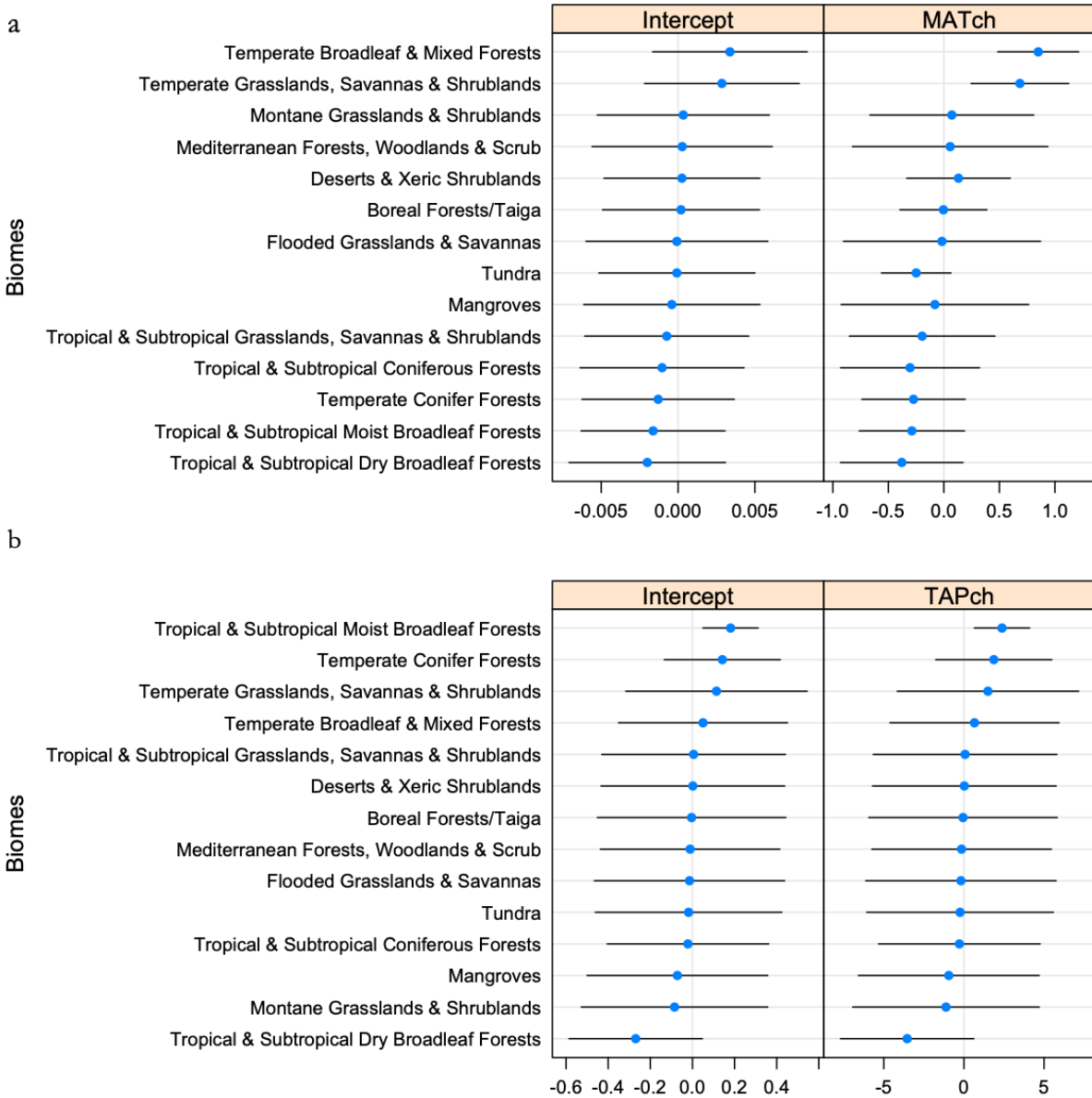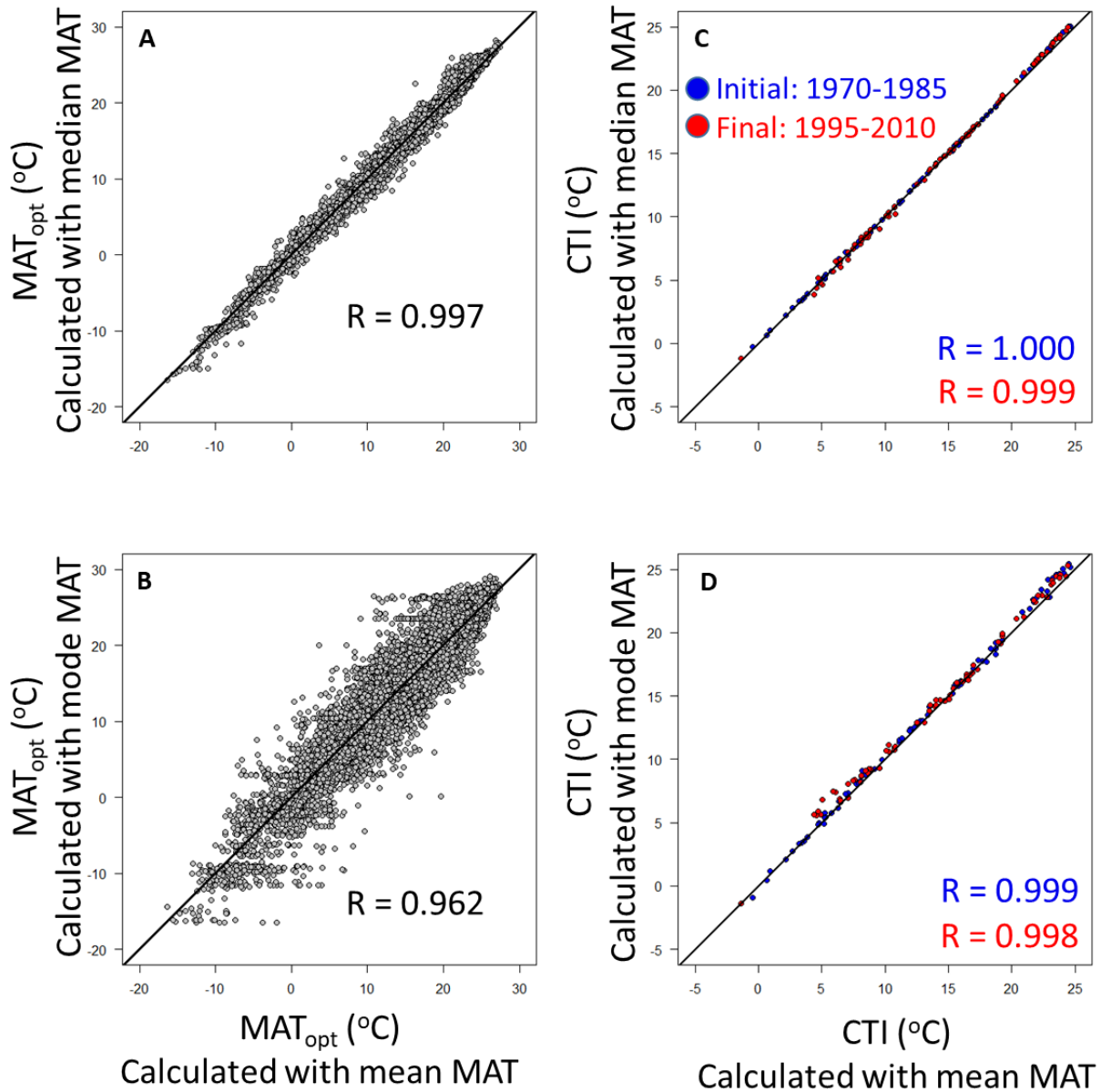