A multivariate phylogenetic comparative method incorporating a flexible function between discrete and continuous traits

Yuki Haba[1] & Nobuyuki Kutsukake[2]

1. Department of Ecology and Evolutionary Biology, Princeton University (corresponding author: yhaba@princeton.edu)
2. Department of Evolutionary Studies of Biosystems, Sokendai, The Graduate University for Advanced Studies, Hayama, Kanagawa 240-0193, Japan, (kutsu@soken.ac.jp)

Abstract

One major challenge of using the phylogenetic comparative method (PCM) is the analysis of the evolution of interrelated continuous and discrete traits in a single multivariate statistical framework. In addition, more intricate parameters such as branch-specific directional selection have rarely been integrated into such multivariate PCM frameworks. Here, originally motivated to analyze the complex evolutionary trajectories of group size (continuous variable) and social systems (discrete variable) in African subterranean rodents, we develop a flexible approach using approximate Bayesian computation (ABC). Specifically, our multivariate ABC-PCM method allows the user to flexibly model an underlying latent evolutionary function between continuous and discrete traits. The ABC-PCM also simultaneously incorporates complex evolutionary parameters such as branch-specific selection. This study highlights the flexibility of ABC-PCMs in analyzing the evolution of phenotypic traits interrelated in a complex manner.

Key words: Phylogenetic comparative method (PCM), approximate Bayesian computation (ABC), multivariate analysis, social evolution, African mole rats

Introduction

Phylogenetic comparative methods (PCMs) provide a powerful statistical framework for investigating the patterns and processes of trait evolution (Felsenstein 1985; Harvey & Pagel 1991; Nunn 2011; Garamszegi 2014). The recent development of PCMs permits analyses of biologically interrelated discrete and continuous variables in a single multivariate statistical framework (Table 1). The development of such multivariate PCMs is crucial for two reasons. First, conducting two separate univariate analyses for a continuous trait and a discrete trait is redundant when the two variables are interrelated. Second, and more importantly, separate univariate analyses will miss the opportunity to consider important biological links between these traits.

The threshold model (Felsenstein 2005, 2012) was among the first PCMs to fully combine both discrete and continuous traits. The idea of the threshold model was originally developed in quantitative genetics by Wright (1934) to understand how multiple underlying genetic loci contribute to categorical traits such as the number of digits in guinea pigs. The threshold model assumes an unobservable continuous trait called 'liability' that underlies a discrete trait of interest. Because the liability is a continuous trait, Brownian motion has been conventionally used to model its evolution. Then, the state of the discrete trait of interest is determined by whether the liability trait value is below or above a particular threshold. This model allows users to incorporate both continuous and discrete traits in a straightforward way, as well as to estimate the covariance between the liability trait and other continuous traits of interest. However, because liability is unobservable, it is impossible to directly infer a latent function between the discrete trait and other observable continuous traits of interest. Moreover, although it is convenient to assume that the discrete trait is determined by a simple threshold (which can be treated as a probit link function in a framework of a phylogenetic generalized linear mixed model, PGLMM; Hadfield 2015; also see Ives and Garland 2014), it is unclear if the assumption is always biologically valid. At the very least, it is desirable for researchers to be able to assume other forms of latent functions (Fig. 1).

In addition to Felsenstein's threshold model, other methods that can link discrete traits and continuous traits have been proposed (Ives & Garland 2010, 2014; Hadfield and Nakagawa 2009; Hadfield 2015; see Table 1). For example, Ives and Garland (2010) developed a phylogenetic logistic regression to test the effects of observable continuous independent traits on a discrete dependent trait (phylogenetic logistic regression, Ives & Garland 2010). Hadfield and Nakagawa proposed a phylogenetic generalized linear mixed model employing a Bayesian approach (MCMCglmm, Hadfield and Nakagawa 2009; Hadfield 2015; also see Ives and

Garland 2014 for other approaches and a comparison of their performance). These approaches successfully extended traditional linear models to enable nonlinear link functions (logit function in a phylogenetic logistic regression; logit or probit functions in MCMCglmm) between discrete and continuous traits. Notably, a probit-GLM can be mathematically equivalent to the threshold model (see Hadfield 2015). However, limitations to these models still exist. For example, the model by Ives and Garland (2010) assumes that the continuous traits are known values measured empirically and does not allow the continuous traits to evolve along the tree (see Felsenstein 2012 and Hadfield 2015 for models that do not have such assumption; but also see Ives and Garland 2014 for analyses of the relatively small effects of phylogenetic signal on continuous traits). Moreover, the link function in the models is predetermined. In most cases, equipped functions (e.g., logit or probit functions) are those that have useful mathematical properties for analyzing a relationship between discrete and continuous traits. Still, cases may exist in which it is biologically valid to establish a more complicated function between discrete and continuous traits. Therefore, it is ideal to prepare a framework that enables researchers to examine a flexible function to test specific hypotheses of interest.

Furthermore, analyses will be even more complex if additional variables of interest are included, such as the presence of branch-specific directional selection (Kutsukake and Innan 2013, 2014). This complication often prevents the description of a likelihood function of the model, which most conventional PCMs require (e.g., the maximum likelihood, Bayesian approach). The aforementioned PCMs (Felsenstein 2005, 2012; Ives and Garland 2010, 2014; Hadfield and Nakagawa 2009; Hadfield 2015) cannot incorporate branch-specific directional selection into their models.

Here, we propose a PCM using approximate Bayesian computation (ABC) to overcome the difficulties discussed above (Fig. 2). ABCs have been shown to facilitate flexible analyses in a comparative framework and therefore have increasingly been applied to PCMs with intricate evolutionary scenarios (Bokma 2010; Slater *et al*., 2012; Kutsukake and Innan 2013, 2014; Janzen *et al*. 2015; Harano & Kutsukake 2018). Briefly, an ABC-PCM estimates parameters of interest by simulating phenotypic evolution without a likelihood function (Beaumont 2010; Csillery *et al*. 2010). The proposed parameters are accepted only when the simulated data and real data are similar, and the accepted data comprise posterior distributions of parameters. Thanks to this flexibility, ABC-PCMs enable researchers to test evolutionary models whose likelihood function is mathematically intractable.

Our initial motivation for extending ABC-PCMs was to analyze a heterogeneous evolutionary pattern between group size and social system in African subterranean mole rats (family

Bathyergidae; Table 2). In this clade, species have varied social structures that span solitary, social, and eusocial states. Eusociality has independently evolved twice: in naked mole rats (*Heterocephalus glaber*) and in Damaraland mole rats (*Fukomys damarensis*) (Jarvis 1981; Sherman *et al*. 1991; Bennett & Faulkes 2000; Faulkes & Bennett 2013). We hypothesized that the probabilities of evolutionary transitions among eu/social and solitary states (discrete trait) change depending on group size (continuous trait). To capture all of these complex evolutionary features, we are required to incorporate (1) the evolutionary trajectory of sociality and group size, (2) the parameters for the trait-dependent latent functions, and (3) the presence of branch-specific directional selection on group size in the two eusocial species. Certain aspects of these features can be analyzed using previous methods; however, it is not possible to incorporate all factors in a single model with those methods. Then, based on the example data, we discuss how this ABC-PCM can be used for inferring other similar complex evolutionary processes.

Methods

ABC-PCM

Our ABC-PCM extends a previously developed framework (Fig. 2; Kutsukake & Innan 2013, 2014) designed to analyze heterogeneous evolutionary modes whose likelihood is not straightforward to describe. We assume knowledge of the species tree Ψ, which consists of information on tree shape and the lengths of all branches in the topology. We also assume that trait data at the tips of the tree $\boldsymbol{D}$ have been observed.

In bivariate analyses of continuous and discrete traits, several causational patterns are possible. For example, a discrete trait can be determined by a continuous trait that evolves on its own, or vice versa. Alternatively, it is also possible to assume no *a priori* causality between the continuous or discrete traits (Hadfield and Nakagawa 2009; Ives and Garland 2010). Any of these cases can be modeled by the ABC-PCM framework proposed in this paper. Moreover, this framework can be used regardless of whether a given value is latent (e.g., liability in the threshold model) or measurable.

We hereafter consider a simple case of interrelated evolution in which a continuous trait determines a discrete trait. In our example study of African subterranean mole rats, both traits were measurable.

