Supporting Information Appendix

Basal Resistance Enhances Warming Tolerance of Alien over Indigenous Species Across Latitude

Charlene Janion-Scheepers, Laura Phillips, Carla M. Sgrò, Grant A. Duffy, Rebecca Hallas, and

Steven L. Chown

School of Biological Sciences, Monash University, Melbourne, VIC 3800, Australia

correspondence to: charlene.janionscheepers@monash.edu

Materials and Methods

Collection, identification, and alien species assignment

The thirty springtail species were collected, by aspiration, leaf litter sieving, or Tullgren extraction of litter (1) between 2013 and 2016 (Table S8). A single species was collected from the island of South Georgia, south of the Antarctic Polar Frontal Zone (Table S8). The focus was on hemiedaphic (litter-dwelling) species. Individuals were initially assigned to species in the field by one of the authors (CJ-S) with taxonomic expertise (e.g. 2, 3); and at least 200 individuals collected (in the case of the *Brachystomella* sp. and *Triacanthella* sp. only \sim 50 individuals of each were collected). Collections were maintained either in litter or in 60 ml or 300 ml pots with moist mixed Plaster-of-Paris:Charcoal powder (9:1) base substrates until their return to the laboratory, typically within one week of collection (two weeks for the remote sites of Macquarie Island and South Georgia).

Species separations were initially verified in the laboratory by CJ-S, then identified to genus, and where species had been described, to species level using keys for the fauna of Australia and the sub-Antarctic islands (e.g. 4-6), keys to the European fauna (7, 8), which are appropriate for many alien species (9, 10), and in consultation with taxonomic experts for specific groups within the springtails. DNA barcoding (11, 12) was used to confirm species identifications. Mitochondrial DNA extraction and sequencing of the cytochrome c oxidase subunit I gene was undertaken by the Biodiversity Institute of Ontario, University of Guelph, Canada, following standard protocols developed for springtails (13, 14). Sequences of 74 specimens from 23 species were compared with the springtail sequences available through the Barcode of Life Data Systems (BOLD) (www.barcodinglife.org; Table S9). Individuals that could not be identified

using available keys and which were not represented in BOLD were examined by one of the authors (CJ-S) and assigned to uniquely identifiable species based on morphological characteristics and/or a barcoding gap of at least 2.5% (15). Sequences are available on BOLD (www.boldsystems.org) as part of Project COLMU (Collembola of Monash University) either identified as indigenous or alien to Australia or to the sub-Antarctic islands in faunal treatments (5, 6, 16-18). Undescribed species not represented in BOLD previously, or represented only from individuals already collected across Australia, New Zealand or south of the Wallace line were considered indigenous. Following previous authors (9), undescribed species that had sequences present in BOLD from other distant tropical regions (such as the Neotropics) or from the Holarctic (typically Europe) were considered alien species (Tables S8, S9).

Site microclimate characteristics

The soil microclimate characteristics of each site were calculated using remote-sensed daytime land-surface temperature data (LST) from the MODIS aqua/terra satellite network (MOD11C2 v006; 30 arcseconds spatial resolution; 8-day temporal resolution from January 2001 to December 2015; doi:10.5067/MODIS/MOD11C2.006), which were linearly transformed to account for the diffusion of heat from the land-surface to 2.5 cm below the soil surface. The slope of this linear transformation was derived from the microclim dataset (19), which contains validated estimates of soil temperature for each hour in a 24-hr cycle of an average day in each month of an average year under varying shade conditions. Soil temperatures from the microclim dataset incorporate a 5 cm 'organic cap' with reduced thermal conductivity and increased heat capacity. A soil depth of 2.5 cm sits within this 'organic cap' and was, therefore, identified as the best approximation of the litter microclimate in which hemiedaphic springtails are found. A

linear model was fitted between microclim LST, as a predictor, and microclim soil temperature at 2.5 cm depth, as a dependent variable, for daytime hours of all 12 months across eastern Australia (> 142 °E; Fig. 1) and across three shade scenarios (25 %, 50 %, 75 %). The strong fit of this combined model (y = $0.8489x + 0.6411$; adjusted R² = 0.9704) meant that it could be used to confidently convert the MODIS LST from our remote sensed time series to an estimate of temperature at 2.5 cm soil depth for every 8-day period between 2001 and 2015. The median (MODIS soil median), 99 % quantile (MODIS soil99), maximum (MODIS soil max; i.e. the warmest 8-day mean), and minimum (MODIS soil min; i.e. the coldest 8-day mean) soil temperature of each site were calculated from our linearly transformed MODIS time series.

