



Speciation by perception

ANDERS BRODIN & FREDRIK HAAS

Department of Ecology, Lund University

(Received 4 July 2005; initial acceptance 10 August 2005;
final acceptance 11 October 2005; published online 24 May 2006; MS. number: 8601R)

During sympatric and parapatric speciation the diverging populations maintain geographical contact throughout the speciation process. In theoretical models an isolating mechanism, such as assortative mating, is usually needed to obtain genetic separation of populations in contact. We simplified the visual appearance of the crows that occur in the crow hybrid zone in Europe, the carrion crow, *Corvus corone*, hooded crow, *C. cornix*, and hybrid crows. We then mimicked the sexual imprinting process by training artificial neural networks to separate their own type from the other two. The network learned pure phenotypes faster and better than the hybrid patterns showing that already at the receptor level there may be signal reception properties that will make speciation under sympatric and parapatric conditions possible. Thus, mate choice preference can be a self-emerging nervous system property not requiring evolutionary selection.

© 2006 The Association for the Study of Animal Behaviour. Published by Elsevier Ltd. All rights reserved.

Allopatric speciation occurs when diverging populations have become geographically separated. Since speciation occurring under such conditions is easy to understand, it has long been the dominating theory (Mayr 1963), but recently models of sympatric and parapatric speciation have received increasing attention (reviewed in Turelli et al. 2001; Kirkpatrick & Ravigné 2002; Coyne & Orr 2004). Sympatric speciation will occur if two populations separate genetically even though they overlap spatially. The strongest argument against this type of speciation has been that it relies on prezygotic isolation to prevent gene flow between the diverging populations, and that such isolation would rarely evolve as long as the populations remain in contact (Mayr 1963).

Parapatric speciation is intermediate between allopatric and sympatric speciation in the sense that the diverging populations meet and hybridize in a narrow hybrid zone. Parapatric speciation is more likely than sympatric speciation (Endler 1977; Turelli et al. 2001; Gavrillets 2004; Polechkova & Barton 2005) but has still been criticized, with the main objection being that gene flow from conspecifics outside the hybrid zone will swamp the effects of selection against hybridization within the zone (Bige-low 1965). Sympatric speciation, but to some extent also parapatric speciation, would be easier to understand if

some kind of prezygotic isolation mechanism commonly occurred under these conditions.

During the last 25 years, at least 100 models that have considered prezygotic isolation mechanisms during speciation have been published (Kirkpatrick & Ravigné 2002). Of these, 34 have focused on assortative mating, that is, when similar phenotypes mate preferentially with each other. The role and evolution of assortativeness in speciation have been analysed under both sympatric and parapatric conditions (e.g. Kondrashov & Shpak 1998; Cain et al. 1999; Dieckmann & Doebeli 1999; Servedio 2000; Doebeli & Dieckmann 2003; Gavrillets 2004; Polechkova & Barton 2005). The empirical evidence for assortative mating among sympatric and parapatric populations is strong. Howard (1993) reported that of 37 hybridizing taxa at least 19 show evidence for assortative mating.

Sexual imprinting is a learning process by which young individuals learn species-specific characteristics that enable them to choose conspecific mates as adults (Bateson 1966; Immelmann 1972; Clayton 1990). The term has primarily been used in birds (Price 1998; ten Cate & Vos 1999) but similar processes occur in other taxa, for example mammals (Kendrick et al. 1998; Penn & Potts 1998). During this process the parents and siblings are normally used as templates (e.g. Seiger 1967; Immelmann 1975; Irwin & Price 1999) and it is probably the most important mechanism for development of mate preferences in birds (ten Cate et al. 1993). Whenever imprinting reduces hybridization it may have a key role in speciation (Irwin & Price 1999). Laland (1994) found that sexual imprinting can act as a barrier to gene flow, but is unlikely to lead to speciation by itself.

Correspondence: A. Brodin, Department of Theoretical Ecology, Ecology Building, Lund University, S-223 62 Lund, Sweden (email: anders.brodin@teorekol.lu.se).

In Europe, the crow *Corvus corone/cornix* occurs in two distinctly separate types, the all black carrion crow, the *corone* type, and the grey and black hooded crow, the *cornix* type. There is some disagreement on whether these should be regarded as full species or as subspecies. Here, we refer to them as types or phenotypes. The types meet and interbreed in a 50–150-km-wide hybrid zone that stretches from Great Britain to northern Italy (Meise 1928). Hybrids are common inside the zone and these have a more or less shaggy appearance with black areas and spots distributed on the parts that are clean grey on the hooded crow (Saino & Villa 1992; Fig. 1). In relation to the maximum dispersal distances for young crows (Siefke 1994), the hybrid zone is relatively narrow and the shape of the cline seems to be stable over time (Haas & Brodin 2005). This suggests that it is a so-called tension zone (Barton & Hewitt 1985) with reduced fitness for hybrids.

As far as we know, only pure phenotypes (i.e. nonhybrids) have been used as templates in studies of sexual imprinting (e.g. ten Cate et al. 1993; Irwin & Price 1999; Slagsvold et al.

2002). We compared the consequences for the learning process of having pure or hybrid phenotypes as templates. We did this by mimicking the sexual imprinting process in a neural network simulation in which we trained the networks on stylized representations of the crow phenotypes. In reality, imprinting may be much more complex than in our simplified model. For example, visual impression may be combined with acoustic and chemical signals (Irwin & Price 1999; Dukas 2004) and social experience with nonrelatives may be important (Freeberg 1996). Since we are primarily interested in the principle rather than in exactly mimicking imprinting, we believe that our approach may produce important insights.