Briefly, our ABC-PCM process is implemented as follows (see Fig. 2 for a visual schematic).

Let $\boldsymbol{\Lambda}$ be the parameter set to be estimated based on a hypothesis.

126 (Step 1) Determine prior distributions for all parameters in $\boldsymbol{\Lambda}$. If prior biological
127 knowledge is available, it can be used to set informative (i.e., strong) prior distributions.
128 (Step 2) Parameters used in the simulation ($\boldsymbol{\Lambda}$**'**) are randomly generated from the prior
129 distributions.
130 (Step 3) Using $\boldsymbol{\Lambda}$**'**, the trait evolution is simulated on the phylogeny Ψ. In the current
131 model, the trait simulation has two parts corresponding to continuous and discrete trait
132 evolution (Fig. 1).
133 (Step 3a) The continuous trait evolves via Brownian motion (Felsenstein 1985; note
134 that other evolutionary models can be used; see Kutsukake and Innan 2013).
135 (Step 3b) Then the discrete trait is determined according to the probability of
136 transition between two states (*A* and *B*) as a function of the continuous trait. Note
137 again that this setting can be relaxed such that causation between continuous and
138 discrete traits can be flexibly changed according to a hypothesis of interest. Here,
139 $p(x)$ is the probability function that state *A* changes to *B* given the continuous trait x.
140 Similarly, $q(x)$ is the probability that state *B* changes to *A* given the continuous trait
141 x (Fig. 1).
142 (Step 4) Calculate the likelihood by comparing simulated data $\boldsymbol{\Theta}$ with the real data $\boldsymbol{D}$ and
143 determine whether the parameter set is accepted or rejected. A joint probability (full
144 likelihood) for the comparison of n species can be calculated. In most cases, this
145 probability is difficult to obtain. In such cases, a composite likelihood that is proportional
146 to the full likelihood can be used as an approximate proxy. Intraspecific variation in the
147 trait data can be considered by assuming a certain distribution for the trait when
148 calculating the likelihood.
149 Then the acceptance or rejection of the parameters can be determined based on the
150 likelihood . Several methods of judgment exist (Marjoram *et al*. 2003; Marjoram and
151 Tavare 2006).
152 (Step 5) Repeat Steps 2–4 until a sufficient number of parameters $\boldsymbol{\Lambda}$**'** is accepted. Then,
153 posterior distributions and credible intervals can be estimated.
154
155 Although the fundamental structure of ABC is straightforward, the number/choice of summary
156 statistics and the width of tolerance for judging the acceptance or rejection of simulated data at
157 Step 4 are controversial, and there is no general consensus on the choice of summary statistics
158 and tolerance (Beaumont *et al*. 2002; Csillery *et al*. 2010; Leuenberger & Wegmann 2010). In

159 this study, we use a combination of perfect match and joint probability as summary statistics (see
160 *Acceptance and summary statistics* below for details).
161
162 *Application—social evolution in African mole rats*
163 One example of complex evolution where continuous and discrete traits are interrelated is
164 social evolution. Sociality in animals can be characterized using a discrete classification based on
165 mating and/or social systems (e.g., Shultz *et al*. 2011). Average group size of a species, a
166 continuous trait, is also an important variable in characterizing sociality in animals (e.g.,
167 comparative analyses: Faulkes *et al*. 1997; Sheehan *et al*. 2015). Loss of sociality is correlated
168 with a decrease in group size (secondary loss of sociality; Wcislo & Danforth 1997; Beauchamp
169 1999; Sheehan *et al*. 2015). Likewise, cooperative social systems are likely to appear as the
170 group size increases (e.g., limited dispersal by ecological constraints: Emlen 1982; Duffy &
171 Macdonald 2010). Thus, the complex social evolution across African mole rats is an ideal system
172 to test our framework. Here, species-specific sociality and the mean group size for each species
173 (x) are the target traits.
174
175 *Dataset*
176 We surveyed the literature for field data on the sociality (solitary, social, or eusocial), mean
177 group size, and/or distribution of group sizes in each species (Table 2). When the distribution of
178 group size was available, we calculated its mean and standard deviation for each species. For
179 solitary species, we regarded the mean group size as one. In reality, however, their group sizes
180 can deviate from this value, because females may have dependent, pre-dispersal offspring; thus,
181 the group size of solitary species can also have a distribution. Therefore, we incorporated a
182 realistic value for the variance of solitary species (Table 2). We used the phylogeny presented in
183 Faulkes *et al*. (2004), who used mitochondrial genes 12s rRNA and cyt *b*. The mean value of the
184 estimated divergent interval in millions of years was used as the length of each branch.
185
186 *Parameter set and trait simulation*
187 The trait simulation included six parameters to be estimated: the group size of the most
188 recent common ancestor (MRCA) (θ), the rate of evolution (μ), the parameters of directional
189 selection (k_n and k_d), and the parameters for latent functions (a and b, or c and d, depending on
190 the model used; see next section). Table 3 shows the notation and settings of the prior
191 distributions.
192 The evolutionary process of the continuous trait was as follows. First, we generated the

values of six parameters from the prior distributions. Then, based on the proposed value of x at the root (i.e., the group size of the MRCA, θ), we randomly determined the state (i.e., sociality) of the MRCA by either $p(x)$ or $q(x)$. The number of evolutionary events that change traits (i.e., either an increase or decrease in group size) on each branch was respectively modeled as a random number from a Poisson distribution with a mean equaling the product of the evolutionary rate, μ (trait change per million years), and branch length τ (millions of years). The degree of trait change by one evolutionary event is a random value from an exponential distribution, with its mean being an arbitrary value φ (set to 0.01). φ corresponds to the average effect of one change on the trait value. We used an exponential distribution because a change in trait due to one evolutionary event would have small effects on traits in most cases, and large effects less frequently. Thus,

$$\Delta^{+} = \mathrm{Pois}(\mu \times \tau_i) \times \exp(\varphi) \ldots \text{ trait increase in a branch}$$

$$\Delta^{-} = \mathrm{Pois}(\mu \times \tau_i) \times \exp(\varphi) \ldots \text{ trait decrease in a branch}$$

where τ_i is the length of the ith branch. The total trait change on the branch is calculated as $\Delta^{+} - \Delta^{-}$.

Changes in group size caused by one evolutionary event would depend on the initial group size. For example, an increase in the group size by one individual will have different biological meanings in solitary and eusocial species. Therefore, we transformed the value of group size to a natural logarithmic scale during trait simulation. Because the mean group size x cannot be less than one, we enforced a lower bound of one. That is, when a trait change resulted in a value less than one on a branch, the change was not implemented.

We also tested whether there were selective pressures for increasing group size in the two branches leading to eusocial species. In our ABC-PCM model, the parameter k represents directional selection (Kutsukake & Innan 2013, 2014). The selection parameter k biases the number of positive or negative trait changes such that the number of trait changes $\mathrm{Pois}(\mu \times \tau_i) \times \exp(\varphi)$ is multiplied or divided by k, respectively (see Kutsukake & Innan 2013 for more details). Thus,

$$\Delta^{+} = \mathrm{Pois}(\mu \times \tau_i \times k) \times \exp(\varphi) \ldots \text{ trait increase in a branch}$$

$$\Delta^{-} = \mathrm{Pois}(\mu \times \tau_i / k) \times \exp(\varphi) \ldots \text{ trait decrease in a branch}$$

and again, the total trait change on the branch is calculated as $\Delta^{+} - \Delta^{-}$.

When $k = 1$, the evolutionary mode of trait evolution is asymptotically identical to the Brownian motion. On the other hand, a significant departure of k from 1 is used as a signature of directional selection.

