Supplementary information to:

The evolution of social learning as phenotypic cue integration

Bram Kuijper, Olof Leimar, Peter Hammerstein,

John M. McNamara, Sasha R. X. Dall

S1 Supplementary Figures

Figure S1   Example simulations in a negatively autocorrelated environment ($1 - p = 0.9$), where each line depicts the average trait value of a single replicate run for over 75000 generations. Sensitivity to individually learned cues a_{ind} evolves to slight negative values (panel B), while both vertical and horizontally learned cues based on conformism (v_{c} and h_{c}) evolve to positive values (panels D,E). See also panel B of Figure S2 at the value $1 - p = 0.9$. Parameters: $q_{\mathrm{ind}} = 1.0$, $q_{\mathrm{mat}} = 0.5$, $n = 5$, $m = 0.1$, $1 - p = 0.9$, $\sigma_{\mathrm{h}} = \sigma_{v} = 0.0$. Number of replicate simulations: 5.
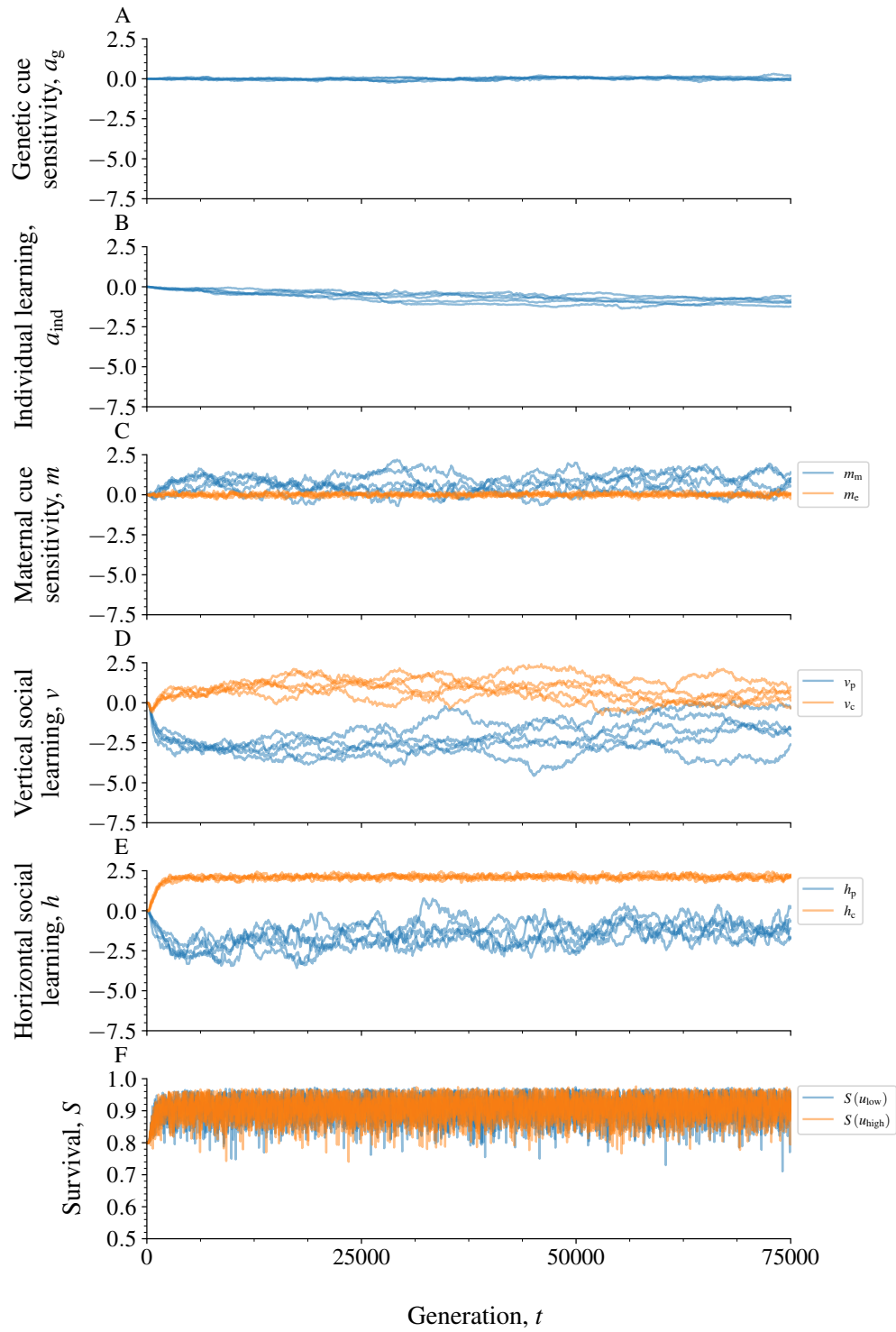
Figure S1:

635 **Figure S2** The evolved values of the sensitivities when varying the probability of environmental

636 change $1 - p$. See Figure 2 for the corresponding proportions of variance explained by the various