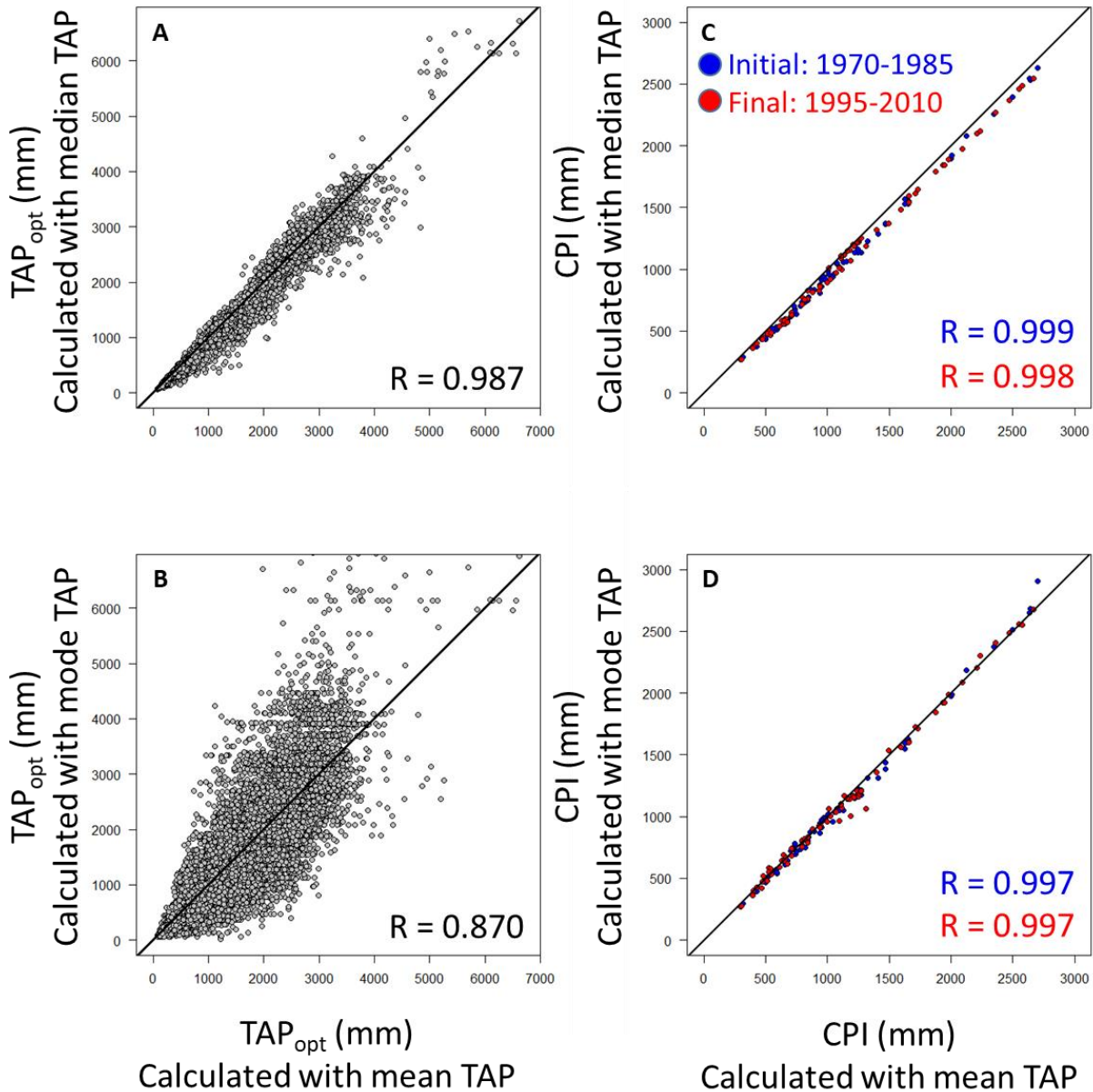
Supplementary Fig. S4.

