

# Vestigial preference functions in neural networks and túngara frogs.

S. M. Phelps\*<sup>†</sup>, M. J. Ryan\*<sup>‡</sup>, and A. S. Rand<sup>†</sup>

<sup>†</sup>Section of Integrative Biology C0930, University of Texas, Austin, TX 78712; and <sup>‡</sup>Smithsonian Tropical Research Institute, Apartado 2074, Balboa, Panama

Edited by Peter Marler, University of California, Davis, CA, and approved September 7, 2001 (received for review June 12, 2001)

**Although there is a growing interest in understanding how perceptual mechanisms influence behavioral evolution, few studies have addressed how perception itself is shaped by evolutionary forces. We used a combination of artificial neural network models and behavioral experiments to investigate how evolutionary history influenced the perceptual processes used in mate choice by female túngara frogs. We manipulated the evolutionary history of artificial neural network models and observed an emergent bias toward calls resembling known ancestral states. We then probed female túngara frogs for similar preferences, finding strong biases toward stimuli that resemble a call hypothesized for a recent ancestor. The data strongly suggest that female túngara frogs exhibit vestigial preferences for ancestral calls, and provide a general strategy for exploring the role of historical contingency in perceptual biases.**

Early ethologists reported that animal signals evolved toward simplicity, specificity, and salience—hallmarks of the minimal stimulus, the “sign stimulus,” that was required to evoke a response from its receiver (1, 2). This observation suggested that signal form was being shaped by the perceptual mechanisms of the receiver, a view that has been rekindled by recent work in sexual selection and sensory ecology (3–7; reviewed in 8–10). The renewed interest in proximate causes of behaviors has prompted a number of workers to return to the methodical titration of receiver decision mechanisms used in classic ethology (e.g., refs. 2, 11, and 12; more recently, refs. 13–17); the resulting generalization gradients (or preference functions, as they are known in mate choice) are thought to strongly affect the fitness of individual signalers (18). Although the shapes of generalization gradients are presumed to influence signal evolution, and to be of interest in their own right, few ethologists have addressed the forces that determine these shapes (but see refs. 19–26).

A number of groups have begun to use artificial neural network models to investigate the evolution of perceptual mechanisms (24–31). Because neural network models distribute the representation of a signal across many “neurons,” these models often generalize as an automatic result of training (32, 33), making them useful tools for the exploration of preference functions. In recent studies, we evolved artificial neural networks along distinct evolutionary trajectories and found that their emergent responses to novel signals were strongly shaped by their selection histories (31, 34, 35). Moreover, those networks with a history approximating that of our focal species, the túngara frog, were better at reproducing female responses to test stimuli. These findings and others (36, 37) indicate that female túngara frogs may exhibit preferences that are remnants of past selection for species recognition. They do not, however, suggest what form vestigial preferences might take.

We suggest that the evolutionary persistence of ancestral recognition mechanisms contributes to current preferences, and that this contribution biases preference functions in favor of ancestor-like signals. To test this hypothesis, we manipulate the histories of artificial neural networks and assess their emergent preference functions. We probe the response biases of networks with stimuli that vary in similarity to the túngara frog call along an axis that passes near the most recent ancestral call of the

networks. We test the empirical validity of the network findings by comparing the patterns of vestigial preference in neural networks to female responses on a similar transect of calls.

We begin the neural network simulations by assigning networks one of two history types (Fig. 1; see also refs. 31, 34, and 35). The first is a “mimetic history,” which roughly mimics the history hypothesized for female túngara frogs. Mimetic history networks are selected to recognize each of four calls in a sequence beginning with the root call reconstructed for the *Physalaemus pustulosus* species group, ascending through a series of nodes hypothesized for túngara frog call evolution, and ending with the call of the túngara frog. The second is a “mirrored history,” which precisely mirrors the trajectory of calls in the mimetic history, but begins in a different region of acoustic space. Networks of both history types are required to recognize the túngara frog call, but are derived from populations selected to recognize distinct ancestral calls. The mimetic history approximates the reconstructed history of túngara frogs; the mirrored history controls for the diversity of calls the networks have been selected to recognize, but has no natural counterpart (ref. 31; reviewed in refs. 34, 35, and 38).

Conventional approaches to unimodal generalization gradients and preference functions predict that, as a stimulus becomes less like the conspecific signal, receiver responses will decline without respect to the region of stimulus space the novel stimulus occupies (39–41); the vestigial-preference hypothesis, however, predicts that responses will persist in the vicinity of ancestral stimuli. Because the ancestral calls of mimetic and mirrored histories differ reliably in one call character—the whine shape, or time to half frequency (THHz, Fig. 1*b*)—the vestigial-preference hypothesis predicts mutually exclusive response biases for networks of the two history types, permitting a strong test of the hypothesis (see Figs. 2 and 3*a*). When presented with calls that deviate from the túngara THHz, networks should respond most to stimuli that resemble ancestral calls: mimetic history ancestors have a short THHz, mirrored history ancestors a long THHz.

