

HOKKAIDO UNIVERSITY

| | |
|---|---|
| Title | Spatiotemporal fluctuations in natural selection acting on the gall-parasitic aphid Tetraneura sorini |
| Author(s) | Muramatsu, Koichi; Akimoto, Shinichi |
| Citation | Journal of Evolutionary Biology, 29(7), 1423-1436<br>https://doi.org/10.1111/jeb.12881 |
| Issue Date | 2016-07 |
| Doc URL | https://hdl.handle.net/2115/82052 |
| Rights | This is the peer reviewed version of the following article: Muramatsu, K. and Akimoto, S. (2016), Spatiotemporal fluctuations in natural selection acting on the gall-parasitic aphid Tetraneura sorini. J. Evol. Biol., 29: 1423-1436, which has been published in final form at https://doi.org/10.1111/jeb.12881. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Self-Archiving. |
| Type | journal article |
| File Information | Tetraneura sorini. J. Evol. Biol._29_1423.pdf |




Hokkaido University Collection of Scholarly and Academic Papers : HUSCAP

# Spatiotemporal fluctuations in natural selection acting on the gall-parasitic aphid *Tetraneura sorini*

K. MURAMATSU & S. AKIMOTO

*Department of Ecology and Systematics, Graduate School of Agriculture, Hokkaido University, Sapporo, Japan*

Running title: Fluctuating selection on an aphid

*Correspondence:* Koichi Muramatsu, Laboratory of Systematic Entomology, Department of Ecology and Systematics, Graduate School of Agriculture, Hokkaido University, Sapporo 060-8589, Japan.
Tel.: +81-11-706-2486; fax: +81-11-706-4939
e-mail: muram-k@res.agr.hokudai.ac.jp

**Abstract**

The measurement of the selection gradient is crucial for understanding the magnitude of selection acting directly on a trait and predicting the evolutionary trajectory of that trait. The present study evaluated the selection gradient acting on the morphology of the gall-parasitic aphid *Tetraneura sorini* during the galling process and compared the strength among populations. Gall formers (first instars) frequently fight with conspecifics or heterospecifics for usurping incipient galls using their well-developed hind legs. First instars that successfully acquired galls were found within galls, whereas those that failed were found dead on leaf surfaces. Selection gradients were estimated using logistic stepwise regression and partial least square (PLS) regression. Calculated selection differentials indicated that first instars that secured galls were larger in body size than failed individuals through all populations. However, selection gradients on weapon traits varied largely among populations or among years in the same population. We confirmed microevolutionary changes in the relationship between traits, which accorded with the expectation from changes in the selection gradients. When gall formers were transferred onto developing buds individually, individuals that successfully induced galls had smaller body size than failed individuals. Available evidence suggests that the selection gradient on body size becomes higher with an increasing proportion of *T. sorini* in the *Tetraneura* species community. Thus, we concluded that more intense fighting with conspecifics leads to stronger selective pressure on body size, but that selective pressure for each trait is variable depending on differences in the tactics and species composition among populations.

*Keywords:*

Eriosomatinae;

fight;

gall;

multiple regression;

selection differential;

selection gradient;

partial least square regression.

## Introduction

One of the main challenges in evolutionary ecology is to elucidate evolutionary processes in which the traits of organisms vary over both space and time in response to changes in the abiotic and biotic environments. To understand the process of trait evolution, it is important to measure the strength and direction of selective pressures. The selection gradient and differential, respectively, represent the direct and total effects of selective pressures on a quantitative trait, and measurement of these indices is critical for understanding the magnitude of selection on each trait and predicting the trajectory of trait evolution (Lande & Arnold, 1983; Arnold & Wade, 1984ab; Mitchell-Olds & Shaw, 1987; Schluter, 1988). Although there are a number of attempts to measure selective pressures and their variation in the wild, the causal mechanisms underlying the variability of selective pressures are poorly understood (Siepielski *et al.*, 2009). Comparisons of selective pressures among different local populations and between populations separated temporally in the same location will provide important information for understanding trait evolution.

An important role of measuring the selection gradient is to predict the direction and rate of evolution for the focal trait. Given a certain value of heritability for a trait, the direction and rate of the evolution of that trait will be determined by the magnitude of the selection gradient (Lande, 1979; Lande & Arnold, 1983; Stinchcombe *et al.*, 2002). However, attempts to associate the measured selection gradient with the microevolutionary changes in that trait in contemporary populations have been exceptional despite the importance of the confirmation of natural selection in the wild (but see Grant & Grant, 1995, 2002). If temporal changes in the selection gradients are investigated coupled with the evaluation of trait relationships through years, such

studies would provide excellent opportunities to confirm microevolutionary changes driven by natural selection.

There are several examples in which an arthropod species genetically differentiates among local populations, as means of coping with local differences in host plant (Komatsu & Akimoto, 1995; Gandon *et al*., 1998; Toju, 2008) or host insects (Baba *et al.*, 2012; Pekár *et al*., 2012). Nevertheless, few studies have directly evaluated the direction and magnitude of selective pressures on the herbivore and parasite populations, as well as their spatiotemporal variability. This is most likely due to the difficulties associated with discerning the reproductive success and failure of individuals in the wild. However, in order to understand the evolutionary process of ecological traits of herbivores and parasites, identifying and evaluating the natural selection pressures is crucial. This is because other factors – specifically abiotic factors and genetic drift – could also drive the diversification of ecological traits among populations despite the equivalence of biotic environments (Nuismer *et al.*, 2010).

In the present study, by using a “parasitic” gall-forming aphid, *Tetraneura sorini* Hille Ris Lambers 1970, we attempt to measure selective pressures acting on the morphological traits and compare the selection gradients among spatially and temporally separated populations. In particular, by focusing on the same location, this study attempts to identify the different actions of selection between populations separated in time. Our principal object is to test the prediction from changes in the selection gradients by evaluating temporal changes in size relationships between the traits in the same location. For this purpose, we repeated sampling over years in two locations.

Several studies have reported fighting behaviour among first instar foundresses of gall-forming aphids over optimal galling sites or incipient galls (Whitham, 1979; Aoki

& Makino, 1982; Akimoto, 1988; Akimoto & Yamaguchi, 1997; Ngakan & Yukawa, 1997; Inbar, 1998). *Tetraneura sorini* is an aphid species that induces a closed pouch gall on the leaf surface of the Japanese elm, *Ulmus davidiana*. The first instar foundress of *T. sorini* is peculiar in morphology, being characterized by well-developed hind legs and large body size (Akimoto & Yamaguchi, 1997). Unlike other *Tetraneura* species, *T. sorini* foundresses frequently fail to induce their incipient galls. However, when failing in initial gall formation, foundresses choose an alternative tactic, that of depriving other conspecific and heterospecific foundresses of incipient galls (Akimoto & Yamaguchi, 1997). Thus, conflicts frequently occur between *T. sorini* foundresses or between those of *T. sorini* and other *Tetraneura* species (Blackman & Eastop, 1994). *Tetraneura sorini* foundresses that fail in galling often attack another gall owner, using a kicking motion with their hind legs or by pushing the gall owner out of the incipient gall. In instances of successful usurpation, *T. sorini* foundresses can develop and transform the incipient galls into mature galls, inside which they reproduce parthenogenetically. We hypothesize that competitive interaction between foundresses has been a primary driver in the evolution of the peculiar morphology of *T. sorini* foundresses. It is most likely that a large body size and well-developed hind legs are advantageous in fighting, and thus, positive directional selection will act on both traits. The most significant advantage of using *T. sorini* for investigating selective pressures is that it is relatively easy to distinguish between foundresses that have survived the galling process and those that have failed; foundresses that succeeded in gall formation or usurpation of galls are found inside galls, whereas those that failed either die or are found wandering on the leaf surface. Thus, we can easily detect the operations of selection imposed by fighting by collecting foundresses from developing leaves at an appropriate point in time.

145 The estimates of the selection gradient based on the Lande-Arnold (1983)
146 regression technique are biased if environmental factors induce covariance between
147 traits and fitness (Alatalo *et al*., 1990; Rausher, 1992; Kruuk *et al*., 2002; Stinchcombe
148 *et al*., 2002). Wade and Kalisz (1990) recommend the reciprocal use of the selection
149 gradient estimated by multivariate selection analysis and manipulative experiments in
150 the field to identify the agent of selection. *Tetraneura sorini* foundresses that
151 successfully survive the galling process experience conflicts, more or less, with other
152 foundresses, so that the selection gradient for weapon traits should include a component
153 from interaction with other foundresses. To evaluate the selection gradient for weapon
154 traits in a non-competitive situation, we conducted an experiment where *Tetraneura*
155 *sorini* foundresses are confined individually on developing host buds and allowed to
156 induce galls singly. By conducting this transplantation experiment, we were able to
157 distinguish selective pressures arising from gall usurpation from those from gall
158 induction.

159 The standardized selection gradients ($\beta$) were evaluated in the traditional manner,
160 through multiple regression analysis (Lande & Arnold, 1983; Arnold & Wade, 1984a, b).
161 However, for analysis of selection on morphological traits, it is expected that estimation
162 errors for selection gradients become excessive due to multicollinearity arising from
163 high correlations between morphological traits (Mitchell-Olds & Shaw, 1987). In the
164 present study, we proposed a new method to overcome this problem, by estimating
165 selection gradients ($\beta$) via partial least square regression (PLS), as well as the use of
166 classical technique. We tested differences in $\beta$ estimated by PLS by checking whether
167 the 95% bootstrap confidence intervals overlap or not.

168 The goal of this study was therefore to address the following five issues: (1)
169 whether positive directional selection is exerted on weapon traits (e.g. hind femur) of *T.*

*sorini* foundresses via gall usurping behavior; (2) whether selective pressures on *T. sorini* foundresses act heterogeneously between study years and among locations; (3) whether spatiotemporal fluctuations in selective pressures, if any, result from variation in species composition of *Tetraneura* aphids; (4) whether temporal fluctuations of selective pressures, if any, lead to changes in the relationship between traits on which selection is acting; and (5) whether *T. sorini* foundresses in a non-competitive environment are subject to selective pressures distinct from those in the wild environment.

