



Title	Spatiotemporal fluctuations in natural selection acting on the gall-parasitic aphid <i>Tetraneura sorini</i>
Author(s)	Muramatsu, Koichi; Akimoto, Shinichi
Citation	Journal of Evolutionary Biology, 29(7), 1423-1436 https://doi.org/10.1111/jeb.12881
Issue Date	2016-07
Doc URL	http://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 <i>Tetraneura sorini</i> . 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	article (author version)
File Information	<i>Tetraneura sorini</i> . J. Evol. Biol._29_1423.pdf



[Instructions for use](#)

1 **Spatiotemporal fluctuations in natural selection acting on the gall-parasitic aphid**

2 *Tetraneura sorini*

3

4 K. MURAMATSU & S. AKIMOTO

5

6 *Department of Ecology and Systematics, Graduate School of Agriculture, Hokkaido*

7 *University, Sapporo, Japan*

8

9 Running title: Fluctuating selection on an aphid

10

11 *Correspondence:* Koichi Muramatsu, Laboratory of Systematic Entomology,

12 Department of Ecology and Systematics, Graduate School of Agriculture, Hokkaido

13 University, Sapporo 060-8589, Japan.

14 Tel.: +81-11-706-2486; fax: +81-11-706-4939

15 e-mail: muram-k@res.agr.hokudai.ac.jp

16

17 **Abstract**

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

39

40

41

42 *Keywords:*
43 Eriosomatinae;
44 fight;
45 gall;
46 multiple regression;
47 selection differential;
48 selection gradient;
49 partial least square regression.
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71

72 **Introduction**

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

87 An important role of measuring the selection gradient is to predict the direction
88 and rate of evolution for the focal trait. Given a certain value of heritability for a trait,
89 the direction and rate of the evolution of that trait will be determined by the magnitude
90 of the selection gradient (Lande, 1979; Lande & Arnold, 1983; Stinchcombe *et al.*,
91 2002). However, attempts to associate the measured selection gradient with the
92 microevolutionary changes in that trait in contemporary populations have been
93 exceptional despite the importance of the confirmation of natural selection in the wild
94 (but see Grant & Grant, 1995, 2002). If temporal changes in the selection gradients are
95 investigated coupled with the evaluation of trait relationships through years, such

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

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

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

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

121 & Makino, 1982; Akimoto, 1988; Akimoto & Yamaguchi, 1997; Ngakan & Yukawa,
122 1997; Inbar, 1998). *Tetraneura sorini* is an aphid species that induces a closed pouch
123 gall on the leaf surface of the Japanese elm, *Ulmus davidiana*. The first instar foundress
124 of *T. sorini* is peculiar in morphology, being characterized by well-developed hind legs
125 and large body size (Akimoto & Yamaguchi, 1997). Unlike other *Tetraneura* species, *T.*
126 *sorini* foundresses frequently fail to induce their incipient galls. However, when failing
127 in initial gall formation, foundresses choose an alternative tactic, that of depriving other
128 conspecific and heterospecific foundresses of incipient galls (Akimoto & Yamaguchi,
129 1997). Thus, conflicts frequently occur between *T. sorini* foundresses or between those
130 of *T. sorini* and other *Tetraneura* species (Blackman & Eastop, 1994). *Tetraneura sorini*
131 foundresses that fail in galling often attack another gall owner, using a kicking motion
132 with their hind legs or by pushing the gall owner out of the incipient gall. In instances of
133 successful usurpation, *T. sorini* foundresses can develop and transform the incipient
134 galls into mature galls, inside which they reproduce parthenogenetically. We
135 hypothesize that competitive interaction between foundresses has been a primary driver
136 in the evolution of the peculiar morphology of *T. sorini* foundresses. It is most likely
137 that a large body size and well-developed hind legs are advantageous in fighting, and
138 thus, positive directional selection will act on both traits. The most significant
139 advantage of using *T. sorini* for investigating selective pressures is that it is relatively
140 easy to distinguish between foundresses that have survived the galling process and those
141 that have failed; foundresses that succeeded in gall formation or usurpation of galls are
142 found inside galls, whereas those that failed either die or are found wandering on the
143 leaf surface. Thus, we can easily detect the operations of selection imposed by fighting
144 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 (β) 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 (β) via partial least square regression (PLS), as well as the use of
166 classical technique. We tested differences in β 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.*