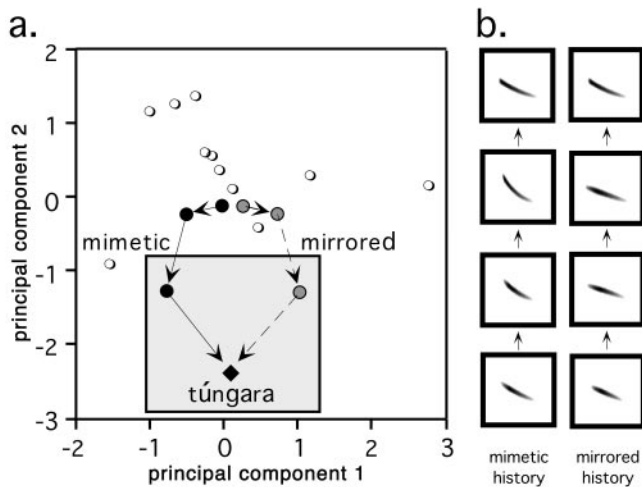
To test the vestigial-preference hypothesis with female túngara frogs, we constructed a set of six calls that varied along an axis passing through the túngara frog call and the reconstructed call of a recent ancestor (Fig. 4*a*). For convenience, we refer to the three stimuli resembling the hypothesized ancestral state as “ancestral” calls, and the remaining stimuli as “anti-ancestral” calls. If female túngara frogs exhibit the vestigial preferences predicted by the neural network models, subjects should prefer calls resembling a hypothesized ancestor to anti-ancestral calls of matched similarity. Such a demonstration would suggest that female túngara frogs retain vestigial preferences, and would

This paper was submitted directly (Track II) to the PNAS office.

Abbreviation: THHz, time to half frequency.

\*To whom reprint requests should be addressed at the present address: Center for Behavioral Neuroscience, Emory University, 954 Gatewood Road NE, Atlanta, GA 30329. E-mail: sphelps@rmy.emory.edu.

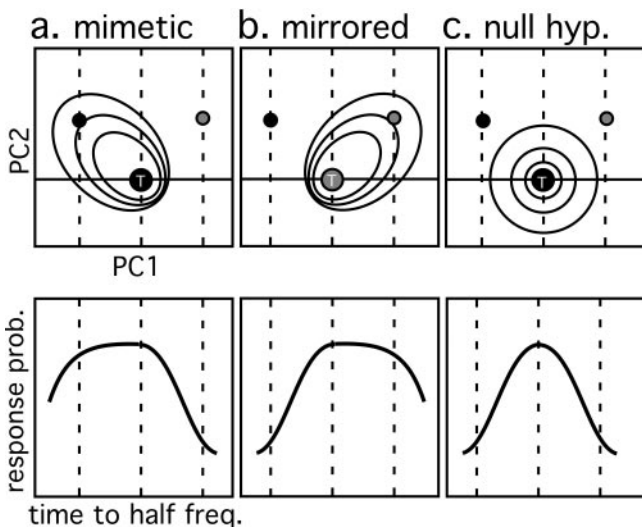
The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. §1734 solely to indicate this fact.



**Fig. 1.** (a) Mimetic and mirrored histories in acoustic space defined by two dimensions of a principal components analysis (see refs. 31 and 36). Open circles represent the calls of extant and “ancestral” taxa. The shaded area is used in subsequent figures. (b) Sonograms of signals used in mimetic and mirrored histories. Root calls are at bottom of sequence, túngara at top. Each sonogram is 600 msec long (x axis), and spans from 0 to 1.5 kHz (y axis).

provide a simple means of probing comparable preferences in other taxa.

By combining artificial neural network models with behavioral experiments, we are able to test the vestigial-preference hypothesis on receivers with a known and controlled history (neural networks). We assess the networks’ external validity by predicting responses in a model species whose history has been reconstructed based on available phylogenetic data. The two approaches provide complementary data that, if concordant, provide a much more compelling view of historical influences than either approach could provide alone.



**Fig. 2.** Schematic depiction of hypotheses for network preference functions. (a) Mimetic history network exhibiting vestigial preferences. (b) Mirrored history network exhibiting vestigial preferences. (c) Null hypothesis of no historical influence for both mimetic and mirrored history networks. The upper row depicts isolines of preferences in the principal-component space given in Fig. 1. Lower row depicts hypothesized preferences varied in a single dimension, time to half frequency, which loads highly on principal component 1 (PC1).