## Materials and methods

### Study organisms

Aphids of the genus *Tetraneura* induce closed galls on the leaves of *Ulmus* species (Blackman & Eastop, 1994). In Hokkaido, northern Japan, galls of six *Tetraneura* species, including *T. sorini*, are found on *U. davidiana* (Akimoto, 1995). In early to mid-May, the first instar foundress of *Tetraneura* species settles on the underside of a young leaf and stimulates one point of the leaf using her stylet, thereby inducing a shallow depression (incipient gall pit). The stimulated plant tissues rapidly proliferate upward, encasing the foundress inside, to form a hollow, bean-shaped gall, 5 to 20 mm tall. When first instar foundresses of *T. sorini* fail to induce gall pits, they try to usurp gall pits induced by other foundresses by pushing the residents away from the pits (Akimoto & Yamaguchi, 1997). When facing resistance by the residents, *T. sorini* attackers grasp and kick them with their well-developed hind legs (Fig. 1). Once removed from gall pits, the residents have few chances to induce galls again because galls can be induced only on young leaves at an appropriate developmental stage (Burstein & Wool, 1993; Akimoto & Yamaguchi, 1994; Yukawa, 2000). Thus, the losers

of fights lose almost all their fitness, and fighting would impose strong selection on body size and the dimensions of weapon traits of *T. sorini*. *Tetraneura sorini* usually coexists with *T. nigriabdominalis, T. triangula*, and *T. radicicola* (Akimoto, 1995; Fig. 1), and the species composition of *Tetraneura* aphids that co-occur with *T. sorini* varies among locations.

**Collection and measurements**

First instar foundresses of *T. sorini* were collected from four locations in Hokkaido, northern Japan and two locations in Honshu, the mainland of Japan. At Forestry Research Institute of Hokkaido, Bibai City, Hokkaido (43.28961N, 141.8531E), we collected first instars three times over years, on 22 May 1987, 10 June 1990, and 17 May 2014, while collection was made once at the following three locations in Hokkaido; Sapporo (43.08956N, 141.4392E) on 20 May 2014, Sunagawa (43.54674N, 141.9251E) on 28 May 2014, and Lake Utonai (42.68131N, 141.70734E) on 27 May 2015. In Honshu, we collected first instars twice at Kashiwa (35.89417N, 139.9415E) on 30 April 2012 and on 29 May 2015, and at Sagae (38.36408N, 140.26542E) on 10 May 2015. Thus, first instar foundresses were collected from six locations, including two collections in Kashiwa and three collections in Bibai. In each collection, developing leaves of *U. davidiana* with incipient galls and foundresses were haphazardly collected from one tree; thus, we collected not only *T. sorini* foundresses but also those of other *Tetraneura* species without identifying the species. Collected leaves were put in plastic bags and transferred to the laboratory within 2 hours after collection. In the laboratory, we classified the first instar foundresses into two categories under binocular microscopes; (1) foundresses found in closed incipient galls (success group), and (2) foundresses found moving or dead on the leaf surface (failure group). The foundresses

were separately preserved in vials of 80% ethanol, and then mounted on glass slides with Canada balsam or Euparal as a mountant after several processes of chemical treatments of the samples (van Emden, 1972). For measurements of body dimensions, the images of the first instars were captured into a computer via a microscope camera (Dino-Eye AM423, AnMo Electronics Corp., Taipei, Taiwan), and the lengths of morphological traits were measured using the software ImageJ version 1.47 (Abràmoff *et al.*, 2004 available from http://rsbweb.nih.gov/ij/). The traits we measured in this study were the lengths of (1) hind femur, (2) hind tibia, (3) fore femur, (4) fore tibia, (5) ultimate rostral segment, and the width of (6) the first abdominal tergite. Measurement was conducted by one researcher (K.M.). The method for measurement is indicated in supplemental information (Fig. S1), and all data for 2548 first instars were placed in Appendix S1. All specimens used for measurement were preserved in Systematic Entomology, Graduate School of Agriculture, Hokkaido University.

**Transplantation of foundresses**

*Tetraneura sorini* foundresses used in the transplantation experiment were hatched in the laboratory from overwintered eggs, which were collected on 13 May 1989 at Forestry Research Institute of Hokkaido, Bibai City. Barks with *T. sorini* eggs were scraped off the trunk of *U. davidiana* trees on which abundant galls were found in the previous year. Incubation was conducted at 17℃ and 16L8D in a climate chamber. A total of 106 first instar foundresses that had just hatched were individually transferred onto a bursting bud of young *U. davidiana* trees by using a small brush. All the first instars were transferred onto different trees on the campus of Hokkaido University from 15 to 18 May 1989. At this time, the leaves were rapidly growing and most appropriate for gall formation. Buds onto which the first instars were transplanted were enveloped

by a plastic bag (140mm×100mm). The base of the plastic bag was sealed by plastic tape to prevent the escape of first instars. On 24 May, all the bagged buds were collected, and the outcomes of gall induction were recorded. First instars confined in initial galls (successes) and dead ones (failures) on leaves were preserved in vials of 80% ethanol and then mounted on glass slides. Mounting and measurement were conducted in the same way as in the wild populations, and selection gradients on the traits were estimated.

**Species composition in populations**

Foundresses collected from wild populations were composed of several *Tetraneura* species in every location. We identified all the collected first instar foundresses based on slide-mounted specimens to know the species composition, except for the samples from Bibai in 1987 and 1990. From these two samples, only *T. sorini* foundresses were slide-mounted. However, in 1988, first instar foundresses were haphazardly collected from developing host leaves at the same locality for recording the species composition. This data was used to represent species composition at Bibai 1987, and the data of Bibai 1990 were omitted from the analysis of the effect of species compositions on selection.

**Statistics**

A problem in any type of regression analysis is that there is no way to know whether a measured trait is the actual target of selection or not (Lande & Arnold, 1983; Endler, 1986; Mitchell-Olds & Shaw, 1987). If only a small number of traits were measured, it is difficult to identify the traits on which selection is actually acting (Kingsolver *et al.*, 2012). Thus, we measured as many morphological traits as possible, including non-weapon traits and tried to find out traits that are actually subject to directional

selection.

To test if there are differences in the action of selection among *T. sorini* populations, logistic regression was first performed. In the model, success or failure in gall acquisition was treated as the response variable, and the six morphological traits, localities, and the interactions between localities and each trait were included as explanatory variables. However, the morphological traits were strongly correlated with one another (Appendix S2), resulting in multicollinearity (for six traits, $0.635 \leq r \leq 0.952$, with a mean of 0.805, $2.40 \leq \text{VIF} \leq 13.7$). Thus, we additionally performed logistic stepwise regression (backward elimination technique) by using the original model. If significant interactions were detected between localities and any trait, it suggests significant differences in the action of selection among populations. Significant interactions were actually detected (see Results and Table 1), and thus we analyzed selective pressures separately in each population.

Our samples came from six locations, two of which included multiple collections in different years. We compared the magnitude of variation in the selection gradients between locations and that between years to know whether collections in different years should be lumped or treated separately. We estimated the variances of the random effects for the slope (selection gradient) and the intercept of the response variable using a linear mixed-effects model. This estimation was conducted by using the "lmer" function in the package *lme4* for R based on the restricted maximum likelihood (REML) method (Bates *et al*., 2014). Success/failure was analyzed for each morphological trait with the model comprising trait length + random effect of trait length between locations + random effect of trait length between years nested in location. In this model, we assumed that both slopes and intercepts vary randomly between locations or years. The dataset of two locations (Bibai and Kashiwa) with multiple collections were used for the

295 REML estimates of the variances. *F* test was used to test the difference between
296 variance among locations and that among years in each location.
297 For evaluating selective pressures in each population, we used three methods:
298 logistic stepwise regression, multiple regression, and partial least square regression
299 (PLS) analyses. In the logistic regression model, success or failure in gall acquisition
300 was used as the response variable, and the standardized values of the six morphological
301 traits were treated as explanatory variables, to which the stepwise method was applied
302 to evaluate the significance of selected traits (Koshio *et al.*, 2007) (Table 2). However,
303 this analysis does not provide exact estimates of standardized selection gradients ($\beta$).
304 Thus, to estimate them, we performed multiple regression for the selected traits, with
305 success (1) or failure (0) in gall acquisition treated as a continuous variable (Table 2)
306 (Morrissey & Sakrejda, 2013).
307 Logistic stepwise regression selects explanatory variables based on a criterion of
308 AIC or BIC but does not consider the magnitudes of selective pressures on all the
309 explanatory variables. To circumvent this drawback, we applied PLS analysis to
310 describe selection gradients for all traits by using package *pls* in R (Mevik *et al.*, 2013).
311 PLS analysis uses latent variables consisting of linear combinations of the original
312 variables (trait lengths) and seeks such latent variables that maximize the covariance
313 between the latent variables and the response variables. PLS analysis can avoid
314 multicollinearity by using latent variables that are uncorrelated with one another and
315 provide the estimates of partial regression coefficients ($\beta$) for all traits stably even if the
316 sample size is limited (Wold *et al.*, 2001; Tobias, 1995; Carrascal *et al.*, 2009;
317 Mehmood *et al.*, 2012). After the latent variables are fixed, the number of latent
318 variables used in the model is determined so as to minimize the root mean squared error
319 of prediction (RMSEP) by using a 10-fold cross validation method (Schluter, 1988;

320 Mehmood *et al.*, 2012; Krstajic *et al.*, 2014). After these PLS analyses were performed,
321 we estimated confidence intervals for the estimated $\beta$ by 1000 times bootstrapping using
322 package "boot" in R (Canty & Ripley, 2015). Statistical differences in the estimated $\beta$
323 between different populations were tested by checking whether the 95% confidence
324 intervals of the $\beta$ values overlap or not.

325 For each trait in a population, selection differentials were calculated as the
326 difference in the standardized trait size between successful and failed individuals (Table
327 2). Significance of the differentials was tested with *t* test. In addition to the estimation of
328 the $\beta$ and selection differential for each morphological trait, we estimated them for
329 general body size by calculating PC1 scores. Principal component analysis for pooled
330 data indicated that PC1 accounted for 87.8% of the variance and that PC1 loadings for
331 the six traits were of almost the same positive values (0.37 ~ 0.43), suggesting that PC1
332 can be used as a good index of body size. The standardized selection gradient ($\beta$) for
333 PC1 was estimated by GLM in each population.

334 We tested the correlation between the standardized selection gradient ($\beta$)
335 estimated from PLS analysis and the absolute trait size of each trait in each population
336 by using Spearman's rank sum correlation. In addition, the correlation between $\beta$ for
337 each trait in a population and the slope of linear regression of that trait on PC1 was also
338 tested. This test was conducted because several studies have reported that the strength
339 of selection is reflected in the allometry coefficient of the traits on which directional
340 selection is acting (Petrie, 1988; Wilkinson, 1993; Simmons & Tomkins, 1996; Emlen,
341 1996; Bonduriansky, 2007).