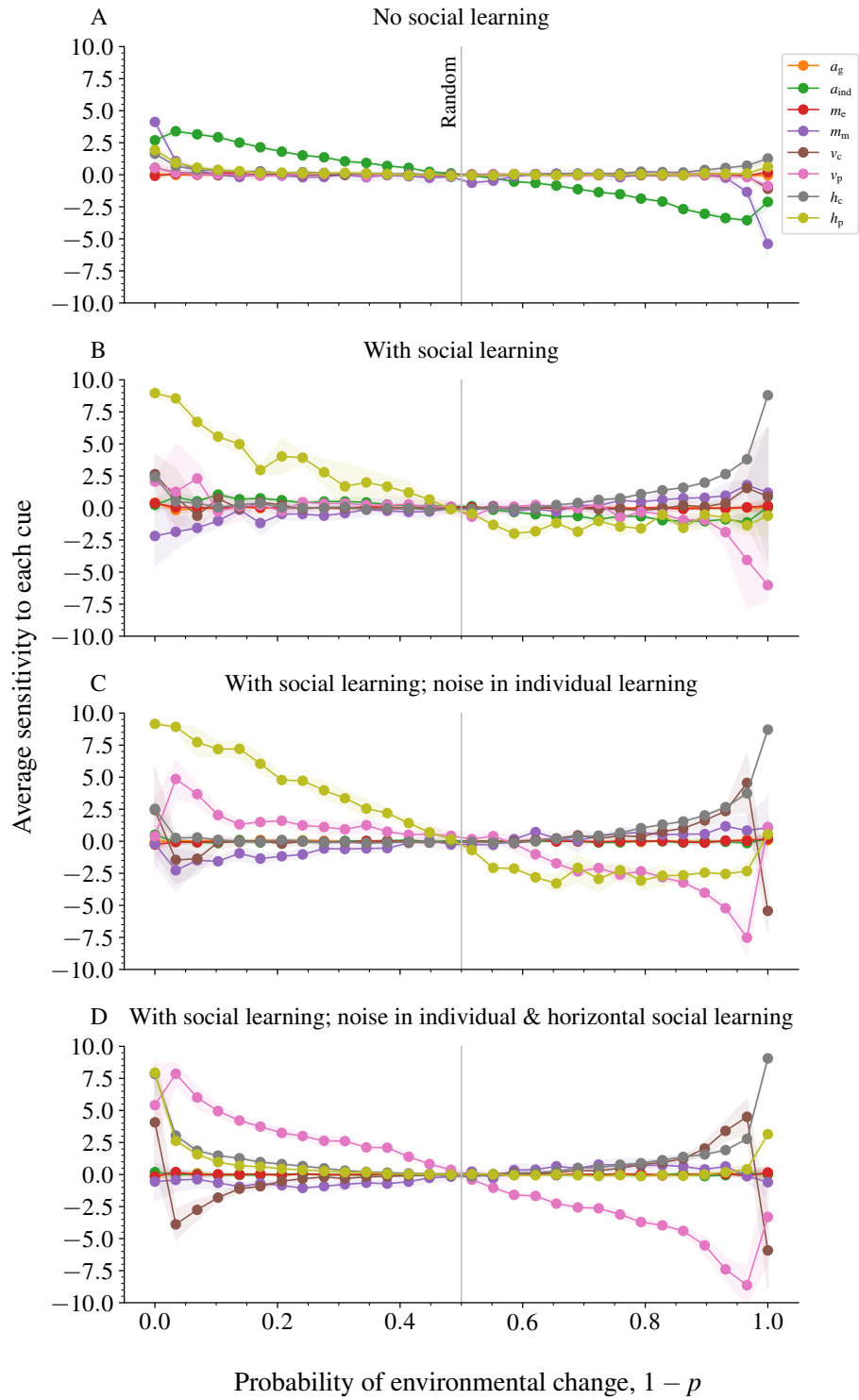637 sensitivity $\times$ cue combinations. See Figure 2 for parameters.

Figure S2:

638 **Figure S3** Proportions of variance explained when horizontal social learning based on prestige
639 or conformism biases coevolves exclusively with genetic cues (panel A), exclusively with maternal
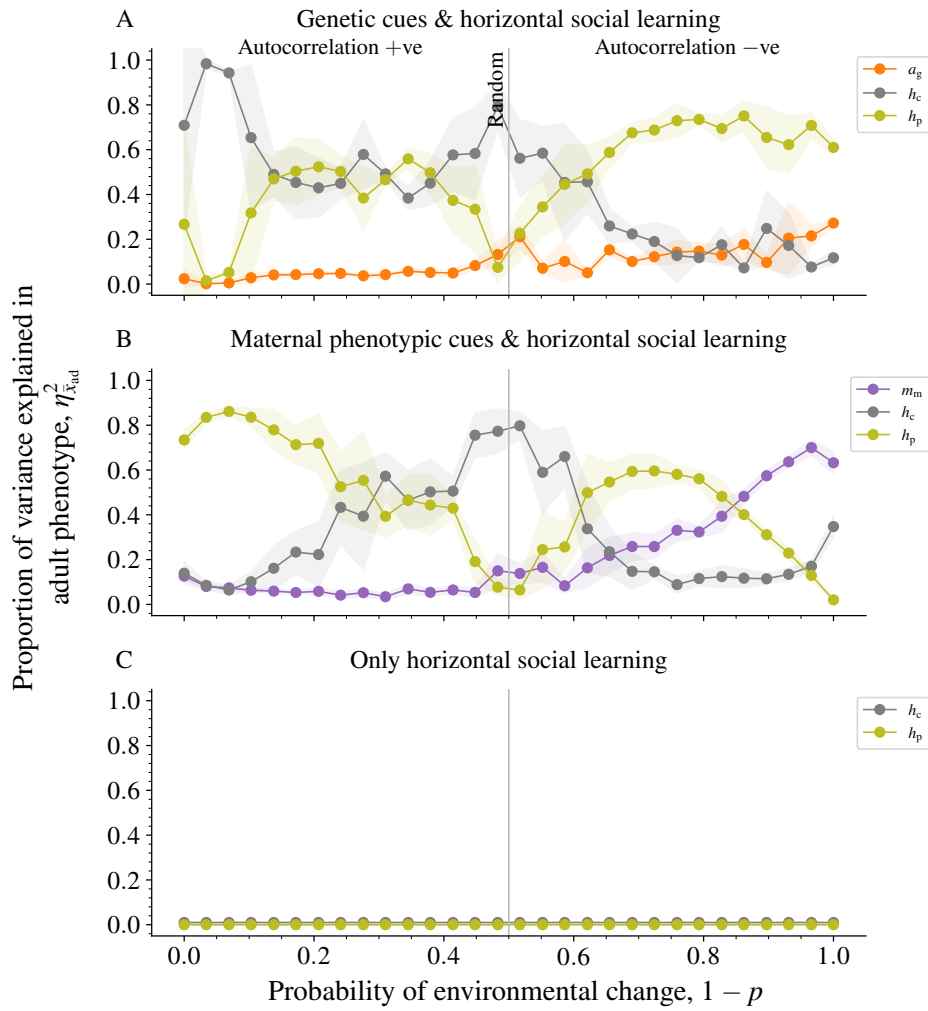640 phenotypic cues (panel B) and without any other cues (panel C). See Figure 2 for parameters.

Figure S3:

Figure S4 Most important cues affecting juvenile phenotype determination (i.e., before horizontal social learning). Vertical prestige biases and individual learning mostly predominate. However, when fidelities of both vertical social learning and individual learning are low (top left corner of each panel), maternal effects typically prevail. Cases where vertical social learning based on conformity evolve (green areas) despite high levels of noise in vertical social learning merely reflect the generation of phenotypic variance: horizontal prestige-based learning (shown in Figure 3) evolves to strong values here, so that other cues are largely superfluous. See Figure 3 for parameters.
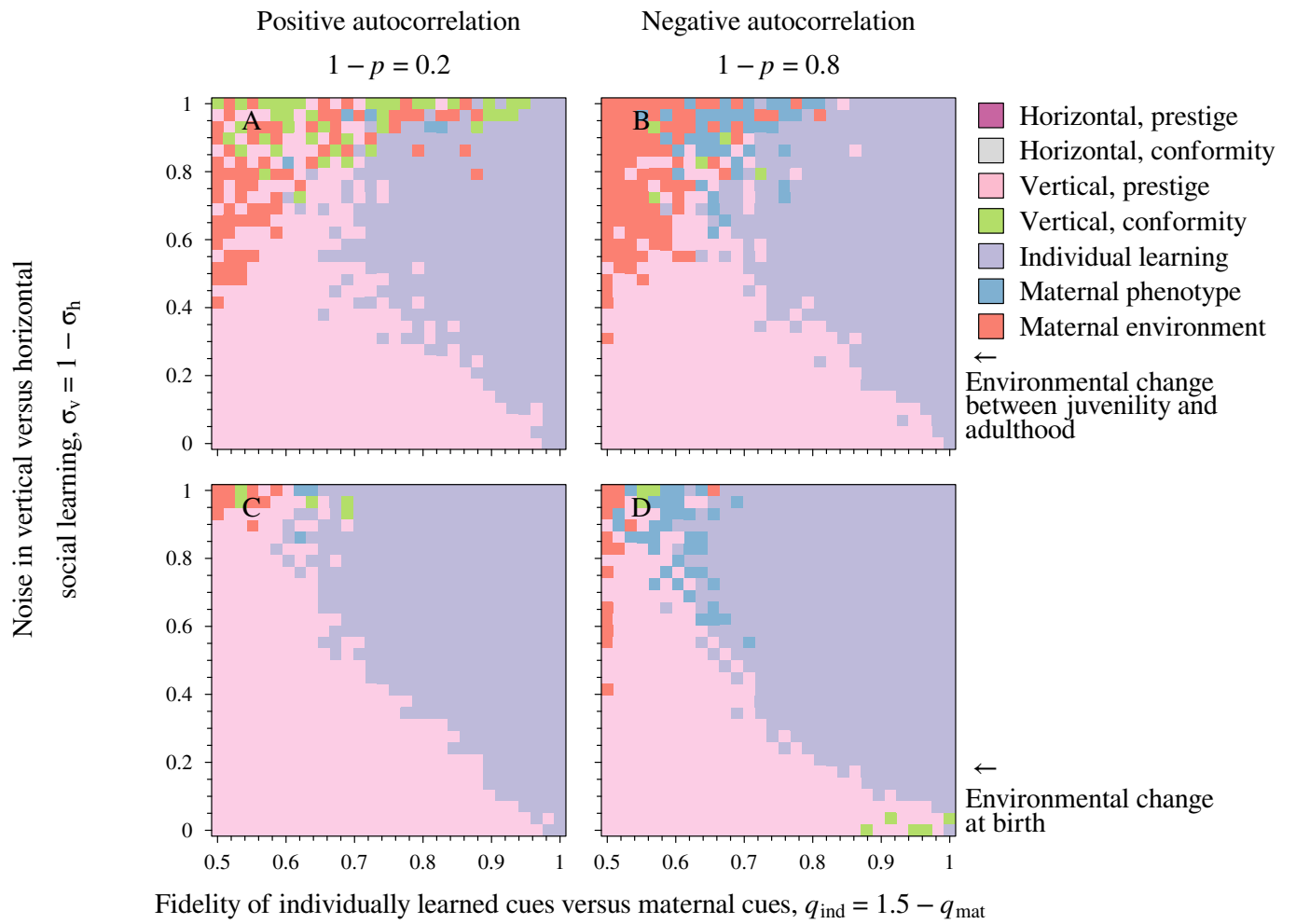
Figure S4:

Figure S5 Proportions of variance explained when environmental change occurs at birth, rather

than between juvenility and adulthood (compare with Figure 2). See Figure 2 for parameters.

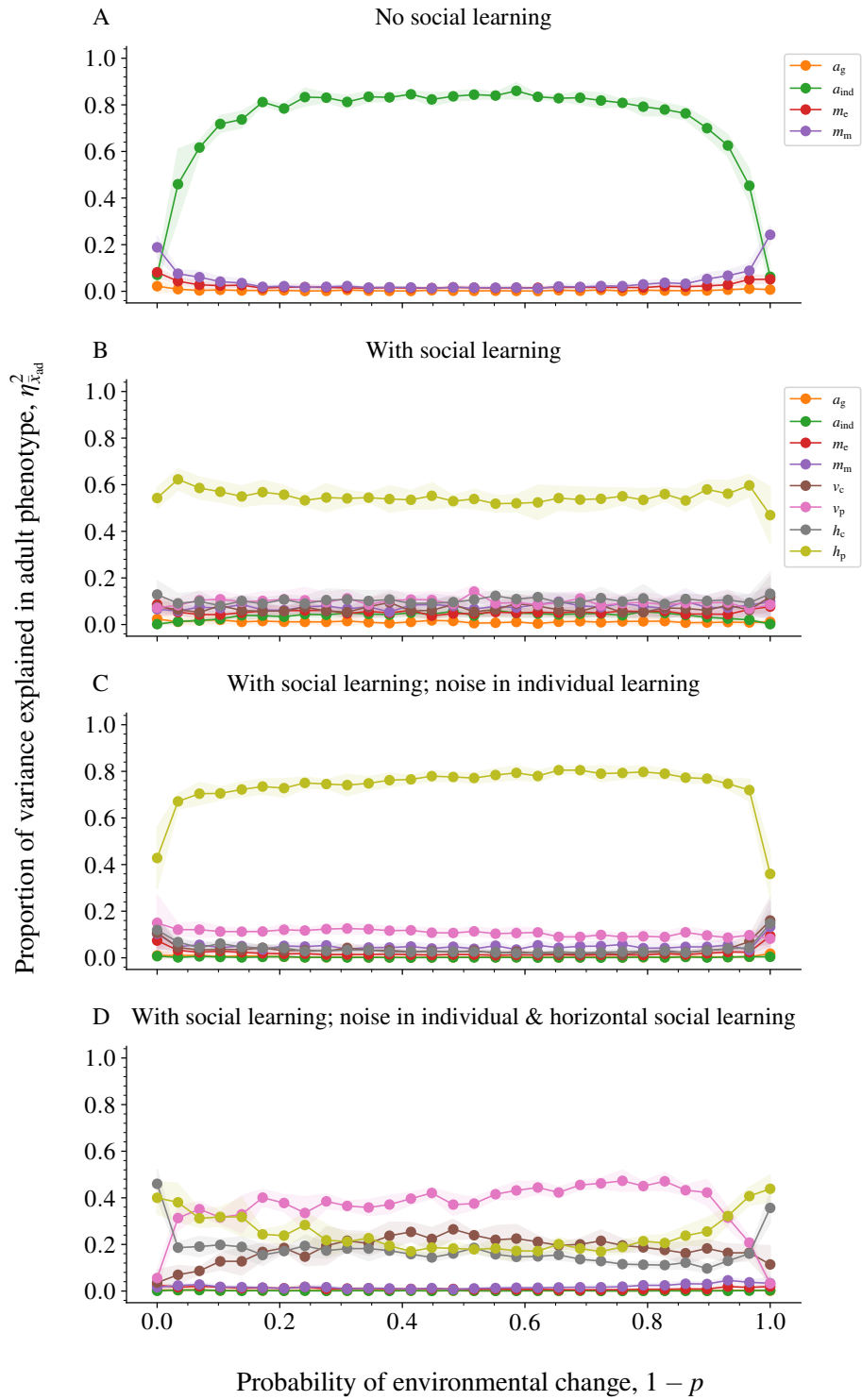Figure S5:

Figure S6 Sensitivity analysis of the timing of vertical social learning and different values of migration *m* (from left to right). The top two rows (panels A - F) consider the scenario discussed in the main text, in which juveniles learn vertically from breeders in the natal patch (i.e., vertical learning *before* migration), while they learn horizontally from other juveniles on the patch to which they have migrated (i.e., social learning *after* migration. By contrast, the bottom two rows (panels G - L) consider a scenario where both vertical and social learning occur from individuals on the patch to which a focal individual has migrated (i.e., learning *after* migration). Coloration again depicts the cue which explains the majority of phenotypic variance in adult phenotype (measured at the logistic scale).

When individuals learn vertically in their natal patch and learn horizontally in their patch of arrival (panels A - F), we find that for high levels of dispersal horizontal social learning prevails in both negatively and positively autocorrelated environments (panels C, F). This occurs because only horizontal learning allows an individual to obtain information about its current local environment.

By contrast, when individuals learn vertically as well as horizontally on their patch of arrival (panels G - L), horizontal social learning only prevails when it is characterised by low levels of noise, while vertical social learning prevails otherwise (compare panels I, L with C, F). In this case, both horizontal and vertical social learning (based on prestige) provide accurate information about the environment on the patch of arrival, so that prevalence of either is determined by the relative amount of noise $\sigma_v = 1 - \sigma_h$ that is varied on the *y* axis. Other parameters as in Figure 3.
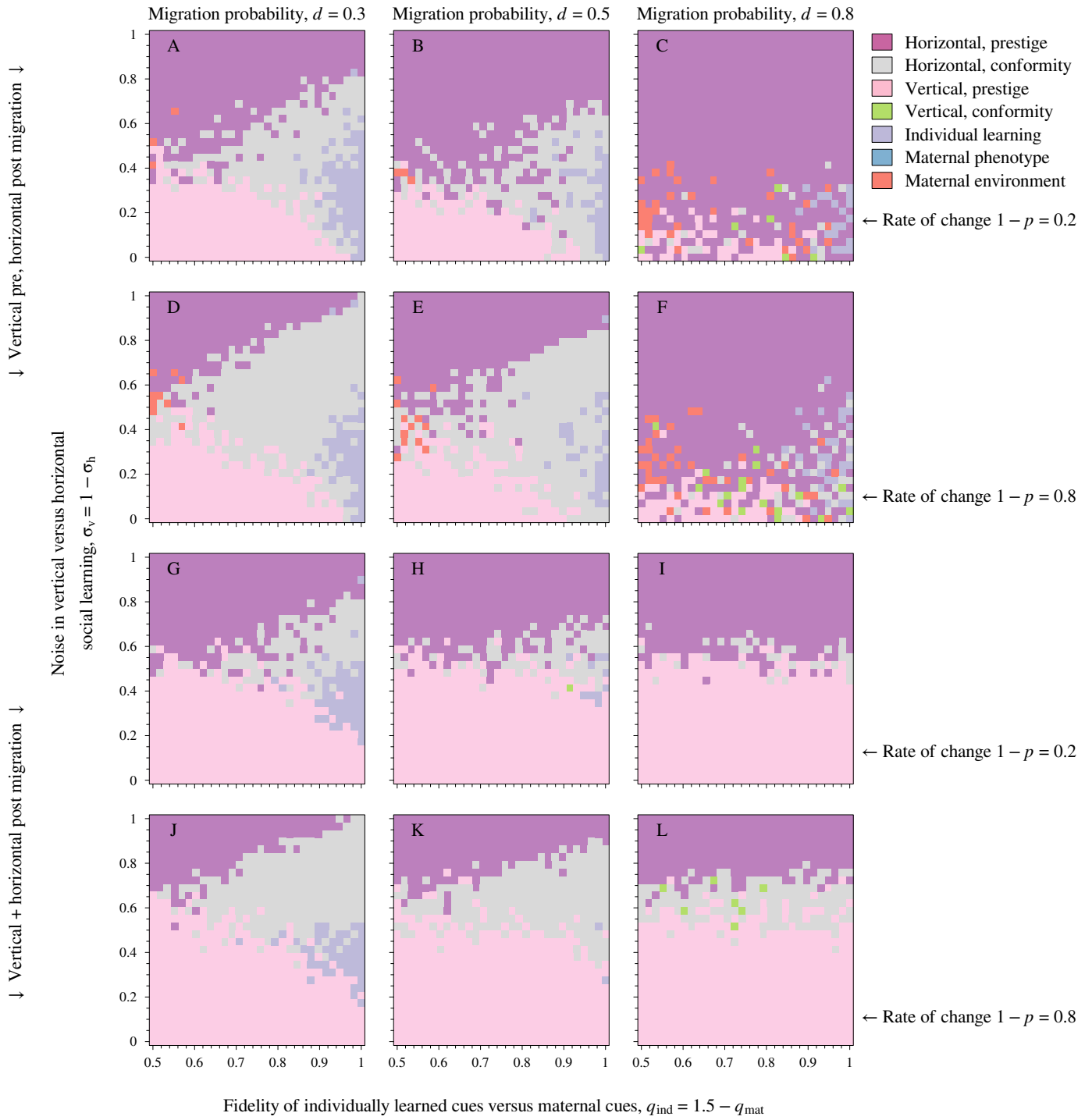
Figure S6:

Figure S7 Sensitivity analysis of the timing of horizontal learning (panels A - F), the timing of individual learning (panels G - L) and different values of migration m (from left to right).

The top two rows (panels A - F) consider a scenario in which individuals perform all their learning before migration: individual learning and vertical social learning is assumed to occur before migration as in the main text, while now also horizontal social learning happens before migration As a consequence, individuals only acquire information about the environmental state of their natal patch. With an increasing probability of migration, individuals thus lack any information about their future environment, resulting in an outcome where all cues are equally informative (panels C, F) as testified by the large variation between individual simulations in the cue that is most informative.

