Appendix S1

Ant sampling

In South Africa and Lesotho, pitfall traps were arranged in a 10 m X 40 m grid. At each elevation, four grids were used, each separated by 300 m. Traps were 55 mm in diameter and were half filled with a 50% ethylene glycol solution to preserve trapped ants. Traps were left out for five days and five nights. In Australia, sampling grids were the same dimensions as in Africa but grids within the same elevation were separated by 100 m. Traps were left out for 7 days and nights in Australia. In Argentina, sampling grids were 9 pitfall traps arranged in a 10 m X 10 m grid. Each trap was separated from the next by 5 m. Only a single grid was used at each elevation. Traps in Argentina were also open for 7 days and nights. Ants were transferred to 70-80% ethanol in the laboratory and identified to species or morphospecies level. All transects were sampled during the austral summer (November – May).

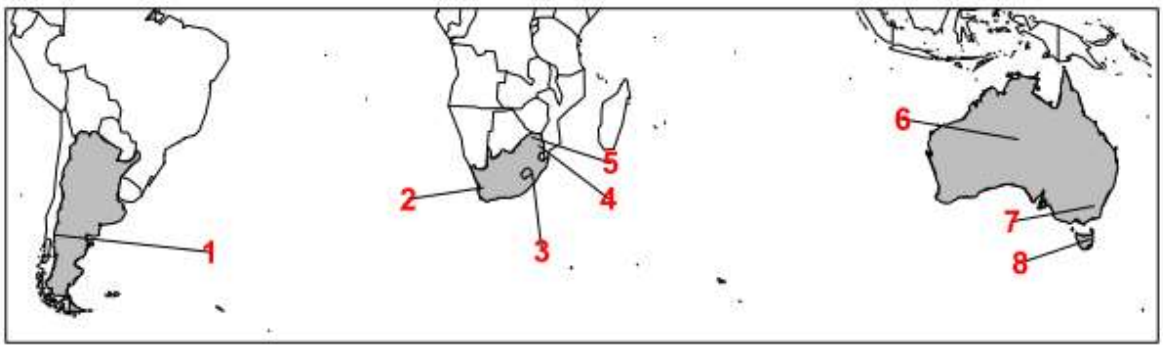

Figure S1.1. *Map of the southern hemisphere showing the study sites. 1 = Patagonia, Argentina; 2 = Cederberg, South Africa; 3 = Maloti-Drakensberg, Lesotho and South Africa; 4 = Mariepskop, South Africa; 5 = Soutpansberg, South Africa; 6 = MacDonnell Ranges, Australia; 7 = Snowy Mountains, Australia; 8 = Ben Lomond Plateau, Australia.*

Colour data

We recorded the colour of each ant species in the dataset using a predetermined set of colours (Parr et al. 2017). The colour of the head, mesosoma and gaster was recorded for 6 individuals, where possible, for every species. We focussed on the colour of the cuticle and ignored colouration from hairs. The most common colour across these three body parts was assigned as the dominant colour for each species. The categorical colours were converted to HSV (hue, saturation, value) format and the lightness value (v, value) was extracted and used for analysis. Lightness ranges from 0 (black) to 1 (white). We assessed observer bias by providing the same set of 71 photographs taken from AntWeb to each of the 5 colour data recorders. Using standardised monitor settings, the standard error of colour estimates for the same photograph was low (SE = 0.04) suggesting that different observers assigned almost identical colours to the same images.

Body size data

The body size of each ant species was recorded as Weber's length. This was measured to the nearest 0.01 mm using ocular micrometers attached to stereomicroscopes and the highest level of magnification able to fit the entire mesosoma of the ant into the view. Six specimens from each species were measured where possible. Only minor workers were measured. Physical specimens were not available for eight species from the Cederberg transect. High resolution images from taxonomic publications (Mbanyana & Robertson, 2008) and AntWeb (www.antweb.org) were measured using the tpsDig2 software instead (http://life.bio.sunysb.edu/morph). Weber's length was not available for the ant species form the MacDonnell Ranges but head width and head length were. We estimated Weber's length for these MacDonnell Ranges species based on the relationship between head length, head

width and Weber's length for the species in the other Australian transects in the Snowy Mountains and the Ben Lomond Plateau. This was done using a multivariate imputation by chained equations procedure (Buuren & Groothuis-Oudshoorn, 2011).