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

178

179 **Materials and methods**

180 **Study organisms**

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

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

200

201 **Collection and measurements**

202 First instar foundresses of *T. sorini* were collected from four locations in Hokkaido,
203 northern Japan and two locations in Honshu, the mainland of Japan. At Forestry
204 Research Institute of Hokkaido, Bibai City, Hokkaido (43.28961N, 141.8531E), we
205 collected first instars three times over years, on 22 May 1987, 10 June 1990, and 17
206 May 2014, while collection was made once at the following three locations in
207 Hokkaido; Sapporo (43.08956N, 141.4392E) on 20 May 2014, Sunagawa (43.54674N,
208 141.9251E) on 28 May 2014, and Lake Utonai (42.68131N, 141.70734E) on 27 May
209 2015. In Honshu, we collected first instars twice at Kashiwa (35.89417N, 139.9415E)
210 on 30 April 2012 and on 29 May 2015, and at Sagae (38.36408N, 140.26542E) on 10
211 May 2015. Thus, first instar foundresses were collected from six locations, including
212 two collections in Kashiwa and three collections in Bibai. In each collection, developing
213 leaves of *U. davidiana* with incipient galls and foundresses were haphazardly collected
214 from one tree; thus, we collected not only *T. sorini* foundresses but also those of other
215 *Tetraneura* species without identifying the species. Collected leaves were put in plastic
216 bags and transferred to the laboratory within 2 hours after collection. In the laboratory,
217 we classified the first instar foundresses into two categories under binocular
218 microscopes; (1) foundresses found in closed incipient galls (success group), and (2)
219 foundresses found moving or dead on the leaf surface (failure group). The foundresses

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

233

234 **Transplantation of foundresses**

235 *Tetraneura sorini* foundresses used in the transplantation experiment were hatched in
236 the laboratory from overwintered eggs, which were collected on 13 May 1989 at
237 Forestry Research Institute of Hokkaido, Bibai City. Barks with *T. sorini* eggs were
238 scraped off the trunk of *U. davidiana* trees on which abundant galls were found in the
239 previous year. Incubation was conducted at 17°C and 16L8D in a climate chamber. A
240 total of 106 first instar foundresses that had just hatched were individually transferred
241 onto a bursting bud of young *U. davidiana* trees by using a small brush. All the first
242 instars were transferred onto different trees on the campus of Hokkaido University from
243 15 to 18 May 1989. At this time, the leaves were rapidly growing and most appropriate
244 for gall formation. Buds onto which the first instars were transplanted were enveloped

245 by a plastic bag (140mm×100mm). The base of the plastic bag was sealed by plastic
246 tape to prevent the escape of first instars. On 24 May, all the bagged buds were
247 collected, and the outcomes of gall induction were recorded. First instars confined in
248 initial galls (successes) and dead ones (failures) on leaves were preserved in vials of
249 80% ethanol and then mounted on glass slides. Mounting and measurement were
250 conducted in the same way as in the wild populations, and selection gradients on the
251 traits were estimated.

252

253 **Species composition in populations**

254 Foundresses collected from wild populations were composed of several *Tetraneura*
255 species in every location. We identified all the collected first instar foundresses based on
256 slide-mounted specimens to know the species composition, except for the samples from
257 Bibai in 1987 and 1990. From these two samples, only *T. sorini* foundresses were
258 slide-mounted. However, in 1988, first instar foundresses were haphazardly collected
259 from developing host leaves at the same locality for recording the species composition.
260 This data was used to represent species composition at Bibai 1987, and the data of Bibai
261 1990 were omitted from the analysis of the effect of species compositions on selection.

262

263 **Statistics**

264 A problem in any type of regression analysis is that there is no way to know whether a
265 measured trait is the actual target of selection or not (Lande & Arnold, 1983; Endler,
266 1986; Mitchell-Olds & Shaw, 1987). If only a small number of traits were measured, it
267 is difficult to identify the traits on which selection is actually acting (Kingsolver *et al.*,
268 2012). Thus, we measured as many morphological traits as possible, including
269 non-weapon traits and tried to find out traits that are actually subject to directional

270 selection.

271 To test if there are differences in the action of selection among *T. sorini*
272 populations, logistic regression was first performed. In the model, success or failure in
273 gall acquisition was treated as the response variable, and the six morphological traits,
274 localities, and the interactions between localities and each trait were included as
275 explanatory variables. However, the morphological traits were strongly correlated with
276 one another (Appendix S2), resulting in multicollinearity (for six traits, $0.635 \leq r \leq$
277 0.952 , with a mean of 0.805 , $2.40 \leq \text{VIF} \leq 13.7$). Thus, we additionally performed
278 logistic stepwise regression (backward elimination technique) by using the original
279 model. If significant interactions were detected between localities and any trait, it
280 suggests significant differences in the action of selection among populations. Significant
281 interactions were actually detected (see Results and Table 1), and thus we analyzed
282 selective pressures separately in each population.

