

SUPPLEMENTARY MATERIAL

Effect of host-switching on the ecological and evolutionary patterns of parasites

¹*Elvira D'Bastiani, ^{2,3}Débora Princepe, ^{2,3}Flavia MD Marquitti, ^{1,4}Walter A Boeger,

^{1,4}Karla M Campião, ^{1,5}Sabrina BL Araujo

¹Laboratório de Interações Biológicas, Programa de Pós-Graduação em Ecologia e Conservação, Universidade Federal do Paraná, UFPR - Curitiba, Paraná, Brasil

²Instituto de Física "Gleb Wataghin", Universidade Estadual de Campinas, UNICAMP - Campinas, São Paulo, Brasil

³Instituto de Biologia, Universidade Estadual de Campinas, UNICAMP - Campinas, São Paulo, Brasil

⁴Departamento de Zoologia, Universidade Federal do Paraná, UFPR - Curitiba, Paraná, Brasil

⁵Departamento de Física, Universidade Federal do Paraná, UFPR - Curitiba, Paraná, Brasil

*Corresponding author: Elvira D'Bastiani e-mail: elviradbastiani@gmail.com

SHORT RUNNING TITLE: Host-switching influences parasite patterns

Contents

S1 MATERIAL AND METHODS.....	03
S1.1 Derrida-Higgs model.....	03
S1.2 Two-islands speciation models.....	06
S1.3 Parameters used.....	09
S1.4 Exploratory analysis.....	12
S1.4.1 Population size test.....	12

S1.4.2 Mutation test.....	15
S1.4.3 Relationship between parasite richness and host-switching intensity for each population size.....	17
S1.5 Details for empirical database.....	18
S2 Details Results.....	28
S2.1 Relationship between variation in parasite composition and intensity of host-switching for the nine empirical parasite communities.....	28
S2.2 Relationship between tree imbalance and intensity of host-switching for the nine empirical parasite communities.....	29
S2.3 Relationship between parasite richness and intensity of host-switching for the nine empirical parasite communities.....	32
S2.4 Example ecological and evolutionary patterns over time.....	33
S3 Others exploratory analysis.....	36
S3.1 Influence of host-switching events on the ecological and evolutionary patterns of simulated parasites with $q_{min}=2q_0$	36
S3.2 Different shapes of host phylogeny (hypothetical host phylogenies) - Influence of host-switching events on the ecological and evolutionary patterns of simulated parasites.....	38

SUPPLEMENTARY MATERIAL - Data SI

S1 MATERIAL AND METHODS

Here we detail the models that Higgs and Derrida (1991) and Manzo and Peliti (1994) proposed. Their models approach speciations of populations inhabiting a unique or two spatial regions (islands) respectively. Our model does not explicitly describe the space, but we consider hosts limiting gene flow. It means that if parasites do not migrate to other host species, the gene flow is broken, in analogy to allopatric speciation. Then, the terminology "sympatry" and "allopatry" used in the cited models must be interpreted here as "in the same host species" and "different host species" respectively.

S1.1 Derrida-Higgs model

The model introduced by Higgs and Derrida (1991) considers a sympatric population of K haploid individuals (population carrying capacity) whose genomes are represented by binary strings of size B , $\{S_1^\alpha, S_2^\alpha \dots S_B^\alpha\}$ where S_i^α can assume the values ± 1 . Each locus of the genome is dubbed as a gene and the values +1 and -1 the corresponding alleles. The number of individuals at each generation is kept constant and the population is characterised by a $K \times K$ matrix q measuring the degree of genetic similarity between pairs of individuals:

$$q_{\alpha\beta} = \frac{1}{B} \sum_{i=1}^B S_i^\alpha S_i^\beta. \quad (1)$$

If the genomes of α and β are identical $q_{\alpha\beta} = 1$ whereas two genomes with random entries will have $q_{\alpha\beta}$ close to zero. Each generation is constructed from the previous one as follows: a first parent P_1 is chosen at random. The second parent P_2 has to be genetically compatible with the first, i.e., their degree of similarity has to satisfy $q_{P_1 P_2} \geq q_{min}$.

Individuals P_2 are then randomly selected until this condition is met with K trials. If no

such individual is found, P_l is discarded and a new first parent is selected. The offspring inherits, gene by gene, the allele of either parent with equal probability (sexual reproduction). The process is repeated until K offspring have been generated.

Individuals are also subjected to a mutation rate μ per gene, which is typically small.

To understand how the similarity matrix changes through generations, consider first an asexual population where each individual α has a single parent P_l in the previous generation. The allele S_i^α will be equal to $S_i^{P(\alpha)}$ with probability $\frac{1}{2}(1 + e^{-2\mu}) \approx 1 - \mu$ and $-S_i^{P(\alpha)}$ with probability $\frac{1}{2}(1 - e^{-2\mu}) \approx \mu$, so that the expected value is $E(S_i^\alpha) = e^{-2\mu} S_i^{P(\alpha)}$. For independent genes, the expected value of the similarity between α and β is, therefore, $E(q_{\alpha\beta}) = e^{-4\mu} q_{P(\alpha), P(\beta)}$. In sexual populations α and β have two parents each, P_1^α, P_2^α and P_1^β, P_2^β , and since each inherits (on average) half the alleles from each parent, we obtain

$$\begin{aligned} E(S_i^\alpha) &= \frac{1}{2} (S_i^{P_1(\alpha)} \frac{1+e^{-2\mu}}{2} - S_i^{P_1(\alpha)} \frac{1+e^{-2\mu}}{2}) + \frac{1}{2} (S_i^{P_2(\alpha)} \frac{1+e^{-2\mu}}{2} - S_i^{P_2(\alpha)} \frac{1+e^{-2\mu}}{2}) \\ &= \frac{e^{-2\mu}}{2} (S_i^{P_1(\alpha)} + S_i^{P_2(\alpha)}). \end{aligned} \quad (2)$$

It follows that, on average, the similarity between α and β is

$$q_{\alpha\beta} = \frac{e^{-4\mu}}{4} (q_{P_1(\alpha), P_1(\beta)} + q_{P_2(\alpha), P_1(\beta)} + q_{P_1(\alpha), P_2(\beta)} + q_{P_2(\alpha), P_2(\beta)}), \quad (3)$$

with $q_{\alpha\alpha} \equiv 1$. In the limit of infinitely many genes, $B \rightarrow \infty$, this expression becomes exact and the entire dynamics can be obtained by simply updating the similarity matrix.

If there is no restriction on mating, $q_{min} = 0$, we can demonstrate that the overlaps $q_{\alpha\beta}$ converge to a stationary distribution (Higgs & Derrida, 1991), as follows.

At generation t , the probability that α and β have one parent in common, $P_1^\alpha = P_1^\beta$,

$P_1^\alpha = P_2^\beta, P_2^\alpha = P_1^\beta$ or $P_1^\alpha = P_2^\beta$, is $4/K$. In this case, the average similarity $\langle q_{\alpha\beta} \rangle$

between α and β is given by $\langle q_{\alpha\beta} \rangle = \frac{e^{-4\mu}}{4} (1 + 3\bar{q})$, where \bar{q} is the average similarity in the previous generation. If α and β do not share a parent, which happens with probability $1-4/K$, then $\langle q_{\alpha\beta} \rangle = e^{-4\mu} \bar{q}$. Therefore, at generation $t+1$ we get

$$\bar{q}_{t+1} = \frac{4}{K} \frac{e^{-4\mu}}{4} (1 + 3\bar{q}_t) + (1 - \frac{4}{K}) e^{-4\mu} \bar{q}_t = e^{-4\mu} [(1 - \frac{1}{K})\bar{q}_t + \frac{1}{K}] . \quad (4)$$

Setting $\bar{q}_{t+1} = \bar{q}_t = q_0$, we find the equilibrium at

$$q_0 = \frac{1}{1+4\mu K} . \quad (5)$$

The approximation holds for μ and $1/K$ much smaller than one, which is always the case for real populations. For example, with $K = 250$ and $\mu = 0.025$, population similarity will reach an equilibrium with $q_0 = 0.038$. If $q_{min} \leq q_0$, the genetic distribution reaches this equilibrium, and no species will arise. On the other hand, if the minimal similarity q_{min} satisfies $q_{min} > q_0$, sympatric speciation will occur and we can estimate the time for it to occur. Subtracting \bar{q}_t from both sides of Equation (4) and approximating

$\bar{q}_{t+1} - \bar{q}_t = \frac{dq}{dt}$, we get the differential equation

$$\frac{dq}{dt} = \frac{1}{K} (1 - \frac{q_t}{q_0}) \quad (6)$$

with solution given by

$$q(t) = q_0 + (1 - q_0) e^{-t/Mq_0} . \quad (7)$$

The time τ to speciation can be estimated as the time $q(t)$ takes to reach q_{min} . Setting

$q(\tau) = q_{min}$ we find

$$\tau = Mq_0 \log(\frac{1-q_0}{q_{min}-q_0}) . \quad (8)$$

S1.2 Two-islands speciation models

Manzo and Peliti (1994) extended the Derrida-Higgs model to describe populations geographically isolated in islands to investigate the possibility of allopatric speciation in the presence of gene flow. They considered two islands with K individuals each following the Derrida-Higgs dynamics with sexual reproduction. The interaction between individuals of different islands occurs only via migration: after reproduction with partners of the same island, individuals can migrate to the other island with probability ϵ . The genetic variation of the populations is measured by the quantity $q_{\alpha\beta}$, which is the similarity between individuals α and β on the same island, and a new quantity $p_{\alpha\beta}$ is introduced to measure the similarity of individuals belonging to different islands. Similarly to the Derrida-Higgs model, the description of dynamical properties of the system, such as the values of the similarities in the equilibrium, q_0 and p_0 , can be obtained in the approximation of infinitely large genomes and in the regime of small mutation rate and large population (Manzo and Peliti, 1994).

In special, Manzo and Peliti (1994) were interested in exploring the feasibility of the regime where $p_0 < q_{min} < q_0$: there is no speciation in the islands and each one bears a single species, but the islands differentiate from each other, i.e., the species are endemic. Therefore, the differentiation occurs only due to the exchange of individuals between the islands. In our model (Fig. 1 in the main text), we are similarly interested in studying the speciation of the parasite purely induced by host-switching (equivalent to migration events in the two islands model). Then, we also adopt $q_{min} < q_0$ to inhibit the speciation of parasites that would occur within hosts in the absence of host-switching (we use $q_{min} = 0.5q_0$). However, our model can not be further described by the expressions of Manzo-Peliti because the host-switching probability depends on time and

multiple hosts (islands) emerge over the simulation. Nevertheless, we still can use this description for the scenario where there is no migration/host-switching, i.e., the allopatric speciation. As follows, we demonstrate how to calculate the time for allopatric speciation, τ_a , which is a parameter to rescale the input phylogenies in our simulations.

In the case of two islands, we need to distinguish the similarities between individuals inhabiting the same island, q , or different islands, p , as stated previously. For individuals α and β that belong to the same island, the average similarity \bar{q}_{t+1} follows equation (4), $\bar{q}_{t+1} = e^{-4\mu}[(1 - \frac{1}{K})\bar{q}_t + \frac{1}{K}] = e^{-4\mu} Q(\bar{q}_t, K)$. Otherwise, if they were born in different islands, their similarity evolves as $\bar{p}_{t+1} = e^{-4\mu} \bar{p}_t$ since they do not have any parents in common. After reproduction, pair of individuals of the same island keep their original geographic relation if they do not migrate or both migrate, with probability $(1 - \epsilon)^2 + \epsilon^2 \equiv a(\epsilon)$. The probability of a pair of individuals changing their geographic relative position is given by $2\epsilon(1 - \epsilon) \equiv b(\epsilon)$, that accounts for one individual staying at the island and the other migrating (note that $a(\epsilon) + b(\epsilon) = 1$). Therefore, the dynamics of q and p is given by

$$q_{t+1} = a(\epsilon) e^{-4\mu} Q(q_t, K) + b(\epsilon) e^{-4\mu} p_t \quad (9)$$

$$p_{t+1} = b(\epsilon) e^{-4\mu} Q(q_t, K) + a(\epsilon) e^{-4\mu} p_t \quad (10)$$

where we omitted the bars for simplicity. For ϵ, μ and $1/K$ all much smaller than 1, the equations can be approximated by

$$q_{t+1} = (1 - 2\epsilon - 4\mu - 1/K)q_t + 2\epsilon p_t \quad (11)$$

$$p_{t+1} = 2\epsilon q_t + (1 - 2\epsilon - 4\mu) p_t \quad (12)$$

We obtain the coupled dynamical equations for $q(t)$ and $p(t)$ from equations 11 and 12 using the approximations $q_{t+1} - q_t = dq/dt$ and $p_{t+1} - p_t = dp/dt$, as done previously:

$$\frac{dq}{dt} = -(2\epsilon + 4\mu + 1/K)q(t) + 2\epsilon p(t) \quad (13)$$

$$\frac{dp}{dt} = 2\epsilon q(t) - (2\epsilon + 4\mu)p(t). \quad (14)$$

When in strict allopatry, $\epsilon = 0$, the similarity between islands is simply given by

$p(t) = e^{-4\mu t}$, i.e., it depends only on the mutation rate. The time for the diversification due to geographical isolation τ_a can be calculated by

$p(\tau_a) = e^{-4\mu\tau_a} = q_{min}$. The time for allopatry is then

$$\tau_a = \frac{1}{4\mu} \log\left(\frac{1}{q_{min}}\right). \quad (15)$$

The time for allopatry decreases with q_{min} : making the reproduction more restricted (increasing q_{min}) facilitates the differentiation between islands, which occurs in a smaller time.