Supplementary Fig. S4 Legend. Random intercept and slope coefficients for relationships between TR and MAT_{ch} (a) and MR and TAP_{ch} (b) for ecoregions within biomes as predicted with linear mixed-effects models (Best Linear Unbiased Prediction). Blue points represent the mean and lines are the 95% confidence interval (conditional deviation). In each panel, biomes are ordered by their random intercept value. The overall relationship between thermophilization rate and temperature change was $TR = 0.01°C + 0.38 \times MAT_{ch}$ (Parametric bootstrap 95% CI for MAT_{ch} = -0.07 - 0.81; Marginal R^2 = 0.03; Conditional R^2 = 0.20; F-value = 2.8739). The overall relationship between mesophilization rate and precipitation change was $MR = 1.46mm + 0.78 \times TAP_{ch}$ (95% CI for TAP_{ch} = -2.55 – 4.41; Marginal R^2 = 0.004; Conditional R^2 = 0.079; F-value = 0.215).

Supplementary Fig. S5.

Supplementary Fig. S5 Legend. The relationships between species' MAT_{opt} ($^{\circ}$C) calculated using the mean MAT at observation locations from 1970-1985 vs. species' MAT_{opt} calculated using A) the median and B) mode MAT at observation locations. The relationship between the Community Temperature Index (CTI, $^{\circ}$C) of ecoregions calculated using MAT_{opt} values estimated from mean observation locations vs. the CTI of ecoregions calculated using MAT_{opt} values estimated from the C) median and D) mode observation locations. In C and D, blue points indicate initial CTI values (1970-1985) and red points indicate final CTI values (1996-2011). The lines show the 1:1 relationship and the Pearson product-moment correlation coefficients for all relationships are indicated on the figure (df = 189).