## Methods

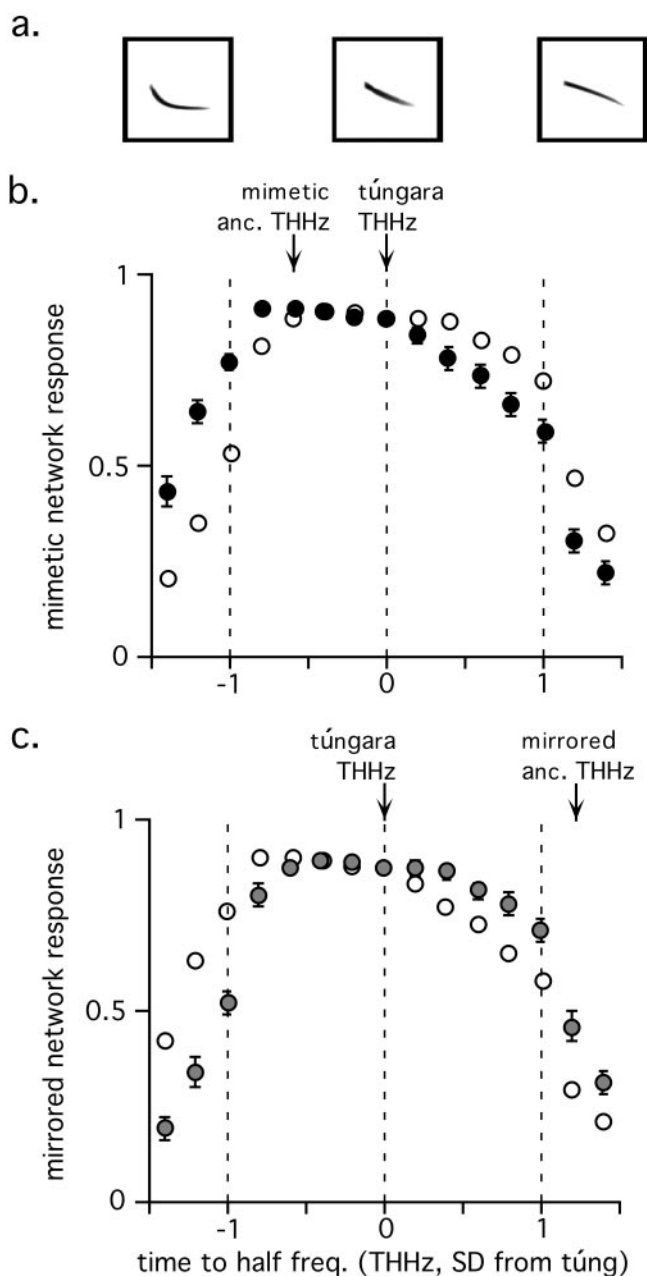
**Call Synthesis.** Call parameters for extant species and reconstructed ancestral states were taken from ref. 36. These authors used a local squared-change parsimony algorithm and branch lengths derived from mitochondrial DNA data to estimate ancestral states (see ref. 37 for this and other algorithms used, and ref. 42 for phylogenetic hypotheses). Estimates based on other phylogenetic models or assumptions yield similar but not identical ancestral states, and consistently suggest a role for phylogenetic history in receiver biases (37). Because we predict enhanced female responses in the vicinity of the ancestral state, it is not necessary for the reconstructed ancestor to agree precisely with a true (but unknowable) ancestor. Inaccuracies in ancestral reconstruction will reduce the proximity of ancestral test stimuli to the true ancestor, and should lead to conservative estimates of vestigial preference. We confirm this result in the neural network models by constructing test stimuli that do not pass through the ancestors of either history type, but that differ in their proximity to the ancestors of each history type.

The túngara frog call is an exponential, descending frequency sweep known as a whine. The shape of the whine is quantified by the time from the call’s onset to its mid-frequency, its time to half frequency (THHz). Network test stimuli were generated by manipulating the whine shape but leaving all other call parameters matched to the call of the túngara frog. THHz varied between 1.4 standard deviations above and below the túngara frog call. (Standard deviations are measured for the variation in THHz of the *Physalaemus pustulosus* species group and three outgroup taxa.) Calls tested on neural networks were processed as previously described (30, 31).

To test female responses, we first synthesized an anti-ancestral call that was as similar to the túngara frog call as was the most recent ancestor, while being as different from the ancestor as possible. Ancestral, anti-ancestral, and túngara frog calls lie along a single line in the multidimensional space defined by our call descriptors: We operationally define the ancestor and anti-ancestor calls to lie at positions +1.0 and –1.0 along this axis (Fig. 4a). To determine the call characters of the anti-ancestor, we (i) took each value for the parameters needed to synthesize a túngara frog call; (ii) subtracted the appropriate call parameter used to define the ancestor; and (iii) added this difference onto the túngara frog call. (The túngara frog whine, for example, has a typical starting frequency of 884 Hz, the ancestral reconstruction 1050 Hz. The anti-ancestral call therefore has a starting frequency of  $(884 - 1050) + 884 = 718$  Hz. Descriptions of all six call variables and how they were measured are provided in ref. 37.) Variations on this procedure led to six test stimuli: +1.5, 1.0, 0.5, –0.5, –1.0, and –1.5 calls. The call pairs 1.5/–1.5, 1.0/–1.0 and 0.5/–0.5 represent ancestral/anti-ancestral pairs of matched similarity. Stimuli were synthesized at a sampling rate of 20 kHz by using a sound synthesis program developed by J. Schwartz (Pace University, Pleasantville, NY).

**Network Evolution and Testing.** The neural network architecture and evolutionary training algorithms are described in detail elsewhere (30, 31, 34). Briefly, we used a recurrent neural network architecture made of four layers of neurons. Networks consisted of an input layer with 15 “neurons,” each responding to a unique range of frequencies between 261 and 1565 Hz; a hidden layer that received input directly from the input layer as well as from a second hidden layer; a second hidden layer, or context layer, that received input from the first hidden layer; and an output neuron that received input from the first hidden layer. The activity of the output neuron at the end of a stimulus window defines how well a network responds to a given test stimulus.

By using a genetic algorithm (detailed in refs. 30 and 31), we trained populations of networks to recognize each call along an



**Fig. 3.** Vestigial preference functions in mimetic and mirrored history networks. (a) Sonograms of calls with a time to half frequency (THHz) 1 SD smaller than, equal to, or 1 SD greater than the túngara frog call (left to right). Each sonogram is 600 ms long (x axis), and spans from 0 to 1.5 kHz (y axis). (b) Mimetic history networks (filled circles) exhibit preferences for steep whines resembling their ancestors. Responses of mirrored networks are shown in open circles for comparison. (c) Mirrored history networks (gray circles) exhibit preferences for shallow whines resembling their ancestors. Mimetic network responses are plotted in open circles.

evolutionary trajectory. Each network was represented as a binary list of the weights between its neurons. Each network was then assigned a fitness based on its ability to respond to the target call without responding to noise. Networks contributed “offspring” to the next generation in proportion to their fitness. Offspring were made to differ from parents by “mutating” a bit in the list of weights, and by “recombining” segments of two parent lists, providing variation that could again be subject to selection.