The bottom two rows (panels G - L) consider the same scenario as in the main text (vertical social learning prior to migration, horizontal social learning after migration), except that individual learning now occurs after dispersal. As a consequence, individual learning (once it has a high fidelity as measured by q_{ind} on the left-hand side of each panel) now prevails over a much larger part of the parameter space, particularly when dispersal is high (i.e., compare panels I, L with Figure S6C, F). Other parameters as in Figure 3.
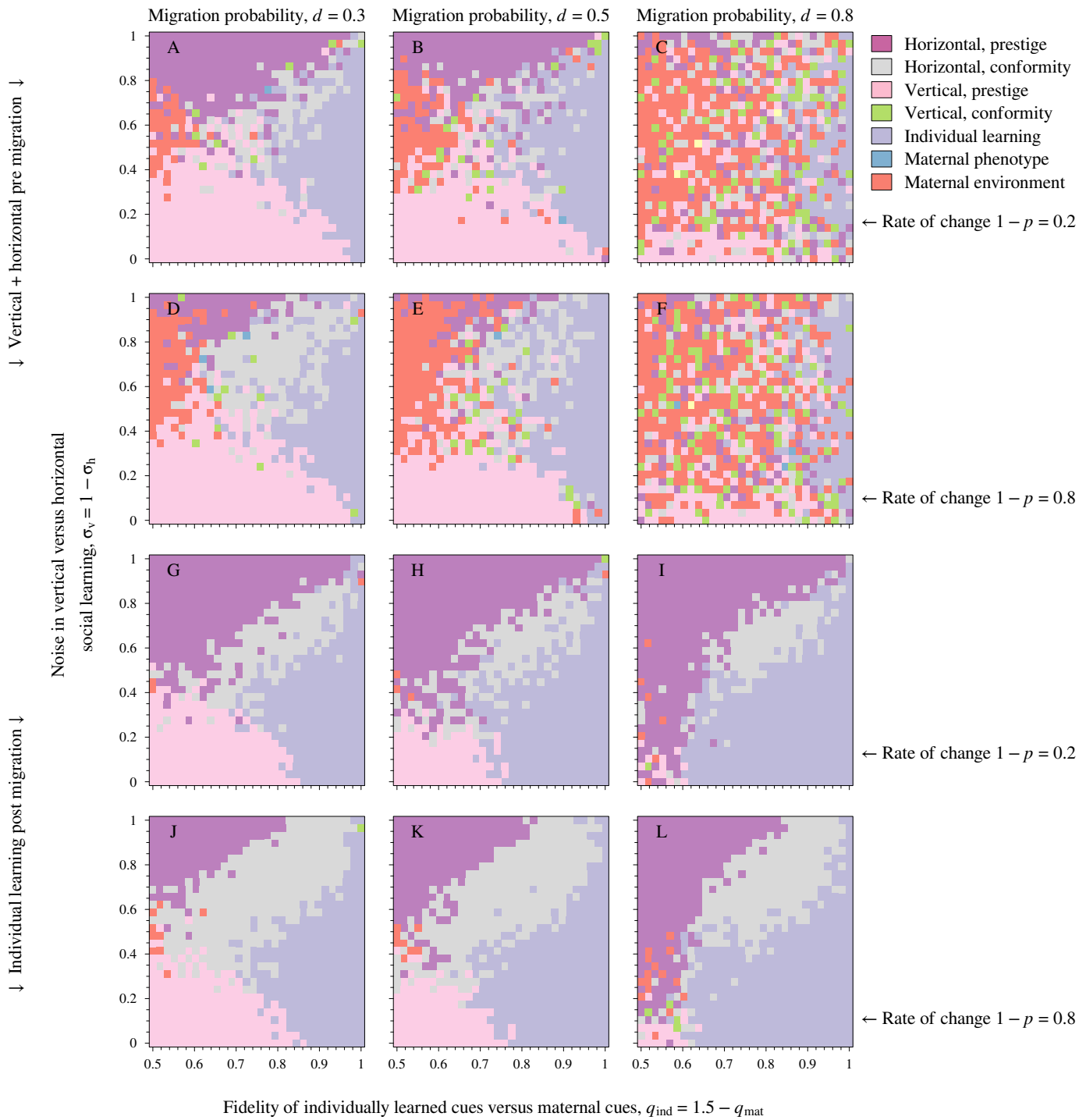
Figure S7:

685 **Figure S8** The effect of varying the number n of models sampled when performing horizontal

686 or vertical social learning. As in Figure 3 in the main text (in which $n = 5$). Panel A: when $n = 1$

687 (i.e., random sampling of models), we find that maternal environmental effects or individual learn-

688 ing prevail, while conformity biases prevail at intermediate values of $q_{ind} = 1.5 - q_{mat}$, effectively

689 reflecting social learning from random individuals. Once n gets larger, we find that maternal effects

690 and individual learning are replaced by prestige-based social learning (either vertical or horizontal).

691 Parameters as in Figure 3.
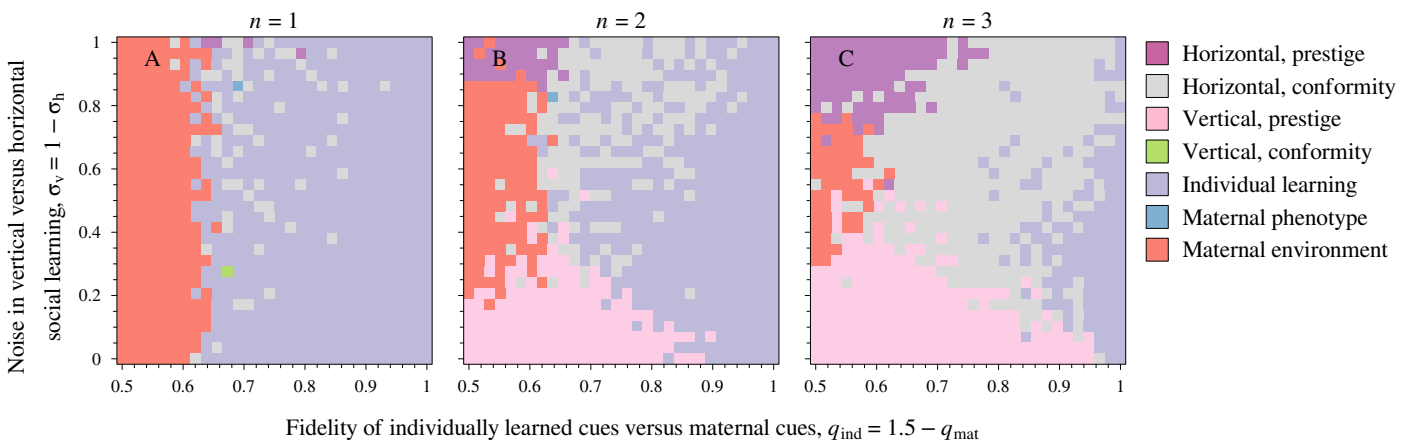
Figure S8:

Figure S9 Cues that explain the majority of phenotypic variance in adult phenotype (see Figure 3 in the main text) when social learning is only based on conformity biases. In comparison to Figure 3, we now find that horizontal prestige-based social learning is replaced by horizontal learning based on conformity biases. By contrast, vertical prestige-based social learning is replaced either by maternal effects (towards left-hand sign of each panel) or individual learning once its fidelity is large enough. Parameters as in Figure 3.