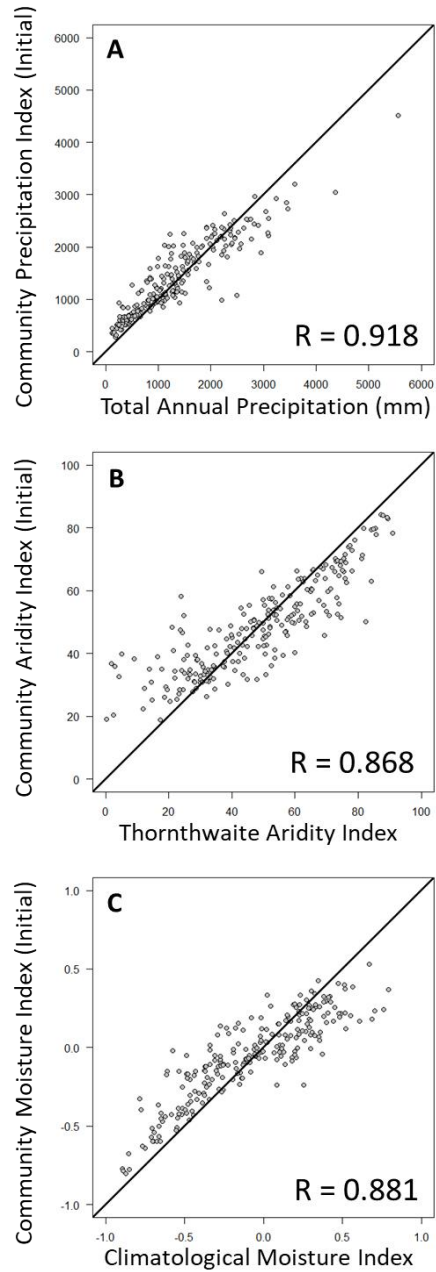**Supplementary Fig. S6.**

Supplementary Fig. S6 Legend. The relationships between species' TAP$_{opt}$ (mm) calculated using the mean TAP at observation locations from 1970-1985 vs. species' TAP$_{opt}$ calculated using A) the median and B) mode TAP at observation locations. The relationship between the Community precipitation Index (CPI, mm) of ecoregions calculated using TAP$_{opt}$ values estimated from mean observation locations vs. the CPI of ecoregions calculated using TAP$_{opt}$ values estimated from the C) median and D) mode observation locations. In C and D, blue points indicate initial CPI values (1970-1985) and red points indicate final CPI values (1996-2011). The lines show the 1:1 relationship and the Pearson product-moment correlation coefficients for all relationships are indicated on the figure (df = 189).

Supplementary Fig. S7.