The evolutionary trajectories of mimetic history networks consisted of selecting populations of networks to recognize each call in a sequence beginning at the root call reconstructed for the *P. pustulosus* species group—including the out-group taxa *P. enesefae*, *P. ephippifer*, and an undescribed species we refer to as species A (36, 37). The mimetic trajectory ascends through each subsequent node en route to the call of the túngara frog. The ancestral states for mirrored history networks were constructed by flipping the coordinates of the mimetic ancestors in the acoustic space defined by a principal components analysis (Fig. 1) and solving for the original call variables (see ref. 31 for details).

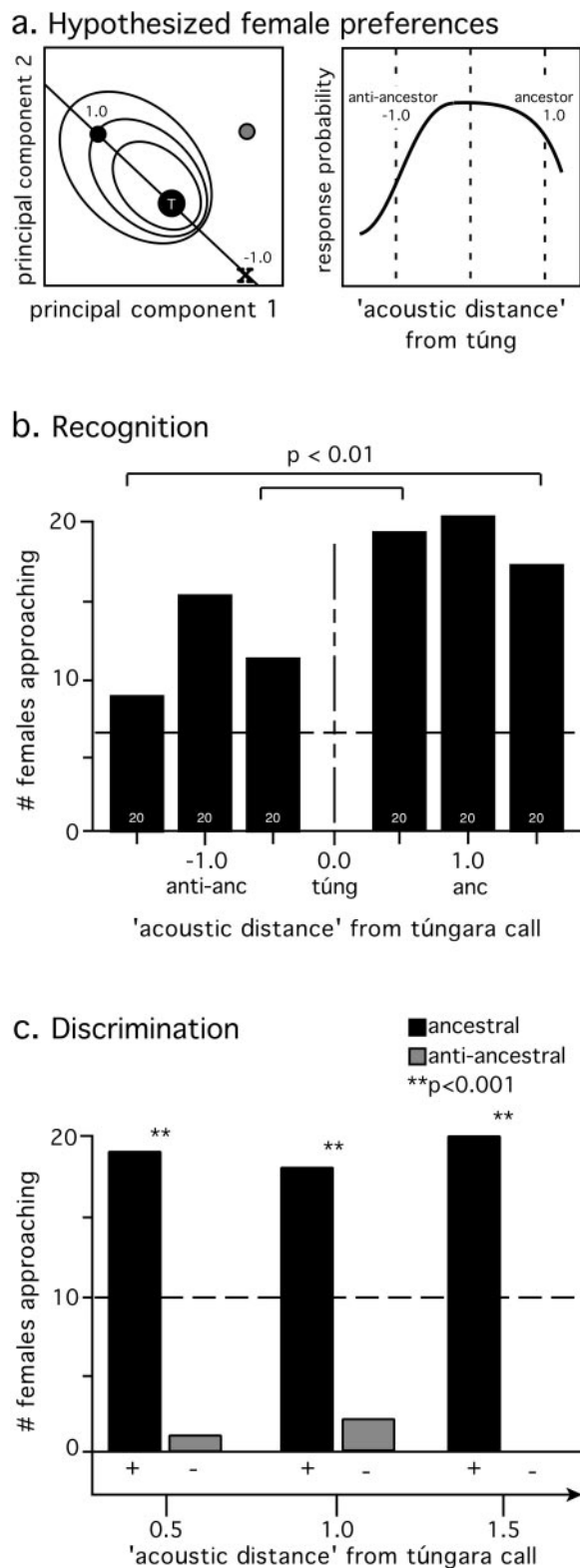
The target call switched to the next stimulus in the historical sequence when the maximum fitness in a population (averaged over five consecutive generations) exceeded 0.90. When this criterion was met for the túngara call, the simulation ended and the architecture of the best network in the population was recorded. Except at high mutation rates or noise levels, this proved to be a more stringent criterion for successful completion than used previously (30, 31). For most simulations, the networks were allowed to evolve until they met this criterion for the last call in the sequence (the túngara frog call), or until a total of 20,000 generations elapsed. In the case of networks with population sizes of greater than 100, the simulation was aborted after 2,000,000 networks had been evaluated: if the population size was 500, the simulations were stopped at 4,000 generations; populations of 1,000 after 2,000 generations. (In practice, only one run from the large population sizes failed to meet this criterion.) If the simulation ended without meeting the criteria for all calls in the evolutionary trajectory, the results were not used in our analysis.

The probability of mutating a given bit in a network was 0.0001, 0.0005, 0.001, 0.005, or 0.01. The probability of recombination between a pair of networks was 0.00, 0.25, 0.50, 0.75, or 1.00. Population sizes varied among the values 10, 50, 100, 500, and 1,000. And finally, “ambient noise” was constructed by randomly permuting each coefficient in the input matrix with a probability of 0.0, 0.02, 0.05, 0.10, or 0.25. When varying any one parameter, the remaining parameters were left as reported in the original studies (refs. 30 and 31; Table 1). Each manipulation was run 10 times for each history type, for a total of 340 replicates of the simulations. Table 1 depicts the precise combinations of parameters used in each simulation, as well as a summary of statistics on the outcomes of the various simulations.

#### Replication of Historical Influences on Neural Network Response Biases.

A neural network’s performance was defined by its ability to predict the responses of female túngara frogs. To determine whether previously reported historical influences (31) were consistent across the many parameters reported here, we investigated the accuracy of current networks on the stimulus set used in prior studies (30, 31). The set comprised 34 stimuli, including the calls of extant species and estimates of ancestral states for the clade. For a given stimulus, we defined the error of a set of networks as the difference between the average network response and the female response recorded in phonotaxis experiments. We took the mean error across the standard 34 novel test stimuli (30, 31) as a measure of network accuracy. We ran 340 simulations—170 of each history type—resulting in 17 combinations of parameters (see Table 1). Of the 17 parametric permutations, 14 resulted in networks of both history types evolving to meet the criteria for successful completion of the simulation (of these 280 simulations initiated, 279 ran to completion, see Table 1). For each of the 14 successful permutations, we compared the average accuracies of networks with mimetic and mirrored histories. We tested the null hypothesis that mimetic and mirrored history networks are equally good at predicting female responses.