Colony maintenance

Species were reared at temperatures that typically reflect the average soil temperatures of the sites at which they were collected (Table S8), though also bearing in mind the need to achieve standardization of conditions (20). Site temperatures were assigned based either on measurements at the sites during the time of collection using a handheld soil temperature meter (IQ150, Spectrum Technologies Inc., IL, USA), or based on data from microclim (19), assuming a soil depth of 2.5 cm and 50 % shade. Species were reared at constant rearing temperatures in controlled-temperature incubators (MIR-154, SANYO Electric Co. Ltd., Osaka, Japan) on a 12 light:12 dark light cycle, with temperatures monitored using Hygrochron iButtons (DS 1923-F5, Maxim Integrated, San Jose, CA, USA) (Table S8).

The F2 generation was the focus of this work (Fig. S4) to minimize any carry-over effects from the environment of origin, including parental effects, and to reduce the possibility of adaptation

to laboratory conditions (21, 22), which might confound interspecific comparisons (23, 24, 25). Between 50 and 200 adults from the collected (F0) individuals were randomly assigned to two to four 60 ml pots lined with moistened Plaster-of-Paris:Charcoal powder (9:1) substrates. Deionised water was added once to twice a week to maintain high humidity and, depending on the species, individuals were fed two to three times a week with algae from the bark of *Platanus* sp. or on slime mould *ad libitum* (26), enabling individuals to select nutrients optimally. Rearrangement of pots among shelves (at each feeding event) and the use of multiple pots ensured that container and shelf effects were minimized. For each species, eggs from all the F0 generation pots were collected twice a week and combined randomly into new pots and reared to adults (i.e. F1). The F1 generation, emerging after 21 to 194 days (average egg to adult development time among the species is 74.16 days \pm 40.56 (sd), depending on the species and conditions), was reared as described above. Eggs from this F1 generation were collected and combined randomly within new pots (typically six to ten pots with a density of 50 to 100 individuals per pot, appropriate given very high densities of springtails under field conditions (27, 28)) to rear the F2 generation. Adults from the F2 generation were used for most experiments (Fig. S4), though in some instances adults from the F3 and F4 generations were used where initial stock numbers were slightly lower than the experimental design required.

Acclimation to assess phenotypic plasticity

Prior to the experimental trials, all species were subject to temperature treatments (referred to 'acclimation' hereafter), undertaken in controlled-temperature incubators (MIR-154, SANYO), with temperatures verified using Hygrochron iButtons (DS 1923-F5, Maxim Integrated, San Jose, CA, USA), and under 12 light:12 dark conditions (Table S8). Acclimation treatments lasted seven days, given that complete responses usually occur within less time in terrestrial arthropods $(29, 30)$. Low, medium and high acclimation temperatures were set 5°C below, and 5 and 10°C above standard rearing temperatures, respectively (Fig. S4). For control temperatures, individuals were subject to the same manipulation as those in the acclimation treatments.

Critical thermal limit and warming tolerance determinations

Critical thermal limits (critical thermal maximum (CT_{max}) and critical thermal minimum (CT_{min})), which represent limits to movement, and provide a proxy for adult survival given that lack of movement leaves individuals unable to feed or escape predators (31, 32), were determined following standard methods (33, 34). Programmable water baths (Model TXF200, Grant Instruments, Cambridge, UK) were used to heat or cool custom-built thermal stages (Monash University Instrument Facility, Clayton Campus, VIC, Australia) into which a 50 ml plastic vial with a moistened Plaster-of-Paris substrate, to preclude desiccation of individuals during trials, was fitted. Assays began at the control (rearing) temperature for each species to ensure that comparisons between the control and acclimated groups could be made. Individuals were held for 15 minutes at the starting temperature prior to ramping. Heating (for CT_{max}) and cooling (for CT_{min}) rates were set at 0.05° C/minute – rates that are within the range, and close to the mean (for temperature increase), of those recorded for tropical to temperate microhabitats (1). Moreover, empirical data and modelling indicate that these rates result in comparable estimates of warming tolerance and acclimation responses across environments (1).