342 For the Kashiwa and Bibai populations from which samples were obtained twice and
343 three times, respectively, we tested the difference in the regression slope of the fore tibia
344 or hind tibia on the hind femur between different years by using ANCOVA. In the

345 model, the length of the fore tibia or hind tibia was treated as the response variable,
346 while the length of the hind femur, years, and the interaction between hind femur and
347 years were treated as the explanatory variables. The difference between the regression
348 slopes was tested by checking whether the interaction term was significant or not.
349 To test if there is difference in species composition among populations, we used
350 generalized linear model with Poisson error structure (GLM-p). The number of
351 individuals of each species was used as the response variable, while localities, species
352 and interaction between them were treated as explanatory variables. The contribution of
353 each explanatory variable to the model was tested with likelihood ratio test.
354 Selective pressures on foundresses in the transplantation experiment were
355 analyzed with the same models as in wild populations; stepwise logistic regression,
356 multiple regression, and PLS analysis were applied to success/failure in gall acquisition.
357 To test if there is any difference in selective pressures between the wild and transplant
358 populations, we used logistic regression analysis. In the analysis, gall acquisition
359 (success/failure) was treated as the response variable, and the six traits, the origin of
360 populations (transplantation/wild), and the interactions between the origin and
361 respective traits were treated as explanatory variables. The significance of each
362 interaction term was tested by using likelihood ratio test. For this analysis, all wild
363 populations were pooled.
364 All analyses were performed on JMP version 9.0.2 (SAS Institute Inc., 2011 Cary,
365 North Carolina) and R version 3.1.1 (R Core Team, 2014)
366
367 **Results**
368 **Natural selection on *T. sorini* foundresses**
369 Selection gradient analysis showed that the direction and magnitude of selection on

each morphological trait varied largely among populations and were unstable (Table 1 and Fig. 2). On the other hand, PC1 scores (body size) exhibited a significantly positive selective gradient ($\beta$) for all populations (Fig. 2 and Appendix S3). Similarly, calculation of the selection differential for PC1 scores indicated that in all populations, foundresses that successfully acquired galls were significantly larger in PC1 scores than those that failed (Table 2).

Body width was subject to significant positive selection in all populations except Kashiwa 2012 and 2015 in respect of the selection gradient (Fig. 2). Hind femur, a weapon trait, was subject to significant positive selection in three populations but significant negative selection in three populations. In the Bibai population, $\beta$ of hind femur fluctuated in the direction from year to year; PLS analysis confirmed that the 95% confidence intervals of $\beta$ did not overlap between samples in 1990 (-0.209 ~ -0.049) and in 2014 (0.111 ~ 0.326) (Appendix S3). A similar tendency was detected in the Kashiwa population, where $\beta$ of hind femur fluctuated between samples in 2012 (0.074 ~ 0.309) and in 2015 (-0.329 ~ -0.036). Standardized selection gradients ($\beta$) estimated by PLS analysis and those by stepwise regression were highly correlated ($\rho$ = 0.863, $P < 0.001$, $n = 28$), suggesting that PLS analysis could provide reliable estimates for $\beta$ of such traits that were not calculated by stepwise regression.

For non-weapon traits, fore femur and fore tibia, the standardized selection gradients ($\beta$) were negative or insignificant through populations. In fore femur and fore tibia, $\beta$ was significant negative in one and four populations, respectively. For these non-weapon traits, $\beta$ (PLS) ranged from -0.264 to 0.070, with a mean of -0.041 ($n$ = 18) through populations. However, for ultimate rostral segment, albeit a putative non-weapon trait, $\beta$ varied markedly; we found significant negative and positive $\beta$ in one and three populations, respectively.

When the Bibai and Kashiwa populations were used as a dataset, variance in the selection gradient between collection years in each location was larger than variance between locations for all traits but body width, although statistical significance was not detected (Table 3). Body width exhibited significantly larger between-location variance than between-year variance, and this tendency was graphically indicated in Figure 2. These results show that except for body width, variance of the selection gradient between years was as large as variance between locations, thus suggesting that samplings of different years from each of the Bibai and Kashiwa populations should be analyzed as independent data. Thus, analysis was applied to nine datasets from six locations.

For every trait, there was no significant correlation between $\beta$ (PLS) and the absolute trait size (Table 4). However, for fore tibia and ultimate rostral segment, we detected a significant positive correlation between $\beta$ (PLS) and the regression slope of the trait on PC1 scores throughout populations (Table 4). This result suggests that stronger selective pressures have accelerated the growth of these traits relative to body size.

**Morphological change and selection within populations**

Temporal changes in the action of selection led to morphological changes within populations. In the Bibai population, selection gradients for the hind femur and fore tibia changed largely between 1987-90 and 2014 (Fig. 2). In 1987-90, the hind femur and fore tibia were subject to negative or insignificant selective pressures, but in 2014 the hind femur and fore tibia were subject to strongly positive and negative selective pressures, respectively. The regression of the fore tibia on the hind femur was calculated and compared among the years (Fig. 3a). ANCOVA indicated that the regression slope

of the fore tibia on the hind femur changed significantly; in 2014 the slope of the fore tibia became much lower than that for 1987 or 1990 (in 2014 *vs* 1987, for the interaction between hind femur and year, $F_{1,1}$ = 25.77, $P$ < 0.0001; in 2014 *vs* 1990, for the interaction between hind femur and year, $F_{1,1}$ = 43.19, $P$ < 0.0001). This change in the trait relationship agreed with the expectation from changes in the direction of selection.

In the Kashiwa population, selection gradients for the hind femur and hind tibia changed markedly between 2012 and 2015 (Fig. 2). ANCOVA for the regression of the hind tibia on the hind femur showed that the regression slope of the hind tibia in 2015 became much higher than in 2012 (Fig. 3b; for the interaction between hind femur and year, $F_{1,1}$ = 10.85, $P$ = 0.0010). This change accorded with the direction of selection that operated negatively on the hind femur but positively on the hind tibia in 2015.

**Transplantation of foundresses**

In the transplantation experiment, successful gall inducers were smaller than failures in body size (PC1 scores), whereas in the wild populations pooled, successful foundresses were larger in body size than failures (Fig. 4). The action of selection also differed significantly between the transplantation experiment and wild populations. Significant positive selection was detected for the hind tibia, ultimate rostral segment and body width in the wild populations pooled, whereas significant negative selection was found in the hind femur in the transplantation experiment (Table 5, Appendix S4 for detailed $P$ values). Likelihood ratio test indicated that in four of the six traits, the magnitude of selection varied significantly between the transplant population and wild populations pooled (for the hind femur, $\chi^2$ = 7.767, $P$ = 0.005; for the hind tibiae, $\chi^2$ = 4.967, $P$ = 0.026; for the fore femur, $\chi^2$ = 9.285, $P$ = 0.002; for the fore tibiae, $\chi^2$ = 6.427 $P$ = 0.011; for the rostrum, $\chi^2$ = 1.043, $P$ = 0.307; for the body width, $\chi^2$ = 3.254, $P$ = 0.071).

445

446 **Species compositions in populations and selection**

447 Species composition varied significantly among populations (GLM-p, $\chi^2 = 2904.1$, $P <$
448 0.001; Fig. 5). The largest species, *T. sorini*, was abundant in Kashiwa 2012, 2015,
449 Bibai 2014, Sagae, Utonai, and Sunagawa. In contrast, small-sized species, *T. triangula*
450 and *T. radicicola* (Fig. 1), were abundant in Sapporo and Bibai 1988. Regression
451 analysis showed that variation in the standardized selection gradient ($\beta$) for body size
452 was readily accounted for by the proportion of *T. sorini* foundresses in the *Tetraneura*
453 species community (Fig. 6a); that is, the higher the proportion of *T. sorini* in the
454 community, the higher $\beta$ for body size (GLM, $t_6 = 3.81$, $P = 0.009$). Similarly, we
455 detected a significant positive regression of $\beta$ for the ultimate rostral segment on the
456 proportion of *T. sorini* foundresses in each population (GLM, $t_6 = 2.78$, $P = 0.032$; Fig.
457 6b).

458

459 **Discussion**

460 **Selective pressures through parasitism**

461 The present study reveals that the selection differentials for all traits and PC1 scores are
462 positive in all populations, although significance was not found in the fore tibia and
463 ultimate rostral segment in two populations. The results from the selection differentials
464 indicate that larger foundresses are more likely to secure incipient galls than are smaller
465 foundresses in all of the wild populations. This result supports Akimoto and Yamaguchi
466 (1997)'s hypothesis that larger body size and longer hind legs are advantageous in
467 fighting. Larger body size has been reported to be advantageous in intraspecific
468 competition in other aphids as well (Whitham, 1978; Aoki & Makino, 1982). However,
469 calculation of selection gradients indicates that the direction and magnitude of selection

($\beta$) on each trait tend to vary spatiotemporally, although $\beta$ for PC1 scores is stable throughout the populations. Notably, there is a drastic change in selective pressure on a weapon trait – hind femur – that was either positively or negatively selected for depending on the study year or location.

A meta-analysis by Siepielski *et al.* (2009) suggests that the direction of selection can fluctuate temporally, whereas Morrissey and Hadfield (2012) argue that selective pressures are rather consistent such that fluctuations of selective pressures observed in several studies can be largely attributed to sampling error. Our analysis of the selection gradients supports temporal variability in selective pressures, whereas analysis of the selection differentials supports consistency in the action of selective pressures. The apparent contradiction of these results can be explained by the fact that selection differentials and gradients represent different aspects of selective pressures (Arnold & Wade, 1984ab; Phillips & Arnold, 1989; Alatalo *et al*., 1990; Geber & Griffen, 2003): a selection differential includes both the direct effects of selective pressures on the focal trait and the indirect effects of selective pressures on other traits; positive selective pressures for general body size would thus enhance selection differentials for all traits. In contrast, the selection gradient on a trait is estimated by controlling indirect effects of selection arising via correlations between that trait and others, so that subtle differences in the action of selection on the focal trait are reflected in selection gradients (Lande & Arnold, 1983; Brodie *et al.*, 1995). Therefore, our results suggest that selection for general body size (PC1) was constantly positive, whereas selection for shape elements fluctuated over space and time.

Regarding hind legs, a weapon trait, femur and tibia lengths are highly correlated with each other ($r$ = 0.951, $n$ = 2863) and equivalent in having large positive values for the selection differential. However, there was no correlation in the selection gradients ($\beta$) in each population (Spearman's $\rho$ = 0.133, $n$ = 9, $P$ = 0.0732). Although these two

segments of hind legs appear to be used integrally in fighting, there may be differences in their function during fighting; for example, the driving power of the hind legs may be generated primarily by the muscles in the hind femurs, whereas the hind tibiae may be used for pulling the gall owner out of the incipient gall in a manner akin to forceps. This example demonstrates that calculation of selection gradients is indispensable for understanding the distinctions between the trait of the selection target and traits that are highly correlated with it.