**Fig. 4.** (a) Schematic depicting vestigial-preference hypothesis for túngara frog females. Note that ancestral calls are defined as positive, and so lie on the right of the túngara frog call. (b) Responses of female túngara frogs to ancestral and anti-ancestral calls in recognition tests. Dashed line represents criterion for a significantly recognized call. Horizontal lines indicate recognition responses that were significantly different from one another ( $P < 0.01$ ). (c) Responses of female túngara frogs to ancestral and anti-ancestral calls in discrimination tests. Filled bars represent responses to ancestral stimuli; gray bars responses to anti-ancestral stimuli. Horizontal line depicts null hypothesis of no preference.

**Preference Tests.** Females were collected while in amplexus from Gamboa, Panama, between 1900 and 2100 h. Females were separated from males, brought into the laboratory, and kept in a dark cooler at room temperature (26°C) until 2200 h, at which time testing began.

The test chamber was a sound resistant and light proof chamber 1.8 m wide by 2.7 m in length. A speaker was placed at either end, and an infrared camera was mounted on the ceiling directly above the center of the arena. A pure tone was played through the speakers, and the playback amplitude from each speaker was set to 82 dB SPL (peak, re: 20  $\mu$ Pa) at the center of arena. The sound level was recalibrated between females. The linoleum floor was sprayed periodically with water.

At the beginning of a test, the subject was removed from the cooler and placed in the center of the test chamber beneath a small funnel. The door was closed and the test stimuli were broadcast from each of the two speakers. Stimuli were presented 180° out of phase, with a period of 2 s. In the recognition tests, one stimulus was the call of interest—one of the six ancestral or anti-ancestral calls—and the second stimulus was white noise in an amplitude envelope matching the call of the túngara frog. In the discrimination tests, one stimulus was a call on the ancestral side of the transect, the second was an anti-ancestral call with the same overall similarity to the túngara frog call. Six stimuli were tested in recognition tests, three ancestral/anti-ancestral pairs in discrimination tests. The stimuli were counterbalanced so that each was broadcast from the opposite speaker in alternating tests.

After 3 min acclimation, the funnel was lifted remotely. Responses of females were observed on a video monitor located outside the test chamber.

Females were scored as responding to a stimulus if they approached within 10 cm of a speaker traversing a path that did not follow the walls of the chamber. If females approached the speaker by following the wall, this was scored as a positive response if and only if the subject passed the speaker and then returned to the speaker. A ring on the floor around the perimeter of the funnel indicated the position of the female at the start of the test. If the female did not leave this ring within 5 min of the onset of the trial, females were scored as not responding. If a female passed the speaker by following the wall but did not return to the speaker; if she stayed in one place for more than 2 min; or if she spent 2 min climbing the walls of the chamber, the subject was scored as not responding.

In the discrimination tests, only positive responses to one of the two stimuli were recorded. Recognition tests were always preceded and followed by tests examining female responses to a whine/whine-chuck discrimination, a pair of stimuli known to elicit reliable responses from females (43). If females did not respond to either stimulus, the females were considered unresponsive, and any intervening recognition tests were discarded.

**Data Analysis.** For each network, we summed responses to signals along the ancestral side of the transect and compared this sum to the responses of networks to stimuli along the anti-ancestral side of the transect. For mimetic history networks, ancestral calls were those calls with a steep whine shape—a short THHz. For mirrored history networks, ancestral calls were those with a shallow whine shape. The resulting 279 entries were analyzed with a sign test to determine whether networks given the two history types consistently exhibited asymmetric biases in favor of calls resembling ancestral states.

Responses of females in recognition tests were first compared with data gathered from a null test. When females were given a choice between a silent speaker and a speaker playing white noise, one female of twenty subjects approached the silent speaker. Consequently, a response to no-response ratio of 1:19 was taken as the null hypothesis. By convention (36), we define females as recognizing a stimulus if significantly more approach