Buuren, S. & Groothuis-Oudshoorn, K. (2011) mice: multivariate imputation by chained equations in R. *Journal of Statistical Software*, 45, 1–67.

Mbanyana, N. & Robertson, H.G. (2008) Review of the ant genus *Nesomyrmex* (Hymenoptera: Formicidae: Myrmicinae) in southern Africa. *African Natural History*, 4, 35–55.

Parr, C. L., et al. (2017). GlobalAnts: a new database on the geography of ant traits (Hymenoptera: Formicidae). *Insect Conservation and Diversity*, 10, 5-20.

Appendix S2

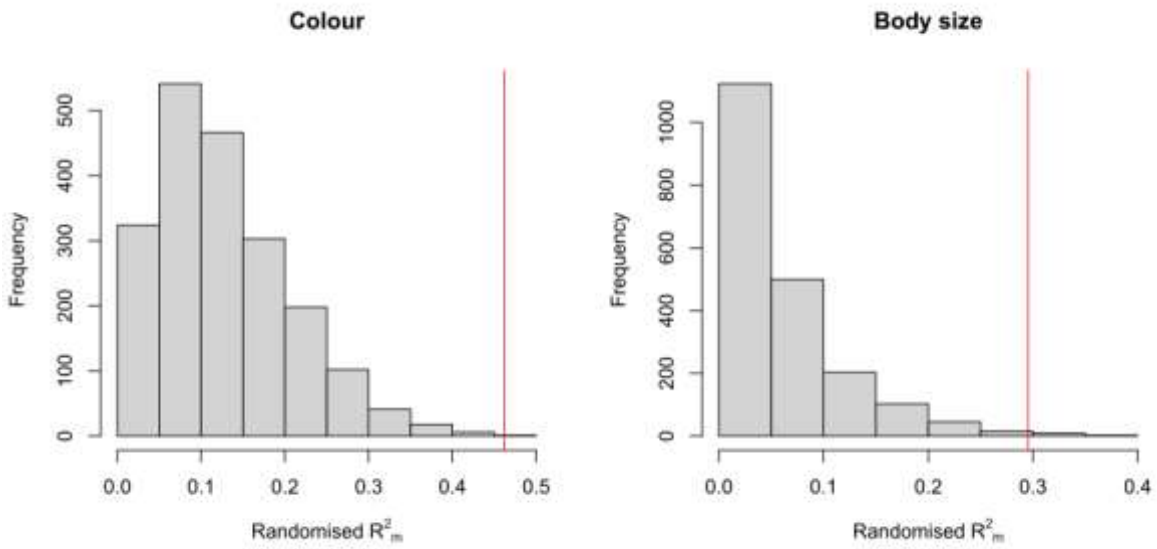

Figure S2. *Histograms showing the distribution of marginal R^2s (R^2_m) calculated from repeated trait-environment models where the species-trait value associations were randomised 2000 times. The observed R^2_ms (using the original data) are displayed as red lines. The observed R^2_m for colour was in the 99th percentile of the randomised distribution. The observed R^2_m for body size was also in the 99th percentile of the randomised distribution.*