Comparison of selective pressures between the competitive (in the wild) and non-competitive (transplantation) environments clearly indicates that gall parasitism entailing fighting is the main selective agent of *T. sorini* foundresses. An alternative hypothesis is that the observed selection on body size may result from the gall formation process, not from gall parasitism. However, the results of the transplantation experiments negated this hypothesis, because foundresses that induced galls successfully were smaller than foundresses that failed to induce galls in terms of their body size, and significant negative selection was exerted on hind femurs (Table 5). In previous studies of selective pressures, few have identified the direct agent of selective pressures observed (MacColl, 2011, but see Grant & Grant, 2002; Craig *et al.*, 2007), but the results of the present study clearly indicate that gall parasitism exerts strong selective pressures on the body size of *T. sorini* foundresses.

**Selection and species composition**

Although it is difficult to fully account for the instability of the standardized selection gradient ($\beta$) for each trait, we were able to detect the link between the magnitude of $\beta$ and environmental elements for two traits. The fact that $\beta$ for body size (PC1) increases with an increasing proportion of *T. sorini* in the *Tetraneura* community suggests that more frequent fighting with conspecific foundresses creates stronger selective pressure

on body size. By contrast, if conspecific foundresses account for a lower proportion in the community, *T. sorini* will fight more frequently with smaller species, from which it is easy to usurp and secure the incipient gall. Thus, fighting with smaller heterospecific foundresses will relax selective pressures on the body size of *T. sorini*. Therefore, it is possible to explain the changes in $\beta$ for body size in terms of the changes in the proportion of *T. sorini* foundresses.

The standardized selection gradient ($\beta$) for the ultimate rostral segment in a population also increased with an increasing proportion of *T. sorini*. The rostrum in aphids is one of the feeding structures, a now-weapon trait, but the stylet within the rostrum is used by *T. sorini* as a weapon in escalated fighting (Akimoto & Yamaguchi, 1997), as *T. sorini* foundresses occasionally use the stylet to pierce competitors. Thus, longer rostrums and stylets may be advantageous in conflicts with conspecifics, and the advantage of a longer ultimate rostral segment may be greater in populations with a higher proportion of *T. sorini* foundresses.

Temporal changes in $\beta$ for the hind femur in the Bibai and Kashiwa populations are also likely to be due to temporal changes in the species composition of the *Tetraneura* community. In the Bibai population, small-sized *Tetraneura* species (*T. radicicola* and *T. triangula*) accounted for 43.5% of the *Tetraneura* foundresses in 1988 (Fig. 5). In this environment, *T. sorini* foundresses were able to push out most host foundresses from their incipient gall pits without using their hind legs (Akimoto & Yamaguchi, 1997). Thus, in a population with a high proportion of small species, weak or negative selection is expected to have operated on the hind femur of *T. sorini*. More than 20 years later, however, the proportion of the two small species had decreased to 13.7 %, whereas large-sized species (*T. nigriabdominalis* and *T. sorini* itself) became more abundant (Fig. 5). A recent increase in the abundance of *T. nigriabdominalis* galls

has already been reported in this region (Muramatsu & Akimoto, 2014). Because large-sized species are better able to resist usurpers, potential gall usurpers have to use their hind legs more frequently during fighting to acquire incipient galls. As a result, the direction of selection on the hind femur may have changed markedly over the course of the past two decades.

In the Kashiwa population, the proportion of *T. sorini* foundresses increased slightly in the community from 30% to 38% from 2012 to 2015. *Tetraneura sorini* usurpers use their hind tibiae as forceps when attempting to extract gall owners from deep gall pits. A drastic change in the target of selection from the hind femur to the hind tibia over the three-year period may be due to changes in tactics when fighting with conspecifics, although the direct cause of the change is unclear.

**Morphological change and selection**

It is noteworthy that temporal changes in selective pressures on two traits led to changes in the allometry between these traits. In the Bibai and Kashiwa populations, in accordance with the drastic changes in $\beta$ on the hind femur and other traits, allometry between the hind femur and other traits was altered in a direction that can be predicted from the changes in $\beta$. The development of hatchling morphology proceeds through utilization of the nutrition in the egg only. Even if the effect of egg size was removed statistically, allometry between morphological traits of aphid hatchlings was altered between outbreeding and inbreeding settings (Akimoto, 2006). This report suggests that allometry between morphological traits is likely to be genetically determined in aphid hatchlings. Therefore, if selection acts positively on one trait and negatively on another trait, then the size relationship between these traits should be expected to change evolutionarily. In fact, the temporal changes in allometry between the hind femur and

572 fore tibia in the Bibai population and between the hind femur and hind tibia in the
573 Kashiwa population accorded well with the changes in the direction of selective
574 pressure. These examples imply that temporal changes in selective pressures have led to
575 microevolutionary changes in allometry between morphological traits in contemporary
576 populations.

577 In addition to changes in allometry within the same population, we observed a
578 positive relationship between $\beta$ for a trait and the regression slope of that trait on body
579 size (PC1) between populations (Table 4). This result suggests that stronger selective
580 pressures have promoted the growth of the trait. This correlation was detected for the
581 fore tibia and ultimate rostral segment, but not for other traits. This result may be
582 because these two traits exhibited a large amount of variation in their $\beta$ among different
583 populations. As abovementioned, there was correlation between the standardized
584 selection gradient ($\beta$) for the ultimate rostral segment in a population and the proportion
585 of *T. sorini* in the *Tetraneura* community. Therefore, regarding the ultimate rostral
586 segment, more intense intraspecific competition is responsible for a higher value of $\beta$,
587 which in turn results in a larger slope value in the regression of that trait on body size.
588 As predicted from these correlations, there is a significant correlation between the
589 proportion of *T. sorini* and regression slope values for the ultimate rostral segment
590 across the populations (Spearman's $\rho = 0.833$, $P = 0.0154$).

591 In conclusion, by regressing six morphological traits on the survival/failure status,
592 we evaluated the selection differential and gradient for each trait and confirmed that
593 body size and some traits are subject to positive directional selection. Notably, the
594 magnitude of selection for body size and the ultimate rostral segment increased with
595 increasing intensity of intraspecific competition. Contrary to the prediction, selective
596 pressure on the hind femur was unstable. This may be due to relaxed or negative

selective pressures on weapons in population with higher proportions of smaller species. Furthermore, it became clear that temporal changes in the action of selection on more than one trait altered the allometry between the traits. The parasitic behavior of *T. sorini* foundresses consistently created directional selection for larger body size, but transfer experiments demonstrated that small-sized foundresses have a greater ability to induce galls. Although the genetic basis of gall-inducing capability is unknown, it is possible that gall formers and usurpers coexist in *T. sorini* populations, and that gall usurpers depend on gall formers of their own species or other smaller species. It is therefore necessary that future research focuses on examining the evolutionary interactions between gall formers and usurpers, and on the evolutionary impacts of gall usurpers on other *Tetraneura* species.

**Acknowledgments**

We are grateful to members of Systematic Entomology lab at Hokkaido University. We also thank S. Kudo for discussion about statistical analyses. This research was supported by Grants-in-Aid (number 17370028, 23370037) for Scientific Research from the Japan Society for the Promotion of Science given to S. Akimoto and number 15J00706 to K. Muramatsu.

**References**

Abràmoff, M. D., Magalhães, P. J. & Ram, S. J. 2004. Image processing with ImageJ. *Biophotonics Int.* **11**: 36–42.

Akimoto, S. 1995. Coexistence and weak amensalism of congeneric gall-forming aphids on the Japanese elm. *Res. Popul. Ecol.* **37**: 81–89.

621 Akimoto, S. 2006. Inbreeding depression, increased phenotypic variance and a trade-off
622 between gonads and appendages in selfed progeny of the aphid *Prociphilus oriens*.
623 *Evolution* **60**: 77–86.

624 Akimoto, S. 1988. The evolution of gall parasitism accompanied by a host shift in the
625 gall aphid, *Eriosoma yangi* (Homoptera: Aphidoidea). *Biol. J. Linn. Soc.* **35**:
626 297–312.

627 Akimoto, S. & Yamaguchi, Y. 1997. Gall usurpation by the gall-forming aphid,
628 *Tetraneura sorini* (Insecta Homoptera). *Ethol. Ecol. Evol.* **9**: 159–168.

629 Akimoto, S. & Yamaguchi, Y. 1994. Phenotypic selection on the process of gall
630 formation of a *Tetraneura* aphid (Pemphigidae). *J. Anim. Ecol.* **63**: 727–738.

631 Alatalo, R., Gustafsson, L. & Lundberg, A. 1990. Phenotypic selection on heritable size
632 traits: environmental variance and genetic response. *Am. Nat.* **135**: 464–471.

633 Aoki, S. & Makino, S. 1982. Gall usurpation and lethal fiting among fundatrices of the
634 aphid *Epipemphigus niisimae* (Homoptera, Pemphigidae). *Kontyu* **50**: 365–376.

635 Arnold, S.J. & Wade, M.J. 1984a. On the measurement of natural and sexual selection:
636 applications. *Evolution* **38**: 720–734.

637 Arnold, S.J. & Wade, M.J. 1984b. On the measurement of natural and sexual selection:
638 theory. *Evolution* **38**: 709–719.

639 Baba, Y.G., Walters, R.J. & Miyashita, T. 2012. Complex latitudinal variation in the
640 morphology of the kleptoparasitic spider *Argyrodes kumadai* associated with host
641 use and climatic conditions. *Popul. Ecol.* **55**: 43–51.

642 Bates, D., Mächler, M., Bolker, B. & Walker, S. 2014. Fitting linear mixed-effects
643 models using lme4. See http://arxiv.org/abs/1406.5823

644 Blackman, R.L. & Eastop, V.F. 1994. *Aphids on the World's Trees: an Identification*
645 *and Information Guide.* CAB International, Wallingford.

646 Bonduriansky, R. 2007. Sexual selection and allometry: A critical reappraisal of the
647 evidence and ideas. *Evolution* **61**: 838–849.

648 Brodie, E.D., Moore, A.J. & Janzen, F.J. 1995. Visualizing and quantifying natural
649 selection. *Trends Ecol. Evol.* **10**: 313–318.

650 Burstein, M. & Wool, D. 1993. Gall aphids do not select optimal galling sites
651 (*Smynthurodes betae*; Pemphigidae). *Ecol. Entomol.* **18**: 155–164.

652 Canty, A. & Ripley, B. 2015. boot: Bootstrap R (S-Plus) Functions. R package version
653 1.3-15. See http://CRAN.R-project.org/package=boot

654 Carrascal, L.M., Galván, I. & Gordo, O. 2009. Partial least squares regression as an
655 alternative to current regression methods used in ecology. *Oikos* **118**: 681–690.

656 Craig, T.P., Itami, J.K. & Horner, J.D. 2007. Geographic variation in the evolution and
657 coevolution of a tritrophic interaction. *Evolution* **61**: 1137–1152.