283 Our samples came from six locations, two of which included multiple collections
284 in different years. We compared the magnitude of variation in the selection gradients
285 between locations and that between years to know whether collections in different years
286 should be lumped or treated separately. We estimated the variances of the random
287 effects for the slope (selection gradient) and the intercept of the response variable using
288 a linear mixed-effects model. This estimation was conducted by using the “lmer”
289 function in the package *lme4* for R based on the restricted maximum likelihood (REML)
290 method (Bates *et al.*, 2014). Success/failure was analyzed for each morphological trait
291 with the model comprising trait length + random effect of trait length between locations
292 + random effect of trait length between years nested in location. In this model, we
293 assumed that both slopes and intercepts vary randomly between locations or years. The
294 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 (β).
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 (β) 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 β by 1000 times bootstrapping using
322 package “boot” in R (Canty & Ripley, 2015). Statistical differences in the estimated β
323 between different populations were tested by checking whether the 95% confidence
324 intervals of the β 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 β 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 (β) for
333 PC1 was estimated by GLM in each population.

334 We tested the correlation between the standardized selection gradient (β)
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 β 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

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

376 Body width was subject to significant positive selection in all populations except
377 Kashiwa 2012 and 2015 in respect of the selection gradient (Fig. 2). Hind femur, a
378 weapon trait, was subject to significant positive selection in three populations but
379 significant negative selection in three populations. In the Bibai population, β of hind
380 femur fluctuated in the direction from year to year; PLS analysis confirmed that the
381 95% confidence intervals of β did not overlap between samples in 1990 (-0.209 ~
382 -0.049) and in 2014 (0.111 ~ 0.326) (Appendix S3). A similar tendency was detected in
383 the Kashiwa population, where β of hind femur fluctuated between samples in 2012
384 (0.074 ~ 0.309) and in 2015 (-0.329 ~ -0.036). Standardized selection gradients (β)
385 estimated by PLS analysis and those by stepwise regression were highly correlated ($\rho =$
386 0.863, $P < 0.001$, $n = 28$), suggesting that PLS analysis could provide reliable estimates
387 for β of such traits that were not calculated by stepwise regression.

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

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

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

411

412 **Morphological change and selection within populations**

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

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

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

431

432 **Transplantation of foundresses**

433 In the transplantation experiment, successful gall inducers were smaller than failures in
434 body size (PC1 scores), whereas in the wild populations pooled, successful foundresses
435 were larger in body size than failures (Fig. 4). The action of selection also differed
436 significantly between the transplantation experiment and wild populations. Significant
437 positive selection was detected for the hind tibia, ultimate rostral segment and body
438 width in the wild populations pooled, whereas significant negative selection was found
439 in the hind femur in the transplantation experiment (Table 5, Appendix S4 for detailed P
440 values). Likelihood ratio test indicated that in four of the six traits, the magnitude of
441 selection varied significantly between the transplant population and wild populations
442 pooled (for the hind femur, $\chi^2 = 7.767$, $P = 0.005$; for the hind tibiae, $\chi^2 = 4.967$, $P =$
443 0.026 ; for the fore femur, $\chi^2 = 9.285$, $P = 0.002$; for the fore tibiae, $\chi^2 = 6.427$, $P = 0.011$;
444 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. radicolica* (Fig. 1), were abundant in Sapporo and Bibai 1988. Regression
451 analysis showed that variation in the standardized selection gradient (β) 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 β for body size (GLM, $t_6 = 3.81$, $P = 0.009$). Similarly, we
455 detected a significant positive regression of β 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

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

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

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

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

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

515

516 **Selection and species composition**

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

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

528 The standardized selection gradient (β) for the ultimate rostral segment in a
529 population also increased with an increasing proportion of *T. sorini*. The rostrum in
530 aphids is one of the feeding structures, a now-weapon trait, but the stylet within the
531 rostrum is used by *T. sorini* as a weapon in escalated fighting (Akimoto & Yamaguchi,
532 1997), as *T. sorini* foundresses occasionally use the stylet to pierce competitors. Thus,
533 longer rostrums and stylets may be advantageous in conflicts with conspecifics, and the
534 advantage of a longer ultimate rostral segment may be greater in populations with a
535 higher proportion of *T. sorini* foundresses.