Appendix S3

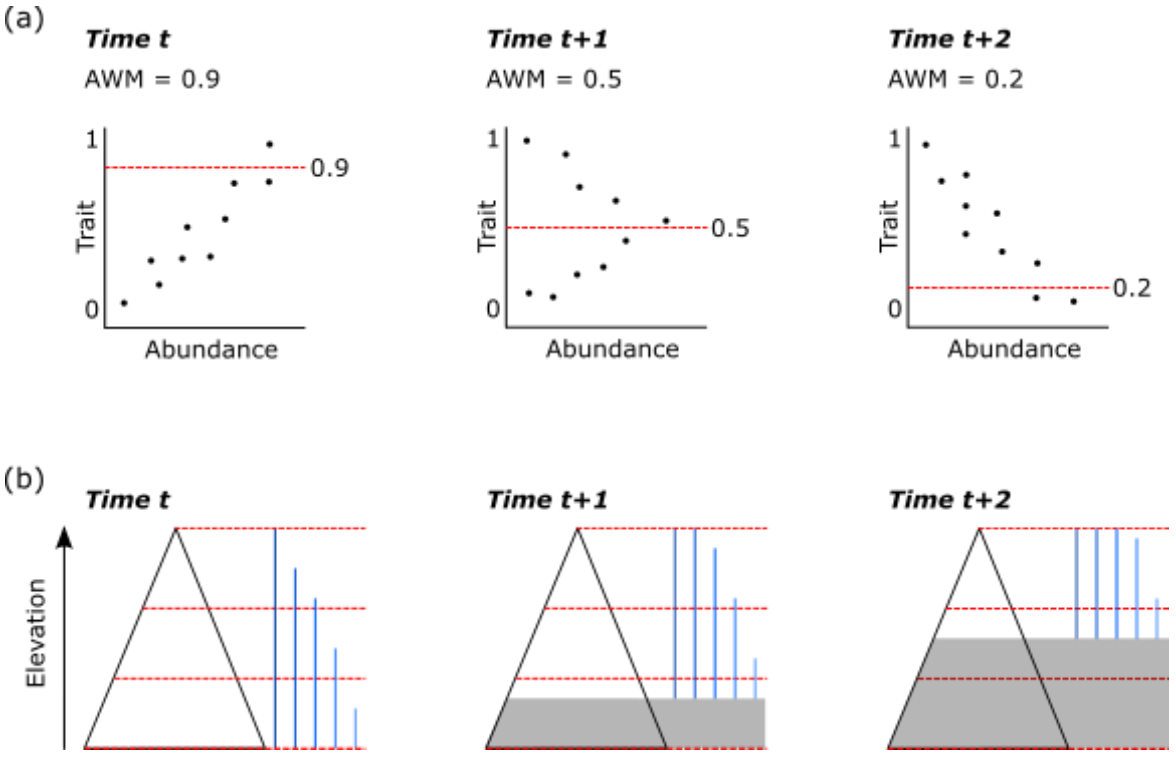

Figure S3. *Schematic showing how (a) the community assembly by trait selection (CATS) model and (b) the elevational range shift model works. In (a), the predicted assemblage weighted mean (AWM, dashed red lines) changes through time. The CATS model uses maximum entropy to find the most likely relative abundance structure given this target AWM and the traits of each species present. As a result, the relative abundances of species changes through time in line with the predicted AWMs. In (b), species ranges (blue) are distributed along an elevational gradient which has four sampling locations (dashed red lines). Species ranges move upslope uniformly though time as temperature increases. The grey area in time step t+1 and t+2 indicates an area of artificial lowland biotic attrition. Future assemblage compositions of sites within these lowland areas are not predicted in order to avoid this artefact.*

Appendix S4

We reran the trait simulation under RCP8.5 8 additional times to test whether different thresholds of predicted relative abundance influenced the results. Originally, we used a threshold of 0.0001 to determine whether a species should be classed as present or not. Values greater than 0.0001 were classed as present.

Our logic for choosing this particular threshold was that it was close to the minimum detection thresholds of ants in the field across out study areas. The smallest relative abundances for each mountain region range from 0.000085 to 0.008475. Our value of 0.0001 is close to the lower end of this range. Consequently, it is a relatively liberal threshold. We argue that a threshold value radically smaller or larger than these values would be ecologically unrealistic.

For each rerun, we used one of the minimum relative abundances recorded from each mountain range as the threshold value. We compared the predicted species richness, Sørensen's dissimilarity and Bray-Curtis dissimilarity from these reruns to our original threshold choice (see table below). We calculated the average difference between predictions made using the original threshold, and those made using a different threshold.

There is almost no difference in species richness, Bray-Curtis dissimilarity or Sørensen's dissimilarity predictions when using a different presence threshold, in comparison to our original threshold. The exception is when using the largest threshold value (0.008475). Consequently, we present the results from our original threshold, 0.0001, in the main text.