658 van Emden, H. F. 1972. *Aphid Technology; With Special Reference to the Study of*
659 *Aphids in the Field*. Academic Press, London.

660 Emlen, D.J. 1996. Artificial selection on horn length-body size allometry in the horned
661 beetle *Onthophagus acuminatus* (Coleoptera: Scarabaeidae). *Evolution* **50**:
662 1219–1230.

663 Endler, J. 1986. *Natural Selection in the Wild*. Princeton University Press, Princeton,
664 N.J.

665 Gandon, S., Ebert, D., Olibieri, I. & Michalakis, Y. 1998. Differential adaptation in
666 spacially heterogeneous environments and host-parasite coevolution. In: *Genetic*
667 *Structure and Local Adaptation in Natural Insect Populations: Effects of Ecology,*
668 *Life History and Behavior* (S. Mopper & S.Y. Strauss, eds), pp. 325–342. Chapman
669 and Hall, New York.

670 Geber, M.A. & Griffen, L.R. 2003. Inheritance and natural selection on functional traits.
671 *Int. J. Plant Sci.* **164**: S21–S42.

672 Grant, P.R. & Grant, B.R. 1995. Predicting microevolutionary responses to directional
673 selection on heritable variation. *Evolution* **49**: 241–251.

674 Grant, P.R. & Grant, B.R. 2002. Unpredictable evolution in a 30-year study of Darwin's
675 finches. *Science* **296**: 707–117.

676 Inbar, M. 1998. Competition, territoriality and maternal defense in a gall-forming aphid.
677 *Ethol. Ecol. Evol.* **10**: 159–170.

678 Kingsolver, J.G., Diamond, S.E., Siepielski, A.M. & Carlson, S.M. 2012. Synthetic
679 analyses of phenotypic selection in natural populations: lessons, limitations and
680 future directions. *Evol. Ecol.* **26**: 1101–1118.

681 Komatsu, T. & Akimoto, S. 1995. Genetic differentiation as a result of adaptation to the
682 phenologies of individual host trees in the galling aphid *Kaltenbachiella japonica*.
683 *Ecol. Entomol.* **20**: 33–42.

684 Koshio, C., Muraji, M., Tatsuta, H. & Kudo, S. 2007. Sexual selection in a moth: effect
685 of symmetry on male mating success in the wild. *Behav. Ecol.* **18**: 571–578.

686 Krstajic, D., Buturovic, L.J., Leahy, D.E. & Thomas, S. 2014. Cross-validation pitfalls
687 when selecting and assessing regression and classification models. *J. Cheminform.*
688 **6**: 10.

689 Kruuk, E.B., Slate, J., Pemberton, J.M., Brotherstone, S., Guinness, F. & Clutton-Brock,
690 T. 2002. Antler size in red deer: heritability and selection but no evolution.
691 *Evolution* **56**: 1683–1695.

692 Lande, R. 1979. Quantitative genetic analysis of multivariate evolution, applied to
693 brain: body size allometry. *Evolution* **33**: 402–416.

694 Lande, R. & Arnold, S. 1983. The measurement of selection on correlated characters.
695 *Evolution* **37**: 1210–1226.

696 MacColl, A.D.C. 2011. The ecological causes of evolution. *Trends Ecol. Evol.* **26**:
697 514–522.

698 Mehmood, T., Liland, K.H., Snipen, L. & Sæbø, S. 2012. A review of variable selection
699 methods in Partial Least Squares Regression. *Chemom. Intell. Lab. Syst.* **118**:
700 62–69.

701 Mevik, B.H., Liland, R.W. & Hovde, K. 2013. pls: partial least squares and principal
702 component regression. R package version 2.4-3. See
703 http://cran.r-project.org/package=pls.

704 Mitchell-Olds, T. & Shaw, R. 1987. Regression analysis of natural selection: statistical
705 inference and biological interpretation. *Evolution* **41**: 1149–1161.

706 Morrissey, M.B. & Hadfield, J.D. 2012. Directional selection in temporally replicated
707 studies is remarkably consistent. *Evolution* **66**: 435–442.

708 Morrissey, M.B. & Sakrejda, K. 2013. Unification of regression-based methods for the
709 analysis of natural selection. *Evolution* **67**: 2094–2100.

710 Muramatsu, K. & Akimoto, S. 2014. Drastic changes in host preference in a
711 gall-parasitic flea weevil, *Orchestes hustachei*, during the last decade. *Ecol. Entomol.*
712 **39**: 531–538.

713 Ngakan, P. & Yukawa, J. 1997. Synchronization with host plant phenology and gall site
714 preference of *Dinipponaphis autumna* (Homoptera: Aphididae). *Appl. Entomol. Zool.*
715 **32**: 81–90.

716 Nuismer, S.L., Gomulkiewicz, R. & Ridenhour, B.J. 2010. When is correlation
717 coevolution? *Am. Nat.* **175**: 525–537.

718 Pekár, S., Smerda, J., Hrušková, M., Sedo, O., Muster, C., Cardoso, P., *et al.* 2012.
719 Prey-race drives differentiation of biotypes in ant-eating spiders. *J. Anim. Ecol.* **81**:
720 838–848.

721 Petrie, M. 1988. Intraspecific variation in structures that display competitive ability:
722 large animals invest relatively more. *Anim. Behav.* **36**: 1174–1179.

723 Phillips, P.C. & Arnold, S.J. 1989. Visualizing Multivariate Selection. *Evolution* **43**:
724 1209–1222.

725 Rausher, M.D. 1992. The measurement of selection on quantitative traits: biases due to
726 environmental covariances between traits and fitness. *Evolution* **46**: 616–626.

727 R Core Team. 2014. R: A language and environment for statistical computing. R
728 Foundation for Statistical Computing, Vienna, Austria. See
729 http://www.R-project.org/.

730 Schluter, D. 1988. Estimating the form of natural selection on a quantitative trait.
731 *Evolution* **42**: 849–861.

732 Siepielski, A., DiBattista, J. & Carlson, S. 2009. It's about time: the temporal dynamics
733 of phenotypic selection in the wild. *Ecol. Lett.* **12**: 1261–1276.

734 Simmons, L.W. & Tomkins, J.L. 1996. Sexual selection and the allometry of earwig
735 forceps. *Evol. Ecol.* **10**: 97–104.

736 Stinchcombe, J.R., Rutter, M.T., Burdick, D.S., Tiffin, P., Rausher, M.D., Mauricio, R.,
737 *et al.* 2002. Testing for environmentally induced bias in phenotypic estimates of
738 natural selection: theory and practice. *Am. Nat.* **160**: 511–523.

739 Tobias, R.D. 1995. An introduction to partial least squares regression. In: *Proc. Ann.*
740 *SAS Users Group Int. Conf., 20th,* pp. 1250-1257, Orlando, Florida, 2-5, April 1995.
741 SAS Institute Inc., Cary, NC, USA

742 Toju, H. 2008. Fine-scale local adaptation of weevil mouthpart length and camellia
743 pericarp thickness: altitudinal gradient of a putative arms race. *Evolution* **62**:
744 1086–1102.

745 Wade, M. & Kalisz, S. 1990. The causes of natural selection. *Evolution* **44**: 1974–1955.

746 Whitham, T.G. 1978. Habitat selection by *Pemphigus* aphids in response to resource
747 limitation and competition. *Ecology* **59**: 1164–1176.

748 Whitham, T.G. 1979. Territorial behaviour of *Pemphigus* gall aphids. *Nature* **279**:
749 324–325.

750 Wilkinson, G.S. 1993. Artificial sexual selection alters allometry in the stalk-eyed fly
751 *Cyrtodiopsis dalmanni* (Diptera: Diopsidae). *Genet. Res.* **62**: 213–222.

752 Wold, S., Sjöström, M. & Eriksson, L. 2001. PLS-regression: a basic tool of
753 chemometrics. *Chemom. Intell. Lab. Syst.* **58**: 109–130.

754 Yukawa, J. 2000. Synchronization of gallers with host plant phenology. *Popul. Ecol.*
755 **42**: 105–113.

756

**Figure legends**

**Figure 1** First instar foundresses of four *Tetraneura* species coexisting on *Ulmus davidiana*. *Tetraneura sorini* is the largest in body size with well-developed hind legs.

**Figure 2** Selection gradient ($\beta$) for PC1 and six morphological traits, respectively, estimated by GLM and partial least square (PLS) regression. Error bars indicate 95% confidence intervals (CI) estimated by 1000 times bootstrapping.

**Figure 3** Chronological changes in the relationship between the lengths of two traits. (a) regression of fore tibial length on hind femoral length in the Bibai population; blue open square, 1987, $n$ = 378; blue circle, 1990, $n$ = 302; orange circle, 2014, $n$ = 232. (b) regression of hind tibial length on hind femoral length in the Kashiwa population; blue circle, 2012, $n$ = 242; orange circle, 2015, $n$ = 352.

**Figure 4** Distribution of body size (PC1 scores) for foundresses that acquired galls and those that failed in the wild populations pooled and in the transplantation experiment. Dashed lines indicate the means for respective groups.

**Figure 5** Species composition in eight populations. TS, *Tetraneura sorini*; TN, *T. nigriabdominalis*; TT, *T. triangula*; TR, *T. radicicola* and TO, *T.* sp. O.

**Figure 6** Regression of the selection gradients ($\beta$) of (a) PC1 score and (b) ultimate rostral segment on the proportion of *T. sorini* within the *Tetraneura* community.