Temperature of the substrate was recorded with a type K thermocouple, using a digital thermometer (Model RDXL 12SD, Omega Engineering, USA) and individuals were monitored every \sim 5 \degree C until behavioural change occurred (e.g. moving considerably faster or slower), after which they were monitored every $\sim 1^{\circ}C$. CT_{min} and CT_{max} were defined as temperature at which a loss of righting response occurred (33, 34) by gently flipping the individuals with a fine brush. Individuals were scored for a loss of righting response every 0.5°C after the knockdown of the first individual. Typically, three replicates of 10-15 individuals were completed for each species and treatment, with a few exceptions where some treatments were excluded owing to low sample sizes. Determining the sex of springtails usually requires mounting specimens on slides to observe the necessary characters under a compound microscope. This cannot be done with live specimens. In consequence, large sample sizes were used $(~40)$ individuals) to incorporate any sex effects. Because body mass may contribute to variation in critical thermal limits (e.g. 24, 35), species' mean body mass (mg) was determined from a randomly selected, separate sample of 40 adult individuals for each species using a high-resolution $(0.1 \mu g)$ microbalance (Mettler-Toledo XP2U, Switzerland) (Table S8).

For each species, basal thermal tolerance was calculated as the mean CT_{min} , for the lower critical thermal limit, and mean CT_{max} , for the upper critical thermal limit, obtained from individuals reared under control conditions and subjected to the same temperature for acclimation. Basal thermal tolerance range was calculated as the difference between these two mean values. To determine the extent of phenotypic plasticity, the acclimation response ratio (ARR) (36) was calculated for each CT_{min} and CT_{max} for each species. Here, the ARR was calculated as the largest difference between mean CT_{min} (or CT_{max}) across any of the acclimation treatments, divided by the maximum temperature range represented by those treatments.

Warming tolerance is widely used as a measure of the likely susceptibility of populations to rising temperatures associated with climate change (37, 38). Here, warming tolerance was calculated as the difference between mean basal CT_{max} (individuals reared under control conditions) of each species and the MODIS 99 % quantile (MODIS soil99).

Selection experiment

Laboratory natural selection (39) was used to investigate the ways in which critical thermal limits respond to elevated rearing temperatures. Four species were selected for this experiment – representing a tropical alien (*Desoria trispinata*), a tropical indigenous (*Ascocyrtus* sp. 2), a temperate alien (*Orthonychiurus* sp.), and a temperate indigenous species (*Lepidocyrtus* sp. 10). Laboratory natural selection was undertaken by exposing individuals to a high temperature treatment (hereafter referred to as the selection group) throughout development for successive generations (39), then consecutively assessing critical thermal limits for phenotypic divergence between the selection group and a control group held at the original rearing temperature. Laboratory natural selection rather than artificial selection was chosen to minimise the risk of sterilization and other cellular damage that can be associated with exposure to extreme temperatures (40, 41). In particular, the effect of extreme temperatures has been shown to lower springtail reproductive success (42), and the design had to ensure the maintenance of large population sizes to reduce the likelihood of genetic drift and inbreeding depression (43).

Selection and control groups were initiated for each species from the F2 generation of field caught individuals. F2 individuals were used to minimize any carry-over effects from the environment of origin, including parental effects, and to reduce the possibility of adaptation to

laboratory conditions (21, 22), which might confound comparisons of adaptive capacity between tropical and temperate species. Each group contained two independent replicate lines starting with 150 individuals divided into two separate vials per replicate line. Control lines were maintained under the original rearing temperatures (temperate = 15° C [mean \pm sd: 14.88 \pm 0.56°C], tropical = 20° C [mean ± sd: $20.16 \pm 0.29^{\circ}$ C], temperatures measured using Hygrochron iButtons model DS 1923-F5, Maxim Integrated, San Jose, CA, USA), while selection lines were maintained under warmer temperatures (temperate = 25° C [mean \pm sd: $24.89 \pm 0.54^{\circ}$ C], tropical $= 27^{\circ}$ C [mean \pm sd: 27.08 \pm 0.61°C]). Temperatures for the selection treatment were based on results from a pilot study, which indicated that 25°C and 27°C were the highest temperatures at which the temperate and tropical species could still reproduce, respectively (Table S10). Throughout the experiment, generations remained discrete, and eggs from replicate vials within each replicate selected and control line were randomly combined within generations to maintain genetic diversity. Population size in each line (control and selection lines) was on average 775 individuals. Critical thermal limits were assessed, as for the interspecific comparisons, for adults of each species prior to selection commencing (at generation 0), then every second generation for individuals in the selection and control groups up to generation ten, and every fourth generation thereafter. Approximately 45 individuals were assessed per replicate line per treatment at each sampling period. These individuals were permanently removed from the control and selection groups.