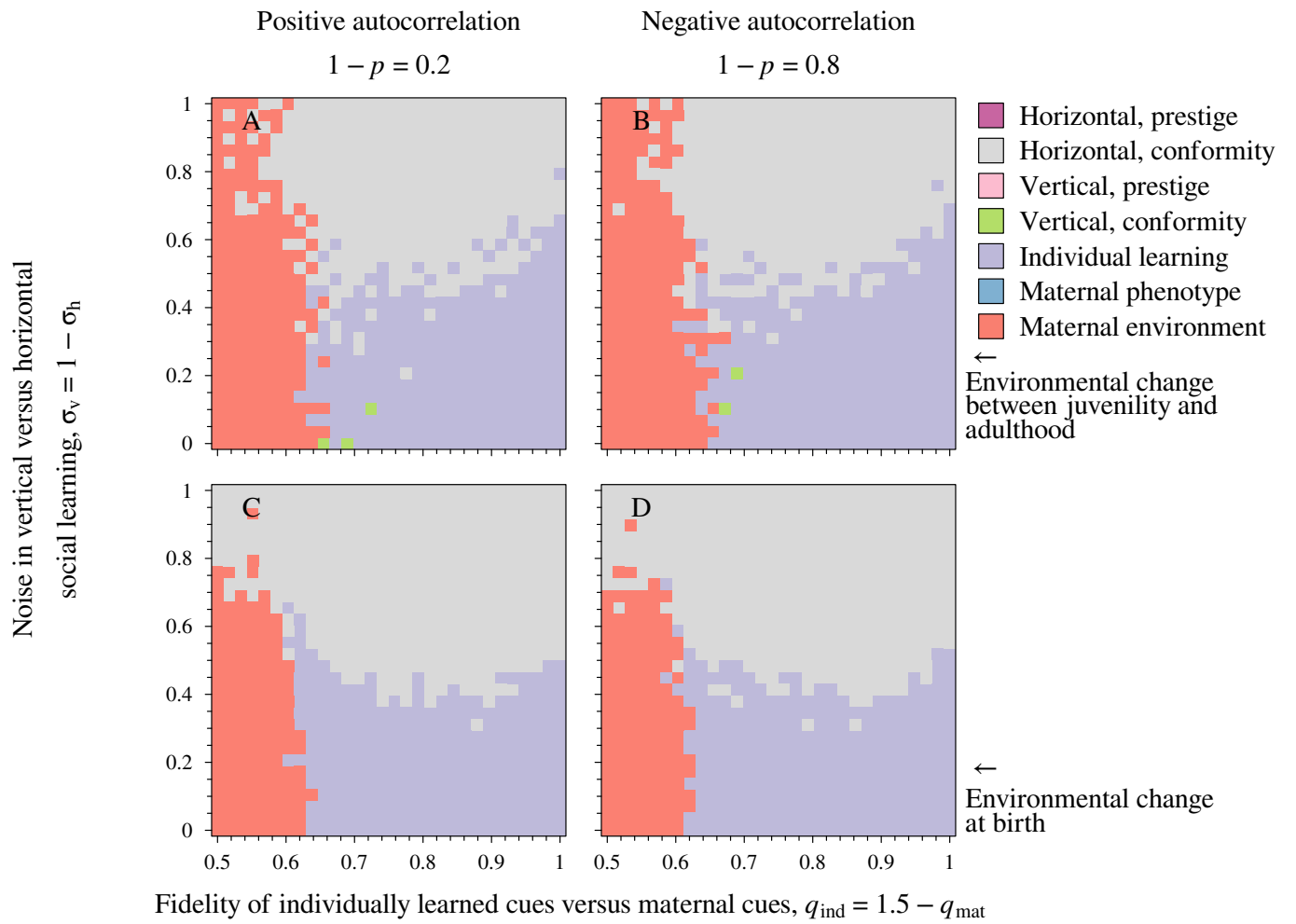Figure S9:

Figure S10 Here we assess the impact of environmental frequencies that are different from $1/2$ on the evolution of cue integration. The figure varies (i) the frequency of the two environments, (ii) the autocorrelation and (iii) the fidelity of individually learned cues (panels A vs B) for the sake of comparison with Figure 2 in the main text. For autocorrelations at or near 0, we find that neither cue is informative (conservative bet-hedging) so that any cue can prevail (see also Figure 2 when $1 - p = 0.5$). By contrast when autocorrelations are positive, we find that horizontal learning based on prestige biases prevails as on the left-hand side in Figure 2B, C. When autocorrelations are negative, we find that vertical social learning typically prevails (in combination with horizontal conformity-based social learning) when the fidelity of individual learning is low (panel A, similar to the right-hand side of Figure 2C), while individual learning prevails otherwise (panel B, similar to the right-hand side of Figure 2B). Hence, these results are very similar to what has been found in Figure 2. In positively autocorrelated environments, horizontal learning based on prestige biases prevails, again as in Figure 2.