781 **Table 1** Difference in selective pressures among populations. The response variable is
782 the survival or mortality of foundresses. Interaction terms indicate among-population
783 differences in the direction of selection.

| | Logistic regression | | | Stepwise logistic regression | | |
|---|---|---|---|---|---|---|
| Explanatory variables | df | $\chi^2$ | *P* | df | $\chi^2$ | *P* |
| Hind femoral length (HFL) | 1 | 0.659 | 0.417 | 1 | 0.971 | 0.324 |
| Hind tibial length (HTL) | 1 | 2.628 | 0.105 | 1 | 2.609 | 0.106 |
| Fore femoral length (FFL) | 1 | 2.639 | 0.104 | 1 | 2.875 | 0.090 |
| Fore tibial length (FTL) | 1 | 5.131 | **0.024** | 1 | 6.505 | **0.011** |
| Rostrum length (RL) | 1 | 11.872 | **< 0.001** | 1 | 12.331 | **< 0.001** |
| Body width (BW) | 1 | 115.091 | **< 0.001** | 1 | 110.585 | **< 0.001** |
| Population | 8 | 61.210 | **< 0.001** | 8 | 61.210 | **< 0.001** |
| HFL * Population | 8 | 41.392 | **< 0.001** | 8 | 64.580 | **< 0.001** |
| HTL * Population | 8 | 14.135 | 0.078 | - | - | - |
| FFL * Population | 8 | 10.561 | 0.228 | - | - | - |
| FTL * Population | 8 | 33.523 | **< 0.001** | 8 | 42.027 | **< 0.001** |
| RL * Population | 8 | 33.260 | **< 0.001** | 8 | 34.241 | **< 0.001** |
| BW * Population | 8 | 41.237 | **< 0.001** | 8 | 40.227 | **< 0.001** |

784

785

786

787 **Table 2** Mean length or width of morphological traits and selection differentials. The
788 selection differential indicates the difference in standardized trait sizes between
789 successful and failed individuals. *P* indicates the result of *t* test for the selection
790 differentials.

| Population (Sample size) | Trait | Mean size (mm) | Selection differential | *P* |
|---|---|---|---|---|
| Bibai 1987 | Hind femoral length | 0.316 | 0.291 | 0.005 |
| (*n* = 359) | Hind tibial length | 0.290 | 0.297 | 0.004 |
| | Fore femoral length | 0.192 | 0.225 | 0.037 |
| | Fore tibial length | 0.148 | 0.194 | 0.065 |
| | Rostrum length | 0.133 | 0.412 | < 0.001 |
| | Body width | 0.413 | 0.728 | < 0.001 |
| | PC1 | - | 0.404 | < 0.001 |
| Bibai 1990 | Hind femoral length | 0.323 | 0.364 | 0.002 |
| (*n* = 220) | Hind tibial length | 0.300 | 0.460 | < 0.001 |
| | Fore femoral length | 0.194 | 0.467 | < 0.001 |
| | Fore tibial length | 0.153 | 0.557 | < 0.001 |
| | Rostrum length | 0.135 | 0.189 | 0.119 |
| | Body width | 0.409 | 0.876 | < 0.001 |
| | PC1 | - | 0.562 | < 0.001 |
| Bibai 2014 | Hind femoral length | 0.294 | 0.996 | < 0.001 |
| (*n* = 201) | Hind tibial length | 0.271 | 0.906 | < 0.001 |
| | Fore femoral length | 0.181 | 0.854 | < 0.001 |
| | Fore tibial length | 0.141 | 0.417 | 0.001 |
| | Rostrum length | 0.132 | 0.773 | < 0.001 |
| | Body width | 0.395 | 0.877 | < 0.001 |
| | PC1 | - | 0.907 | < 0.001 |
| Sunagawa 2014 | Hind femoral length | 0.292 | 0.728 | < 0.001 |
| (*n* = 270) | Hind tibial length | 0.267 | 0.700 | < 0.001 |

| | | | | |
|---|---|---|---|---|
| | Fore femoral length | 0.180 | 0.563 | < 0.001 |
| | Fore tibial length | 0.137 | 0.561 | < 0.001 |
| | Rostrum length | 0.130 | 0.547 | < 0.001 |
| | Body width | 0.402 | 0.919 | < 0.001 |
| | PC1 | - | 0.740 | < 0.001 |
| Sapporo 2014 | Hind femoral length | 0.324 | 0.530 | 0.001 |
| (*n* = 84) | Hind tibial length | 0.295 | 0.559 | 0.001 |
| | Fore femoral length | 0.190 | 0.591 | 0.001 |
| | Fore tibial length | 0.149 | 0.583 | 0.001 |
| | Rostrum length | 0.135 | 0.417 | 0.025 |
| | Body width | 0.423 | 0.736 | < 0.001 |
| | PC1 | - | 0.521 | 0.025 |
| Utonai 2015 | Hind femoral length | 0.327 | 0.735 | < 0.001 |
| (*n* = 486) | Hind tibial length | 0.291 | 0.678 | < 0.001 |
| | Fore femoral length | 0.195 | 0.673 | < 0.001 |
| | Fore tibial length | 0.145 | 0.535 | < 0.001 |
| | Rostrum length | 0.134 | 0.720 | < 0.001 |
| | Body width | 0.397 | 0.836 | < 0.001 |
| | PC1 | - | 0.798 | < 0.001 |
| Sagae 2015 | Hind femoral length | 0.355 | 0.452 | < 0.001 |
| (*n* = 380) | Hind tibial length | 0.316 | 0.572 | < 0.001 |
| | Fore femoral length | 0.205 | 0.478 | < 0.001 |
| | Fore tibial length | 0.155 | 0.566 | < 0.001 |
| | Rostrum length | 0.142 | 0.690 | < 0.001 |
| | Body width | 0.429 | 0.566 | < 0.001 |
| | PC1 | - | 0.627 | < 0.001 |
| Kashiwa 2012 | Hind femoral length | 0.369 | 0.546 | < 0.001 |
| (n = 239) | Hind tibial length | 0.331 | 0.423 | 0.002 |
| | Fore femoral length | 0.216 | 0.337 | 0.015 |
| | Fore tibial length | 0.163 | 0.170 | 0.226 |

| | | | | |
|---|---|---|---|---|
| | Rostrum length | 0.140 | 0.209 | 0.144 |
| | Body width | 0.450 | 0.313 | 0.018 |
| | PC1 | - | 0.379 | 0.006 |
| Kashiwa 2015 | Hind femoral length | 0.363 | 0.402 | < 0.001 |
| (n = 357) | Hind tibial length | 0.330 | 0.504 | < 0.001 |
| | Fore femoral length | 0.216 | 0.485 | < 0.001 |
| | Fore tibial length | 0.166 | 0.436 | < 0.001 |
| | Rostrum length | 0.143 | 0.283 | 0.013 |
| | Body width | 0.435 | 0.488 | < 0.001 |
| | PC1 | - | 0.525 | < 0.001 |

791

792

**Table 3** REML estimates for the variances of the selection gradients on morphological traits between locations and between years in each location. Variations in the selection gradient were treated as random effects. The Bibai and Kashiwa populations are used.

| Random effect | df | Variance | *F* | *P* |
|---|---|---|---|---|
| Hind femoral length (HFL): slope | | | | |
| HFL*Locality | 1 | 0.169 | - | - |
| HFL*Locality*Year | 3 | 7.788 | 46.221 | 0.108 |
| Hind tibial length (HTL): slope | | | | |
| HTL*Locality | 1 | 0.099 | - | - |
| HTL*Locality*Year | 3 | 6.511 | 65.824 | 0.090 |
| Fore femoral length (FFL): slope | | | | |
| FFL*Locality | 1 | 0.286 | - | - |
| FFL*Locality*Year | 3 | 17.719 | 61.974 | 0.093 |
| Fore tibial length (FTL): slope | | | | |
| FTL*Locality | 1 | 0.077 | - | - |
| FTL*Locality*Year | 3 | 2.132 | 27.823 | 0.138 |
| Rostrum length (RL): slope | | | | |
| RL*Locality | 1 | 15.544 | - | - |
| RL*Locality*Year | 3 | 49.860 | 3.208 | 0.384 |
| Body width (BW): slope | | | | |
| BW*Locality | 1 | 4.422 | 6.728e+14 | < 0.001 |
| BW*Locality*Year | 3 | 6.573e-15 | - | - |
| PC1: slope | | | | |
| PC1*Locality | 1 | 0.492 | - | - |
| PC1*Locality*Year | 3 | 1.863 | 3.785 | 0.357 |

798 **Table 4** Spearman's rank correlation between the selection gradient $\beta$ (PLS), the mean
799 absolute trait size and the regression slope of each trait on PC1.

| | Absolute size | | Regression slope | |
|---|---|---|---|---|
| | $\rho$ | $P$ | $\rho$ | $P$ |
| Hind femoral length | -0.183 | 0.644 | 0.583 | 0.108 |
| Hind tibial length | 0.250 | 0.521 | -0.200 | 0.613 |
| Fore femoral length | 0.217 | 0.581 | 0.050 | 0.912 |
| Fore tibial length | 0.333 | 0.385 | 0.733 | 0.031 |
| Rostrum length | -0.317 | 0.410 | 0.766 | 0.021 |
| Body width | -0.467 | 0.213 | -0.142 | 0.715 |

800

801

802 **Table 5** Selection gradients for foundresses in the transplantation experiment and wild
803 populations pooled. * $P < 0.05$, ** $P < 0.01$ and *** $P < 0.001$.
804

| Traits | Transplantation | | | Wild population (pooled) | | |
|---|---|---|---|---|---|---|
| | Selection differential | $\beta$ (stepwise) | $\beta$ (PLS) | Selection differential | $\beta$ (stepwise) | $\beta$ (PLS) |
| Hind femoral length | -0.560** | -0.246** | -0.246 | 0.426*** | - | 0.014 |
| Hind tibial length | -0.372 | 0.190* | 0.202 | 0.453*** | 0.058** | 0.020 |
| Fore femoral length | -0.261 | 0.255** | 0.250 | 0.425*** | - | -0.015 |
| Fore tibial length | -0.628** | -0.290*** | -0.260 | 0.388*** | -0.071*** | -0.045 |
| Rostrum length | -0.430* | - | -0.037 | 0.476*** | 0.040** | 0.044 |
| Body width | -0.525* | - | -0.009 | 0.705*** | 0.130*** | 0.140 |

805
806

807 **Fig. S1** Measurement of dimensions of six traits of *Tetraneura sorini*. a, hind femoral
808 length; b, hind tibial length; c, fore femoral length; d, fore tibial length; e, ultimate
809 rostral segment length; f, width of first abdominal tergite (body width).