Supplementary Fig. S7 Legend. The relationships between different indicators of water availability and corresponding community climate indices: A) the mean TAP vs. the Community Precipitation Index (mm), B) the mean Thornthwaite Aridity Index vs. the Community Aridity Index (unitless, 0 – 100), and C) the mean Climatological Moisture Index vs. the Community Moisture Index (unitless, -1 – 1). The lines show the 1:1 relationship and the Pearson product-moment correlation coefficients for all relationships are indicated on the figure. The Thornthwaite Aridity Index and Community Aridity Index were not available for all ecoregions (df = 172).
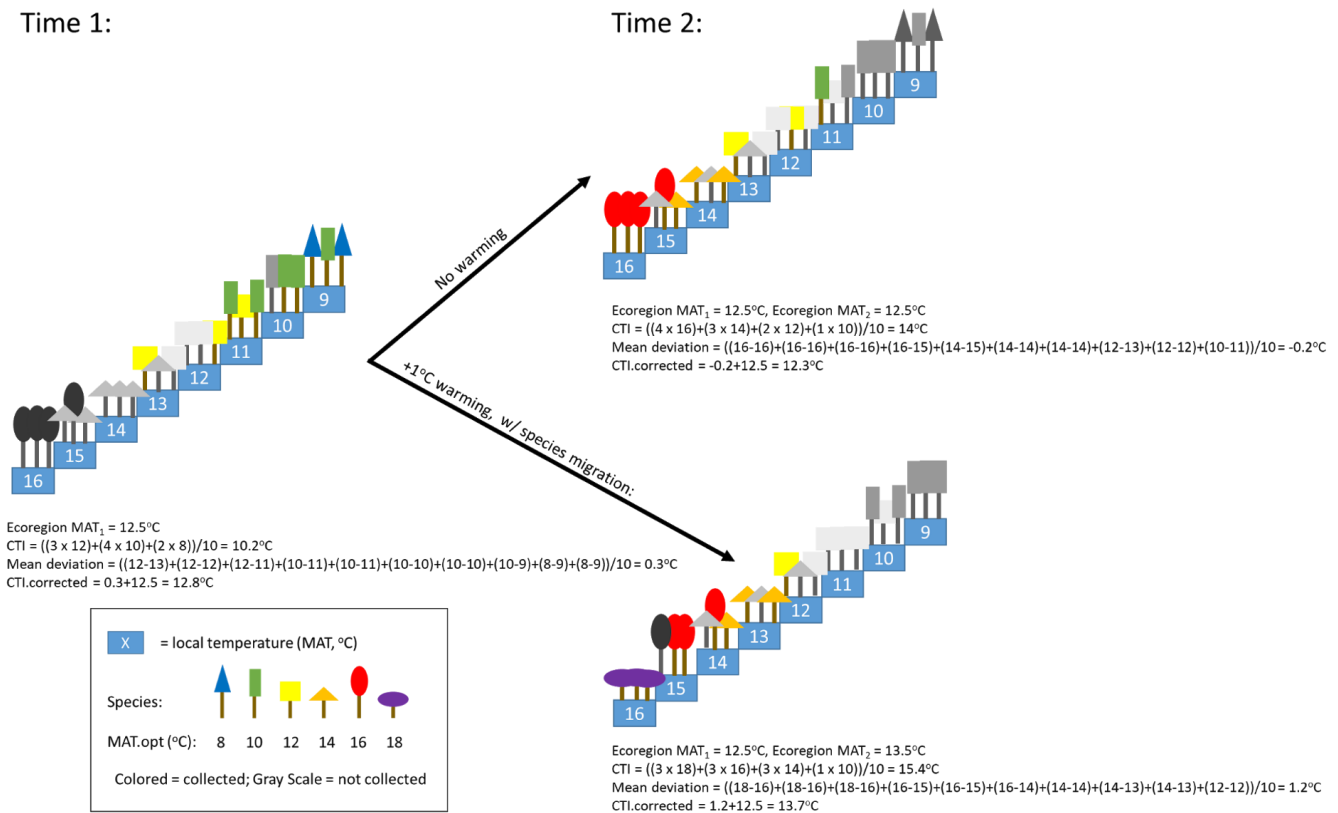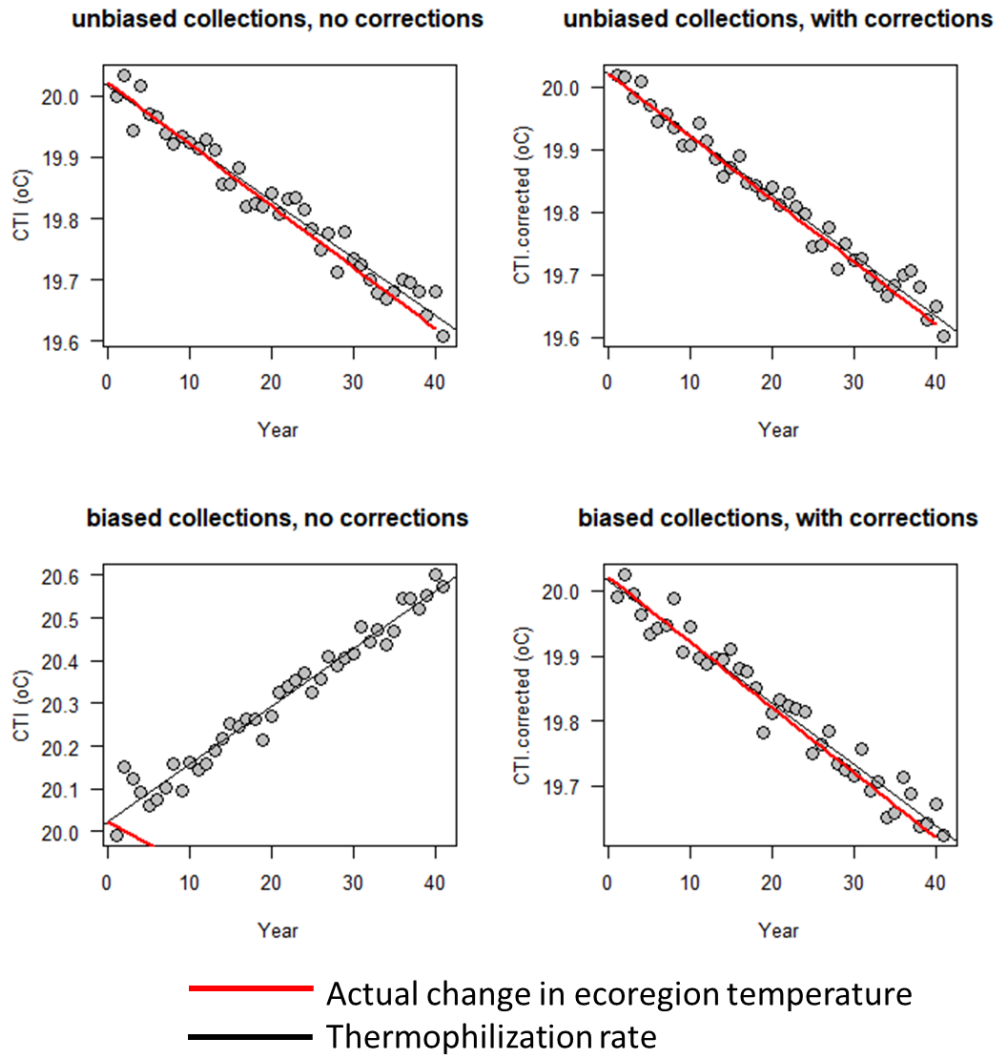
Supplementary Fig. S8.