The degree of plasticity associated with any phenotypic changes observed during the selection experiment was assessed using a reciprocal transplant experiment, investigating developmental plasticity. This involved switching individuals from the selection conditions to the control

conditions and *vice versa*. Individuals were switched within one day of hatching, at generation four for the temperate species and generation six for the tropical species. These different generations were used because of the slower development time of the temperate species. Critical thermal limits were assessed as soon as the switched individuals had reached adulthood. If phenotypic changes in critical thermal limits reflect genotypic change, rather than phenotypic plasticity, the thermal tolerances of the reciprocally transplanted individuals should reflect that of their original treatment group, even after development at the alternative temperature (44).

Statistical analyses

All analyses were conducted in R version 3.3.1 (45), with figures and plots developed using the ggplot2 package (https://cran.r-project.org/web/packages/ggplot2/ ggplot2.pdf). Because species cannot be considered independent units in any comparative analysis (46), and phylogenetic signal has been found in thermal tolerance traits (47), Phylogenetic Generalised Least Squares (PGLS) (48), as implemented in the caper v0.5.2 (49) and APE (50) packages was used. A phylogeny for the species was constructed based on joint considerations of two recent molecular phylogenies for the group (51, 52) with species relative positions based on the cytochrome c oxidase subunit I gene phylogeny, or in a few cases on morphological similarity adjudicated by one of us (CJ-S). The barcoding placements were obtained from a neighbour-joining tree (53) constructed using MEGA6 (54) with the Kimura-2 parameter model (55). For the final tree, branch lengths were assigned using Grafen's method (56), and the tree (Fig. S5) is available as a Newick file. Initially, two covariance matrices were constructed following either Brownian motion or Ornstein–Uhlenbeck models of evolution. Akaike Information Criterion (AIC) values of Brownian motion (BM) and Ornstein–Uhlenbeck (OU) models were compared to identify

which model of evolution provided the best fit to observed data. Phylogenetically-corrected models based on Brownian motion co-variance matrices were a consistently better fit than those based on other evolutionary assumptions (Table S11), thus the outcomes of these models are reported primarily, though for comparative purposes we provide the OU outcomes too (Table S12). In the BM approach, the covariance matrix was constructed following a Brownian motion model of evolution (57) assuming proportional branch lengths in the phylogeny. A maximum likelihood approach provided Pagel's λ (58), which indicates the degree of phylogenetic correlation in the data ($\lambda = 0$ indicates no phylogenetic effect, while $\lambda = 1$ indicates a strong phylogenetic effect equivalent to that expected under the Brownian motion model).

PGLS was used to investigate relationships between species mean critical thermal limits (either CT_{min} , CT_{max} or thermal tolerance range), environmental characteristics (MODIS soil median), springtail species mean mass, and species status (alien or indigenous). The same approach was used to investigate relationships between warming tolerance, maximum soil temperature (MODIS soil maximum), species mass and species status, and to investigate relationships between CT_{min} and CT_{max} . Analyses were repeated using ordinary least squares approaches as implemented in the linear model function of R version 3.3.1, and coefficients were typically similar to those found in the PGLS models (Table S13). Throughout, mass did not appear as a significant term in the models, and in no cases did slopes of the relationships between critical thermal limit traits and environmental features differ between the alien and indigenous species groups (i.e. no interaction terms were significant). For investigations of the ARR and its relationship with mean trait values, ordinary least squares methods indicated no significant relationships and PGLS bore out these conclusions.

For the selection experiment, to analyse differences in critical thermal limits between selection and control lines, nested mixed effect model analyses were conducted using the lmer function in the lme4 package (version $1.1 - 13$) (59) in R version 3.3.1. 'Treatment' (control or selection) and 'generation number' were treated as fixed effects, and 'replicate line' was nested within treatment as a random effect (60). Nested mixed effect analyses were also undertaken to analyse data from the reciprocal transplant experiment examining developmental plasticity. This involved comparing the critical thermal limits of the selection, control and reciprocally transplanted lines at the respective generation of the reciprocal transplant experiment, with 'replicate line' nested within treatment (control, selected, reciprocally transplanted) as a random effect. Separate analyses were performed for each of the four species.

Permit Information

Collection permits were provided by the New South Wales National Parks and Wildlife Services, Queensland Department of Environment and Heritage Protection, Victoria Department of Environment and Water Planning, Tasmania Department of Primary Industries and Water, and the Government of South Georgia and the South Sandwich Islands.