Table S4. Table showing mean differences between predictions made using the original presence threshold and a series of alternative presence thresholds. Differences reported in units of species richness, Bray-Curtis dissimilarity and Sørensen's dissimilarity.

Threshold	Mean difference in species richness	Mean difference in Bray-Curtis	Mean difference in Sørensen's dissimilarity
0.000085	0	<0.001	0
0.000241	0.210526	<0.001	0.003074
0.000292	0.421053	<0.001	0.02912
0.000396	0.526316	<0.001	0.05719
0.000637	0.578947	<0.001	0.05719
0.000766	0.578947	<0.001	0.05654
0.001647	1	<0.001	0.066825
0.008475	4.526316	<0.001	0.076492

Appendix S5

Table S5. Lapse rates (°C per m gained in elevation) for each elevational transect.

Continent	Mountain range	Transect	Lapse rate
Africa	Cederberg	Cederberg	-0.00370
Africa	Mariepskop	Mariepskop	-0.00645
Africa	Soutpansberg	Soutpansberg	-0.00482
Africa	Maloti-Drakensberg	Sani Pass	-0.00410
Australia	Snowy Mountains	Back Perisher	-0.00410
Australia	Ben Lomond Plateau	Stacks Bluff	-0.00655
Australia	MacDonnell Ranges	Mt Zeil	-0.00396
South America	Patagonia	La Mona	-0.00450
South America	Patagonia	Lopez	-0.00485
South America	Patagonia	Pelado	-0.00358
South America	Patagonia	Bayo	-0.00354
South America	Patagonia	Chall-Huaco	-0.00298