536 Temporal changes in β for the hind femur in the Bibai and Kashiwa populations
537 are also likely to be due to temporal changes in the species composition of the
538 *Tetraneura* community. In the Bibai population, small-sized *Tetraneura* species (*T.*
539 *radicicola* and *T. triangula*) accounted for 43.5% of the *Tetraneura* foundresses in 1988
540 (Fig. 5). In this environment, *T. sorini* foundresses were able to push out most host
541 foundresses from their incipient gall pits without using their hind legs (Akimoto &
542 Yamaguchi, 1997). Thus, in a population with a high proportion of small species, weak
543 or negative selection is expected to have operated on the hind femur of *T. sorini*. More
544 than 20 years later, however, the proportion of the two small species had decreased to
545 13.7 %, whereas large-sized species (*T. nigriabdominalis* and *T. sorini* itself) became
546 more abundant (Fig. 5). A recent increase in the abundance of *T. nigriabdominalis* galls

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

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

558

559 **Morphological change and selection**

560 It is noteworthy that temporal changes in selective pressures on two traits led to changes
561 in the allometry between these traits. In the Bibai and Kashiwa populations, in
562 accordance with the drastic changes in β on the hind femur and other traits, allometry
563 between the hind femur and other traits was altered in a direction that can be predicted
564 from the changes in β . The development of hatchling morphology proceeds through
565 utilization of the nutrition in the egg only. Even if the effect of egg size was removed
566 statistically, allometry between morphological traits of aphid hatchlings was altered
567 between outbreeding and inbreeding settings (Akimoto, 2006). This report suggests that
568 allometry between morphological traits is likely to be genetically determined in aphid
569 hatchlings. Therefore, if selection acts positively on one trait and negatively on another
570 trait, then the size relationship between these traits should be expected to change
571 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 β 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 β among different
583 populations. As abovementioned, there was correlation between the standardized
584 selection gradient (β) 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 β ,
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

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

608

609 **Acknowledgments**

610 We are grateful to members of Systematic Entomology lab at Hokkaido University. We
611 also thank S. Kudo for discussion about statistical analyses. This research was
612 supported by Grants-in-Aid (number 17370028, 23370037) for Scientific Research
613 from the Japan Society for the Promotion of Science given to S. Akimoto and number
614 15J00706 to K. Muramatsu.

615

616 **References**

- 617 Abràmoff, M. D., Magalhães, P. J. & Ram, S. J. 2004. Image processing with ImageJ.
618 *Biophotonics Int.* **11**: 36–42.
- 619 Akimoto, S. 1995. Coexistence and weak amensalism of congeneric gall-forming aphids
620 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 fitting 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 (*Smynthuroides 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 spatially 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 autumnna* (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 galls with host plant phenology. *Popul. Ecol.*
755 **42**: 105–113.
- 756

757 **Figure legends**

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

760

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

764

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

770

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

774

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

777

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

780

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.

Explanatory variables	Logistic regression			Stepwise logistic regression		
	df	χ^2	<i>P</i>	df	χ^2	<i>P</i>
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	<i>P</i>
Bibai 1987 (<i>n</i> = 359)	Hind femoral length	0.316	0.291	0.005
	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 (<i>n</i> = 220)	Hind femoral length	0.323	0.364	0.002
	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 (<i>n</i> = 201)	Hind femoral length	0.294	0.996	< 0.001
	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 (<i>n</i> = 270)	Hind femoral length	0.292	0.728	< 0.001
	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 (<i>n</i> = 84)	Hind femoral length	0.324	0.530	0.001
	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 (<i>n</i> = 486)	Hind femoral length	0.327	0.735	< 0.001
	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 (<i>n</i> = 380)	Hind femoral length	0.355	0.452	< 0.001
	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 (<i>n</i> = 239)	Hind femoral length	0.369	0.546	< 0.001
	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

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

Random effect	df	Variance	<i>F</i>	<i>P</i>
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

796

797

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

	Absolute size		Regression slope	
	ρ	P	ρ	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	β (stepwise)	β (PLS)	Selection differential	β (stepwise)	β (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 (β) for six traits in the wild populations. β estimated
817 by the stepwise regression method and β estimated by PLS are shown along with SE and
818 95% confidence intervals (CI) based on 1000 times bootstrapping.