In our simulations, we find the relation between the length of the branches in the host phylogeny and the time in generations by assuming that the smaller branch corresponds to the minimal time to parasites speciate due to geographical isolation. τ_a is calculated by equation (15) with the input parameters and $q_{min} = 0.5 q_0$, with q_0 given by equation (5). Also, we consider that the first host species (the root of the phylogeny) must evolve for τ_a generations for the parasite populations to accumulate genetic diversity before the first splitting event.

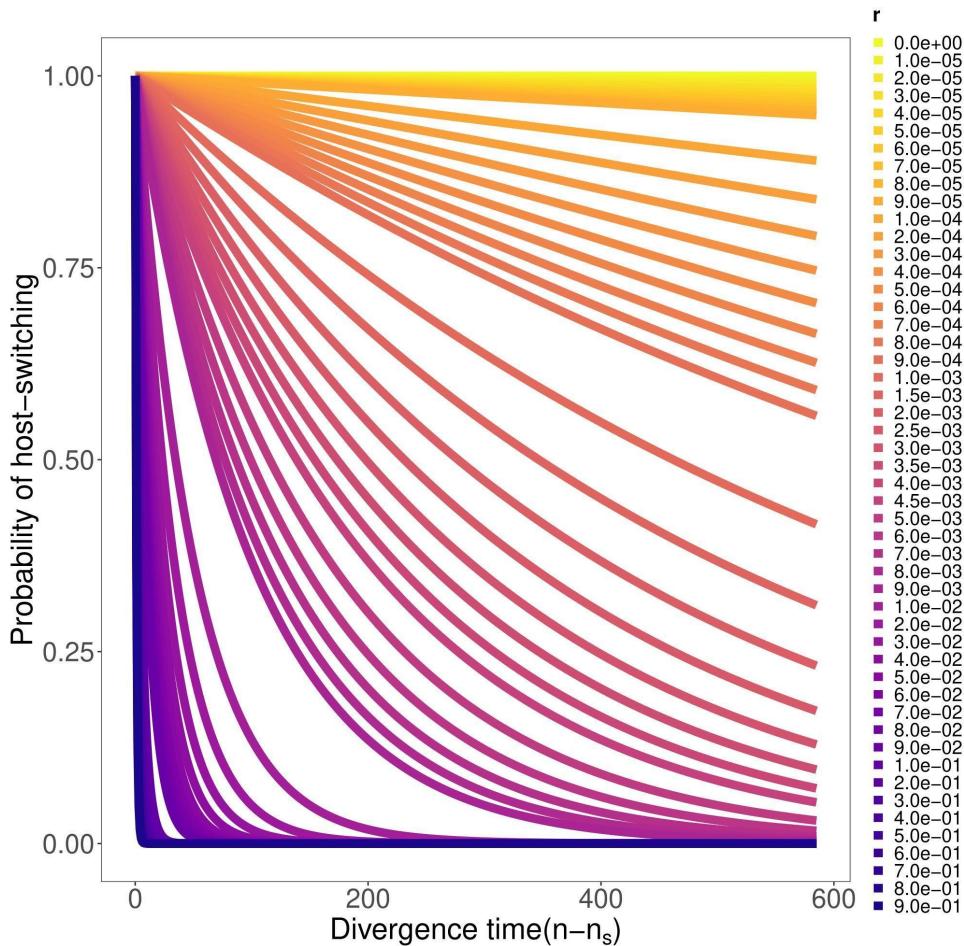


Figure SI1. Probability of a parasite individual successfully host-switching from a host to another host species as a function of the divergence time (how long the two host species had diverged). If $r=0$ the probability is equal to 1 regardless of the divergence time, however, this probability decreases if higher values of r are considered.

S1.3 Parameters used

The parameters involved in the model are described in Table S1. We considered four different carrying capacities $K = \{50, 250, 500, 1000\}$. When we vary the population size (Figure SI2 and S3) we do not observe differences in the patterns of beta diversity (β) and tree imbalance (normalised Sackin index - I_n) between the host-switching intensity, therefore we do not include the variation in the parasite species richness in our

main results (see the details in Fig. S2-S4). Due to computational cost limitations, we selected the least rich host case (ID. 2) to show the influence of all population sizes and we had no qualitative difference between them (see species richness in Figure SI4). The model is not too sensitive in response to changes in parasites' population size, which allows us to fix the values. Then, for the parasite population, we selected $K = 250$. Similarly, we also choose the $\mu = 0.025$, to carry out the simulations with different scenarios of host-switching intensity. We replicated each combination of the parameters for each empirical community 50 times. We included an initial transition time using the value of allopatric time in the model to eliminate the effect of the initial condition on evolutionary patterns. t represents the evolutionary time of the parasites and was parameterized according to the evolutionary time of the hosts and the time of allopatric speciation for each empirical case/community. Simulations were performed separately for each host phylogeny (Table S1).

Table S1. Model parameters with a short description and the investigated values.

Parameters	Short definition	Investigated values
K	Carrying capacity of parasite species per host species.	50, 250 , 500, 1000
μ	Mutation rate.	0.001, 0.025
q_{min}	Minimal genetic similarity.	0.5 * q_0
r	Intensity of decline in host-switching probabilities as the phylogenetic distance increases between hosts.	0-1
	Number of simulation repetitions for a given set of parameters.	50
	Total number of iterations	ID 1. 1812* ID 2. 1491* ID 3. 1343* ID 4. 702* ID 5. 624* ID 6. 541* ID 7. 2614*

ID 8. 339*
ID 9. 248*

Note: The bold values are the fixed values used in the presented results, while the other parameters used for the sensibility test.*Total time for each empirical community.

S1.4 Exploratory analysis

S1.4.1 Population size test

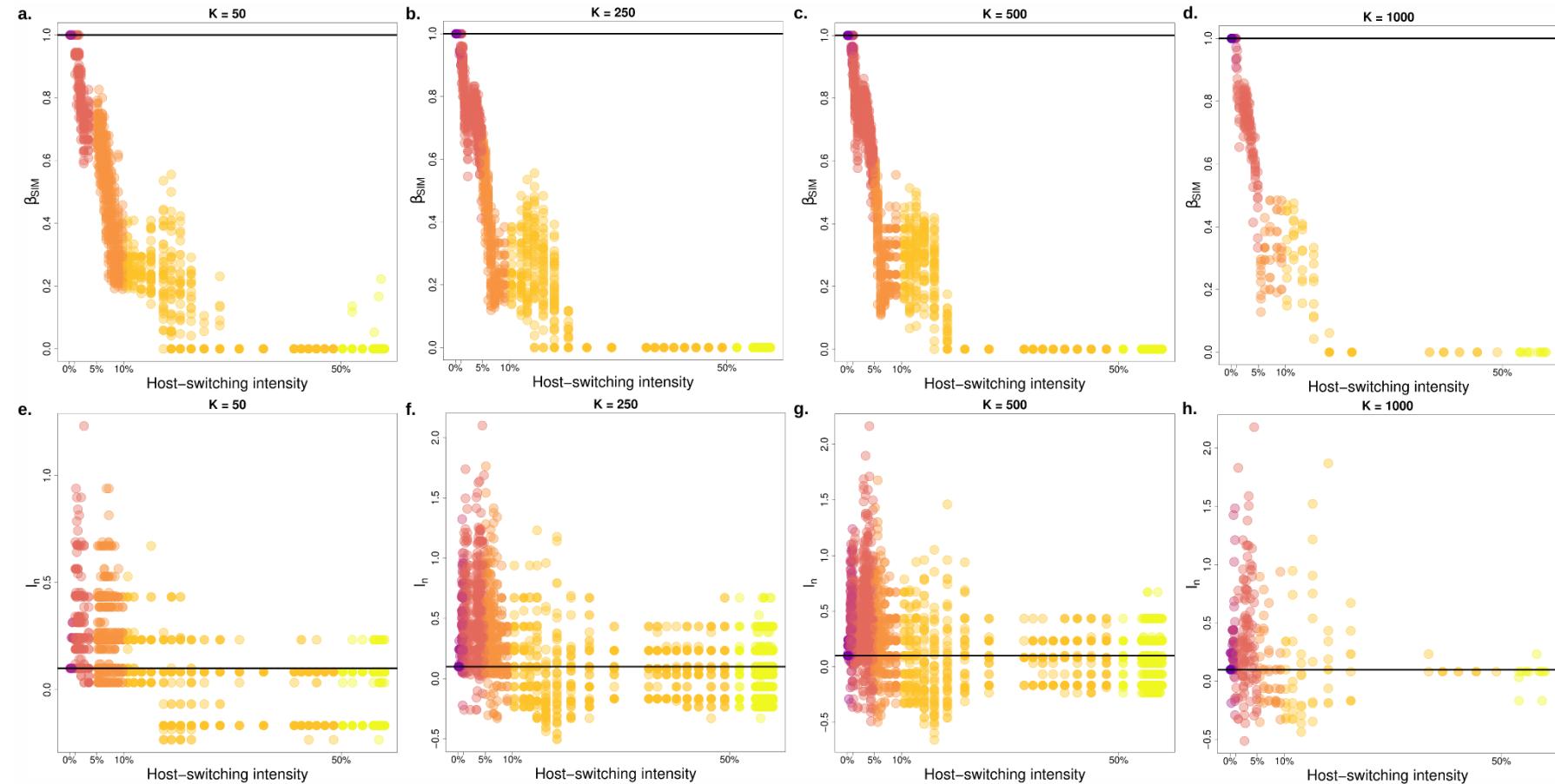


Figure SI2. Population size test. Relationship between variation in the composition and host-switching intensity, and the tree imbalance (normalised Sackin index - I_n) and host-switching intensity (e-h) for different population sizes for empirical community ID. 2 ($K = \{50, 250, 500, 1000\}$). Each K represents the population size tested. Empirical values: the lines refer to empirical information on the parasite (continuous) and host (dotted). In this case, these lines overlap. The scaling of the x-axis is in log scale. A total of 50 replications were performed for each host-switching intensity for each carrying capacity, with the exception of $K = 1000$, which demanded more time for running. For $K = 1000$, a total of 10 runs were performed.