226 In this analysis, we set selection parameters for the branches for both naked mole rats and
227 Damaraland mole rats as k_n and k_d, respectively (Table 3), to test our hypothesis that branch-
228 specific selective pressure has increased group size in eusocial species (Jarvis & Bennett 1993;
229 Young *et al*. 2015). When the 95% CI of the posterior distribution of k was larger than 1, it was
230 considered a signature of directional selection for larger group size.
231
232 *Models of latent evolutionary functions*
233 At each evolutionary step, the discrete trait value was determined by the continuous trait value
234 according to the latent functions (see Step 3b of *ABC-PCM*). Because the shape of this function
235 between sociality and group size was unknown, we took advantage of the flexibility of our ABC-
236 PCM framework and tested two examples of evolutionary models with different transition
237 functions, $p(x)$ and $q(x)$. We note that this framework is not limited to these two models but can
238 use any function, depending on the focal biological traits and hypotheses of interest. For
239 simplicity, we set the latent evolutionary functions such that $p(x) + q(x) = 1$, but this is not
240 required if one desires otherwise.
241
242 *Model 1*: Logistic function
243 In this model, we used a logistic function as the latent evolutionary function, which was similar
244 to a previous model using a generalized linear model for a binary dependent term (Ives and
245 Garland 2010). The transition functions $p(x)$ and $q(x)$ were defined as follows:
246 $$p(\mathrm{x}) = 1–1/(1 + \exp(–a \times (x–b)))$$
247 $$q(\mathrm{x}) = 1/(1 + \exp(–a \times (x–b)))$$
248 The parameter a determined the curvature of $p(x)$ and $q(x)$, i.e., the effects of the group size on
249 sociality (Fig. 1). When a is equal to 0, $p(x)$ and $q(x)$ are flat and the transition between social and
250 solitary is independent of group size (Fig. 1, right). When a is positive, a species with large group
251 size is likely to be social; when a is negative, a species with large group size is likely to be solitary.
252 Importantly, when $|a|$ is large enough, the transition between social and solitary is determined by
253 a certain group size, which virtually behaves like a step function (Fig. 1, left). The parameter b is
254 the x value at which $p(x)$ and $q(x)$ equal 0.5, and determines the group size at which the probability
255 of transitioning from solitary to social becomes larger than the probability of the reverse. Because
256 the curvature and midpoint of $p(x)$ and $q(x)$ (i.e., a and b) were unknown *a priori*, we set broad
257 prior distributions of a and b (Table 3).
258
259 *Model 2*: Exponential function

In the second model, we used an exponential function as the latent evolutionary function. We set the transition functions $p(x)$ and $q(x)$ as follows:

$$When\ \mathrm{d} > 0 \begin{cases} p(x) = \begin{cases} 1 - \exp\left(-c \times (x - d)\right), & x > d \\ 0, & x \leq d \end{cases} \\ q(x) = \begin{cases} \exp\left(-c \times (x - b)\right), & x > d \\ 1, & x \leq d \end{cases} \end{cases}$$

$$When\ d \leq 0 \begin{cases} p(x) = \begin{cases} 1 - \exp\left(c \times (x - d)\right), & x > d \\ 0, & x \leq d \end{cases} \\ q(x) = \begin{cases} \exp\left(c \times (x - d)\right), & x > d \\ 1, & x \leq d \end{cases} \end{cases}$$

These functions describe an exponential decrease or diminishing increase in the transition probability, and a given species with a mean group size of one is always a solitary species. Similar to a and b in model 1, c and d determine the curvature of the exponential function and the point at which the exponential decrease/increase begins, respectively. Note that one side of the exponential function used here is asymptotically equal to either 0 or 1 when x is sufficiently large, but that of the other side is not; therefore, we set the probability of state change as 0 or 1 when x is smaller than d, at which the functions reach 0 or 1, respectively.

Acceptance and summary statistics

Among ABC-PCMs, no clear consensus has been reached concerning the number/choice of summary statistics or the width of tolerance for judging the acceptance or rejection of simulated data. In the present analyses, we used two summary statistics to assess the fit of the simulated data to actual data. First, for the discrete variable (i.e., sociality), we only accepted simulations in which the simulated data were a perfect match to the real data. For the continuous variable, we used a conventional method that uses a joint probability (Kutsukake & Innan 2013, 2014); we first calculated the probability that the real trait value is gained under a simulated trait value for all 10 species. When calculating the probability, group size (back-transformed to an arithmetic scale from a log-transformed value) was assumed to be normally distributed with a mean and standard deviation equal to those of the real data. Then we calculated the product of those 10 probabilities and used the joint probability as a summary statistic. Simulated parameter sets were accepted in proportion to the joint probability. For example, if the joint probability was 0.8 for a simulation, the parameter set of the simulation has an 80% chance of being accepted. Because the joint probability can be considered a direct likelihood, this method is superior to the standard

291 ABCs, which depend on the choice of summary statistics and arbitrary tolerance (discussed in
292 Kutsukake & Innan 2013, 2014).
293 In total, 500 parameter sets were collected for estimating posterior distributions. All
294 results were visualised using R version 3.5.3 (R Core Team 2017). The simulation code was
295 written in C and is available in a public repository (https://github.com/YukiHaba/ABC-PCM).
296
297 **Results**
298 *Model 1*: Logistic functions
299 When the latent evolutionary function was logistic, the social system of the MRCA was
300 estimated to be solitary in 61.4% of the cases (Fig. 3a, node A). The accepted functions of the
301 transition probabilities varied widely (Fig. 4a), but showed several consistent patterns. First, the
302 curvature parameter of functions, a, was positive in all cases (Table 3, Fig. 4b; see
303 Supplementary material for a low risk of type I error), indicating that the probability of transition
304 from a solitary to social species increased as group size increased. In addition, $p(x)$ typically
305 increased rapidly around $x = 2$ to 4, which corresponded to the peak of the posterior distribution
306 of b (Table 3, Fig. 4c).
307 At each internal node before reaching the common ancestor of social species, the inferred
308 social system was predominantly solitary (Fig. 3a). The common ancestor of solitary species was
309 likely to be solitary (Fig. 3a, node C), and a similar pattern was evident at the node of the
310 common ancestor of *B. janetta* and *B. suillus*. By contrast, at the nodes leading to the clade of
311 *Fukomis*, predominantly social states were inferred (Fig. 3a, node B). The estimated group sizes
312 were consistent with the inferred social systems at internal nodes (Fig. 3b–d); the more social the
313 system was likely to be, the larger the group size.
314 We detected marginal directional selection for increasing group size in the branch that led
315 to naked mole rats (the proportion of $k_n < 1$ was 6.4%, Fig. 3f). By contrast, we did not detect
316 significant directional selection in the branch leading to another eusocial species, Damaraland
317 mole rats (the proportion of $k_d < 1$ was 15.8%, Fig. 3g). Based upon this non-significant result
318 for k_d, we repeated the estimation after excluding k_d (i.e., directional selection was assumed only
319 for the branch leading to naked mole rats), but the results did not change qualitatively (data not
320 shown).
321
322 *Model 2*: Exponential functions
323 With the exponential latent evolutionary functions, the MRCA was inferred to be solitary in
324 60.4% of the cases (Fig. 5a, node A), similar to model 1. Again, although the accepted functions

325 of the transition probabilities between social and solitary varied (Fig. 6a), the curvature of $p(x)$,
326 c, was always positive (Table 3, Fig. 6b; see Supplementary material for a low risk of type I
327 error).
328 Group size and sociality at each internal node were almost identical to those in model 1.
329 Within the clade of *Fukomis*, we inferred predominantly social states (Fig. 5a, node B) and
330 relatively large group sizes (Fig. 5c). Within the clade of *Bathyergus* and *Georychus*, in contrast,
331 the social system at each node was consistently solitary (Fig. 5a, node C) and the group size was
332 around 2 to 4 (Fig. 5d).
333 We detected significant directional selection for increasing group size in the branch
334 leading to naked mole rats (the proportion of $k_n < 1$ was 3.4%; Fig. 5f). However, we did not
335 detect significant directional selection in the branch to eusocial Damaraland mole rats (the
336 proportion of $k_d < 1$ was 16.2%; Fig. 5g). Again, based on this non-significant result for k_d, we
337 repeated the estimation after excluding k_d (i.e., directional selection was assumed only for the
338 branch leading to naked mole rats); however, the result did not change qualitatively (data not
339 shown).
340
341 **Discussion**
342 This study proposes an extension of a previously described ABC-PCM (Kutsukake & Innan
343 2013, 2014) to analyze complex evolutionary scenarios in which discrete and continuous traits
344 are biologically interwoven. This is the first ABC-PCM framework that allows simultaneous
345 analyses of the interdependent evolution of discrete and continuous traits. In addition, our model
346 offers at least two advantages over other existing methods.
347 First, this study incorporated a feature that has rarely been included in a PCM framework:
348 branch-specific directional selection (Kutsukake & Innan 2013, 2014; Harano & Kutsukake
349 2018). This inclusion was possible thanks to the flexibility of the ABC-PCM, which does not
350 require the mathematical expression or analytic solution of a likelihood function. Second, our
351 multivariate model can incorporate flexible user-defined functions that describe the evolutionary
352 relationship between continuous and categorical traits. Previous methods can incorporate
353 nonlinear functions such as logit and probit functions (Ives and Garland 2010; Hadfield &
354 Nakagawa 2009; Hadfield 2015), but the choice of a latent function is less flexible than in our
355 present method. Another well-established method for simultaneously analyzing both continuous
356 and categorical traits is the threshold model (Felsenstein 2005, 2012). The threshold model is
357 equivalent to a phylogenetically controlled linear model with a probit latent function (see
358 Supplementary Material in Hadfield 2015 for a useful graphical representation). Although the