**Table 1. Parametric manipulations of historical simulations**

| Simulation parameters |             |              |             | Mimetic history networks |       |    | Mirrored history networks |       |    | Most accur. |
|-----------------------|-------------|--------------|-------------|--------------------------|-------|----|---------------------------|-------|----|-------------|
| Mut prob              | Rec prob    | Pop sz       | Noise       | Avg gen                  | SE    | N  | Avg gen                   | SE    | N  |             |
| <b>0.0001</b>         | 0.50        | 100          | 0.02        | 6,540                    | 1,245 | 10 | 8,151                     | 969   | 10 | Mir         |
| <b>0.0005</b>         | 0.50        | 100          | 0.02        | 2,015                    | 190   | 10 | 1,866                     | 121   | 10 | Mim         |
| <b>0.0010</b>         | 0.50        | 100          | 0.02        | 1,515                    | 280   | 10 | 1,423                     | 126   | 10 | Mim         |
| <b>0.0050</b>         | 0.50        | 100          | 0.02        | 796                      | 85    | 10 | 882                       | 75    | 10 | Mim         |
| <b>0.0100</b>         | 0.50        | 100          | 0.02        | 2,205                    | 391   | 10 | 6,634                     | 2,039 | 10 | Mim         |
| 0.0010                | <b>0.00</b> | 100          | 0.02        | 2,671                    | 291   | 10 | 2,594                     | 352   | 10 | Mim         |
| 0.0010                | <b>0.25</b> | 100          | 0.02        | 1,383                    | 138   | 10 | 1,430                     | 108   | 10 | Mim         |
| 0.0010                | <b>0.50</b> | 100          | 0.02        | 1,515                    | 280   | 10 | 1,423                     | 126   | 10 | Mim         |
| 0.0010                | <b>0.75</b> | 100          | 0.02        | 1,165                    | 157   | 10 | 1,389                     | 167   | 10 | Mim         |
| 0.0010                | <b>1.00</b> | 100          | 0.02        | 883                      | 108   | 10 | 1,294                     | 171   | 10 | Mim         |
| 0.0010                | 0.50        | <b>10</b>    | 0.02        | 20,000                   | 0     | 0  | 20,000                    | 0     | 0  |             |
| 0.0010                | 0.50        | <b>50</b>    | 0.02        | 2,214                    | 180   | 10 | 2,374                     | 325   | 10 | Mim         |
| 0.0010                | 0.50        | <b>100</b>   | 0.02        | 1,515                    | 280   | 10 | 1,423                     | 126   | 10 | Mim         |
| 0.0010                | 0.50        | <b>500</b>   | 0.02        | 356                      | 22    | 10 | 467                       | 35    | 10 | Mim         |
| 0.0010                | 0.50        | <b>1,000</b> | 0.02        | 566                      | 271   | 9  | 409                       | 51    | 10 | Mim         |
| 0.0010                | 0.50        | 100          | <b>0.00</b> | 8,215                    | 2,000 | 10 | 20,000                    | 0     | 0  |             |
| 0.0010                | 0.50        | 100          | <b>0.02</b> | 1,515                    | 280   | 10 | 1,423                     | 126   | 10 | Mim         |
| 0.0010                | 0.50        | 100          | <b>0.05</b> | 1,196                    | 154   | 10 | 1,526                     | 141   | 10 | Mir         |
| 0.0010                | 0.50        | 100          | <b>0.10</b> | 2,761                    | 650   | 10 | 3,869                     | 692   | 10 | Mim         |
| 0.0010                | 0.50        | 100          | <b>0.25</b> | 12,133                   | 2,402 | 7  | 20,000                    | 0     | 0  |             |

Ten simulations of each network type and parameter combination were initiated. N, the number of simulations completed. Most accur., the history type that was best at predicting female responses. Mut prob, mutation probability; rec prob, recognition probability; pop sz, population size; avg gen, average number of generations; SE, standard error; Mir, mirror; Mim, mimetic. Bold type denotes parameters that were varied.

it than approach a silent speaker (using a Fisher's exact test with  $\alpha = 0.05$ , recognition occurs if 7 or more females approach the stimulus). Responses between ancestral/anti-ancestral pairs presented in separate recognition tests were also compared with Fisher's exact test.

To compare female preferences observed in discrimination tests, we compared the number of females that responded to the ancestral and similarity-matched anti-ancestral calls by using a two-tailed Fisher's exact test, with a null expectation that 10 of 20 females would respond to ancestral calls. Females that did not respond in the discrimination paradigm were excluded from analysis.

## Results

**Parametric Influences on Network Responses.** None of the parameters investigated (mutation probability, recombination probability, population size, ambient noise) significantly influenced the overall error rates of the networks (Kendall Rank  $\tau$  adjusted for ties,  $P > 0.05$  after Bonferroni correction). Of the 14 sets of simulations in which we were able to compare the error rates of mimetic and mirrored history networks, 12 of the comparisons (86%) demonstrated that mimetic history networks made smaller errors than mirrored history networks ( $P < 0.02$ , sign test). This result demonstrates that the predictive advantage displayed by mimetic history networks is robust to manipulations of evolutionary parameters.

**Vestigial Preferences in Neural Networks.** Tests of network responses revealed that 99 of 139 mimetic history networks preferred calls with a steep, ancestor-like whine to those with a shallow whine ( $P < 0.001$ , sign test). Similarly, 88 of 140 networks given a mirrored history preferred ancestral calls with a shallow, ancestor-like whine to those with a steep whine ( $P < 0.001$ , sign test). Pooling these data indicates that 198 of 279 ( $P < 0.001$ , sign test) networks preferred ancestral calls (Fig. 3 *b* and *c*).

**Vestigial Preferences in Female Túngara Frogs.** Female *P. pustulosus* responded to all six of the ancestral and anti-ancestral calls

significantly more often than to the null stimulus ( $P < 0.05$ , Fig. 4*b*), indicating that all stimuli were recognized as acceptable calls. Nevertheless, there is an overall asymmetry in the preferences of females, with a general bias toward the reconstructed ancestral calls. The +0.5 ancestral call elicits significantly stronger responses than the -0.5 anti-ancestral call (19/20 ancestral vs. 11/20 anti-ancestral, Fisher's exact test,  $P < 0.01$ ), and the +1.5 call is recognized significantly better than the -1.5 call (18/20 ancestral vs. 9/20 anti-ancestral, Fisher's exact test,  $P < 0.01$ ). A similar trend holds for the +1.0-ancestor and -1.0-anti-ancestor (15/20 ancestral vs. 11/20 anti-ancestral, Fisher's exact test,  $P > 0.05$ ).

In discrimination tests, females exhibited very strong preferences for all calls on the ancestral portion of the transect (+1.5/-1.5 : 20/0,  $P < 0.01$ , sign test; +1.0/-1.0: 18/2,  $P < 0.01$ ; +0.5/-0.5: 19/1,  $P < 0.01$ , Fig. 4*c*).