Appendix S6

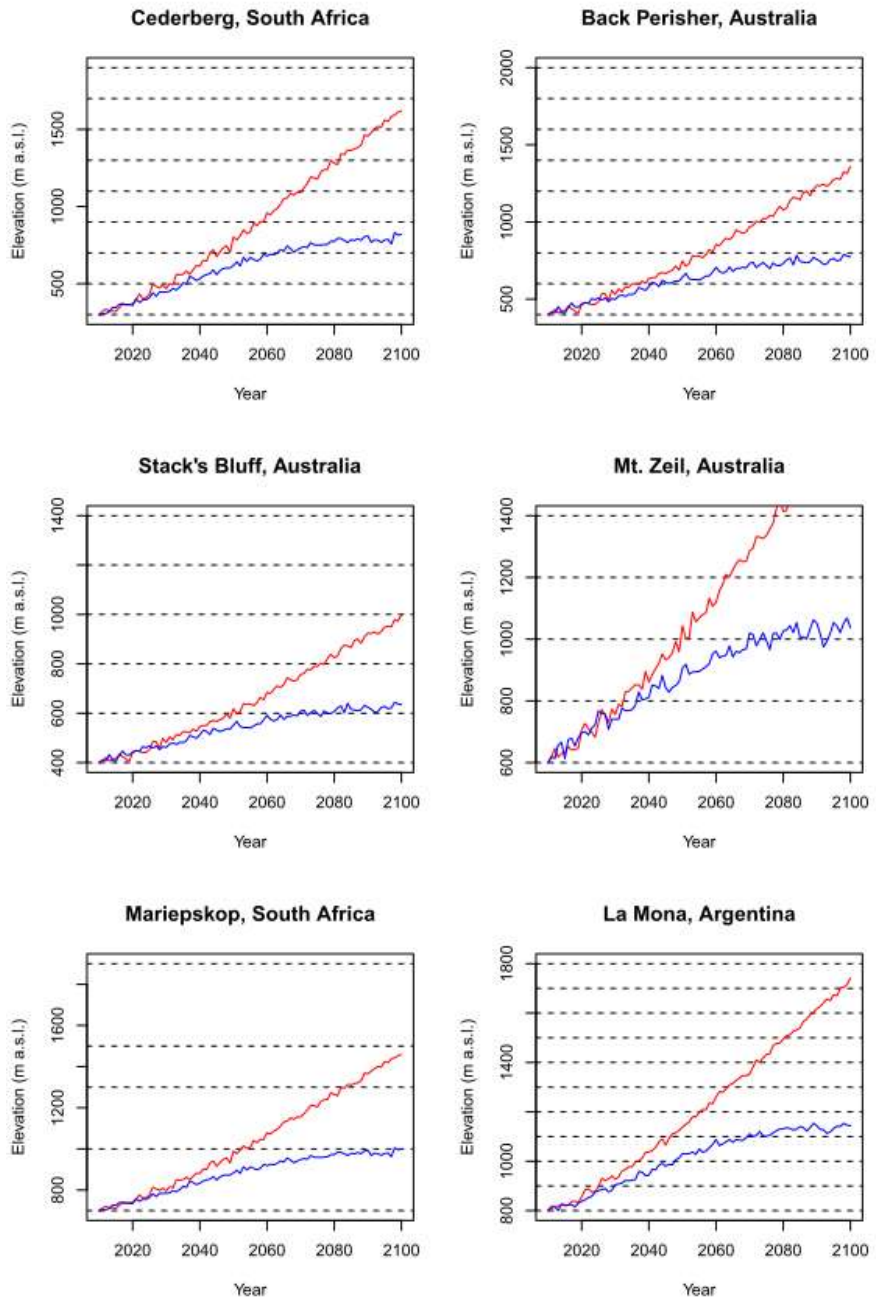

Figure S6.1. *Plots showing the lower predictive limit under each scenario (red = RCP8.5, blue = RCP2.6) through time. Sampling sites are shown as dashed black lines. Only sites above the lower predictive limit for a given year can be forecast. For example, in the Cederberg under RCP8.5, only sites above 1500 m a.s.l. can be predicted to the year 2100.*