359 assumption is mathematically convenient for analyses, it may be an oversimplification of the
360 evolutionary relationship between the continuous and categorical traits of interest.
361 Applying this new approach to data on African subterranean rodents, we estimated the
362 intricate evolutionary trajectory of sociality and group size. Two models with different latent
363 evolutionary functions were tested: logistic and exponential. Although the exponential function
364 has not been used in previous PCMs, we believe that this function is suitable for our study
365 species for two reasons. First, group size should have a minimum value of one, and second, a
366 species whose group size is one must be a solitary species. Although these functions have
367 different mathematical characteristics and are not nested, they have common features such as a
368 monotonic increase and an asymptotical approach from one to zero. Potentially due to these
369 common features, the estimated parameters showed similar posterior distributions. In both
370 models the social state of the MRCA was not decisively solitary or social, and its estimated
371 group size varied (node A in Figs. 3a and 5a).
372 The latent evolutionary functions $p(x)$ and $q(x)$ had a consistent pattern in both models
373 (Figs. 4 and 6). Namely, most of the accepted functions had a steep transition around a group
374 size of 2 to 4. This indicates that this particular window of group size was an evolutionary
375 tipping point of the transition between solitary and social states. This study is the first to
376 quantitatively infer the range of group size that is crucial to the evolution of sociality in this
377 clade.
378 We tested differential selective pressure on group size in the branches leading to the two
379 eusocial species in this clade (k_n and k_d); we detected directional selection (k_n) for larger group
380 size in the branch leading to naked mole rats in model 2 (Fig. 5f) but not in model 1 (Fig. 3f). In
381 both models, a selective pressure was not detected in the branch leading to Damaraland mole rats
382 (k_d, Figs. 3g and 5g). Thus, although both species are deemed eusocial, the two species may have
383 undergone qualitatively different evolutionary paths (Burda *et al*. 2000). Future studies should
384 explore the differences in the ecological and evolutionary causes of the evolution of these two
385 eusocial species.
386 One critical limitation of this case study was the relatively low sample size (10 species;
387 see Supplementary material). This sample size is caused by three inevitable limitations. First, the
388 family Bathyergidae is a monotypic group composed of less than 20 OTUs, which is an
389 inevitable constraint on sample size. Second, detailed data for group size are not available for all
390 species of this family, which further limits the available data. Finally, and most importantly,
391 expanding the number of study species to include closely related non-subterranean species is
392 questionable, as it is highly likely that the underground ecological niche has affected predation

393 pressure and consequently sociality. More broadly, it remains unclear whether heterogeneity of
394 traits other than the traits of interest would affect biological interpretation in PCM studies based
395 on a large sample size. It is generally believed that larger sample sizes (a larger number of
396 species, or more precisely, a larger number of evolutionary transitions) would produce a more
397 powerful test in PCMs. While this is true, it remains largely undiscussed how the inclusion of
398 species that have fundamentally different ecological or behavioral traits would affect the results.
399 In this sense, in addition to a large-scale comparison, it is important to conduct PCMs that focus
400 on a taxon that shares fundamental traits, even if the sample size is not very large.
401 One major challenge in ABC-PCMs is that they are computationally intensive. In our
402 study, for example, it took several days to accept one parameter set (MacPro, OS 10.6.7, 2 *x* 2.93
403 GHz Quad-Core Intel Xeon; also see Kutsukake & Innan 2013). To compare the performance of
404 our model to an available method, we ran a bivariate model on our dataset and phylogeny using
405 MCMCglmm (Hadfield & Nakagawa 2009; Hadfield 2015) assuming that group size and
406 sociality follow Gaussian and probit ("threshold") families, respectively. For simplicity, we ran
407 the program without incorporating intraspecific variation, branch-specific selection, or the
408 bounding of group size. The MCMCglmm results qualitatively agreed with our results (e.g., 95%
409 CI of the covariance between group size and sociality > 0), yet the process took only a few
410 minutes to an hour on a standard laptop, depending on the parameters (results not shown). Thus,
411 the scope of the present ABC-PCM algorithm is currently limited to analyses of relatively small
412 to moderate numbers of species (29 species in Harano & Kutsukake 2018), but it is not suitable
413 for a large number of species. Improvements in both efficient simulation algorithms and
414 computational power will enable analyses based on a larger number of species.
415 In summary, we have developed a flexible multivariate ABC-PCM that has great potential
416 for testing biologically intricate scenarios of trait evolution. Although our analyses considered the
417 evolutionary transition of sociality in a relatively small number of species, our method can be
418 applied to other topics and to larger datasets. Despite the fact that our analyses focused on a simple
419 case in which a continuous trait affects the state of a discrete trait, other causational patterns can
420 be dealt with by a flexible setting of an evolutionary simulation. Furthermore, our ABC-PCM
421 framework can also be extended to model more complex evolutionary trajectories, such as
422 asymmetric transitions between states and/or more than two states of discrete traits with different
423 transition functions. Thus, our method allows evolutionary biologists to explore various
424 hypotheses of interest concerning the evolution of interrelated traits.
425
426 **Acknowledgements**

We would like to thank Dustin R. Rubenstein, Rafael Maia, and Margaret E. O'Brien at Columbia University and Jessica Zung at Princeton University for useful comments and advice. Masahito Tsuboi at University of Oslo also provided useful insights on the earlier version of manuscript. Two reviewers gave detailed and helpful comments on the manuscript. This study was financially supported by MEXT (No. 25711025) to K. N.

Author contributions

Y. H. and K. N. conceived, designed, and performed the analysis. Both discussed the result and wrote the final manuscript.

Conflict of interest

None declared.

Literature cited

Bennett, N.C., Jarvis, J.U.M. & Cotterill, F. P. D. 1994. The colony structure and reproductive biology of the afrotropical Mashona mole-rat, Cryptomys darlingi. *J. Zool*. **234**: 477-487.

Beaumont, M.A. 2010. Approximate Bayesian computation in evolution and ecology. *Ann. Rev. Ecol. Evol. Syst.* **41**:379–406.

Beaumont, M. A., Zhang, W. & Balding, D. J. 2002. Approximate Bayesian computation in population genetics. *Genetics* **162**:2025–2035.

Beauchamp, G. 1999. The evolution of communal roosting in birds: origin and secondary losses. *Behav. Ecol.* **10**: 675-687.

Bennett, N.C. & Faulkes, C.G. 2000. African mole-rats: ecology and eusociality. Cambridge, UK: Cambridge University Press.

Bokma, F., 2010. Time, species, and separating their effects on trait variance in clades. *Syst. Biol.* **59**:602–607.

Burda, H., Honeycutt, R.L., Begall, S., Locker-Grütjen, O. & Scharff, A. 2000. Are naked and common mole-rats eusocial and if so, why? *Behav. Ecol. Sociobiol.* **47**: 293-303.

Burda, H. & Kawalika M. 1993. Evolution of eusociality in the Bathyergidae: the case of the giant mole-rat (*Cryptomys mechowi*). *Naturwissenschaften* **80**:235-237.

Csillery, K., Blum, M.G., Gaggiotti, O.E. & François, O. 2010. Approximate Bayesian computation (ABC) in practice. *Trends. Ecol. Evol.* **25**:410–418.