S1.4.2 Mutation test

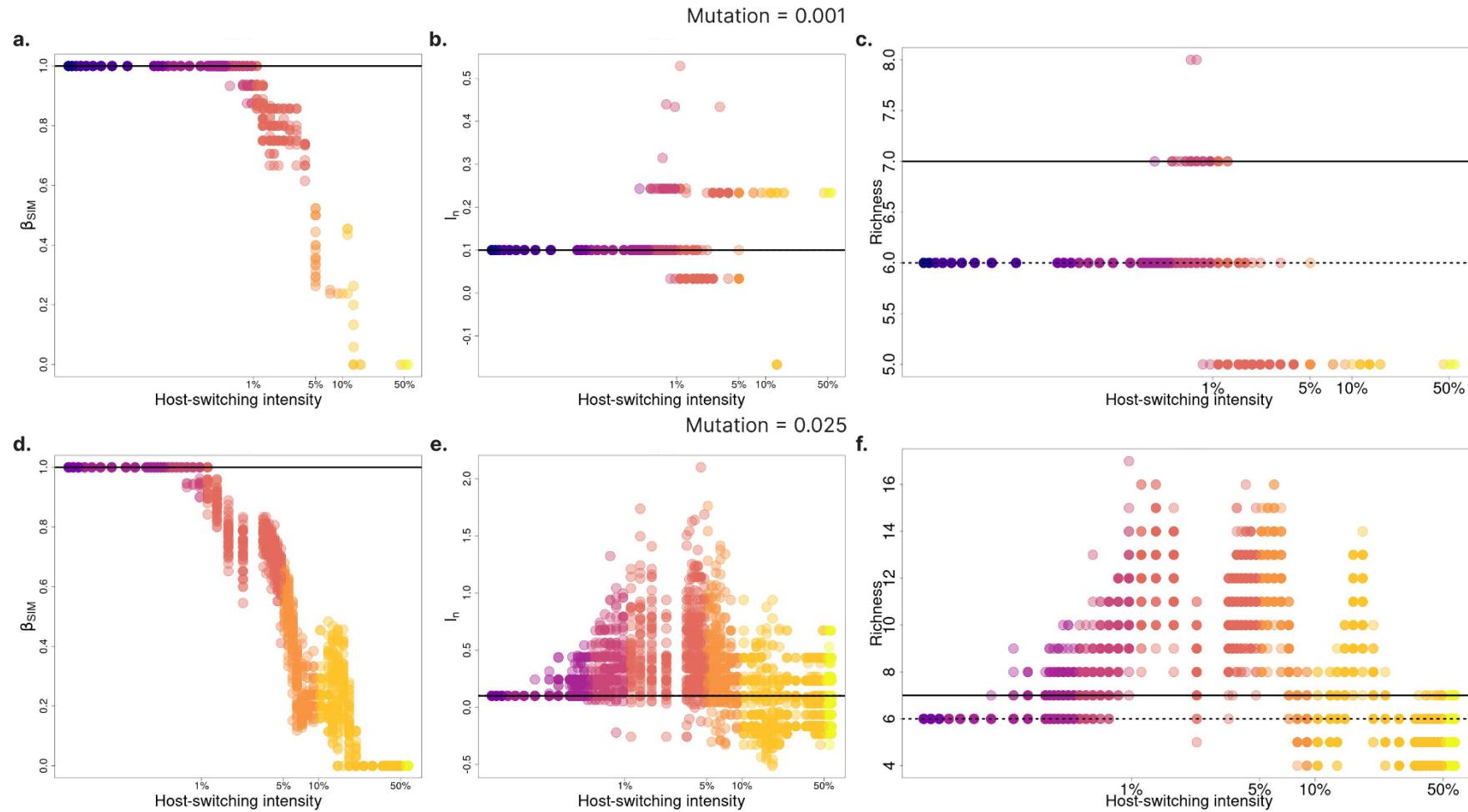


Figure SI3. Mutation test. Relationship between variation in the composition and host-switching intensity and the tree imbalance (normalised Sackin index - I_n), parasite richness and host-switching intensity (e-h) for two mutation rates (μ - 0.001 and 0.025) for empirical community ID. 2. Empirical values: the lines refer to empirical information on the parasite (continuous) and host (dotted). In this case, these lines overlap. A total of 50 replicates were performed for each host-switching intensity.

S1.4.3 Relationship between parasite richness and host-switching intensity for each population size

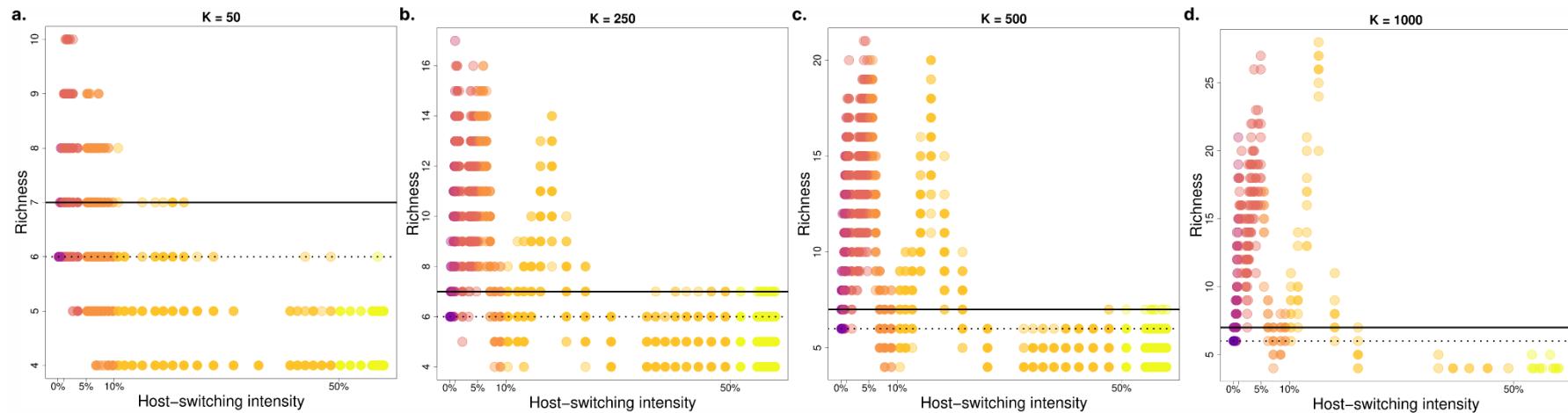


Figure SI4. Relationship between species richness and host-switching intensity (a-d) for empirical community ID. 2. Each K represents the population size tested. Empirical values: the lines refer to empirical information on the parasite (continuous) and host (dotted). A total of 50 replications were performed for each host-switching intensity for each carrying capacity, with the exception of $K = 1000$, which demanded more time for running. For this case, a total of 10 runs were performed.

S1.5 Details for empirical database

To test our model, the communities need to have empiric phylogenies for hosts and parasites and their interactions. A literature search was performed using the phrase "*phylogeny host-parasite*" and was carried out using Google Scholar between December 1, 2019, and January 2021, which identified more than 10.000 works. Among these, a total of 100 articles were selected for feasibility analyses for testing the model. Articles that did not contain phylogeny were immediately excluded. Studies focussing on the population level were also excluded. Studies that included less than six taxa were excluded because, as they have a low sampling effort, they can lead to misinterpretations. Additionally, studies with species that have asexual reproduction were excluded from consideration, as our model is sexed. Finally, for inclusion in our analyses, we extracted nine cases (Table 1 and Figure 2 in the main manuscript and Figure S15-S13) with knowledge of the associations and phylogenies that exist between parasites and host tips. Then, we studied the effects of host-switching and host evolution signatures on parasite speciation patterns (phylogenetic trees and variation in species composition) only for surviving species. Extinct species were not included in the analyses. All datasets are available at
[https://github.com/elviradbastiani/host_switching_model.](https://github.com/elviradbastiani/host_switching_model)

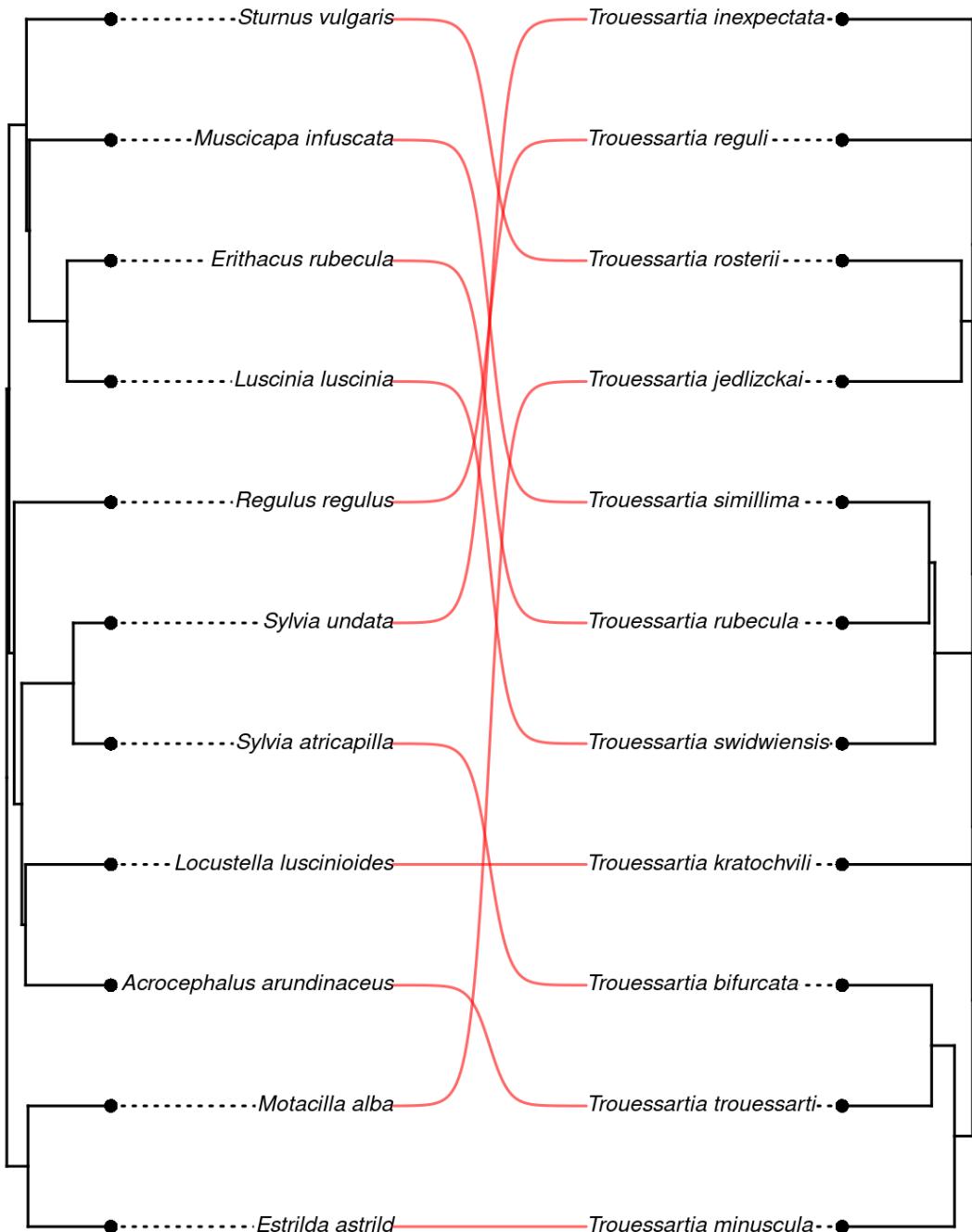


Figure SI5. Phylogenetic trees that correspond to the empirical data of Bird and feather mites (*Trouessartia* spp.) corresponding to ID. 1 extracted from Doña et al. 2017.