## Discussion

We set out to test whether evolutionary history could distort the patterns of generalization exhibited by receivers. Previous data demonstrated that neural networks with distinct histories respond differently to novel stimuli (30, 31). In our current study, we found that neural networks selected to follow a trajectory approximating that of the túngara frog showed emergent response biases toward stimuli resembling ancestral calls. Mirrored history networks exhibited analogous preferences for calls resembling mirrored ancestors. The fact that the mimetic and mirrored history networks display preferences for opposite stimuli indicates that their asymmetric generalization gradients are indeed attributable to history, rather than to any fundamental asymmetry required for the recognition of the túngara frog call. Moreover, these historical differences were robust to changes in simulation parameters, and did not require the stimulus transect to pass through the true ancestors of either history type. These data suggest that distortion of preference functions in favor of ancestor-like stimuli should be detectable even when there are errors in the estimation of ancestral states.

Female túngara frogs responded to all six of the test stimuli significantly more often than to the null stimulus (white noise), indicating that all six stimuli fall within the range of signal variation females classify as acceptable conspecific mates. Nevertheless, the recognition data indicate that females clearly respond more often to ancestral than anti-ancestral calls. These data demonstrate a response bias in the direction predicted by the neural network data. Discrimination of equidistant pairs of stimuli revealed that these preferences were very strong in all stimulus pairs—the weakest preference was measured when 18 of 20 females approached the +1.5-ancestral call rather than the -1.5-anti-ancestral call.

Prior investigations demonstrate that inter-specific mating preferences can be predicted by phylogenetic relationships (e.g., refs. 44–46). The vestigial-preference hypothesis is consistent with these findings, makes novel predictions, and avoids some of the shortcomings of earlier approaches to historical influences. By looking for asymmetric preferences within a species, we eliminate confounding variation in species permissiveness or motivation—a potential problem with comparisons across species. The hypothesis also makes few assumptions about speciation mechanisms or the persistence of ancestral taxa. Lastly, probing for vestigial preferences seems simpler and more generally applicable than multiple regression methods for relating phylogeny to mate choice: Because one would expect signal similarity and phylogenetic relatedness to be correlated, co-linearity assumptions in multiple regression will often be violated.

Although it is beyond the scope of the current investigation, we anticipate more exhaustive studies documenting the distribution of vestigial preferences across a clade. A comparison of vestigial preferences between sister taxa could serve as a control analogous to comparing mimetic and mirrored history networks. Parsimonious reconstructions of ancestral nodes typically fall between trait values of extant sister taxa (47). As a result, vestigial preferences in a species pair are predicted to lie in opposite directions relative to their conspecific signals. Moreover, because this assay for vestigial preferences emphasizes the relative proximity of stimuli to ancestral states rather than absolute parameters of the ancestral signal, the predicted asymmetries should be relatively robust to errors in ancestral reconstructions.

Classic work in mathematical psychology posits that simple generalization gradients may sum to produce a complex composite gradient (48, 49); some ethologists have used this frame-

work to explain the emergence of receiver biases toward super-normal or symmetric stimuli during learning and evolution (8, 19, 20, 23–25, 50, 51). A similar approach can be applied to the data gathered in this study. Vestigial preferences may result from the persistence of a generalization gradient centered at an ancestral signal. Such preferences may decay over time, or if ancestral recognition mechanisms contribute to contemporary species recognition, may persist indefinitely. The first case predicts a negative correlation between the elapsed time since a speciation event and the strength of the vestigial preference. The second assumes that vestigial preferences are maintained because of the pleiotropic effects of preference genes, and should not depend on branch lengths. Analytic models of gradient structure and signal evolution rates could provide quantitative predictions about the distribution and strength of vestigial preferences within a clade. Relating the degree of preference asymmetry to branch lengths would provide strong tests of these alternative hypotheses.

Like other patterns detected by the neural network models of receiver evolution (e.g., peak-shift and directional intensity preferences, reviewed in refs. 8, 24, and 25), the hypotheses that emerge from our data can be examined and tested quite independently of the simulations that brought the patterns to light. In the current study, the neural network data led to novel predictions regarding the behavior of females—predictions that were verified experimentally. The data also suggest an avenue by which neural networks can be used to develop mathematical models of vestigial preferences for general use. These findings highlight the potential for computational approaches to animal decision-making to contribute to both experimental and mathematical analyses of behavioral evolution. By reuniting proximate and ultimate approaches to animal behavior, the computational methods of psychology and neuroscience promise to renew and extend the ambitions of traditional ethology. The integration of these tools with the methods of behavioral ecology and population genetics will promote a deeper and more integrative understanding of animal behavior and how it evolves.

The authors would like to acknowledge W. Wilczynski for review of early drafts of the manuscript, as well as study design. This work was supported by National Institute of Mental Health Grant F31 MH11194 and a Smithsonian Postdoctoral Fellowship to S.M.P., a Guggenheim Fellowship and National Science Foundation Grant IBN 98 16564 to M.J.R., and a Smithsonian Institution Scholarly Studies grant to A.S.R.