819

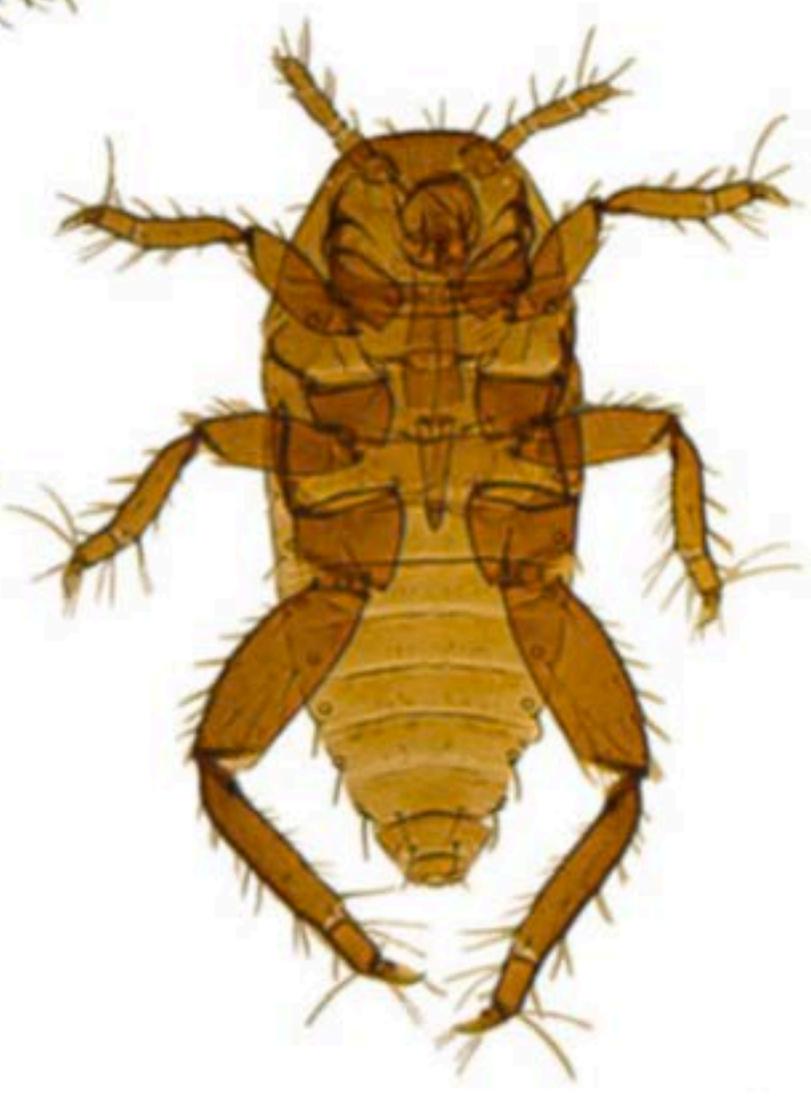
820 **Appendix S4** Exact P values for Table 5.