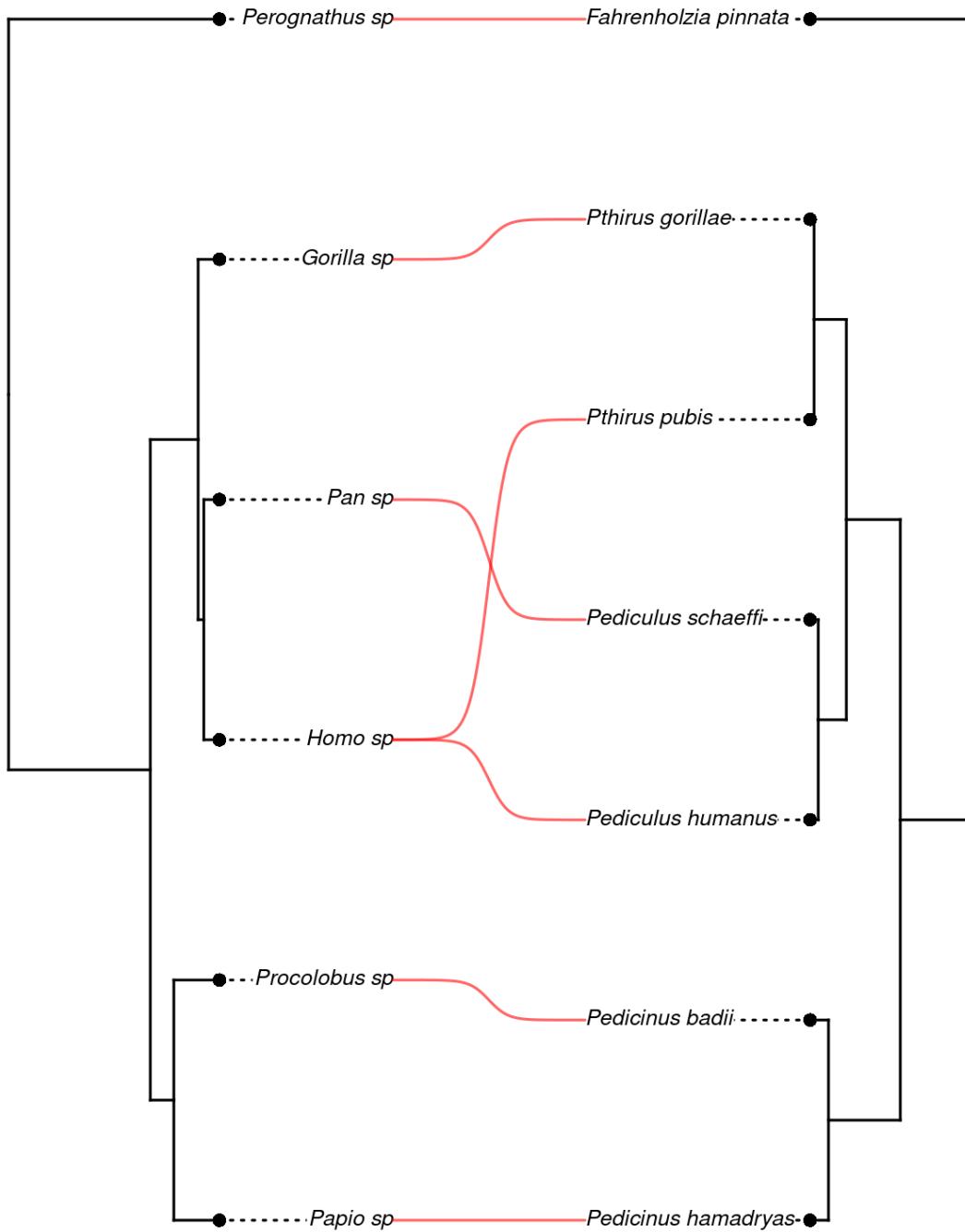


Figure SI6. Phylogenetic trees that correspond to the empirical data of Mammals and Lice (*Pediculus* spp. and *Pthirus* spp.) corresponding to ID. 2 extracted from Reed et al. 2007.

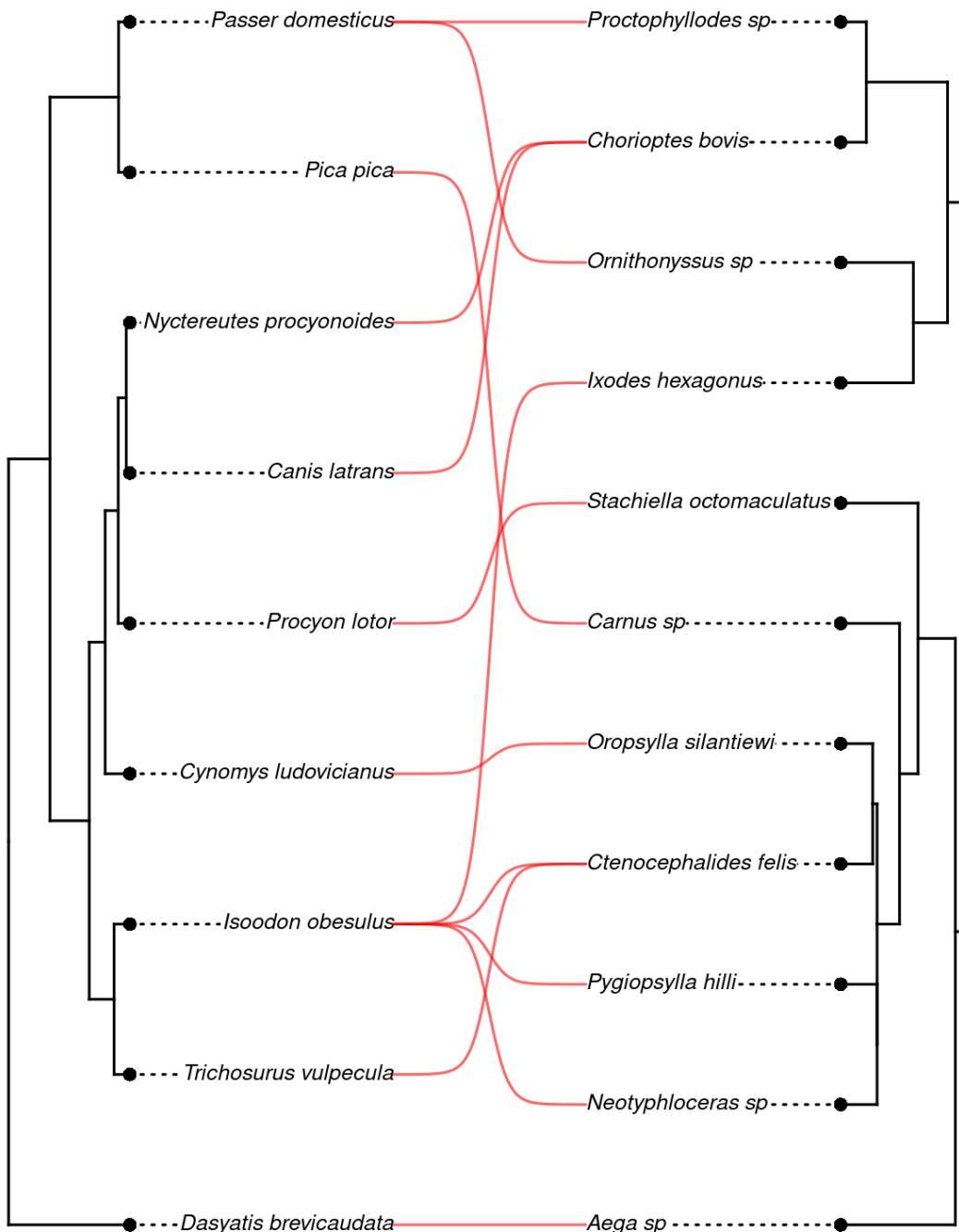


Figure SI7. Phylogenetic trees that correspond to the empirical data of wildlife and arthropods corresponding to ID. 3 extracted from Becker et al. 2018.

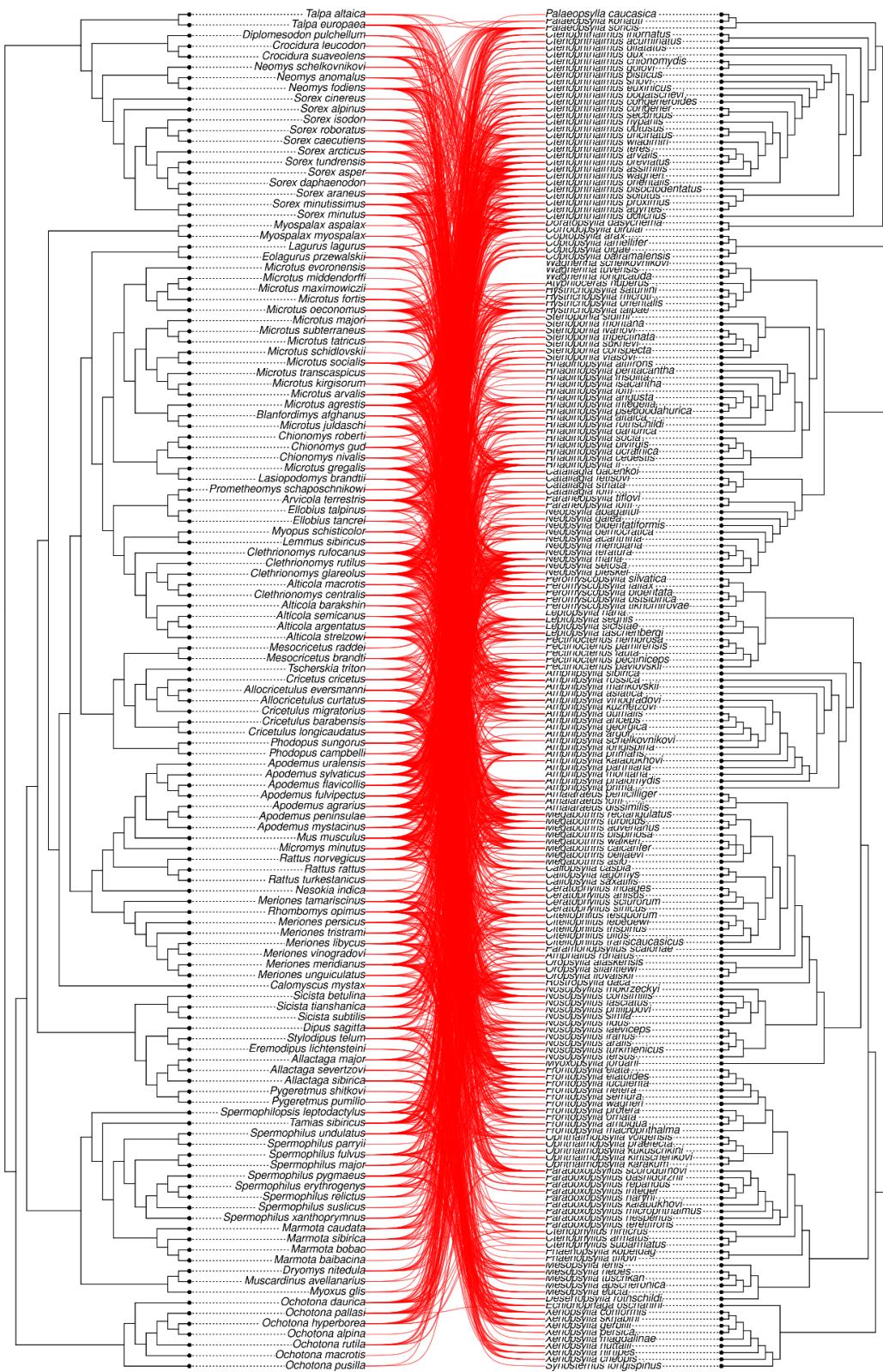


Figure SI8. Phylogenetic trees that correspond to the empirical data of rodents and fleas corresponding to ID. 4 extracted from Krasnov et al. 2016.

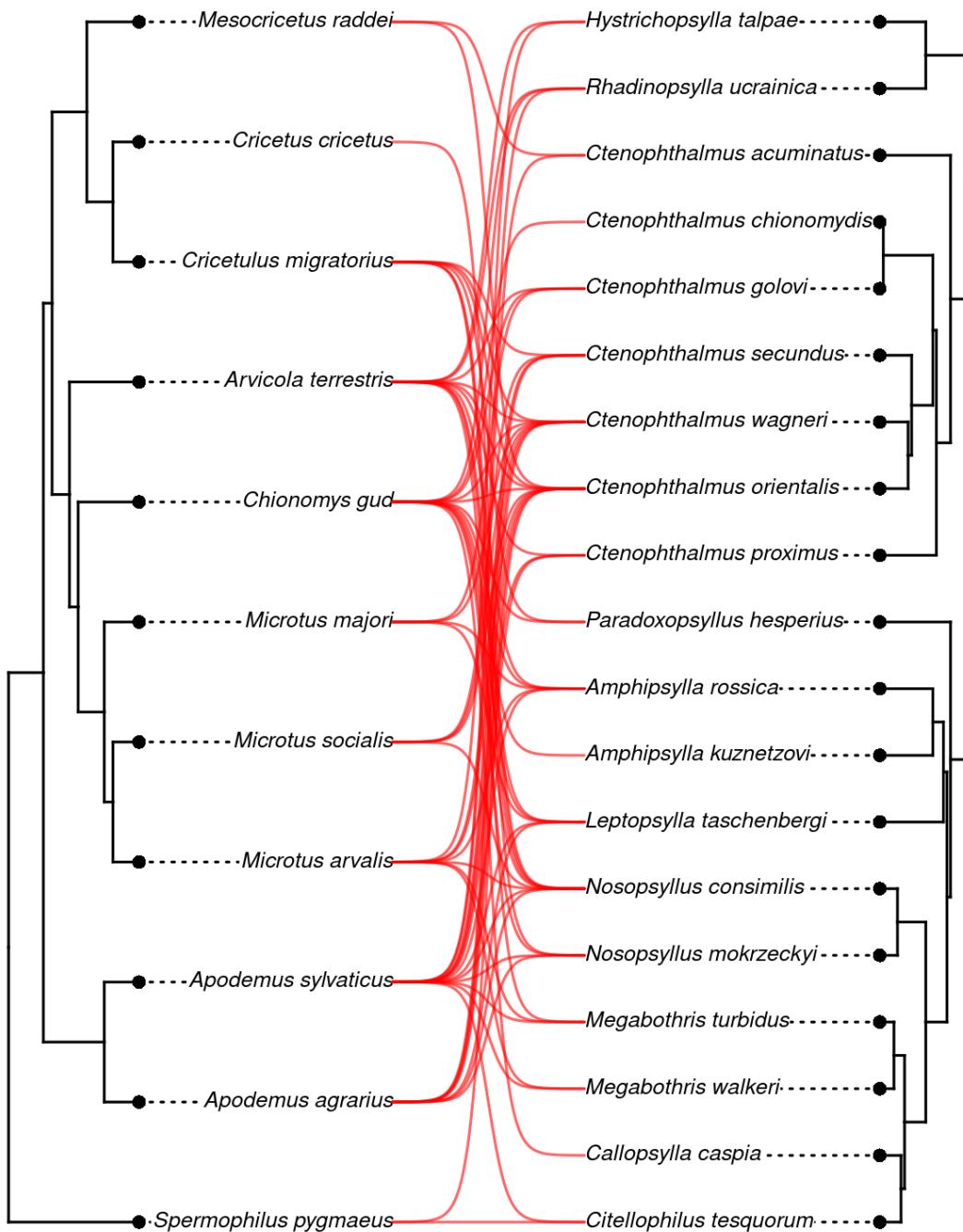


Figure SI9. Phylogenetic trees that correspond to the empirical data of rodents and fleas corresponding to ID. 5 extracted from Krasnov et al. 2016.