References

- 1. Allen JL, Chown SL, Janion-Scheepers C, Clusella-Trullas S (2016) Interactions between rates of temperature change and acclimation affect latitudinal patterns of warming tolerance. *Conserv Physiol* 4:cow053.
- 2. Janion C, Deharveng L, Weiner WM (2013) Synonymy of *Spicatella* Thibaud, 2002 with *Delamarephorura* Weiner & Najt, 1999, and description of two new species (Collembola: Tullbergiidae). *Raffles Bull Zool* 61:657-663.
- 3. Potapov M, Janion-Scheepers C, Deharveng L (2017) Taxonomy of the *Cryptopygus* complex. II. Affinity of austral *Cryptopygus s.s.* and *Folsomia*, with the description of two new *Folsomia* species (Collembola, Isotomidae). *ZooKeys* 658:131-146.
- 4. Greenslade P (2006) *The invertebrates of Macquarie Island* (Australian Antarctic Division, Kingston).
- 5. Greenslade P, Ireson J, Skarzynski D (2014) Biology and key to the Australian species of *Hypogastrura* and *Ceratophysella* (Collembola: Hypogastruridae). *Austral Entomol* 53:53- 74.
- 6. Mateos E, Greenslade P (2015) Towards understanding *Lepidocyrtus* Bourlet, 1839 (Collembola, Entomobryidae) I: diagnosis of the subgenus *Setogaster*, new records and redescriptions of species. *Zootaxa* 4044:105-129.
- 7. Fjellberg A (1998) *Fauna Entomologica Scandinavica Volume 42. The Collembola of Fennoscandia and Denmark. Part I: Poduromorpha* (Brill, Leiden).
- 8. Potapov M (2001) *Synopses on Palaearctic Collembola, Volume 3. Isotomidae*. (Abhandlungen und Berichte des Naturkundemuseums Görlitz).
- 9. Cicconardi F *et al*. (2017) MtDNA metagenomics reveals large-scale invasion of belowground arthropod communities by introduced species. *Mol Ecol* 26:3104-3115.
- 10. Janion-Scheepers C, Deharveng L, Bedos A, Chown SL (2015) Updated list of Collembola species currently recorded from South Africa. *ZooKeys* 503:55-88.
- 11. Hebert PDN, Cywinska A, Ball SL, DeWaard JR (2003) Biological identifications through DNA barcodes. *Proc R Soc B* 270:313-321.
- 12. Ratnashingham S, Hebert PDN (2007) BOLD: The Barcode of Life Data System (www.barcodinglife.org). *Mol Ecol Notes* 7:355-364.
- 13. Hogg ID, Hebert PDN (2004) Biological identification of springtails (Hexapoda: Collembola) from the Canadian Arctic, using mitochondrial DNA barcodes. *Can J Zool* 82:749–754.
- 14. Porco D, Bedos A, Greenslade P, Janion C, Skarżyński D, Stevens MI, Jansen van Vuuren B, Deharveng L (2012) Challenging species delimitation in Collembola: cryptic diversity among common springtails unveiled by DNA barcoding. *Invert Syst* 26:470-477.
- 15. Meyer CP, Paulay G (2005) DNA barcoding: error rates based on comprehensive sampling. *PLoS Biol* 3:2229-2238.
- 16. Greenslade P (1994) Collembola. *Zoological Catalogue of Australia*. Volume 22. *Protura, Collembola, Diplura*, ed WWK Houstin (CSIRO, Melbourne), pp 19-138.
- 17. Greenslade P, Convey P (2012) Exotic Collembola on subantarctic islands: pathways, origins and biology. *Biol Invas* 14:405-417.
- 18. Bellinger PF, Christiansen KA, Janssens F (1996-2017). Checklist of the Collembola of the World. http://www.collembola.org.
- 19. Kearney MR, Isaac AP, Porter WP (2014) microclim: Global estimates of hourly microclimate based on long-term monthly climate averages. *Sci Data* 1:140006.
- 20. Chown SL, Gaston KJ (2016) Macrophysiology progress and prospects. *Funct Ecol* 30: 330-344.
- 21. Sgrò CM, Partridge L (2000) Evolutionary responses of the life history of wild-caught *Drosophila melanogaster* to two standard methods of laboratory culture. *Am Nat* 156:341- 353.
- 22. Blackburn S, van Heerwaarden B, Kellermann V, Sgrò CM (2014) Evolutionary capacity of upper thermal limits: beyond single trait assessments. *J Exp Biol* 217:1918-19244.
- 23. Araújo MB *et al.* (2013) Heat freezes niche evolution. *Ecol Lett* 16:1206-1219.
- 24. Ribeiro PL, Camacho A, Navas CA (2012) Considerations for assessing maximum critical temperatures in small ectothermic animals: insights from leaf-cutting ants. *PLoS One* 7:e32083.
- 25. Pintor AF, Schwarzkopf L, Krockenberger AK (2016) Extensive acclimation in ectotherms conceals interspecific variation in thermal tolerance limits. *PLoS One* 11:e0150408.
- 26. Hoskins JL, Janion-Scheepers C, Chown SL, Duffy GA (2015) Growth and reproduction of laboratory-reared neanurid Collembola using a novel slime mould diet. *Sci Rep* 5:11957.
- 27. Greenslade P, Kitching RL (2011) Potential effects of climatic warming on the distribution of Collembola along an altitudinal transect in Lamington National Park Queensland, Australia. *Mem Queensl Mus – Nature* 55:333-347.
- 28. Terauds A, Chown SL, Bergstrom DM (2011) Spatial scale and species identity influence the indigenous–alien diversity relationship in springtails. *Ecology* 92:1436-1447.
- 29. Weldon CW, Terblanche JS, Chown SL (2011) Time-course for attainment and reversal of acclimation to constant temperature in two *Ceratitis* species. *J Thermal Biol* 36:479-485.