Duffy, J.M. & Macdonald, K.S. 2010. Kin structure, ecology and the evolution of social organization in shrimp: a comparative analysis. *P. R. Soc. B*. **277**: 575-584.

Emlen, S.T. 1982. The evolution of helping. I. An ecological constraints model. *Am. Nat.* **119**: 29–39.

Faulkes, C.G., Bennett, N.C., Bruford, M.W., O'brien, H.P., Aguilar, G.H. & Jarvis, J.U.M. 1997. Ecological constraints drive social evolution in the African mole–rats. *P. R. Soc. B.* **264**: 1619-1627.

465 Faulkes, C.G., Verheyen, E., Verheyen, W., Jarvis, J.U.M. & Bennett, N.C. 2004.
466 Phylogeographical patterns of genetic divergence and speciation in African mole-rats
467 (Family: Bathyergidae). *Mol. Ecol.* **13**: 613-629.
468 Faulkes, C.G. & Bennett, N.C. 2013. Plasticity and constraints on social evolution in African
469 mole-rats : ultimate and proximate factors. *Philos. T. R. Soc. B*. **368**:20120347.
470 Felsenstein, J. 1985. Phylogenies and the comparative method. *Am. Nat.* **125**: 1–15.
471 Felsenstein, J. 2005. Using the quantitative genetic threshold model for inferences between and
472 within species. *Philos. T. R. Soc. B*. **360**: 1427-1434.
473 Felsenstein, J. 2012. A comparative method for both discrete and continuous characters using the
474 threshold model. *Am. Nat.* **179**: 145-156.
475 Garamszegi, L.Z. 2014 *Modern Phylogenetic Comparative Methods and their Application in*
476 *Evolutionary Biology - Concepts and Practice,* (L.Z. Garamszegi, ed.). Springer,
477 Heidelberg.
478 Hadfield, J.D. & Nakagawa, S. 2009. General quantitative methods for comparative biology:
479 phylogenies, taxonomies and multi-trait models for continuous and categorical characters.
480 *J. Evol. Biol.* **23**: 494-508.
481 Hadfield, J. D. 2015. Increasing the efficiency of MCMC for hierarchical phylogenetic models of
482 categorical traits using reduced mixed models. *Methods in Ecology and Evolution*
483 6.6:706-714.
484 Harano, T. & Kutsukake, N. 2018. Directional selection in the evolution of elongated upper
485 canines in clouded leopards and sabre-toothed cats. *J. Evol. Biol*., **31**: 1268-1283.
486 Harvey P.H. & Pagel, M.D. 1991. The comparative method in evolutionary biology. Oxford
487 University Press, Oxford
488 Ives, A. R. & Garland, T. 2010. Phylogenetic logistic regression for binary dependent variables.
489 *Syst. Biol.* **59**: 9-26.

490 Ives, A. R. & Garland, T. 2014. Phylogenetic Regression for Binary Dependent Variables.
491 In: *Modern Phylogenetic Comparative Methods and their Application in Evolutionary*
492 *Biology - Concepts and Practice* (L.Z. Garamszegi, ed.), pp. 231-261. Springer,
493 Heidelberg.
494 Janzen, T., Hoehna, S. & Etienne, R. S. 2015. Approximate Bayesian Computation of
495 diversification rates from molecular phylogenies: introducing a new efficient summary
496 statistic, the nLTT. Methods Ecol. Evol. **6**: 566-575.
497 Jarvis, J.U.M. & Bennett, N.C. 1993. Eusociality has evolved independently in two genera of
498 bathyergid mole-rats—but occurs in no other subterranean mammal. *Behav. Ecol.*
499 *Sociobiol.* **33**: 253-260.
500 Jarvis, J.U.M. 1981. Eusociality in a mammal: cooperative breeding in naked mole-rat colonies.
501 *Science* **212**: 571–573.
502 Jones, K., Bielby, J., Cardillo, M. & Fritz, S. 2009. PanTHERIA: a species-level database of life
503 history, ecology, and geography of extant and recently extinct mammals. *Ecology* **90**:
504 2648.
505 Kutsukake, N. & Innan, H. 2013. Simulation-based likelihood approach for evolutionary models
506 of phenotypic traits on phylogeny. *Evolution* **67**: 355-367.
507 Kutsukake, N. & Innan, H. 2014. Detecting phenotypic selection by approximate Bayesian
508 computation (ABC) in phylogenetic comparative methods. In: *Modern Phylogenetic*
509 *Comparative Methods and their Application in Evolutionary Biology - Concepts and*
510 *Practice* (L.Z. Garamszegi, ed.), pp. 409-424. Springer, Heidelberg
511 Leuenberger, C. & Wegmann, D. 2010. Bayesian computation and model selection without
512 likelihoods. *Genetics* **184**: 243-252.
513 Lewis, P. O. 2001. A likelihood approach to estimating phylogeny from discrete morphological
514 character data. Syst. Biol. **50**: 913–925.
515 Marjoram P, Tavare S. 2006. Modern computational approaches for analysing molecular genetic

516 variation data. Nat. Genet. Rev. **7**:759–770

517 Marjoram P, Molitor J, Plagnol V, Tavare S. 2003. Markov chain Monte Carlo without

518 likelihoods. Proc. Natl. Acad. Sci. USA **100**:15324–15328

519 Nunn, C.L. 2011. The comparative approach in evolutionary anthropology and biology.

520 University of Chicago Press, Chicago

521 Pagel, M. 1994. Detecting correlated evolution on phylogenies: a general method for the

522 comparative analysis of discrete characters. *P. R. Soc. B.* **255**: 37-45.

523 R core team. 2017. R: A language and environment for statistical computing. *R Found. Stat.*

524 *Comput. Vienna, Austria.*

525 Sheehan, M.J., Botero, C.A., Hendry, T.A., Sedio, B.E., Jandt, J.M., Weiner, S., Toth, A.L. &

526 Tibbetts, E.A. 2015. Different axes of environmental variation explain the presence vs.

527 extent of cooperative nest founding associations in *Polistes* paper wasps. *Ecol. Lett.* **18**:

528 1057-67.

529 Sherman, P.W., Jarvis, J.U.M. & Alexander, R.D. 1991. The biology of the naked mole-rat.

530 Princeton, NJ: Princeton University Press.

531 Shultz, S., Opie, C. & Atkinson, Q.D. 2011. Stepwise evolution of stable sociality in primates.

532 *Nature* **479**: 219-222.

533 Sichilima, A.M., Faulkes, C.G. & Bennett, N.C. 2008. Field evidence for aseasonality of

534 reproduction and colony size in the Afrotropical giant mole-rat *Fukomys mechowii*

535 (Rodentia: Bathyergidae). *Afr. Zool*. **43**: 144-149.

536 Sichilima, A.M., Bennett, N.C., Faulkes, C.G. & Bronner, G.N. 2011. Field evidence for colony

537 size and aseasonality of breeding and in Ansell's mole-rat, *Fukomys anselli* (Rodentia:

538 Bathyergidae). *Afr. Zool.* **46**: 334-339.

539 Slater, G. J., Harmon, L.J., Wegmann, D., Joyce, P., Revell, L.J. & Alfaro, M.E. 2012. Fitting

540 models of continuous trait evolution to incompletely sampled comparative data using

541 approximate Bayesian computation. *Evolution* **66**:752–762 .