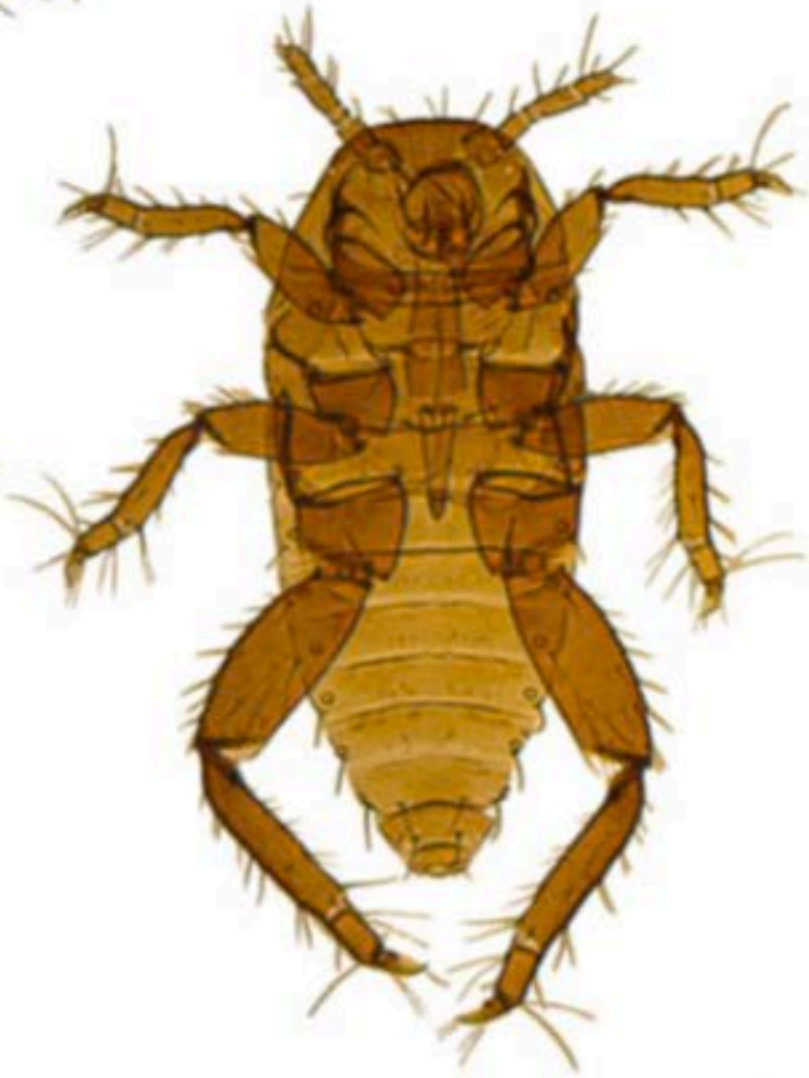
810 **Appendix S1** All data of foundresses including survival/death, locations and
811 measurements of six traits.
812
813 **Appendix S2** Spearman's rank sum correlation coefficient matrix of the six traits. All
814 populations including the transplant experiment were pooled.
815
816 **Appendix S3** Selection gradients ($\beta$) for six traits in the wild populations. $\beta$ estimated
817 by the stepwise regression method and $\beta$ estimated by PLS are shown along with SE and
818 95% confidence intervals (CI) based on 1000 times bootstrapping.
819
820 **Appendix S4** Exact *P* values for Table 5.

Fig. 1

0.1 mm



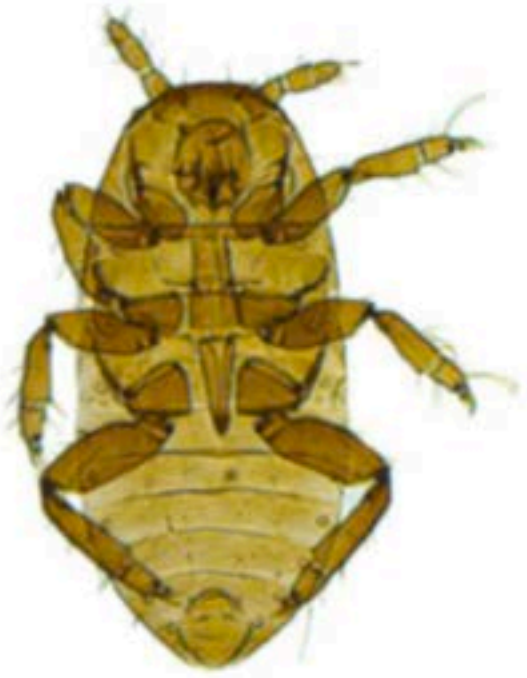
*T. sorini*



*T. nigriabdominalis*



*T. triangula*



*T. radicicola*

Fig. 2

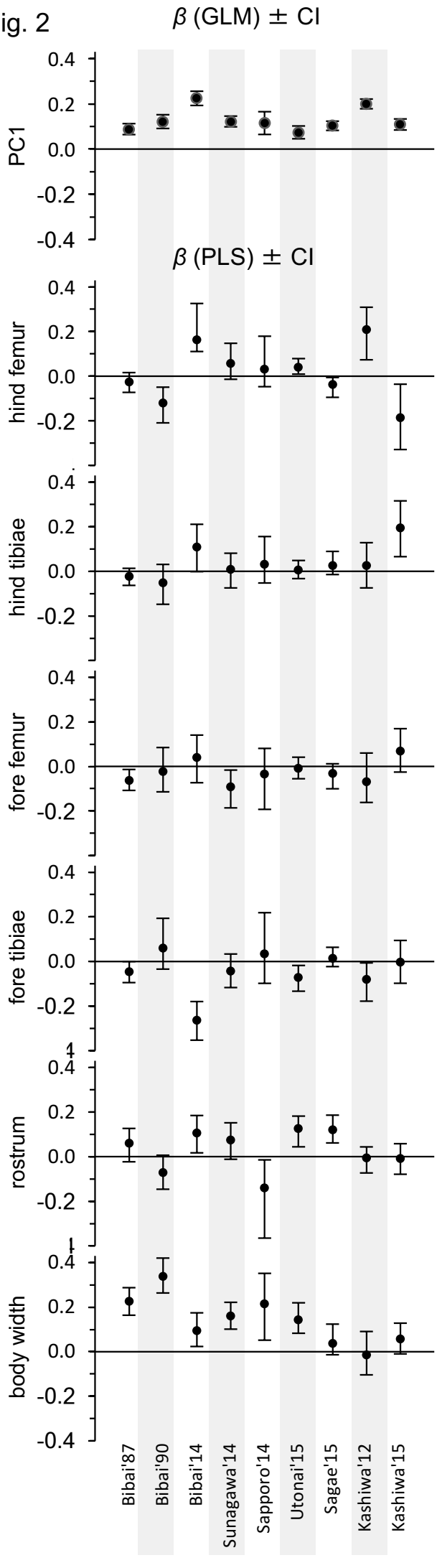

β (GLM) ± CI
PC1
β (PLS) ± CI
hind femur
hind tibiae
fore femur
fore tibiae
rostrum
body width
0.4
0.2
0.0
-0.2
-0.4
Bibai'87
Bibai'90
Bibai'14
Sunagawa'14
Sapporo'14
Utonai'15
Sagae'15
Kashiwa'12
Kashiwa'15

Fig.3

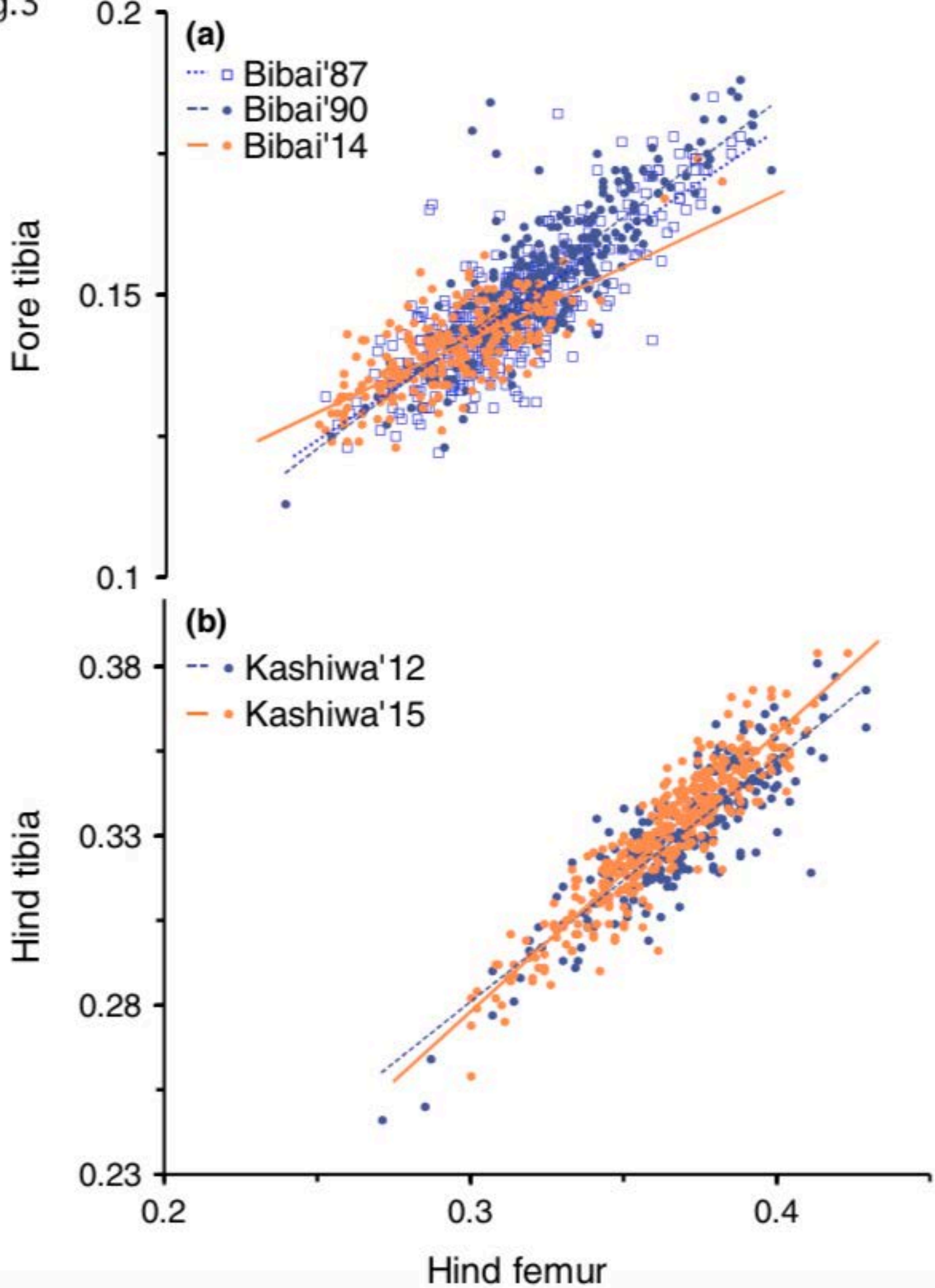

(a)
Bibai'87
Bibai'90
Bibai'14
Fore tibia
0.2
0.15
0.1
(b)
Kashiwa'12
Kashiwa'15
Hind tibia
0.38
0.33
0.28
0.23
0.2
0.3
0.4
Hind femur