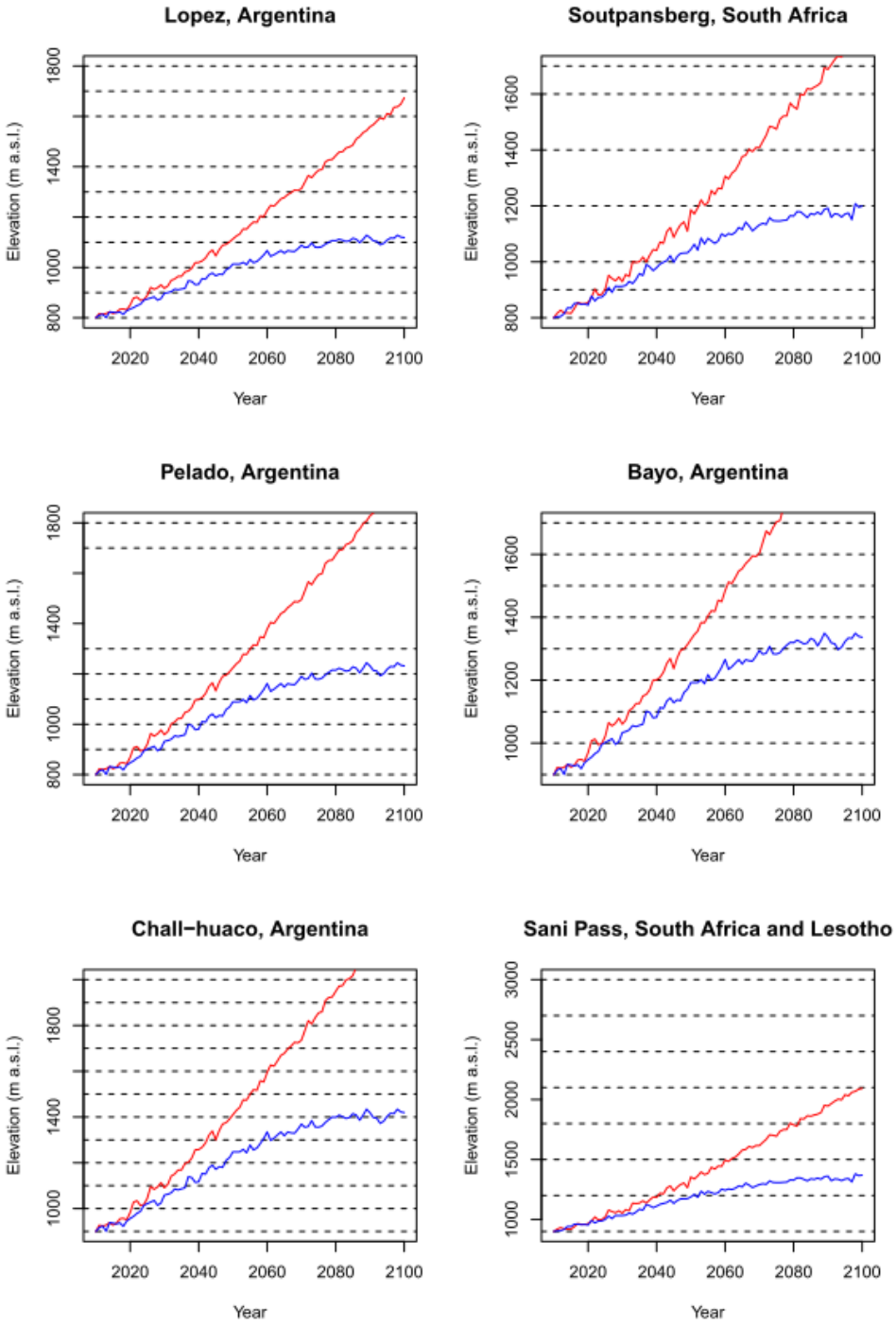

Figure S6.2. *Plots showing the lower predictive limit under each scenario (red = RCP8.5, blue = RCP2.6) through time. Sampling sites are shown as dashed black lines. Only sites above the lower predictive limit for a given year can be forecast.*

Appendix S7

For colour lightness, the best model as determined by AICc was:

Lightness ~ temperature + temperature^2 + UV + temperature*UV + temperature^2*UV

For body size, the best model as determined by AICC was:

Body size ~ temperature + temperature^2

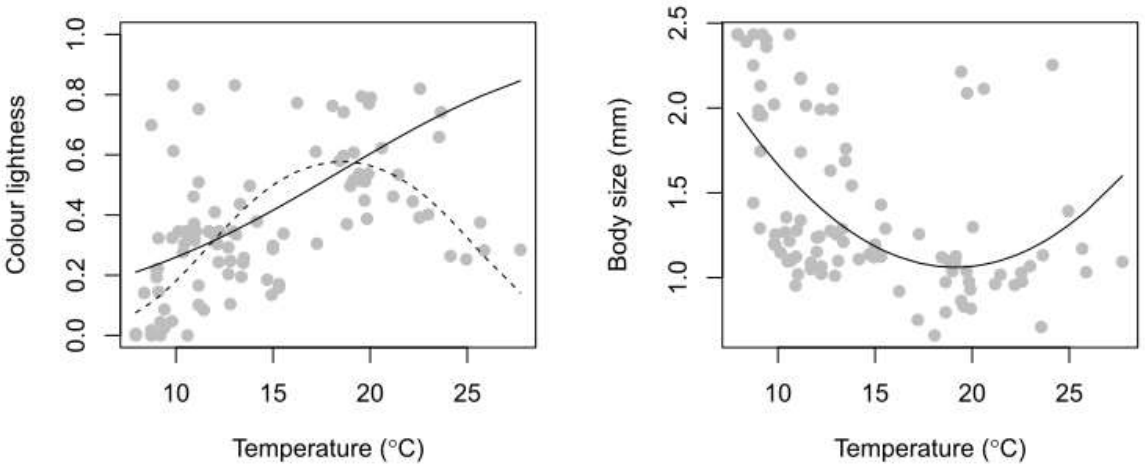

Figure S7. *Plots showing the trait-environment relationship between temperature, colour and body size. Lines represent mixed model predictions. For colour, the dashed line is for high (80th percentile) levels of UV-B and the solid line is for low levels of UV-B (20th percentile)*