542 Van Daele, P.A., Blonde, P., Stjernstedt, R. & Adriaens, D. 2013. A new species of African
543 mole-rat (Fukomys, Bathyergidae, Rodentia) from the Zaire-Zambezi watershed. *Zootaxa*
544 **3636**: 171-189.
545 Wcislo, W.T. & Danforth, B.N. 1997. Secondarily solitary: the evolutionary loss of social
546 behavior. *Trends Ecol. Evol.* **12**: 468-474.
547 Wright, S. 1934. An analysis of variablity in number of digits in an inbred strain of guniea pigs.
548 *Genetics.* **19**(6): 506-536.
549 Young, A.J., Jarvis J.U.M., Barnaville, J. & Bennett, N.C. 2015. Workforce effects and the
550 evolution of complex society in wild Damaraland mole rats. *Am. Nat.* **186**: 302-311.
551

552 **Table 1**. **Summary of previous methods and the originality of this study.**

	Approach	
Trait	Univariate models	Multivariate models (link function)
Discrete	Mk model[1] Threshold model[2*] PLogReg[3] (w/o independent variables)	Mk model[1] Threshold model[2*]
Continuous	PLogReg[3] MCMCglmm[4,5]	MCMCglmm[4,5]
Both discrete and continuous	-	Threshold model[2] (determined by a threshold)[*] PLogReg[3] (logit link) MCMCglmm[4,5] (logit and probit link) This study (researcher-defined function)

553 [1]Pagel 1994; Lewis 2001
554 [2]Felsenstein 2012
555 [3]Phylogenetic logistic regression, Ives and Garland 2010
556 [4]Hadfield and Nakagawa 2009
557 [5]Hadfield 2015
558 [*]continuous trait (liability) is unobservable
559

560 **Table 2**. **Data used in this study.**

Species	Social system	Group size, mean (SD)
Naked mole rat, *Heterocephalus glaber*[1]	eusocial	75 (48.65)[A]
Damaraland mole rat, *Fukomys damarensis*[1]	eusocial	11 (6.26)
Whyte's mole rat, *Fukomys hottentotus*[1]	social	5.16 (2.62)[B]
Mechow's mole rat, *Fukomys mechowii*[2]	social	9.91 (2.49)
Mashona mole rat, *Fukomys darlingi*[3]	social	7 (not available[C])
Ansell's mole rat, *Fukomys anselli*[4]	social	8.72 (2.15)
Cape mole rat, *Georychus capensis*[5]	solitary	1[D]
Cape dune mole rat, *Bathyergus suillus*[5]	solitary	1[D]
Namaqua dune mole rat, *Bathyergus janetta*[5]	solitary	1[D]
Silvery mole rat, *Heliophobius argenteocinereus*[5]	solitary	1[D]

561
562 [1]Bennett & Faulkes 2000; [2]Sichilima *et al*. 2008; [3]Bennett *et al*. 1994; [4]Sichilima *et al*. 2011; [5]Burda & Kawalika
563 1993; see Van Daele *et al*. (2013) for a recent classification.
564 [A]Because group sizes were categorized at an interval of 25 individuals, we used the mean value of each range for
565 calculating the mean and SD (e.g., 37.5 for groups of 25 to 50 individuals).
566 [B]When calculating the mean group size of this species, we used a group size of 11 when it exceeded 10, because the
567 original data for group size pooled group sizes larger than 10 into ">10" (p. 92 in Bennett & Faulkes 2000).
568 [C]Because no data on intraspecific variation were available, we used the mean value of the SD, 2.42, of the other three
569 social species.
570 [D]The mean litter size ranges from 2.46 to 5.94 in solitary species (Jones *et al*. 2009). To account for this, we set a
571 value of 2.56 as their SD to cover the range of possible group sizes.

572 **Table 3. Parameters estimated in this study.**

Parameter	Notation	Prior distribution	Posterior distribution (95% CI)	
			Model 1 (logistic)	Model 2 (exponential)
Most recent common ancestor (MRCA)	θ	U(1.001, 10)	1.1–5.22	2.45–9.15
Baseline evolutionary rate (per million years)	μ	U(0.001, 30)	6.4–29.5	3.62–28.3
Directional selection (naked mole rats)	k_n	U(0.001, 30)	0.99–3.57	1.01–1.93
Directional selection (Damaraland mole rats)	k_d	U(0.001, 5)	0.82–4.22	0.82–4.64
Curvature of transition functions (solitary ⇄ social)	*a* or *c*	U(-15,15)	*a*: 1.52–14.50	*c*: 1.28–14.60
Position of transition functions	*b or d*	U(0,10)	*b*: 1.52–4.94	*d*: 1.28–5.07

573

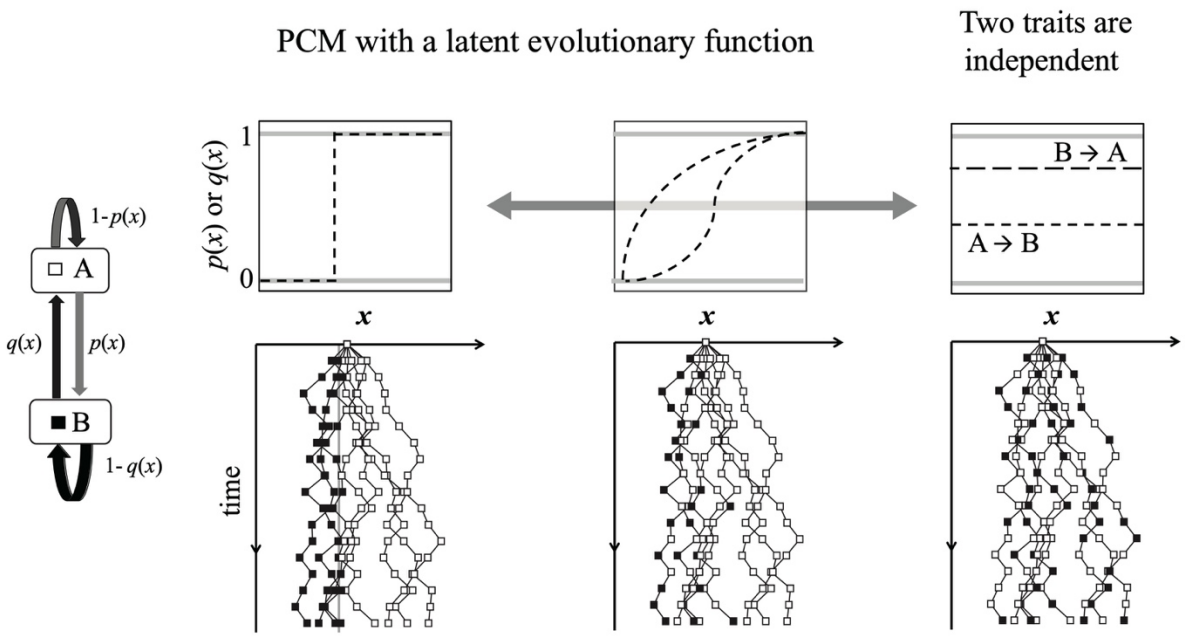

Fig. 1. Variation in multivariate PCMs with a latent function between discrete and continuous traits. Conventional PCMs can explore various latent evolutionary functions between continuous and discrete traits (Table 1). Here, we consider a simple example of the evolutionary link between a continuous and a discrete trait. The continuous trait (x) evolves by a process similar to Brownian motion. The discrete trait, on the other hand, evolves according to the transition probabilities $p(x)$ and $q(x)$, i.e., the latent evolutionary functions. The state of the discrete trait at each time step is shown as a white (A) or black (B) square. At one extreme, the transition of discrete states is determined by a certain "threshold" value of the continuous trait. At the other extreme, the discrete trait and the continuous trait are independent. In such a case, the transition probabilities do not vary with the continuous traits (right). In our framework, any functions between the two extreme cases can be incorporated based on biological hypotheses (middle). Furthermore, x can be either a measurable trait or a latent value. Here, logistic and exponential functions are shown as examples.