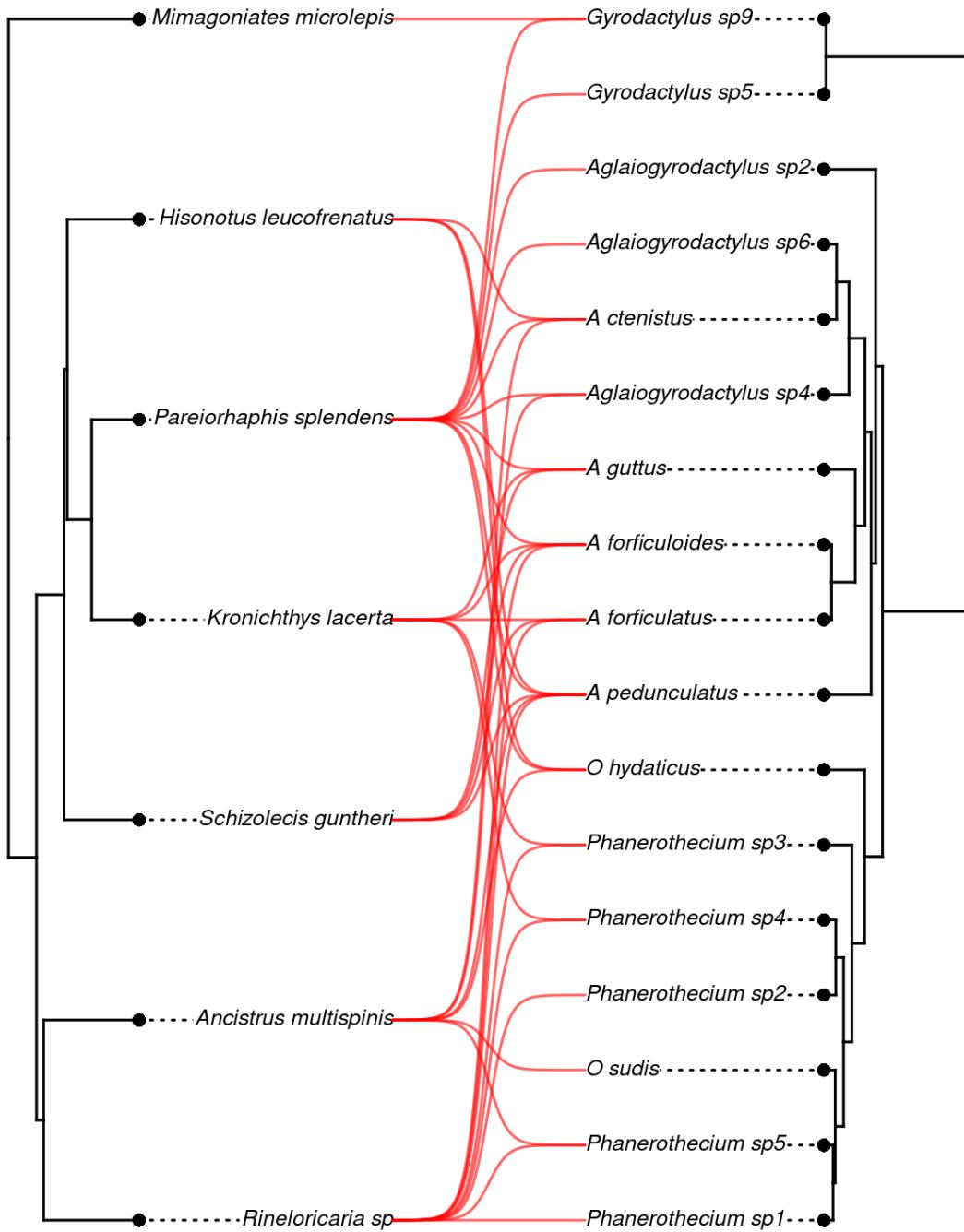


Figure SI10. Phylogenetic trees that correspond to the empirical data of fish and platyhelminthes (Gyrodactylidae) corresponding to ID. 6 extracted from Patella et al. 2017.

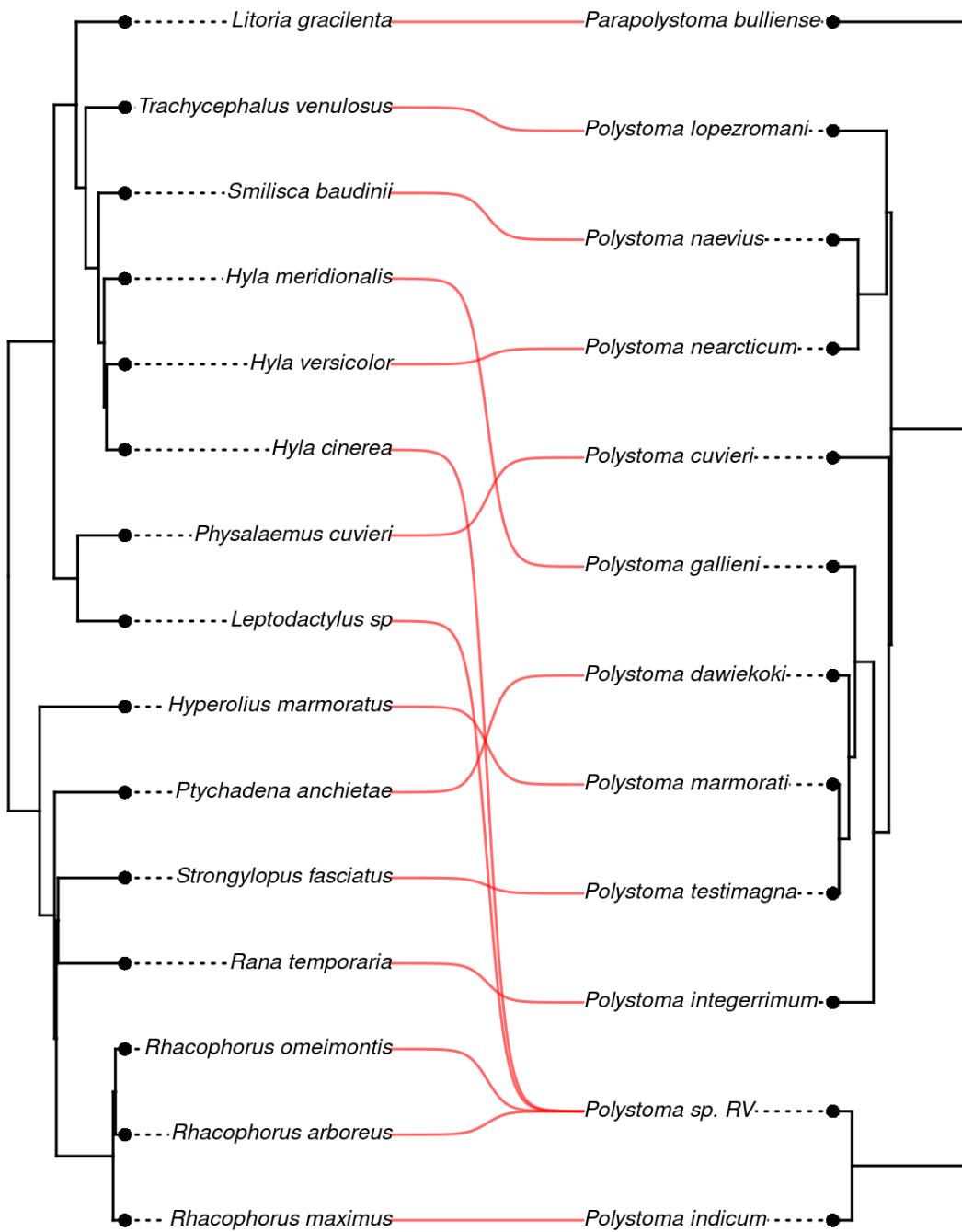


Figure SI11. Phylogenetic trees that correspond to the empirical data of frogs and polystomes (Polystomatidae) corresponding to ID. 7 extracted from Badets et al. 2011.

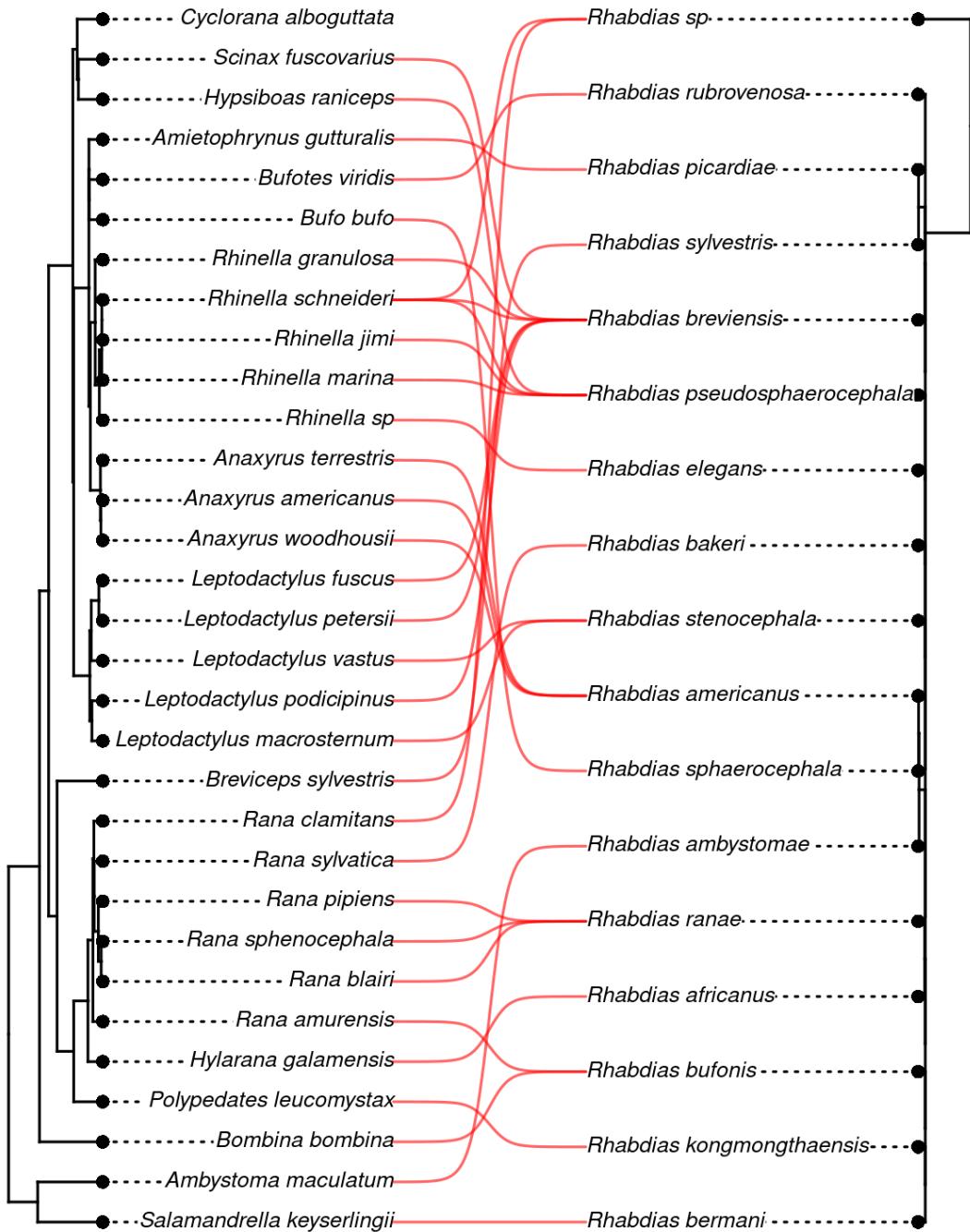


Figure SI12. Phylogenetic trees that correspond to the empirical data of frogs and Nematodes (*Rhabdias* spp.) corresponding to ID. 8 extracted from Müller et al. 2018.