- 30. Allen JL, Clusella-Trullas S, Chown SL (2012) The effects of acclimation and rates of temperature change on critical thermal limits in *Tenebrio molitor* (Tenebrionidae) and *Cyrtobagous salviniae* (Curculionidae). *J Insect Physiol* 58:669-678.
- 31. Chown SL, Nicolson SW (2004) *Insect physiological ecology. Mechanisms and patterns* (Oxford University Press, Oxford).
- 32. Sinclair BJ *et al*. (2016) Can we predict ectotherm responses to climate change using thermal performance curves and body temperatures? *Ecol Lett* 19:1372-1385.
- 33. Everatt MJ, Bale JS, Convey P, Worland MR, Hayward SAL (2013) The effect of acclimation temperature on thermal activity thresholds in polar terrestrial invertebrates. *J Insect Physiol* 59:1057-1064.
- 34. Lutterschmidt WI, Hutchison VH (1997) The critical thermal maximum: data to support the onset of spasms as the definitive end point. *Can J Zool* 75:1553-1560.
- 35. van Dooremalen C, Berg MP, Ellers J (2013) Acclimation responses to temperature vary with vertical stratification: implications for vulnerability of soil-dwelling species to extreme temperature events. *Global Change Biol* 19:975-984.
- 36. Gunderson AR, Stillman JH (2015) Plasticity in thermal tolerance has limited potential to buffer ectotherms from global warming. *Proc R Soc B* 282:20150401.
- 37. Deutsch CA *et al.* (2008) Impacts of climate warming on terrestrial ectotherms across latitude. *Proc Natnl Acad Sci USA* 105:6668-6672.
- 38. Kingsolver JG, Diamond SE, Buckley LB, Grindstaff J (2013) Heat stress and the fitness consequences of climate change for terrestrial ectotherms. *Funct Ecol* 27:1415-1423.
- 39. Gibbs AG (1999) Laboratory selection for the comparative physiologist. *J Exp Biol* 202:2709-2718.
- 40. Vollmer JH, Sarup P, Kaersgaard CW, Dahlgaard J, Loeschcke V (2004) Heat and coldinduced male sterility in *Drosophila buzzatii*: genetic variation among populations for the duration of sterility. *Heredity* 92:257–262.
- 41. Jørgensen KT, Sørensen JG, Bundgaarda J (2006) Heat tolerance and the effect of mild heat stress on reproductive characters in *Drosophila buzzatii*. *J Thermal Biol* 31:280-286.
- 42. Zizzari ZV, Ellers J (2011) Effects of exposure to short-term heat stress on male reproductive fitness in a soil arthropod. *J Insect Physiol* 57:421-426.
- 43. Sgrò CM, Blows MW (2004) The genetic covariance among clinal environments after adaptation to an environmental gradient in *Drosophila serrata*. *Genetics* 167:1281-1291.
- 44. Conner JK, Hartl DL (2004) *A primer of ecological genetics* (Sinauer Associates, Sunderland, Massachusetts, USA).
- 45. R Core Team (2016) R: A language and environment for statistical computing (R Foundation for Statistical Computing, Vienna, Austria).
- 46. J. Felsenstein J (1985) Phylogenies and the comparative method. *Am Nat* 125:1-15.
- 47. Kellermann KV *et al.* (2012) Upper thermal limits of *Drosophila* are linked to species distributions and strongly constrained phylogenetically. *Proc Natnl Acad Sci USA* 109:16228-16233.
- 48. Garland T, Ives AR (2000) Using the past to predict the present: Confidence intervals for regression equations in phylogenetic comparative methods. *Am Nat* 155:346-364.
- 49. Orme CDL, Freckleton RP, Thomas GH, Petzoldt T, Fritz SA, Isaac NJB, Pearse WD CDL (2013) The caper package: comparative analysis of phylogenetics and evolution in R. https://cran.r-project.org/web/packages/caper/vignettes/caper.pdf. Accessed 2017-04-20.
- 50. Paradis E, Claude J, Strimmer K (2004) APE: Analyses of phylogenetics and evolution in R language. *Bioinformatics* 20:289–290.
- 51. D'Haese C (2002) Were the first springtails semi-aquatic? A phylogenetic approach by means of 28S rDNA and optimization alignment. *Proc R Soc B* 269:1143-1151.
- 52. Malcicka M, Berg MP, Ellers J (2017) Ecomorphological adaptations in Collembola in relation to feeding strategies and microhabitat. *Eur J Soil Biol* 78:82-91.
- 53. Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4:406-425.
- 54. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol Biol Evol* 30:2725-2729.
- 55. Kimura M (1980) A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *J Mol Evol* 16:111-120.
- 56. Grafen A (1989) The phylogenetic regression. *Phil Trans R Soc B* 326:119-157.
- 57. Rohlf FJ (2001) Comparative methods for the analysis of continuous variables: geometric interpretations. *Evolution* 55:2143–2160.
- 58. Pagel M (1999) Inferring the historical patterns of biological evolution. *Nature* 401:877-884.
- 59. Bates D, Maechler M, Bolker B, Walker S (2015) Fitting Linear Mixed-Effects Models Using lme4. *J Statist Software* 67:1-48.
- 60. Zuur A, Leno EN, Walker N, Saveliev AA, Smith GM (2009*) Mixed effects models and extensions in ecology with R.* (Springer-Verlag New York).
- 61. Falconer DS (1989*) Introduction to quantitative genetics*. (Longmans, New York).
- 62. Hangartner S, Hoffmann AA (2016) Evolutionary potential of multiple measures of upper thermal tolerance in *Drosophila melanogaster*. *Funct Ecol* 30:442-452