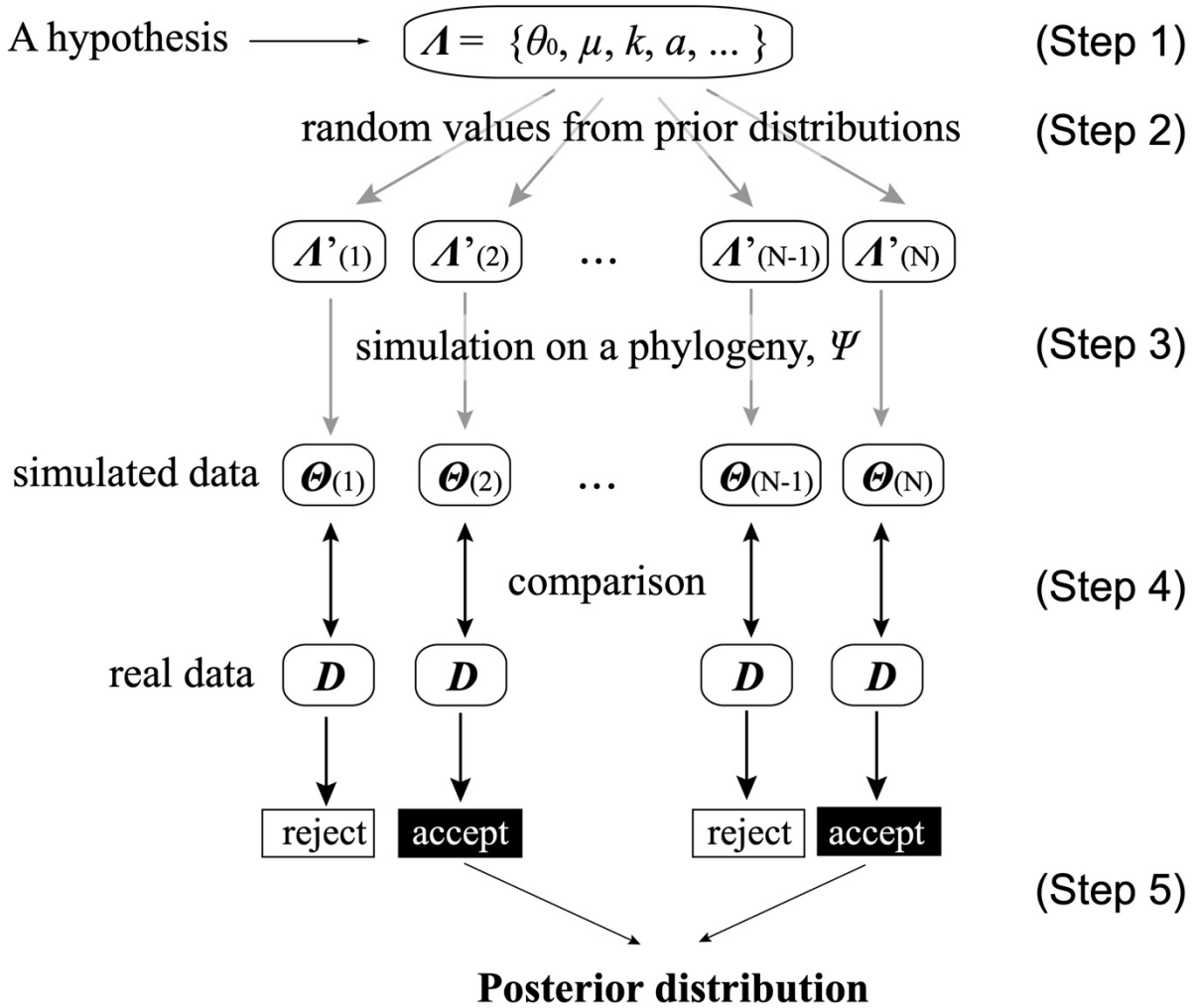

Fig. 2. Schematic of our ABC-PCM approach. First, determine parameters of interest and set prior distributions (Step 1). Example parameters include phenotype of the MRCA (θ_0), evolutionary rate (μ), directional selection parameter (k), latent function parameter (a), and so on. Next, generate random values for parameters from the prior distributions (Step 2). Trait simulations on a phylogeny Ψ are then conducted (Step 3), and the simulated data $\boldsymbol{\Theta}$ are compared to real data $\boldsymbol{D}$ to determine whether the data are accepted or rejected (Step 4). A number (N) of simulations are conducted until enough samples are collected to infer the posterior distribution (Step 5).

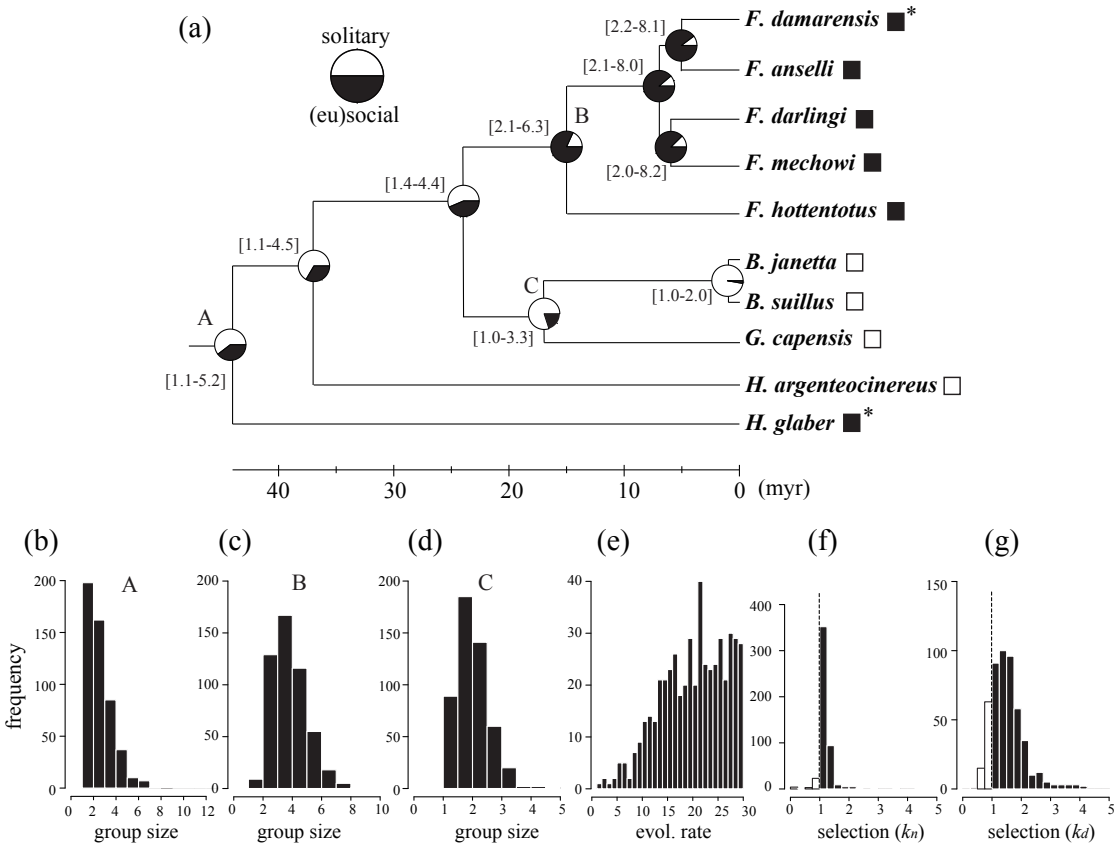

Fig. 3. The estimated evolutionary trajectory with the latent function being logistic (model 1). (a) The estimated parameters on the phylogeny. The social state of each tip species is shown as black (solitary) or white (social) squares. Asterisks indicate eusocial species. The proportion of solitary/social states and the 95% CI of the posterior distribution of group size (within brackets) are shown at each internal node. (b–d) The distributions of simulated group size at nodes A (MRCA), B, and C in the phylogeny. (e) The posterior distribution of the baseline evolutionary rate. (f, g) The posterior distributions of the selection coefficients k_n and k_d. The dashed line at $k = 1$ indicates the proportion of simulations in which directional selection for larger group size was detected (black histograms).