Time 1:

Time 2:

Ecoregion MAT_1 = 12.5°C, Ecoregion MAT_2 = 12.5°C
CTI = ((4 x 16)+(3 x 14)+(2 x 12)+(1 x 10))/10 = 14°C
Mean deviation = ((16-16)+(16-16)+(16-16)+(16-15)+(14-15)+(14-14)+(14-14)+(12-13)+(12-12)+(10-11))/10 = -0.2°C
CTI.corrected = -0.2+12.5 = 12.3°C

Ecoregion MAT_1 = 12.5°C
CTI = ((3 x 12)+(4 x 10)+(2 x 8))/10 = 10.2°C
Mean deviation = ((12-13)+(12-12)+(12-11)+(10-11)+(10-11)+(10-10)+(10-10)+(10-9)+(8-9)+(8-9))/10 = 0.3°C
CTI.corrected = 0.3+12.5 = 12.8°C

X	= local temperature (MAT, °C)

Species:

MAT.opt (°C):   8   10   12   14   16   18

Colored = collected; Gray Scale = not collected

Ecoregion MAT_1 = 12.5°C, Ecoregion MAT_2 = 13.5°C
CTI = ((3 x 18)+(3 x 16)+(3 x 14))/10 = 15.4°C
Mean deviation = ((18-16)+(18-16)+(18-16)+(16-15)+(16-15)+(16-14)+(14-14)+(14-13)+(14-13)+(12-12))/10 = 1.2°C
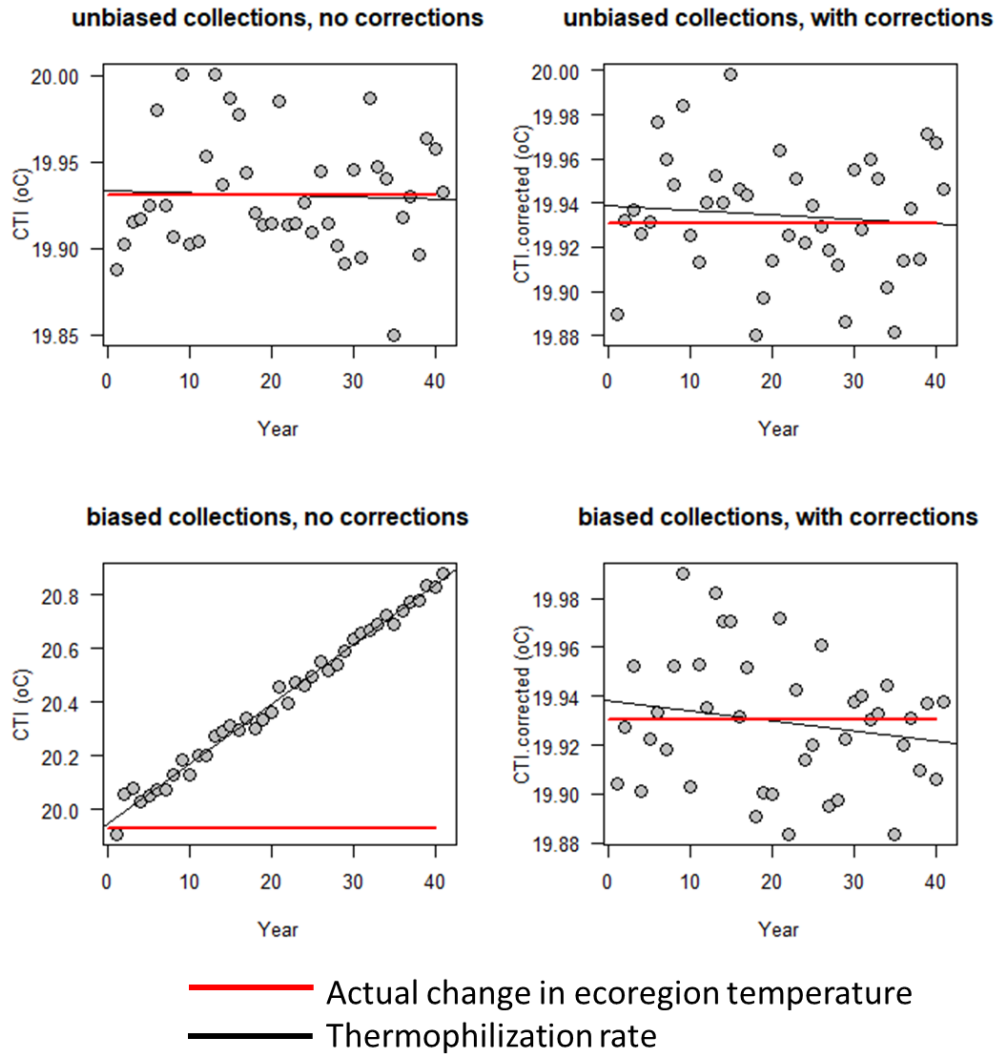CTI.corrected = 1.2+12.5 = 13.7°C