Fig. S1

Plot of the relationship between CT_{max} (°C) and CT_{min} (°C) illustrating the difference among the indigenous (green circles) and alien (orange circles) species. Statistics for the ordinary least squares regressions provided in Table S1. Gray shading represents 95% confidence intervals.

Fig. S2

Mean CT*min* (± SE) for temperate species *Lepidocyrtus* sp. 10, and *Orthonychiurus* sp., and tropical species *Ascocyrtus* sp. 2, and *Desoria trispinata* for selected lines evolved under an elevated temperature (red: tropical = 27° C, temperate = 25° C) and control lines under control temperature (blue: tropical = 20° C, temperate = 15° C). Open symbols to the right indicate the outcomes of a reciprocal transplant experiment at generations four or six, determining the contribution of developmental plasticity to CT_{max} . Here, springtails from selection and control groups were reared under either their standard acclimation temperature (Acc.) or transplanted to the thermal environment of the opposing group (Trans.) and reared for one generation.

Variation between indigenous and alien springtails in the relationship between maximum environmental temperature and critical thermal maxima (CTmax; A), minimum environmental temperature and critical thermal minima (CTmin; B), and critical thermal ranges (CTrange) and maximum (C) and minimum (D) environmental temperatures. Lines represent the fits of ordinary least squares regression for indigenous (green symbols) and alien (orange symbols) species separately with 95% confidence bands (grey shading).

Fig. S4

Schematic of the rearing and experimental design used for this study.

Fig. S5

Phylogeny of the springtails used in this study. The tree was developed based on previous, molecular marker-based assessments of phylogenetic relationships between the major springtail taxa (Refs 51, 52), with species placements made on the basis of the mt COI (mitochondrial gene cytochrome c oxidase subunit I) data collected for this study, or on morphological similarity where barcodes were not available.

Table S1.

Outcome of Phylogenetic Generalised Least Squares (PGLS) and ordinary least squares (OLS) squares analyses showing the relationship between CT_{max} (°C) and CT_{min} (°C) and the difference among indigenous and alien species (model form $CT_{max} \sim CT_{min} +$ status).

Table S2.

Results from nested mixed effects models assessing the main and interactive effects of temperature treatment (control or selection) and generation on the CT_{max} of springtails that had undergone selection for thermal tolerance. Replicate lines are nested within treatment as a random effect.

Table S3.