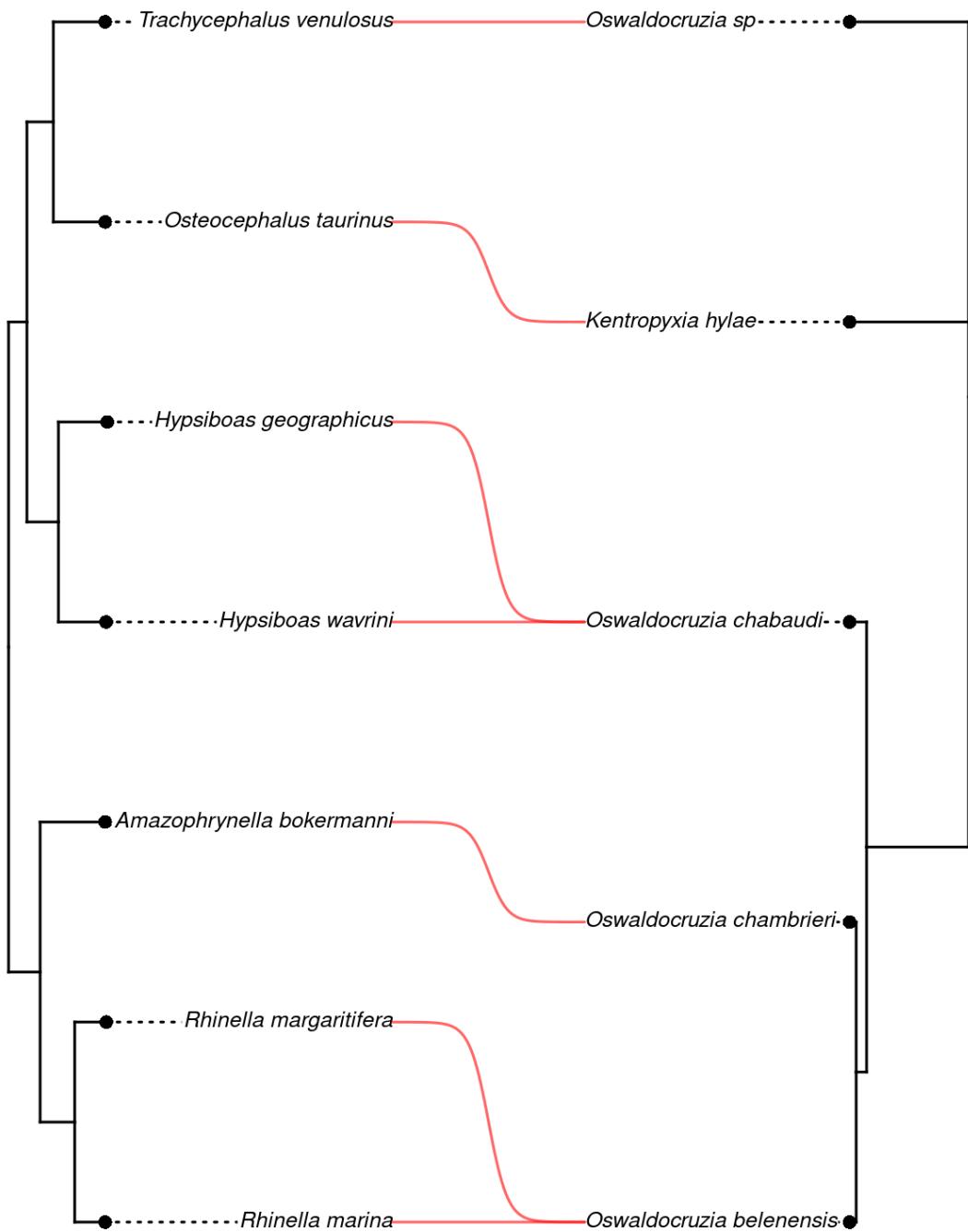


Figure SI13. Phylogenetic trees that correspond to the empirical data of frogs and Nematodes (*Oswaldocruzia* spp.) corresponding to ID. 8 extracted from Willkens et al. 2021.

S2 Details results

S2.1 Relationship between variation in parasite composition and intensity of host-switching for the nine empirical parasite communities

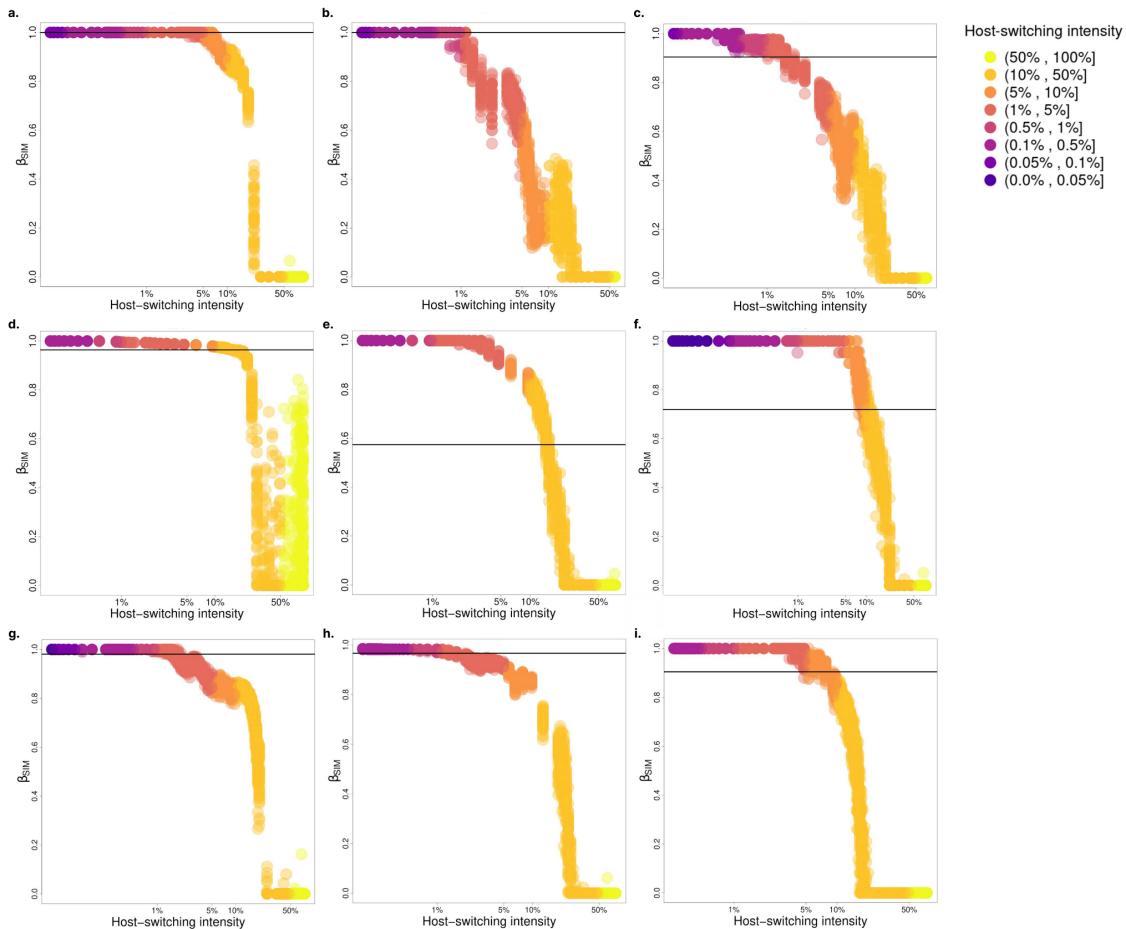


Figure SI14. Relationship between variation in parasite composition (measured by the metric beta diversity - β_{SIM} - y-axis) and intensity of host-switching (x-axis and colours) for the nine empirical communities. Each plot represents an empirical community, (a) ID. 1 - Birds and feather mites, (b) ID. 2 - Mammals and lice, (c) ID. 3 - Wildlife and ectoparasites, (d) and (e) ID. 4 and 5 - Rodents and fleas, (f) ID. 6 - Fish and Monogeneans (Gyrodactylidae), (g) ID. 7 - Frogs and monogeneans (Polystomatidae), (h) ID. 8 - Frogs and lungworms (*Rhabdias* spp.) and (i) ID. 9 - Frogs and gut worms (*Oswaldocruzia* spp.). Empirical values: the lines refer to empirical data on the parasites (continuous) and hosts (dotted). The predicted values of host-switching intensity are the (continuous) and hosts (dotted). The predicted values of host-switching intensity are the

ones which are closer to the crossed continuous line. Colour scales represent each expected percentage range of host-switching intensity. A total of 50 replications were performed for each host-switching intensity.

S2.2 Relationship between tree imbalance and intensity of host-switching for the nine empirical parasite communities

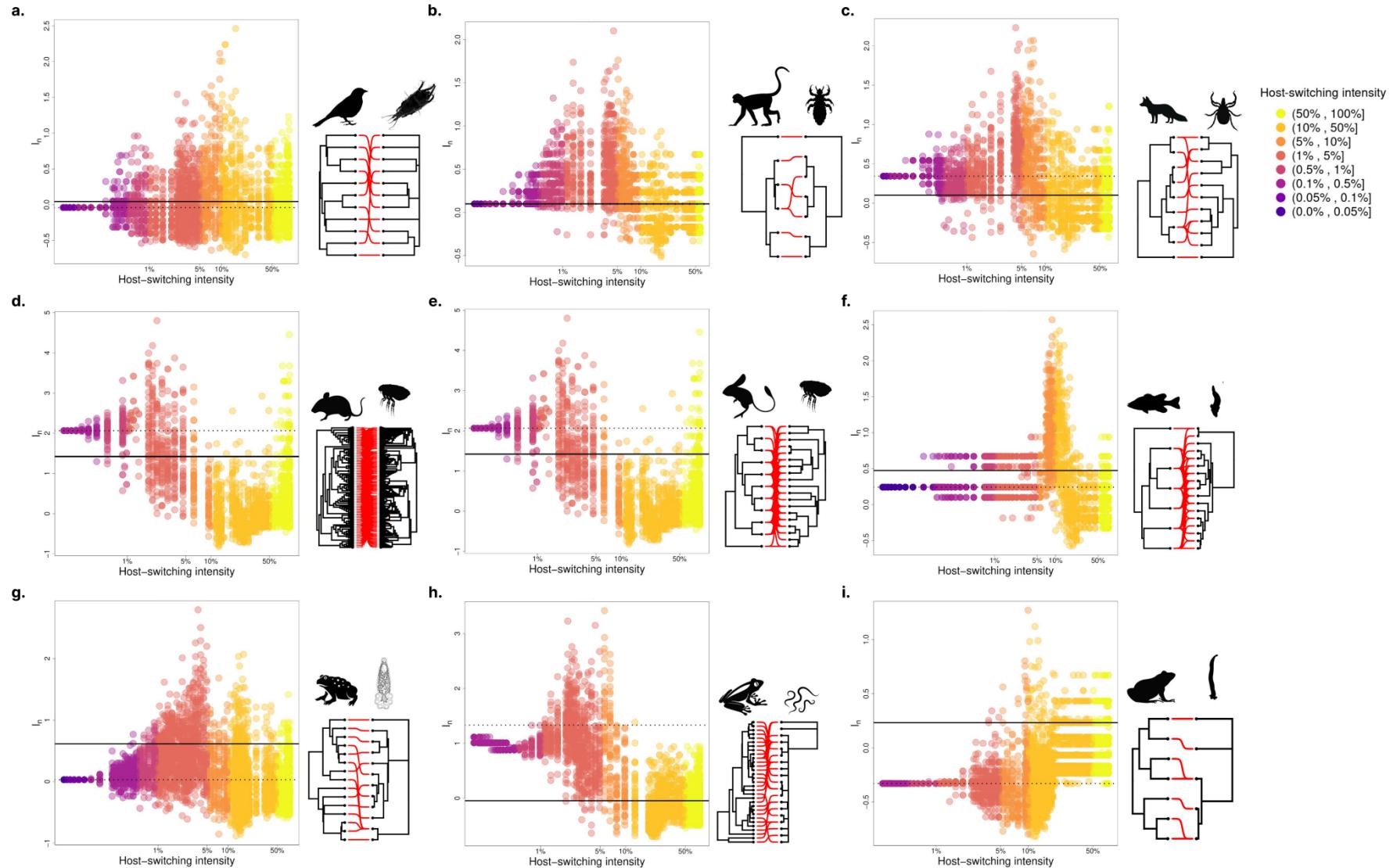


Figure SI15. Relationship between tree imbalance (measured by the metric normalised Sackin index - I_n - y-axis) and intensity of host-switching (x-axis and colours) for the nine empirical communities. Each plot represents an empirical community, (a) ID. 1 - Birds and feather mites, (b) ID. 2 - Mammals and lice, (c) ID. 3 - Wildlife and ectoparasites, (d) and (e) ID. 4 and 5 - Rodents and fleas, (f) ID. 6 - Fish and Monogeneans (Gyrodactylidae), (g) ID. 7 - Frogs and monogeneans (Polystomatidae), (h) ID. 8 - Frogs and lungworms (*Rhabdias* spp.) and (i) ID. 9 - Frogs and gut worms (*Oswaldocruzia* spp.). Empirical values: the lines refer to empirical data on the parasites (continuous) and hosts (dotted). The predicted values of host-switching intensity are the ones which are closer to the crossed continuous line. Colour scales represent each expected percentage range of host-switching intensity. A total of 50 replications were performed for each host-switching intensity.