Supplementary Fig. S8 Legend. Diagram depicting the effects of biased sampling on the Community Temperature Index with and without the application of the correction procedure (CTI and CTI.corrected, respectively) under scenarios with and without warming (and corresponding species migrations). The first, left-hand, panel depicts a single hypothetical ecoregion with 8 grid cells distributed evenly across a thermal gradient from 9 to 16°C. The local MAT of each grid cell is shown in the blue boxes. Each grid cell is populated with a community of plant species with different optimal temperatures (MAT.opt). Species with low MAT.opt are more likely to occur in cold cells; species with high MAT.opt are more likely to occur in hot cells. Ten of the plants are sampled (colored) and the CTI is calculated as the mean of the sampled plants' MAT.opts. In this case, sampling is biased towards cold cells so the CTI is low (10.2°C) relative to the actual mean ecoregion temperature (12.5°C). To correct for sampling bias, we 1) subtract the local grid cell temperature from each plant's MAT.opt, 2) take the average of the temperature differences/deviations for sampled plants and 3) add the mean deviation to the mean ecoregion temperature. The CTI.corrected (12.8°C) is close to the actual mean ecoregion temperature. The two right-hand panels show how CTI would change in the ecoregion if collection biases shift towards hotter areas when there is no warming (top) vs. when there is warming of +1.0°C and a corresponding shift in species' ranges (bottom). In both cases, CTI and CTI.corrected are calculated following the same procedures as above. Importantly, when calculating CTI.corrected, the original grid cell temperatures and original mean ecoregion temperature are still used even when temperatures change through time. In both cases, the collection bias towards hot areas inflates CTI but the CTI.corrected accurately represents the ecoregion temperatures and the amount of warming/thermophilization.

10

Supplementary Fig. S9.

A) Warming rate = -0.01oC yr-1

B) Warming rate = 0.0oC yr-1

unbiased collections, no corrections

unbiased collections, with corrections

biased collections, no corrections

biased collections, with corrections

Actual change in ecoregion temperature
Thermophilization rate

C) Warming rate = 0.01oC yr-1

unbiased collections, no corrections

unbiased collections, with corrections

biased collections, no corrections

biased collections, with corrections

Actual change in ecoregion temperature
Thermophilization rate

Supplementary Fig. S9 Legend.
A simulation illustrating the effectiveness of the correction procedure applied to minimize effects of collection biases on estimates of an ecoregions' CTI and TR. In the simulation, a hypothetical ecoregion is created with a set number of grid cells (in this case 1000) and each cell is assigned a random local temperature (MAT). Hypothetical species are then created with random MAT.opts. In each year of a set study period (in this case 40 years), the ecoregion grid cells are probabilistically populated with plants such that the probability of a species occurring in a grid cell increases as MAT.opt approaches the local grid cell temperature. The local grid cell temperature, and thus the probability of a species occurring there, changes through time as determined by a set of user-selected warming rates (in this case A) – 0.01°C yr^{-1}, B) 0.0°C yr^{-1}, and C) +0.01°C yr^{-1}) . In each year, a set number of plants (in this case 10000) are randomly "collected" from the ecoregion. The number of collections from each individual grid cell is set as being either random (unbiased) or strongly biased