Fig. 4

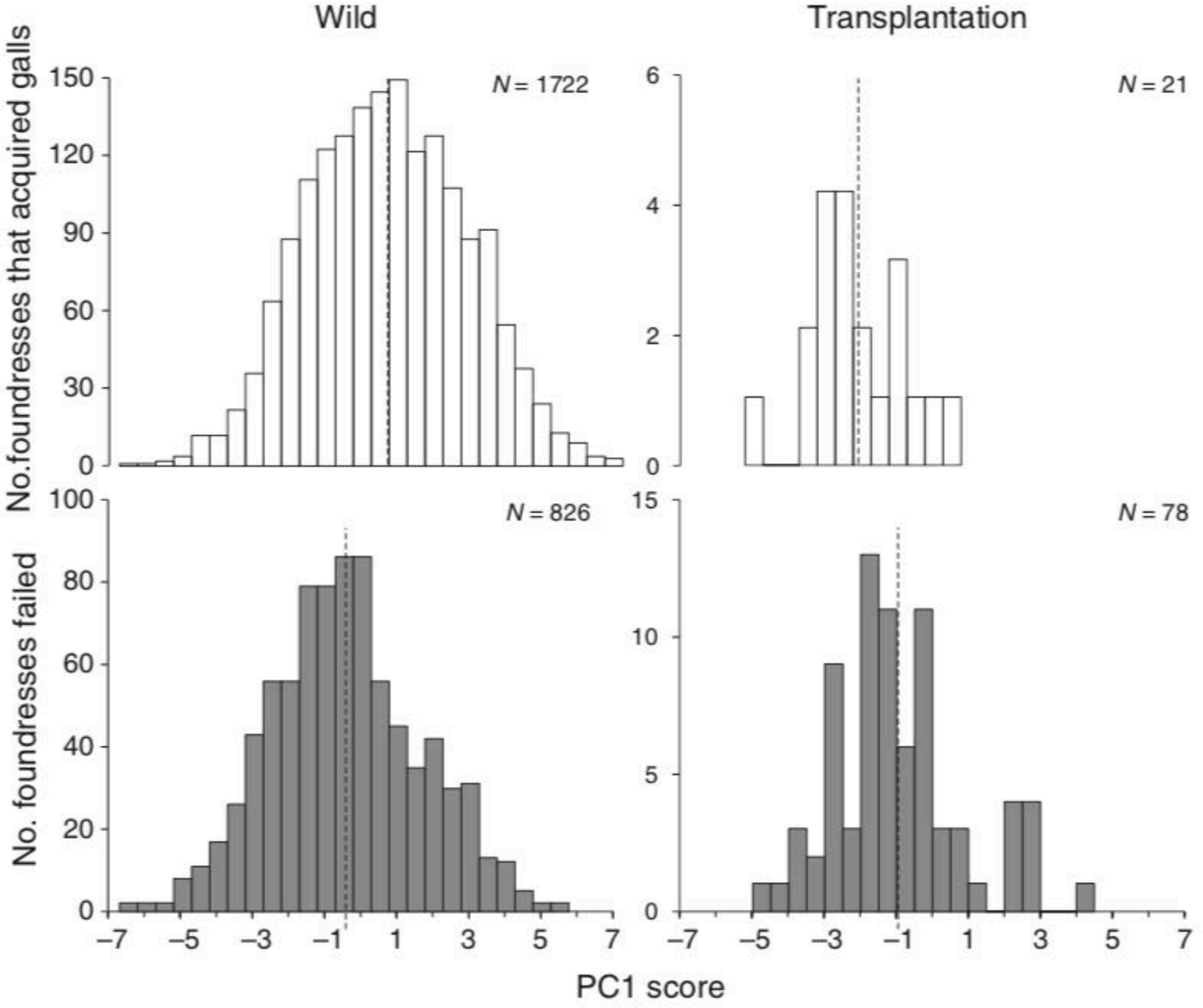

Wild
Transplantation
No.foundresses that acquired galls
No. foundresses failed
N = 1722
N = 21
N = 826
N = 78
150
120
90
60
30
0
6
4
2
0
100
80
60
40
20
0
15
10
5
0
−7
−5
−3
−1
1
3
5
7
PC1 score

Fig. 5

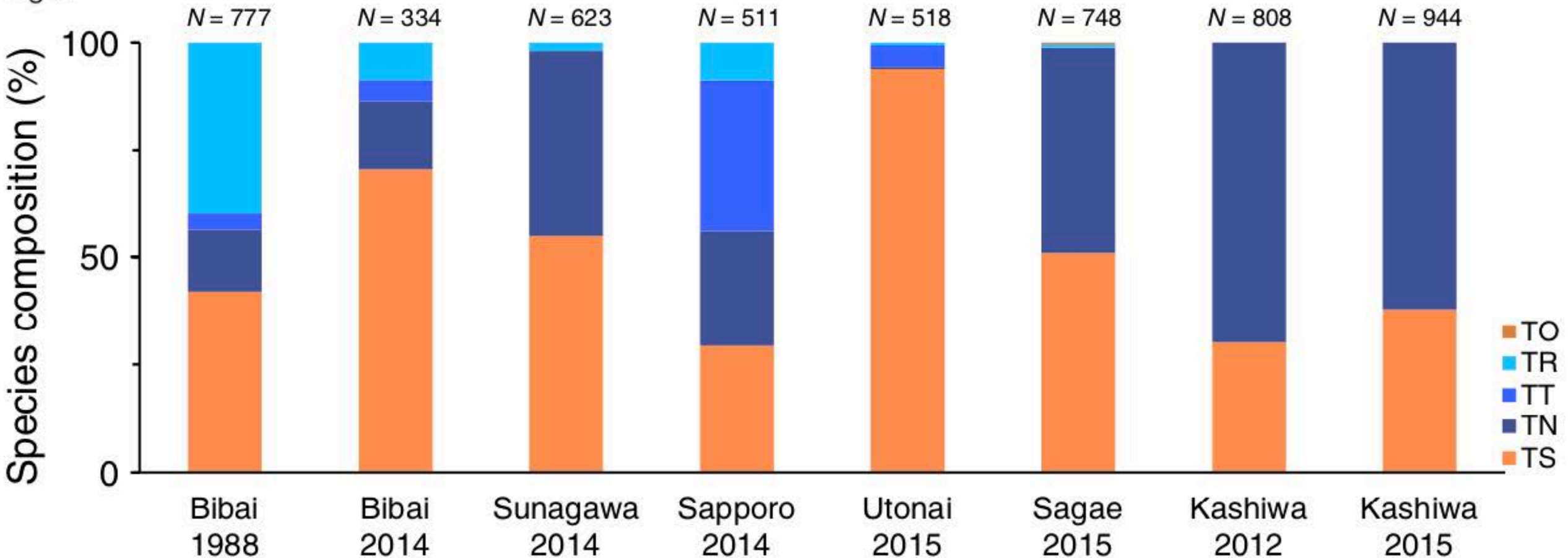
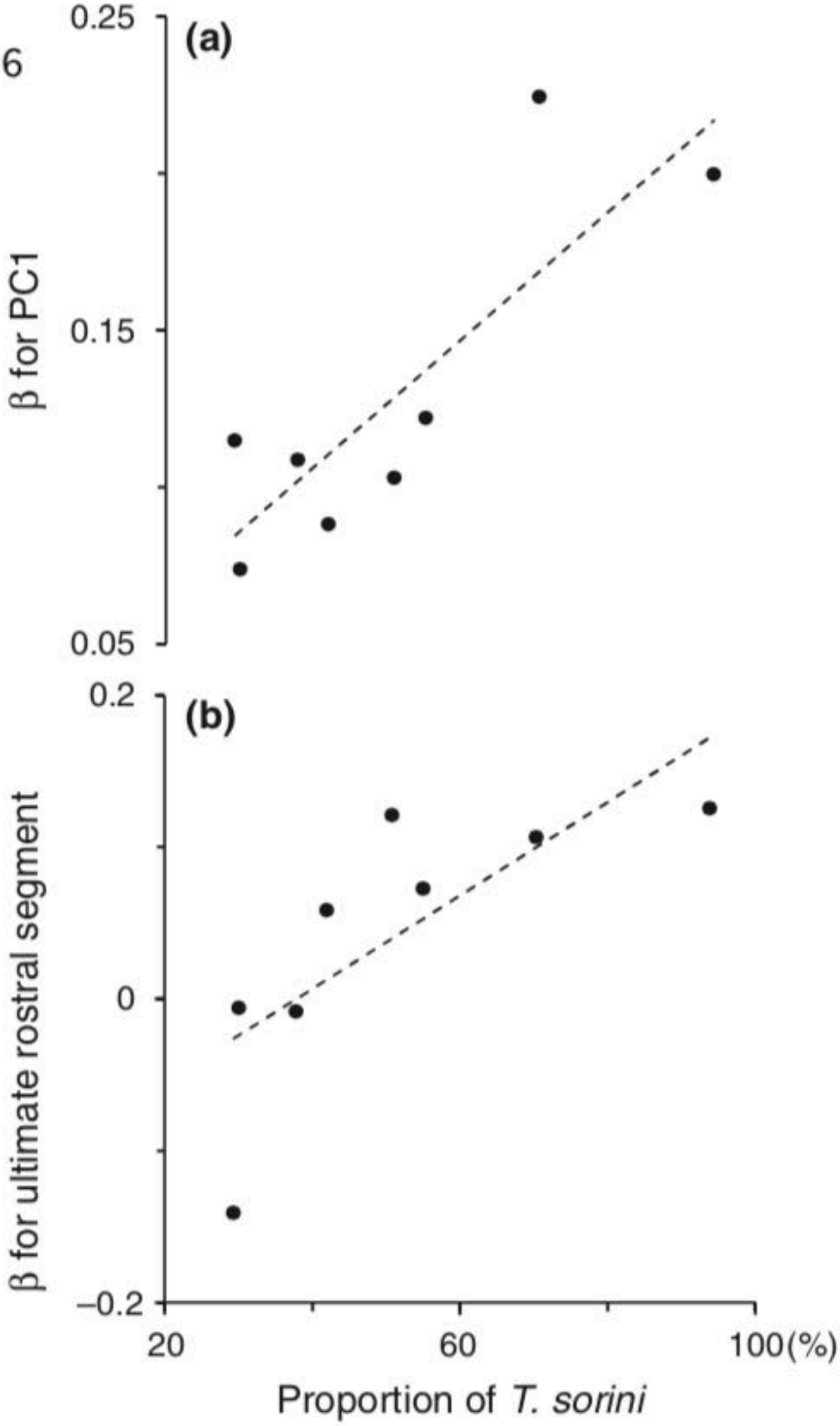

N = 777
N = 334
N = 623
N = 511
N = 518
N = 748
N = 808
N = 944
Species composition (%)
100
50
0
TO
TR
TT
TN
TS
Bibai 1988
Bibai 2014
Sunagawa 2014
Sapporo 2014
Utonai 2015
Sagae 2015
Kashiwa 2012
Kashiwa 2015

Fig. 6


(a)
0.25
0.15
0.05
β for PC1
(b)
0.2
0
−0.2
β for ultimate rostral segment
20
60
100(%)
Proportion of T. sorini