Fig. 1

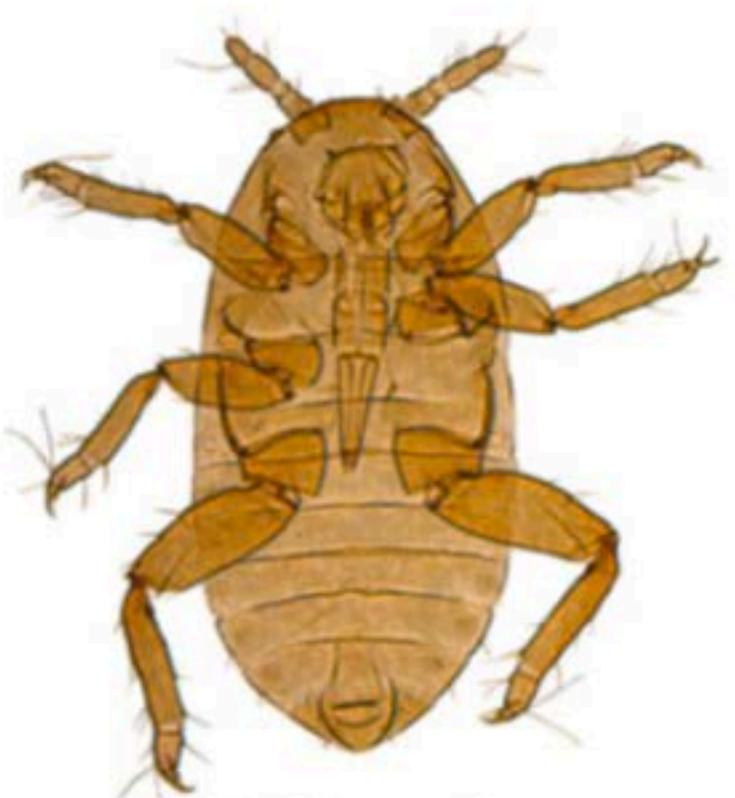
0.1 mm



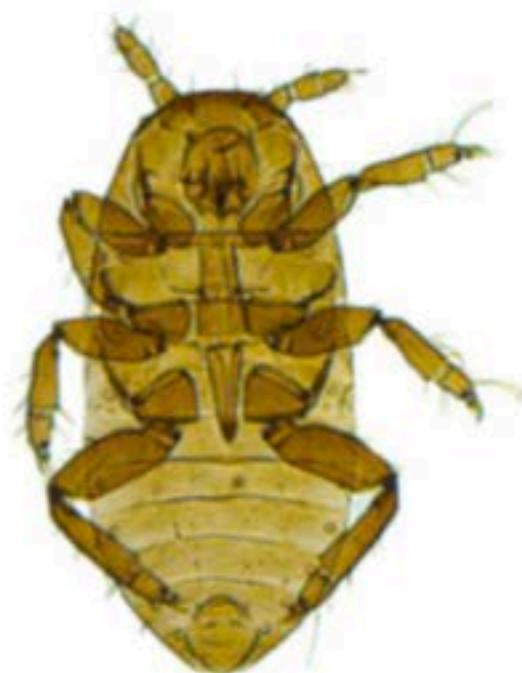
T. sorini



T. nigriabdominalis