1. Russell, E. S. (1943) *Proc. Linn. Soc. Lond.* **154**, 195–216.
2. Tinbergen, N. (1951) *The Study of Instinct*. (Oxford Univ. Press, Oxford).
3. Ryan, M. J., Fox, J. H., Wilczynski, W. W. & Rand, A. S. (1990) *Nature (London)* **343**, 66–67.
4. Basolo, A. (1990) *Science* **250**, 808–810.
5. Guilford, T. & Dawkins, M. S. (1991) *Anim. Behav.* **42**, 1–14.
6. Endler, J. A. (1992) *Am. Nat.* **139**, S125–S153.
7. Christy, J. H. (1995) *Am. Nat.* **146**, 171–181.
8. Enquist, M. & Arak, A. (1998) in *Cognitive Ecology*, ed. Dukas, R. (Chicago Univ. Press, Chicago), pp. 21–87.
9. Ryan, M. J. (1998) *Science* **281**, 1999–2003.
10. Endler, J. A. & Basolo, A. L. (1998) *Trends Ecol. Evol.* **13**, 415–420.
11. Hailman, J. P. (1962) *Auk* **79**, 89–98.
12. Hailman, J. P. (1971) *Anim. Behav.* **19**, 328–335.
13. Wagner, W. E., Murray, A. & Cade, W. H. (1995) *Anim. Behav.* **49**, 1269–1281.
14. Basolo, A. (1998) *Proc. R. Soc. London Ser. B* **265**, 2223–2228.
15. Murphy, C. G. & Gerhardt, H. C. (2000) *Evolution* **54**, 660–669.
16. Shaw, K. L. (2000) *Evolution* **54**, 1303–1312.
17. Shaw, K. L. & Herlihy, D. P. (2000) *Proc. R. Soc. London Ser. B* **267**, 577–584.
18. Jennions, M. D. & Petrie, M. (1997) *Biol. Rev.* **72**, 283–327.
19. Staddon, J. E. R. (1975) *Am. Nat.* **109**, 541–545.
20. ten Cate, C. & Bateson, P. (1988) *Evolution* **42**, 1355–1358.
21. ten Cate, C. & Bateson, P. (1989) *Anim. Behav.* **38**, 356–358.
22. Nelson, D. A. & Marler, P. (1989) *Science* **244**, 976–978.
23. Weary, D. M., Guilford, T. C. & Weisman, R. G. (1993) *Evolution* **47**, 333–336.
24. Ghirlanda, S. & Enquist, M. (1998) *Anim. Behav.* **56**, 1383–1389.
25. Ghirlanda, S. & Enquist, M. (1999) *Anim. Behav.* **58**, 695–706.
26. Ryan, M. J. & Getz, W. (2000) *Brain Behav. Evol.* **56**, 45–62.
27. Enquist, M. & Arak, A. (1993) *Nature (London)* **361**, 446–448.
28. Krakauer, D. C. & Johnstone, R. A. (1995) *Philos. Trans. R. Soc. London B* **348**, 355–361.
29. Hurd, P. L., Wachtmeister, C.-A. & Enquist, M. (1995) *Proc. R. Soc. London Ser. B* **259**, 201–205.
30. Phelps, S. M. & Ryan, M. J. (1998) *Proc. R. Soc. London Ser. B* **265**, 279–285.
31. Phelps, S. M. & Ryan, M. J. (2000) *Proc. R. Soc. London Ser. B* **267**, 1633–1639.
32. Rumelhart, D. E., McClelland, J. L. & the PDP Research Group. (1986) *Parallel Distributed Processing: Explorations in the Microstructure of Cognition* Foundations (MIT Press, Cambridge, MA), Vol. 1.
33. Rumelhart, D. E., McClelland, J. L. & the PDP Research Group. (1986) *Parallel Distributed Processing: Explorations in the Microstructure of Cognition: Psychological and Biological Models* (MIT Press, Cambridge, MA), Vol. 2.
34. Phelps, S. M. (2001) in *Anuran Communication*, ed. Ryan, M. J. (Smithsonian Press, Washington, DC), pp. 167–180.
35. Ryan, M. J., Phelps, S. M. & Rand, A. S. (2001) *Trends Cog. Sci.* **5**, 143–148.
36. Ryan, M. J. & Rand, A. S. (1995) *Science* **269**, 390–392.
37. Ryan, M. J. & Rand, A. S. (1999) *Anim. Behav.* **57**, 945–956.
38. Jennions, M. D. & Brooks, R. (2001) *Trends Ecol. Evol.* **16**, 113–115.
39. Fisher, R. A. (1958) *The Genetical Theory of Natural Selection* (Dover Publications, New York), 2nd Ed.
40. Pfennig, K. S. (1998) *Proc. R. Soc. London Ser. B* **265**, 1743–1748.
41. Reeve, H. K. (1989) *Am. Nat.* **133**, 407–435.
42. Cannatella, D. C., Hillis, D. M., Chippindale, P. T., Weight, L., Rand, A. S. & Ryan, M. J. (1998) *Syst. Biol.* **47**, 311–335.
43. Rand, A. S. & Ryan, M. J. (1981) *Z. Tierpsych.* **57**, 209–214.
44. Kaneshiro, K. (1976) *Evolution* **3**, 740–745.
45. McLennan, D. A. & Ryan, M. J. (1997) *Anim. Behav.* **54**, 1077–1088.
46. McLennan, D. A. & Ryan, M. J. (1999) *Evolution* **53**, 880–888.
47. Harvey, P. H. & Pagel, M. D. (1991) *The Comparative Method in Evolutionary Biology* (Oxford Univ. Press, New York).
48. Spence, K. W. (1937) *Psychol. Rev.* **44**, 430–444.
49. Mackintosh, N. J. (1974) *The Psychology of Animal Learning* (Academic, New York).
50. Arak, A. & Enquist, M. (1993) *Philos. Trans. R. Soc. London B* **340**, 207–213.
51. Enquist, M. & Johnstone, R. (1997) *Proc. R. Soc. London Ser. B* **264**, 1345–1348.