S2.3 Relationship between parasite richness and intensity of host-switching for the nine empirical parasite communities

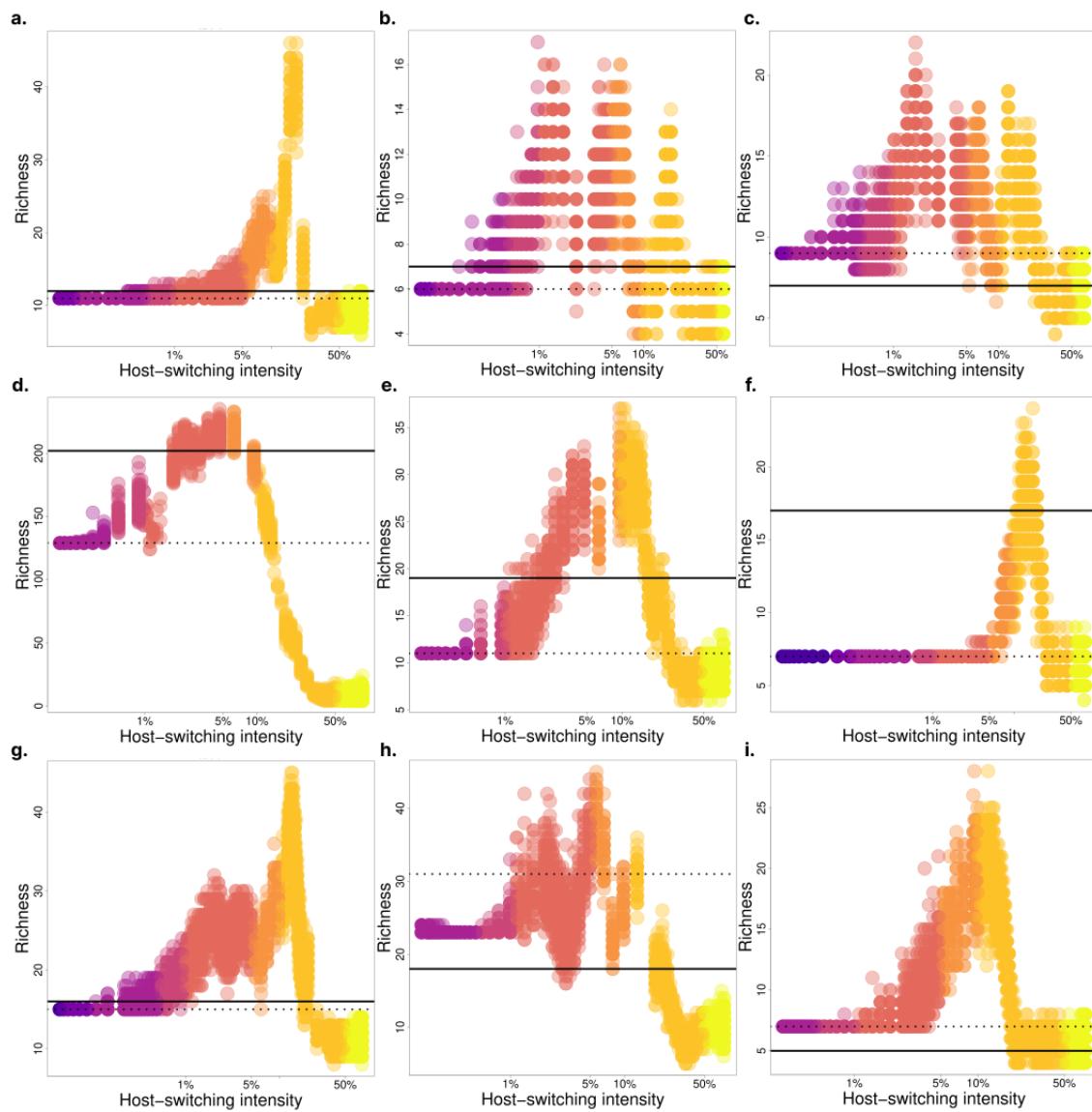


Figure SI16. The relationship between parasite richness and intensity of host-switching (x-axis and colours) for the nine empirical communities. Each plot represents an empirical community, (a) ID. 1 - Birds and feather mites, (b) ID. 2 - Mammals and lice, (c) ID. 3 - Wildlife and ectoparasites, (d) and (e) ID. 4 and 5 - Rodents and fleas, (f) ID. 6 - Fish and Monogeneans (Gyrodactylidae), (g) ID. 7 - Frogs and monogeneans (Polystomatidae), (h) ID. 8 - Frogs and lungworms (*Rhabdias* spp.) and (i) ID. 9 - Frogs and gut worms (*Oswaldocruzia* spp.). Empirical values: the lines refer to empirical data

on the parasites (continuous) and hosts (dotted). The predicted values of host-switching intensity are the ones which are closer to the crossed continuous line. Colour scales represent each expected percentage range of host-switching intensity. A total of 50 replications were performed for each host-switching intensity.

S2.4 Example ecological and evolutionary patterns over time

Here we show the complete temporal results (Fig. S17 and S18) for the empirical community of the feather mites associated with birds to explain the ecological and evolutionary trajectories over time. In the empirical community of feather mites associated with birds (ID. 1), all host species arose only in evolutionary time 1500. Only after this time, do resource limitations become more evident and interactions start to become more restricted. In this case, as hosts diversify, they arrive at an end time (1812) with an ecological and evolutionary pattern similar to the interactions in the empirical study. In the other cases, we observe similar patterns. During the dynamics of the simulation, we can observe host-switching, duplication and extinction events.

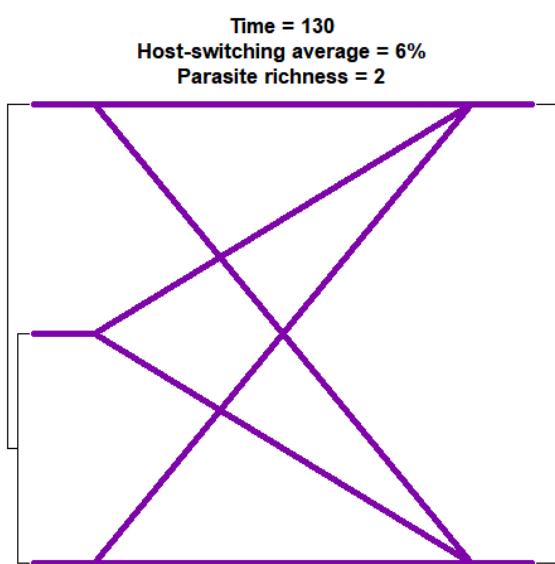


Figure SI17. Complete temporal results (GIF) show the patterns of host-switching over evolutionary time for feather mites associated with birds (Empirical community ID. 1).

Left side - host phylogeny
Right side - parasite phylogeny

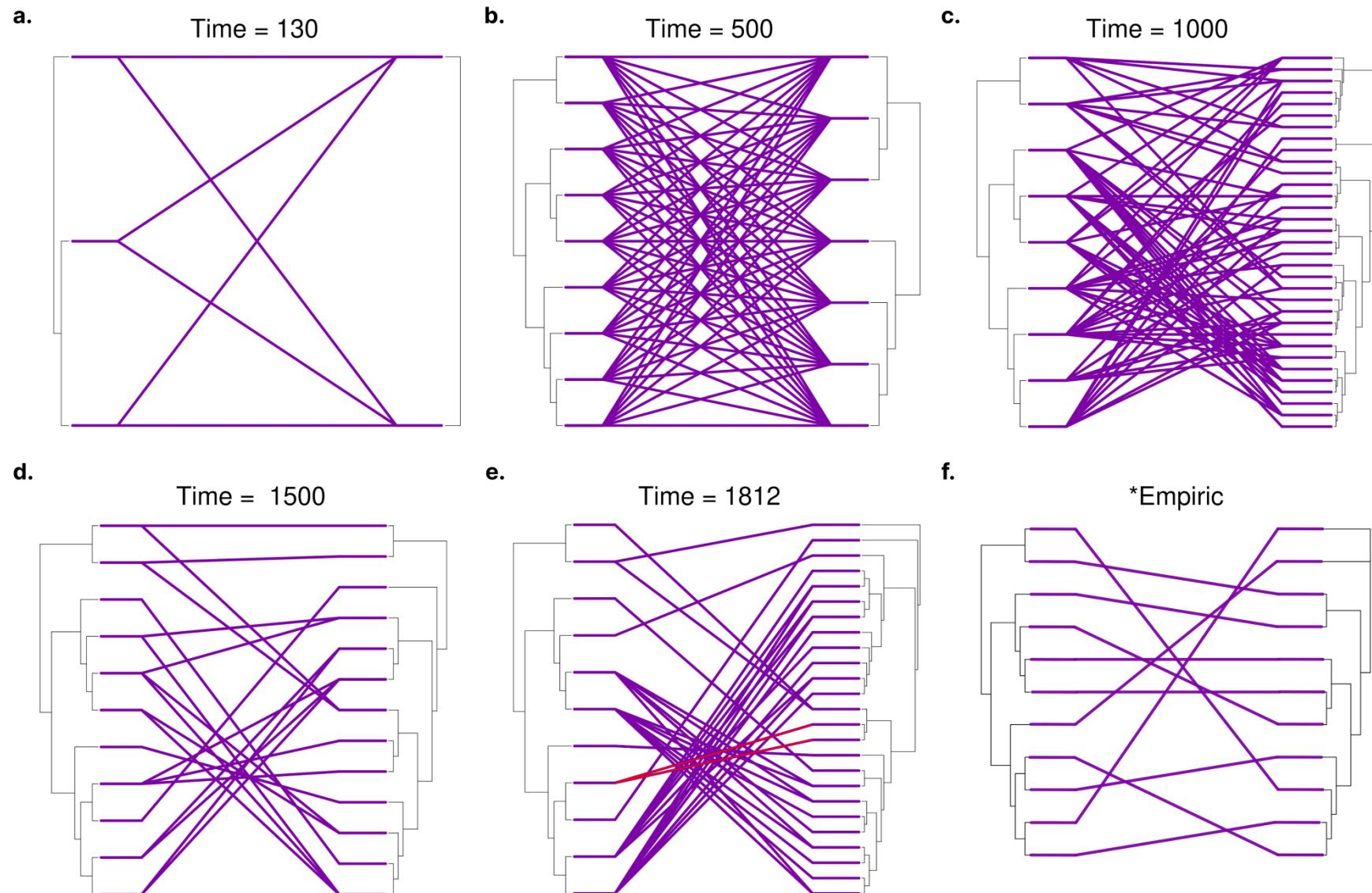


Figure SI18. Ecological and evolutionary patterns of host and parasite to empirical study (case ID1 - feather mites associated with birds)

over time (i-vi) with host-switching intensity between 1% - 5%. The end time has similar structures with empirical community ID 1.