T. triangula



T. radicicola

Fig. 2

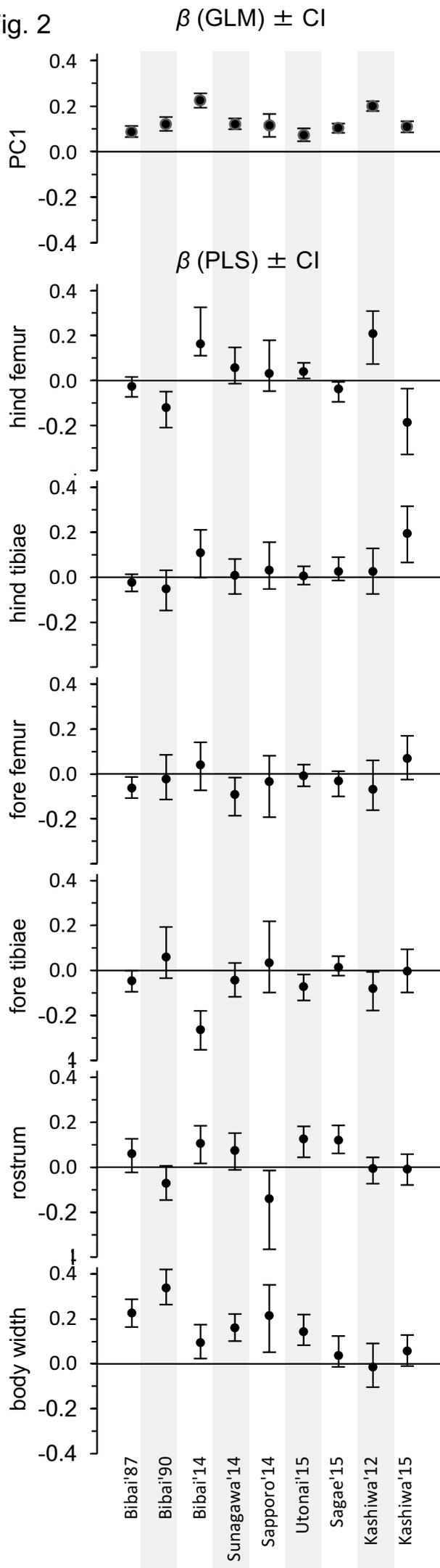


Fig.3

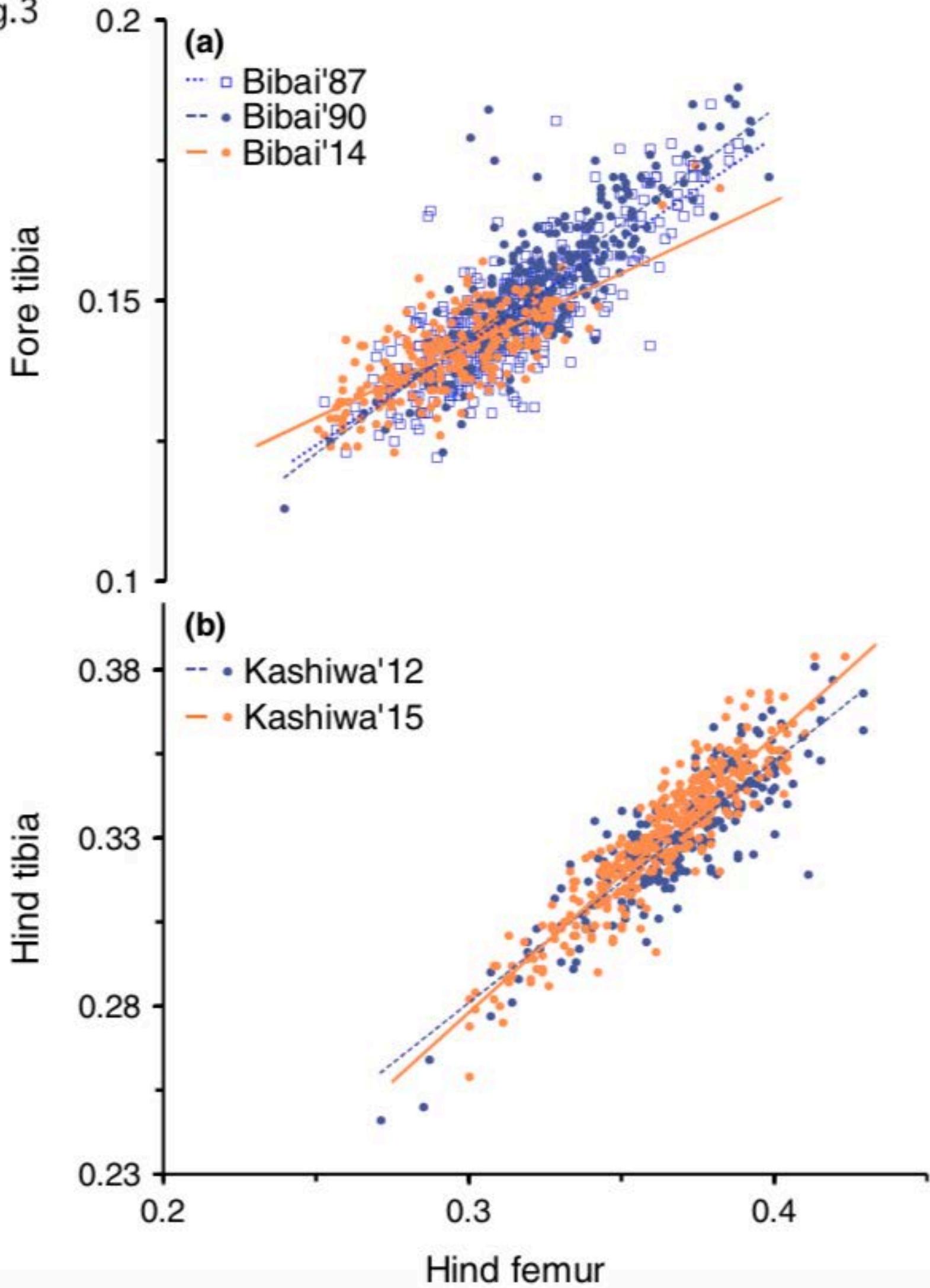