Observed and expected cumulative responses of CT_{max} to selection. Observed responses (R_{obs}) are calculated from the difference between control and selected lines. Expected responses (R_{exp}) are estimated from the equation $R = h^2 i_p$, assuming a heritability of 20% and 30%. *R* is the expected response, h^2 is the heritability for CT_{max} , *i* is the cumulative intensity of selection, and β is the phenotypic standard deviation. The cumulative response to selection was estimated based on the proportion adults surviving at each of the selection temperature treatments (Table S10) (following ref. 61). The expected response to selection was estimated assuming a heritability of 20% and 30%, which reflect the range of heritability values for this trait (62).

Table S4.

Results from nested mixed effects models assessing the effect of developmental plasticity on the *CT_{max}* of springtails. The trial involved rearing individuals from the selection conditions under the control conditions and *vice versa*, and then comparing their traits. Replicate line is nested within treatment as a random effect. $CL =$ Control lines; $SL =$ Selected lines; $CLT =$ Control lines transplanted to high temperature; SLT = Selected lines transplanted to low temperature.

Alien temperate	Treatment (SLT vs	-0.15 ± 0.08	1.98	0.195
	CL)			
Orthonychiurus sp.	Treatment (CLT vs	0.09 ± 0.12	0.72	0.544
	SL)			
	Random Effects	Percentage of		
		variation explained		
	Replicate line (SLT vs	0.54%		
	CL)			
	Residual (SLT vs CL)	99.46%		
	Replicate line (CLT	4.29%		
	vs SL)			
	Residual (CLT vs SL)	95.71%		
Alien tropical	Treatment (SLT vs	-0.04 ± 0.09	-0.45	0.696
	CL)			
Desoria trispinata	Treatment (CLT vs	0.16 ± 0.06	2.92	0.004
	SL)			
	Random Effects	Percentage of		
		variation explained		
	Replicate line (SLT vs	4.10%		
	CL)			
	Residual (SLT vs CL)	95.90%		
	Replicate line (CLT	<0.01%		
	vs SL)			

Residual (CLT vs SL) 86.37%

Table S5.

Results from nested mixed effects models assessing the main and interactive effects of temperature treatment (control or selection) and generation on the CT_{min} of springtails that had undergone selection for thermal tolerance. Replicate lines are nested within treatment as a random effect.

Table S6.

Results from nested mixed effects models assessing the effect of developmental plasticity on the CT_{min} of springtails. The trial involved rearing individuals from the selection conditions under the control conditions and *vice versa*, and then comparing their traits. Replicate line is nested within treatment as a random effect. $CL =$ Control lines; $SL =$ Selected lines; $CLT =$ Control lines transplanted to high temperature; SLT = Selected lines transplanted to low temperature.

Table S7.

Lack of variation in thermal tolerance phenotypic plasticity with either basal trait values or with microhabitat temperature. In all cases $df = 2.27$ and $R^2 \sim 0$ for the linear models.

Table S8.

List of species used in this study, latitude, control and acclimation temperatures, and body mass. Species marked with an asterisk (*) have been successfully barcoded (Table S9), sd = standard deviation, $n = 40$ throughout for mass determinations.

Table S9.

List of species successfully barcoded.

Table S10.

Results from a three-week pilot study assessing springtail mortality, fecundity and egg viability at two high temperatures (tropical = 27° C and 30° C, temperate = 25° C and 27°C) to determine suitable thermal conditions for high temperature treatment lines in the selection experiment. Sample size of 30 individuals per species, per temperature treatment.

Table S11. Akaike Information Criterion (AIC) of Phylogenetic Generalised Least Squares models using a variance-covariance matrix following either a Brownian Motion (BM) or Ornstein-Uhlenbeck (OU) phylogenetic correlation structure. The AIC value of the preferred structure for each model, as indicated by the lowest AIC value, is in bold.

Table S12. Outcome of Phylogenetic Generalised Least Squares analyses using an Ornstein-Uhlenbeck model of evolutionary change showing change in thermal tolerance (°C) with median (or maximum for warming tolerance) daytime soil surface temperature (°C) and the difference among indigenous and alien species. Alpha is the maximum likelihood estimate of the α-parameter, a measure of phylogenetic effect under an Ornstein-Uhlenbeck model of evolutionary change. Note the similar outcomes to those undertaken assuming a Brownian motion model of evolutionary change (Table 1 of the main text).

Warming tolerance

Table S13. Outcome of the Ordinary Least Squares analyses showing change in thermal tolerance (°C) with median daytime soil surface temperature (°C) and the difference among indigenous and alien species. Note the small change in significance of Status in CT_{min} from the PGLS model, and some changes in the estimates in both models, but otherwise consistency with the PGLS model outcomes.