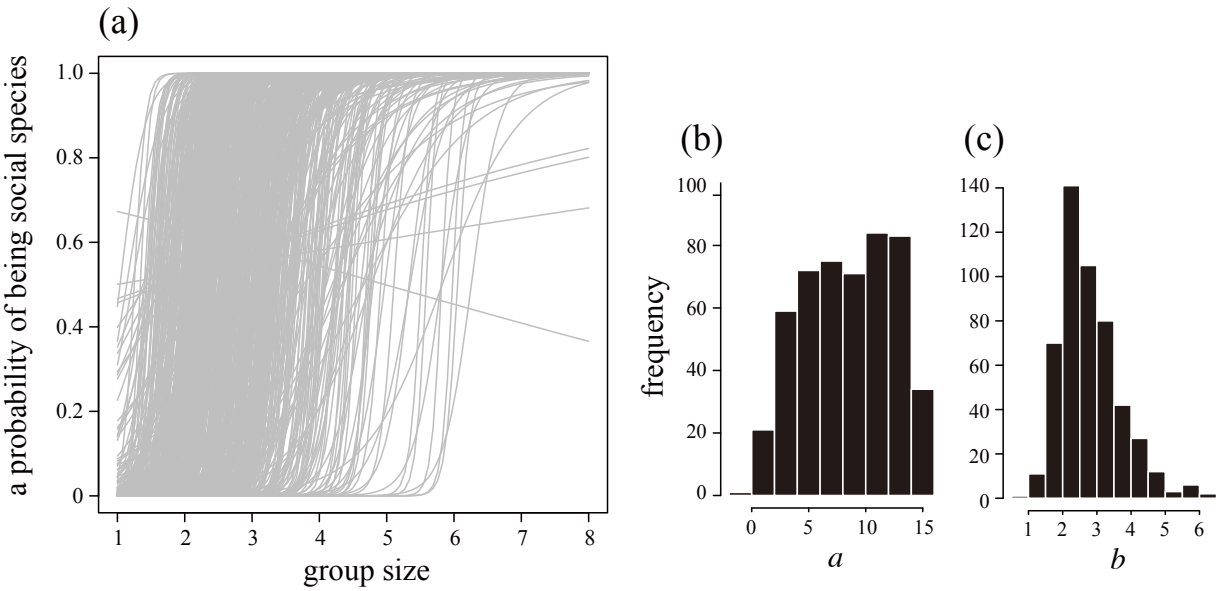

Fig. **4. Accepted latent functions in model 1**. (a) All 500 accepted latent functions $p(x)$, i.e., the probability of transition from a solitary state to a social state as a function of group size. Most accepted functions increased steeply at group sizes of 2 to 4. (b, c) The posterior distributions of a and b, the curvature of the function and the group size at which $p(x)$ equals 0.5.

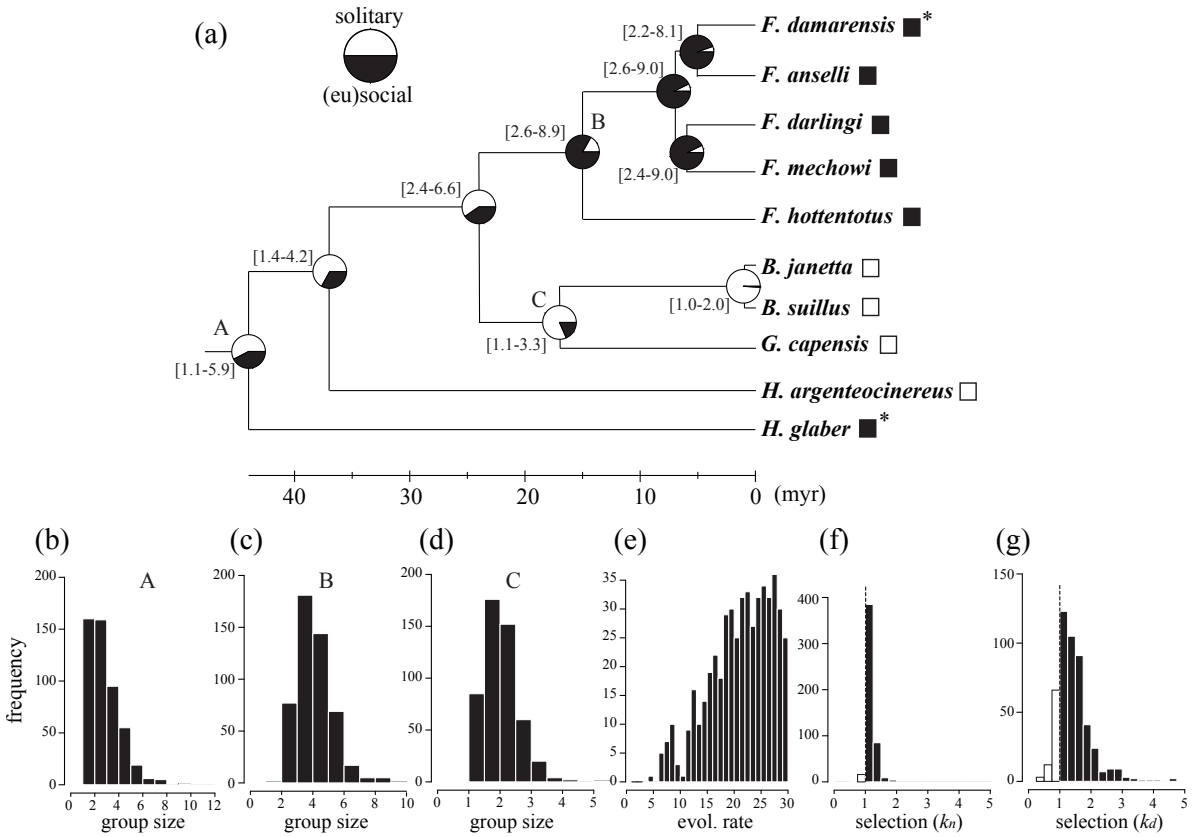

Fig. 5. The estimated evolutionary trajectory with the latent function being exponential (model 2). (a) The estimated parameters on the phylogeny. The social state of each tip species is shown as black (solitary) or white (social) squares. Asterisks indicate eusocial species. The proportion of solitary/social states and the 95% CI of the posterior distribution of group size (within brackets) are shown at each internal node. (b–d) The distributions of simulated group size at nodes A (MRCA), B, and C in the phylogeny. (e) The posterior distribution of baseline evolutionary rate. (f, g) The posterior distributions of the selection coefficients k_n and k_d. The dashed line at $k = 1$ indicates the proportion of simulations in which directional selection for larger group size was detected (black histograms).

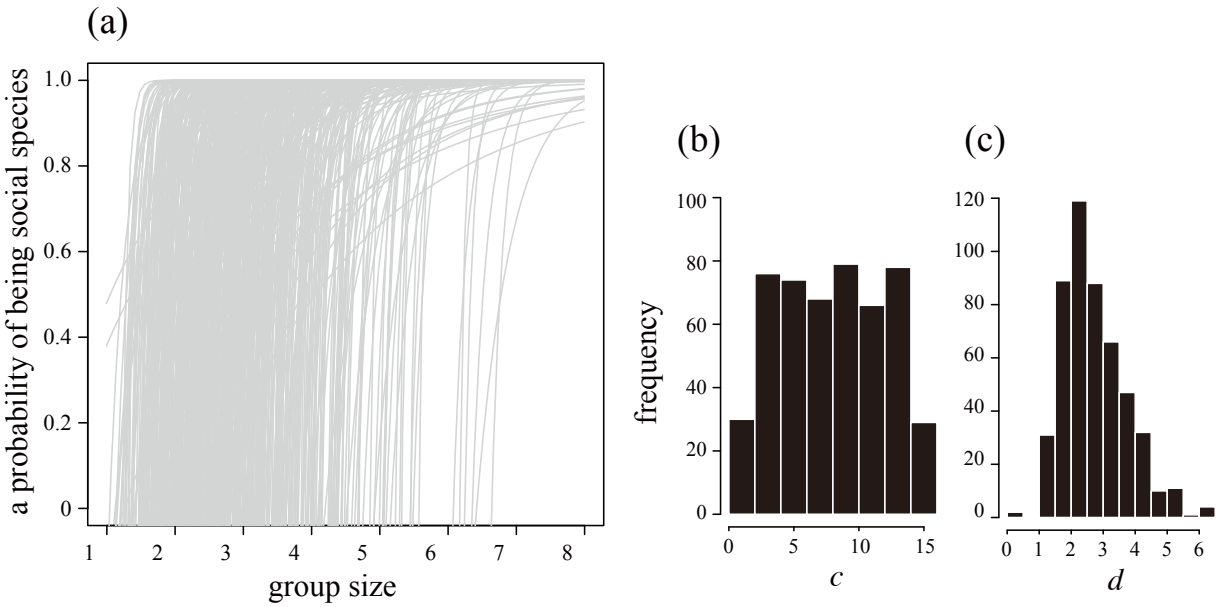

Fig. **6. Accepted latent functions in model 2.** (a) All 500 accepted latent functions corresponding to $p(x)$. Similar to the case in model 1, most of the accepted functions increased at a group size of 2 to 4. (b, c) The posterior distributions of c and d, the curvature of the function and the group size at which $p(x)$ equals 0.