Fig. 4

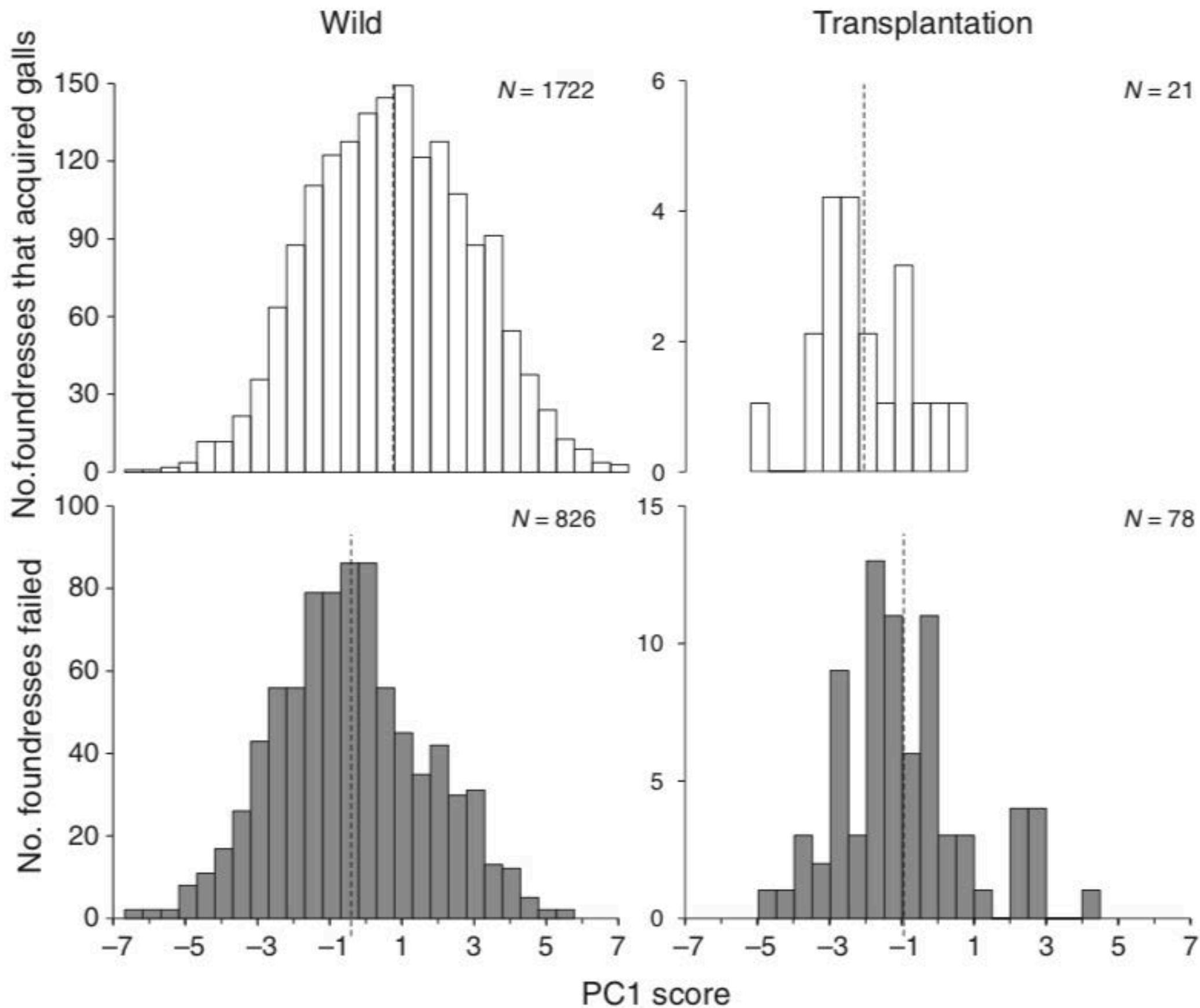


Fig. 5

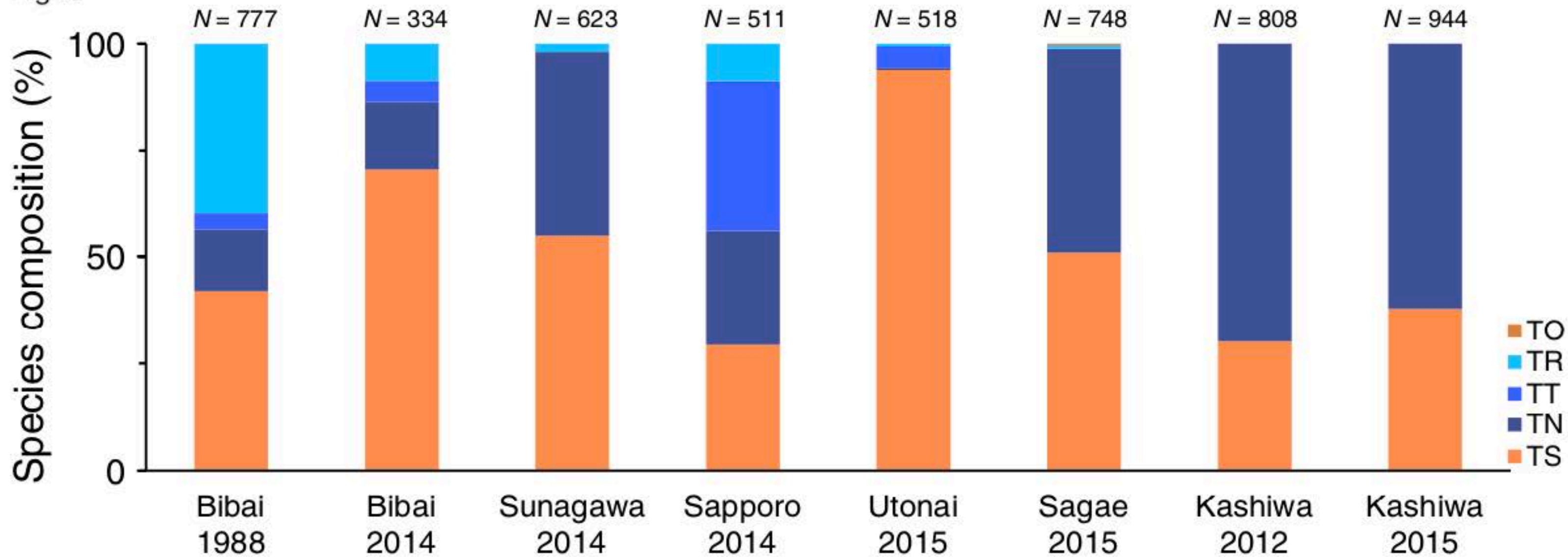


Fig. 6