To generate this figure, we have replaced the probability of environmental change $1 - p$ used in the main text by environment-specific probabilities of environmental change, so that for an individual patch in environments 1 and 2 the per-generation probabilities of change to a different environmental state are given by $\mathcal{S}_{1 \to 2}$ and $\mathcal{S}_{2 \to 1}$ respectively. Consequently, the expected frequency of environment 1 is given by $\mathcal{S}_{2 \to 1}/(\mathcal{S}_{1 \to 2} + \mathcal{S}_{2 \to 1})$ while the autocorrelation of any local environment between the current generation and the next is given by $1 - (\mathcal{S}_{1 \to 2} + \mathcal{S}_{2 \to 1})$. Combinations in which there is a low frequency of either environment, while a strongly negative autocorrelation are not feasible as they result in values of $\mathcal{S}_{i \to j}$ that are outside of the $[0, 1]$ range, as indicated by white areas.
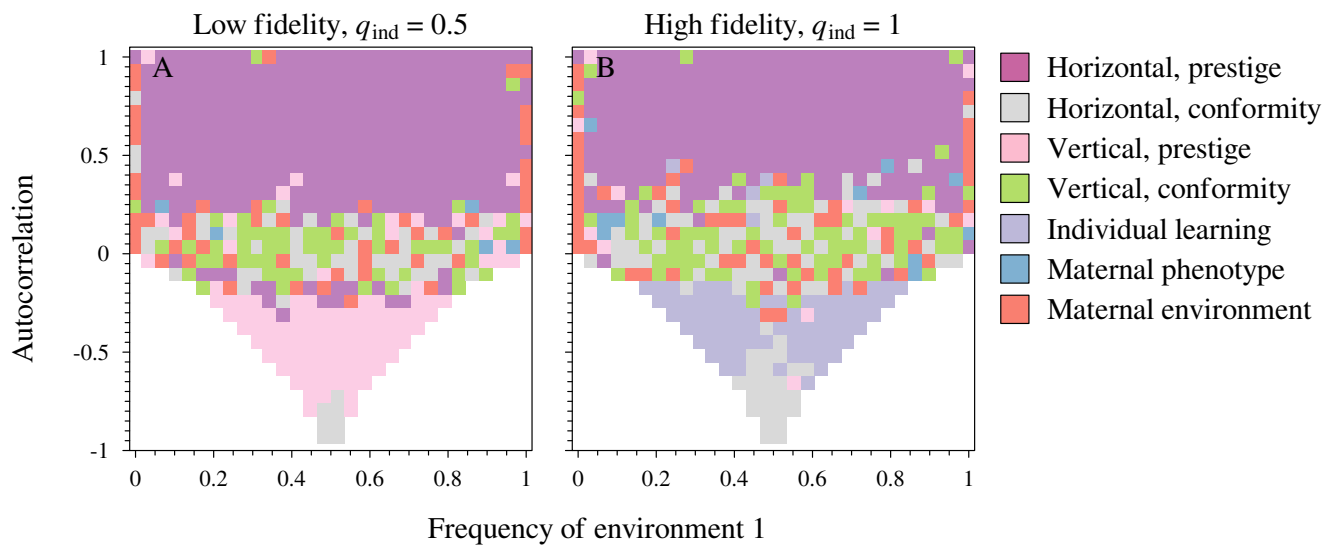
Figure S10:

S2 Supplementary model description

S2.1 Survival selection

We assume that survival selection $S(u, \theta_i)$ of adults is given by the function (see also eqns. [1,2] in [17]):

$$S(u, \theta_i) = \begin{cases} 1 - 0.8u^2 & \theta_i = \theta_{\text{low}} \\ 1 - 0.8(1-u)^2 & \theta_i = \theta_{\text{high}} \end{cases}, \tag{S1}$$

which is the function depicted in Figure 1B.

S2.2 Juvenile phenotype determination

To recap from the main text, the juvenile phenotype u_{juv} that is developed after individual learning is a logistic function of a weighted sum $\bar{x}_{\text{juv}}$ of different cues an individual has received. We have

$$u_{\text{juv}} = \frac{1}{1 + \exp\left(-\bar{x}_{\text{juv}}\right)} \tag{S2}$$

$$\bar{x}_{\text{juv}} = a_{\text{g}} x_{\text{gen}} + a_{\text{ind}} x_{\text{ind}}$$

$$+ m_{\text{m}} x_{\text{mat,phen}} + m_{\text{e}} x_{\text{mat,envt}}$$

$$+ v_{\text{p}} x_{\text{vert,prestige}} + v_c x_{\text{vert,conformity}}, \tag{S3}$$

where the x_is in the equation above are the values of each of the different cues, each of which is weighed by a sensitivity locus that can flexibly evolve. Regarding the evolving sensitivity loci, a_{g} is the sensitivity to the genetic cue x_{gen} (see section S2.2.1 below). The locus a_{ind} reflects evolving sensitivity to individually learned cues x_{ind} that inform about the state of local environment at birth (see section S2.2.2 below). The locus m_{m} reflects evolving sensitivity to the maternal phenotype $x_{\text{mat,phen}}$ as a cue, while the locus m_{e} reflects evolving sensitivity to the maternal environment $x_{\text{mat,envt}}$ as a cue (see section S2.2.3). Finally, the juvenile phenotype is also influenced by vertical

social learning: here, v_p reflects evolving sensitivity to vertically learned phenotypic cues based on prestige biases $x_{\text{vert,prestige}}$ (see section S2.2.4). The locus v_c reflects evolving sensitivity to vertically learned phenotypic cues based on conformity biases (see section S2.2.4).

S2.2.1 Genetic cues

Following similar models by [12, 15, 17], the value of the genetic cue x_{gen} is given by $x_{\text{gen}} = \sum_{i=0}^{n_g} g_i$. That is, it is the sum of the allelic values of $n_g = 3$ unlinked, diploid genetic cue loci g_i. Each allele can have values in the range of $-1 \leq g_i \leq 1$. Consequently, limited dispersal can result in scenarios where alleles become associated with the local environment, so that alleles become informative. Note that alleles of the genetic cue themselves are not involved in local adaptation, they purely have an informational function. The role of genetic cues is further discussed in [13].

S2.2.2 Individually learned cues

Juveniles perform individual learning about the state of their local environments by observing a juvenile environmental cue x_{ind}. To this end, each individual independently observes a binary cue reflecting the state of the local environment. The cue is equal to the actual environmental state (value $-1/2$ in environment θ_{low}, value $1/2$ in environment θ_{high}) with a probability given by the cue fidelity parameter $0.5 \leq q_{\text{juv}} \leq 1$, whereas with probability $1 - q_{\text{juv}}$ an individual receives a cue associated with the opposite environmental state.

S2.2.3 Maternal cues

Juveniles receive two maternal cues: the first is a maternal phenotypic cue $x_{\text{mat,phen}} = u_{\text{ad}}(t-1) - 1/2$ that reflects the value of the maternal phenotype $0 \leq u_{\text{ad}}(t-1) \leq 1$. By subtracting $1/2$ from $u_{\text{ad}}(t-1)$ we standardise the maternal phenotype between $-1/2$ and $1/2$ so that the maternal phenotypic cue has the same range as other plastic cues (and hence evolved cue sensitivities can be directly compared). As $u_{\text{ad}}(t-1)$ is itself a function of her mother's phenotype $u_{\text{ad}}(t-2)$ (a so-called cascading parental effect: [16, 46, 47]), a maternal phenotypic cue can potentially give rise to grandmaternal

766 effects. See [17] for an assessment of noise in maternal phenotypic cues, which we do not consider

767 here for the sake of brevity.