towards having an increasing proportion of collections coming from hotter cells through time. Based on the collections in each year, the ecoregion's uncorrected CTI is calculated as the mean MAT.opt of all collected individuals, and the ecoregion's uncorrected TR is the slope of the regression CTI ~ year. The ecoregion's corrected CTI is calculated as the mean of the differences between each plant's MAT.opt and the local grid cell temperature where it was collected (always using the original year=0 grid cell temperatures). The ecoregion's corrected TR is the slope of the regression CTI.corrected ~ year.

```
#####R code for simulation showing the effects of strongly biased sampling for hotter
areas of a hypothetical ecoregion through time on estimates of Community Temperature
Index (CTI) and thermophilization rates (TR) with and without the correction
procedures applied. The code is currently set for a 40 year study period.

warming = c(-.01, 0.0, 0.01)                ## set desired warming rates (oC yr-1); 0
indicates stable temperatures

for(z in 1:length(warming)){
 CTI.rand = CTI.corrected.rand = CTI.biased = CTI.corrected.biased = c() # create
empty vectors to hold results

 eco.temp = rnorm(n=1000, mean=20, sd=3)              ##create "map" of random
temperatures (MAT) per grid cell in simulated ecoregion

 for(yr in 0:40){         ## for each year over the 40 year study period...
 MAT.opt = runif(n=5000, min=5, max=35)        ## create species with random MAT.opts
between 5 and 35oC
 plants = matrix(NA, nrow=length(eco.temp), ncol=5000)           ## create empty
matrix to hold simulated plant communities within each grid cell

  for(i in 1:length(eco.temp)){          ## for each grid cell in the ecoregion, sample
individuals of the simulated species at random with a probability of being sampled
proportional to MAT.opt such that the probability of a species being sampled is
greatest when eco.temp equals MAT.opt; consequently grid cells will mostly contain
individuals of species 'optimized' to the local temperature
 p = dnorm(eco.temp[i]+yr*warming[z], mean=MAT.opt, sd=2) ## p is the probability of
sampling species with changes through time based on set warming rate
 plants[i,] = sample(MAT.opt, 5000, prob=p, replace=T)            ## sampled
individuals are assigned their MAT.opt for use in calculating CTI
  }


 delta = plants-eco.temp        ## for the correction, individuals are assigned the
difference of their MAT.opt and the local MAT of the grid cell where it occurs
(eco.temp)
 delta2 = delta            ## create replicates for use in comparing random vs biased
sampling
 plants2 = plants

 n.samps.rand = rnorm(n=length(eco.temp), mean=10, sd=2)          ## collect n.samps
random individuals from each grid cell in the ecoregion without bias
 n.samps.rand[n.samps.rand<0] = 0       ## number of sampled individuals cannot be
negative
 n.samps.rand = (n.samps.rand/sum(n.samps.rand))*10000           ## scale n.samps so
that total samples within the ecoregion is 10,000

 n.samps.bias = (eco.temp)^(yr/20)           ## this creates a directional bias such
that more individuals are collected from hot cells and the magnitude of bias increases
each year; in yr 0 there is random sampling
 n.samps.bias[n.samps.bias<0] = 0       ## number of sampled individuals cannot be
negative
```

```
 n.samps.bias = (n.samps.bias/sum(n.samps.bias))*10000          ## scale n.samps so
that total samples within ecoregion is 10,000

 for(i in 1:length(eco.temp)){          ## "collect" the set number of individuals
from the community in each grid cell of the simulated ecoregion
 s.rand = sample(1:5000, n.samps.rand[i])
 s.bias = sample(1:5000, n.samps.bias[i])
 plants[i,-s.rand] = NA          ## uncollected plants are eliminated as NAs
 plants2[i,-s.bias] = NA
 delta[i,-s.rand] = NA
 delta2[i,-s.bias] = NA
 }

 CTI.rand[yr+1] = mean(plants, na.rm=T)          ## CTI in each year with unbiased
sampling using the uncorrrected MAT.opts
 CTI.corrected.rand[yr+1] = mean(delta, na.rm=T)+mean(eco.temp)          ## using the
corrected MAT.opts
 CTI.biased[yr+1] = mean(plants2, na.rm=T)          ## CTI in each year with biased
sampling using the uncorrrected MAT.opts
 CTI.corrected.biased[yr+1] = mean(delta2, na.rm=T)+mean(eco.temp)          ## using the
corrected MATopts and converting back to absolute temperature units instead of
deviations
 }


##make figures of results
 win.graph()
 par(mfrow=c(2,2))
 plot(CTI.rand, xlab='Year', ylab='CTI (oC)', las=1, main="unbiased collections, no
corrections", pch=21, bg="gray", cex=1.5)
 abline(lm(CTI.rand~seq(1,41,1)))          ## plot estimate of thermophilization rate
 points(seq(0,40,1), mean(eco.temp)+(warming[z]*seq(0,40,1)), type='l', col=2, lwd=2)
      ## plots line with change in ecoregion temp per year

 plot(CTI.corrected.rand, xlab='Year', ylab='CTI.corrected (oC)', las=1,
main="unbiased collections, with corrections", pch=21, bg="gray", cex=1.5)
 abline(lm(CTI.corrected.rand~seq(1,41,1)))          ## plot estimate of
thermophilization rate
 points(seq(0,40,1), mean(eco.temp)+(warming[z]*seq(0,40,1)), type='l', col=2, lwd=2)
      ## plots line with change in ecoregion temp per year

 plot(CTI.biased, xlab='Year', ylab='CTI (oC)', las=1, main="biased collections, no
corrections", pch=21, bg="gray", cex=1.5)
 abline(lm(CTI.biased~seq(1,41,1)))          ## plot estimate of thermophilization
rate
 points(seq(0,40,1), mean(eco.temp)+(warming[z]*seq(0,40,1)), type='l', col=2, lwd=2)
      ## plots line with change in ecoregion temp per year

 plot(CTI.corrected.biased, xlab='Year', ylab='CTI.corrected (oC)', las=1,
main="biased collections, with corrections", pch=21, bg="gray", cex=1.5)
 abline(lm(CTI.corrected.biased~seq(1,41,1)))          ## plot estimate of
thermophilization rate
 points(seq(0,40,1), mean(eco.temp)+(warming[z]*seq(0,40,1)), type='l', col=2, lwd=2)
      ## plots line with change in ecoregion temp per year

 }
```