*empirical community = phylogenies and host-parasite interaction of ID 1. The red lines in figure e show a case of duplication (equal: speciation intrahost species or sympatry).

In our model, the host-switches promote host repertoire oscillation, as hypothesised by Janz and Nylin (2008) and favour parasites to speciate (see the dynamics in the movie available in S17 and Fig. SI18, and speciation in the same host species in Fig. SI18e). Although the mean argument behind the difference in speciation rate between host and parasite is the parasite's shorter life cycle, we support the idea that the use of different hosts (hence, resources) may represent another important mechanism to parasite diversification (Boeger et al. 2022).

S3 Others exploratory analysis

S3.1 Influence of host-switching events on the ecological and evolutionary patterns of simulated parasites with $q_{min}=2q_0$

In this case with $q_{min}=2q_0$, similarly to the main results, we observed that the turnover decreases as the host-switching intensity increases (SI19a) and for each host-switching intensity value, there is a small variation in the turnover. On the other hand, the I_n metric is irrelevant, because all intensities of host-switching generate parasite phylogenies with very different structures. In the main results, we fixed $q_{min}=0.5*q_0$ in order not to lose the possibility of cospeciation.

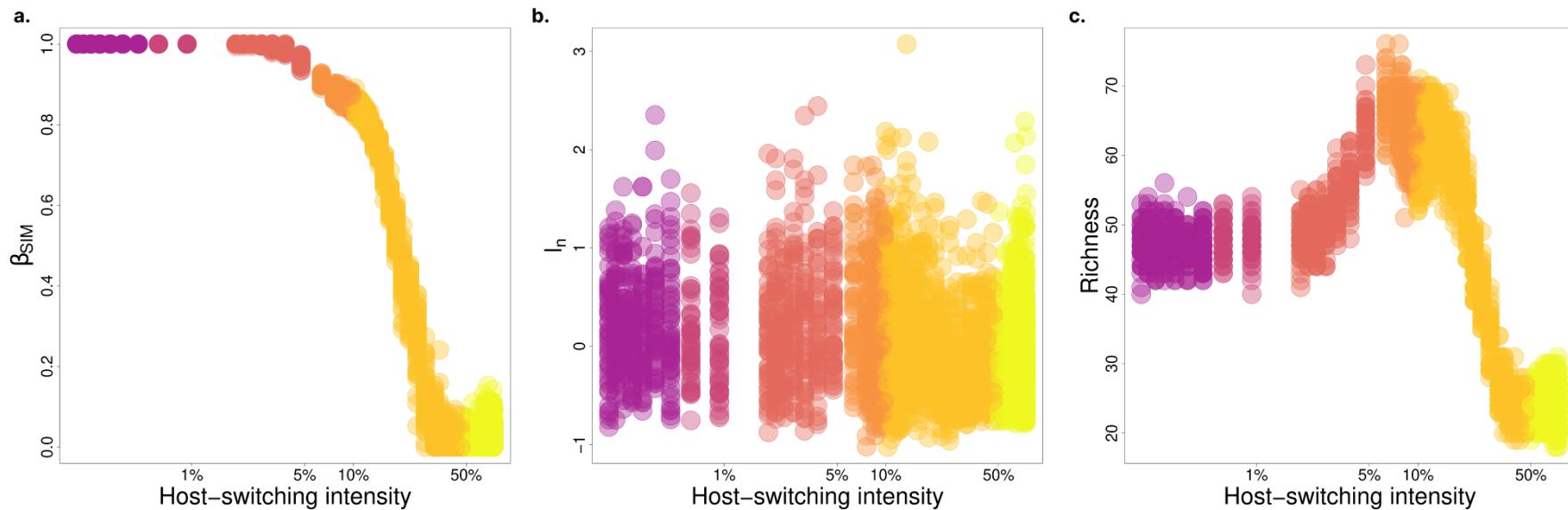


FIGURE SI19. Influence of host-switching events on the ecological and evolutionary patterns of simulated parasites for fleas associated with rodents (ID. 5) with $q_{min} = 2q_0$. Here we demonstrate the relationship between: a. Host-switching intensity and turnover of parasite species (β_{SIM}) between host species; b. Host-switching intensity and parasite normalised Sackin index (I_n); c. Host-switching intensity and parasite richness. The scaling of the x-axis is in log scale. A total of 50 replicates were performed for each host-switching intensity.

S3.2 Different shapes of host phylogeny (hypothetical host phylogenies) - Influence of host-switching events on the ecological and evolutionary patterns of simulated parasites

To understand the effect of the shape of the phylogeny of the hosts on the ecological and evolutionary patterns of the parasites, we varied the phylogenies of the hosts in terms of balance and branch length following four possibilities: stemmy unbalanced; stemmy balanced; tippy unbalanced; and tippy balanced (FigS20). In stemmy trees, the parasite species are specialists regarding the use of the resource/host (high values of beta turnover) for higher values of host-switching when compared to a tippy tree. On the other hand, the effect of host-switching on the imbalance of the parasite tree is more sensitive to the imbalance of the host tree. For all formats of hypothetical phylogenies, parasites' phylogeny tends to zero for high values of host-switching intensity.

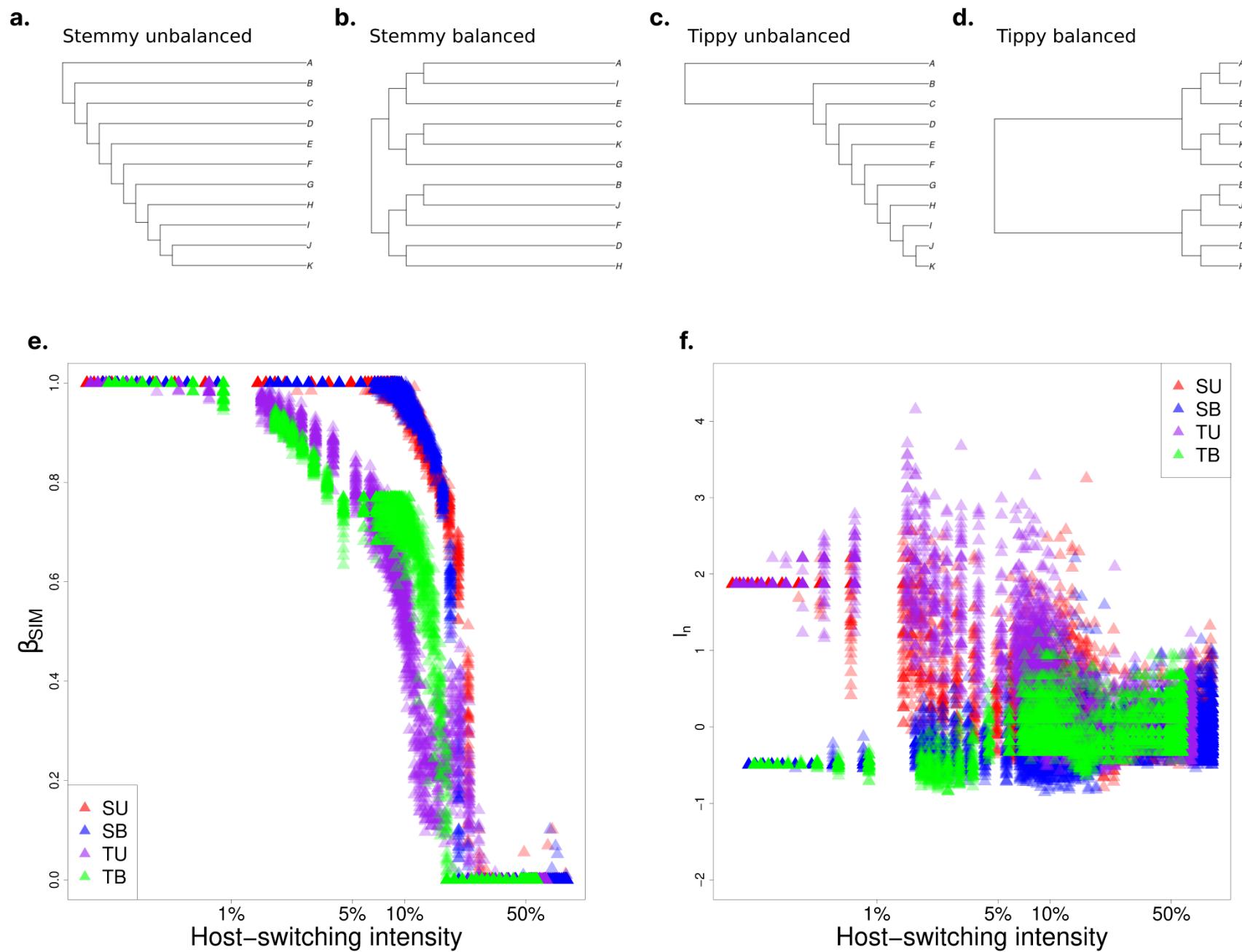


FIGURE SI20. Influence of host-switching events on the ecological and evolutionary patterns of simulated for hypothetical host phylogenies: **a.** Stemmy unbalanced, **b.** Stemmy balanced, **c.** Tippy unbalanced, **d.** Tippy balanced. Here we show the relationship between: **e.** Host-switching intensity and turnover of parasite species (β_{SIM}) between hypothetical host species; **f.** Host-switching intensity and parasite normalised Sackin index (I_n); The colours represent each shape of hypothetical host phylogenies: red - Stemmy unbalanced (SU), blue - Stemmy balanced (SB), purple - Tippy unbalanced (TU), green - Tippy balanced (TB). A total of 50 replicates were performed with a carrying capacity of 250 individuals for each host-switching intensity.

REFERENCES

- Higgs P. G., & Derrida B. 1991. Stochastic models for species formation in evolving populations. *Journal of Physics A: Mathematical and General*, 24(17), L985.
- Manzo F., & Peliti L. 1994. Geographic speciation in the Derrida-Higgs model of species formation. *Journal of Physics A: Mathematical and General*, 27(21), 7079.
- [Dataset] Badets M., Whittington I., Lalubin F., Allienne J. F., Maspimby J. L., Bentz S. Verneau O. 2011. Correlating early evolution of parasitic platyhelminths to Gondwana breakup. *Systematic Biology*, 60, 762-781. Accession number: EDB_10.1093/sysbio/syr078
- [Dataset] Becker D. J., Streicker D. G. Altizer S. 2018. Using host species traits to understand the consequences of resource provisioning for host–parasite interactions. *Journal of Animal Ecology*, 87, 511-525. Accession number: EDB_10.1111/1365-2656.12765
- [Dataset] Doña J., Sweet A. D., Johnson K. P., Serrano D., Mironov S. Jovani R. 2017. Cophylogenetic analyses reveal extensive host-shift speciation in a highly specialized and host-specific symbiont system. *Molecular phylogenetics and evolution*, 115, 190-196. Accession number: EDB_10.1016/j.ympev.2017.08.005
- [Dataset] Krasnov B. R., Shenbrot G. I., Khokhlova I. S. Degen A. A. 2016. Trait-based and phylogenetic associations between parasites and their hosts: a case study with small mammals and fleas in the Palearctic. *Oikos*, 125, 29-38. Accession number: EDB_10.1111/oik.02178

[Dataset] Müller M. I., Morais D. H., Costa-Silva G. J., Aguiar A., Avila R.W. da Silva R. J. 2018. Diversity in the genus *Rhabdias* (Nematoda, Rhabdiasidae): Evidence for cryptic speciation. *Zoologica Scripta*, 47, 595-607. Accession number: EDB_10.1111/zsc.12304

[Dataset] Patella L., Brooks D. R. Boeger W. A. 2017. Phylogeny and ecology illuminate the evolution of associations under the Stockholm paradigm: *Aglaioxyrodactylus* spp. (Platyhelminthes, Monogenoidea, Gyrodactylidae) and species of Loricariidae (Actinopterygii, Siluriformes). *Vie Et Milieu*, 67, 91-102. Accession number: EDB_20183235432

[Dataset] Reed D. L., Light J. E., Allen J. M. Kirchman J. J. 2007. Pair of lice lost or parasites regained: the evolutionary history of anthropoid primate lice. *Bmc Biology*, 5, 1-11. Accession number: EDB_10.1186/1741-7007-5-7

[Dataset] Willkens Y., Furtado A. P., Dos Santos J. N. de Vasconcelos Melo F. T. 2021. Do host habitat use and cospeciation matter in the evolution of *Oswaldoocruzia* (Nematoda, Moleneidae) from neotropical amphibians? *Journal of Helminthology*, 95, e33. Accession number: EDB_10.1017/S0022149X21000250