768   The second maternal cue is a maternal environmental cue $x_{mat,envt}$, which reflects the value of

769 the maternal environment [48] as provided to the mother by the adult environmental cue (see Figure

770 1). The cue is equal to the actual environmental state (value $-1/2$ in environment θ_{low}, value $1/2$

771 in environment θ_{high}) with a probability given by the cue fidelity parameter $0.5 \leq q_{mat} \leq 1$, whereas

772 with probability $1 - q_{mat}$ an individual receives a cue associated with the opposite environmental

773 state. We assume that all offspring born form the same mother obtain the same value of $x_{mat,envt}$

774 (i.e., errors act at the level of the brood, rather than at the level of the individual offspring).

S2.2.4   Vertical social learning: prestige biases

776 Following previous models of prestige-based learning [e.g., 49, 50], we assume that individuals

777 are able to evaluate and rank the performance of observed models. To this end, learners rank the

778 potential survival payoffs $S(u_{ad,i})$ of each individual i in a random sample of n_p adult models after

779 survival selection has taken place in the previous generation (see Figure 1). The potential survival

780 payoff is evaluated based on ith model's adult phenotype $u_{ad,i}$ that is observed by the learner (see

781 eq. [3] below) in the learner's environment at time of birth (see Figure 1). The cue $x_{vert,prestige}$ is

782 then a function of the phenotype $u_{ad,max}$ of the sampled individual which has the highest ranked

783 survivorship value, namely $x_{vert,prestige} = u_{ad,max} + \xi_{vert,prestige} - 1/2$. Here $\xi_{vert,prestige}$ is a sample drawn

784 from a Gaussian noise distribution with mean 0 and standard deviation $\sigma_{vert,prestige}$, and subtracting

785 $-1/2$ standardises the cue range to the same scale as for other cues.

S2.2.5   Vertical social learning: conformity-biases

787 We follow previous models of conformity-based social learning [e.g., 41] where individuals eval-

788 uate phenotypes u_i of each individual belonging to randomly chosen subset of n_c surviving adult

789 models from the parental generation in the local patch. Individuals take account of the number of

790 individuals $n_{c,lo} \leq n_c$ with a phenotypes u_i corresponding to the low environment (i.e., $u_i < 0.5$),

th

Figure S11:

whereas $n_{c,hi} = n_c - n_{c,lo}$ reflects the number of individuals with phenotypes $u_i \geq 0.5$. We then have

$$x_{vert,conformity} = \begin{cases} -1/2 & n_{c,lo} > n_{c,hi} \text{ (low-matching phenotype predominates)} \\ 0 & n_{c,lo} = n_{c,hi} \text{ (no phenotype predominates)} \\ 1/2 & n_{c,lo} < n_{c,hi} \text{ (high-matching phenotype predominates)} \end{cases} .$$

Moreover, we add noise to the cue by adding a random deviate from a Gaussian distribution with mean 0 and standard deviation $\sigma_{vert,conformity}$. We have also considered alternative configurations where $x_{vert,conformity} = n_{c,hi}/(n_{c,hi} + n_{c,lo})$, but these give similar results (results not shown).

S2.3 Adult phenotype determination

To recap from the main text, we have

$$u_{ad} = \frac{1}{1 + \exp(-\bar{x}_{ad})} \tag{S4}$$

$$\bar{x}_{ad} = \bar{x}_{juv} + h_p x_{horiz,prestige} + h_c x_{horiz,conformity}, \tag{S5}$$

where h_p (horizontal social learning; prestige bias) and h_c (horizontal social learning; conformity bias) again reflect unlinked and evolving diploid loci (bounded between [-10,10]) that reflect sensitivity to both horizontally learned social cues $x_{horiz,prestige}$ and $x_{horiz,conformity}$ respectively.

S2.3.1 Horizontal social learning: prestige biases

Horizontally learning juveniles rely on the same mechanism as vertical social learning, where they rank the potential survival payoffs $S(u_{juv,i})$ of each individual i in a random sample of n_p juvenile models (no survival selection has yet taken place; see Figure 1). The potential survival payoff is evaluated based on i model's juvenile phenotype $u_{juv,i}$ (see eq. S2) that is observed by the juvenile

S23

810   learner in the learner's environment at time of birth (see Figure 1). The cue $x_{\text{horiz,prestige}}$ is then a

811   function of the phenotype $u_{\text{juv,max}}$ of the sampled individual which has the highest ranked survivor-

812   ship value, namely $x_{\text{horiz,prestige}} = u_{\text{juv,max}} + \xi_{\text{horiz,prestige}} - 1/2$. Here $\xi_{\text{horiz,prestige}}$ is a sample drawn from

813   a Gaussian noise distribution with mean 0 and standard deviation $\sigma_{\text{horiz,prestige}}$, and subtracting $-1/2$

814   standardises the cue range to the same scale as for other cues.

S2.3.2   Horizontal social learning: conformity-biases

816   We follow previous models of conformity-based social learning [e.g., 41] where individuals eval-

817   uate phenotypes u_i of each individual belonging to randomly chosen subset of n_c juvenile models

818   from the current generation in the local patch. Individuals take account of the number of individu-

819   als $n_{c,\text{lo}} \leq n_c$ with a phenotypes u_i corresponding to the low environment (i.e., $u_i < 0.5$), whereas

820    $n_{c,\text{hi}} = n_c - n_{c,\text{lo}}$ reflects the number of individuals with phenotypes $u_i > 0.5$. We then have

$$
x_{\text{horiz,conformity}} = \begin{cases} -1/2 & n_{c,\text{lo}} > n_{c,\text{hi}} \text{ (low-matching phenotype predominates)} \\ 0 & n_{c,\text{lo}} = n_{c,\text{hi}} \text{ (no phenotype predominates)} \\ 1/2 & n_{c,\text{lo}} < n_{c,\text{hi}} \text{ (high-matching phenotype predominates)} \end{cases} .
$$

823   Moreover, we add noise to the cue by adding a random deviate from a Gaussian distribution with

824   mean 0 and standard deviation $\sigma_{\text{horiz,conformity}}$.

S2.4   Inheritance

826   For the sake of simplicity, the gene loci coding for the cue sensitivities a_i, m_i, h_i and v_i and the ge-

827   netic cue loci g_i are all considered diploid, autosomal and unlinked. Upon inheritance, alleles at all

828   loci independently mutate with probability $\mu = 0.01$, which involves adding a random value drawn

829   from a Laplace(μ, b) distribution with parameters $\mu = 0$ and $b = \sigma_\mu / \sqrt{2} = 0.0141$, corresponding to

830   a mean of 0 and a variance of $\sigma_\mu^2 = 0.0004$. To ensure the genetic cue locus can accumulate suffi-

831   cient genetic variation in, we assume an increased mutational variance $\sigma_{\mu_g}^2 = 0.0625$ for the alleles

832 at the genetic cue locus g_i.