METHODS

To ensure that we trained artificial neural networks (ANNs) on the features that we considered most relevant, that is, colour patterns rather than, for example, silhouettes, we simplified the visual representation of the crows

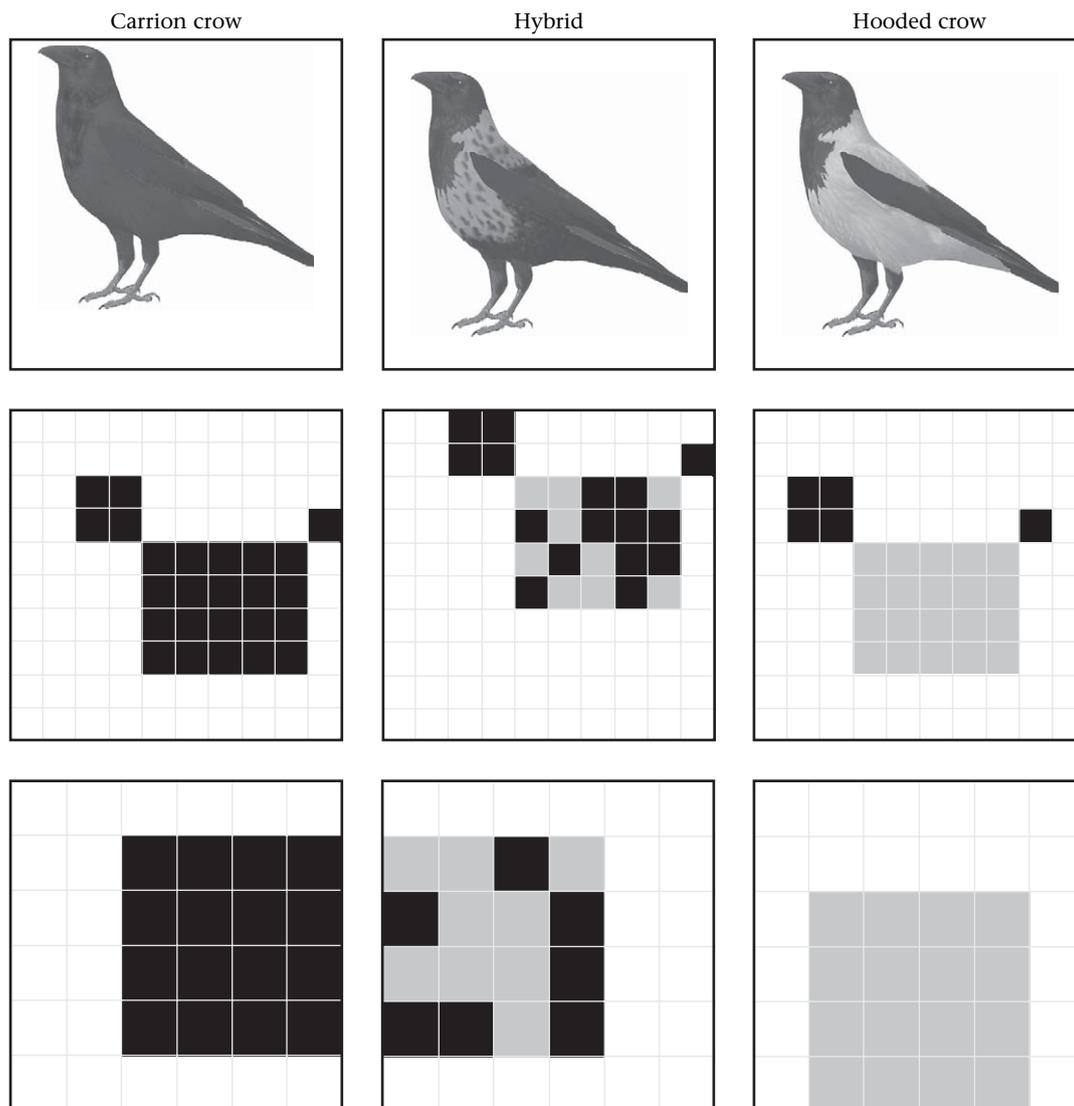


Figure 1. Carrion, hybrid and hooded crows and two levels of simplified representations of these that were used as input to the artificial neural network.

to the stylized representations shown in the central and bottom rows of Fig. 1. This had the additional benefit of reducing the requirements for computer memory compared to using the realistic pictures of crows shown in the top row. To verify that this simplification would not corrupt our results and conclusions, we also trained networks on pictures that were more crow-like than the ones in the central row (to decrease simulation time we did not train networks on pictures that were as detailed as the ones in the top row). Since an increase in the resolution of the pictures strengthened our conclusions (see below), we considered it safe to use the simplest representations (6×6 input neurons) in most simulations.

Our ANNs consisted of one layer of input neurons, one layer of hidden neurons and one output neuron (Fig. 2). The pictures that we presented as input to the networks consisted of square grids constructed of smaller square cells. Each cell was then the signal to an input neuron. The pictures in the central row of Fig. 1 corresponded to 100 (10×10) input neurons and each picture in the bottom row to 36 (6×6) neurons. If n_1 was the number of input neurons, we restricted the hidden layer to consist of $\sqrt{n_1}$ cells. Too many hidden neurons will make networks prone to overtraining (Haykin 1999, see below for an explanation of this term).

The representation of the crow was a smaller part of the picture that constituted the input to the network. This made it possible for us randomly to alter the position of the crows between pictures (Fig. 1). For example, in

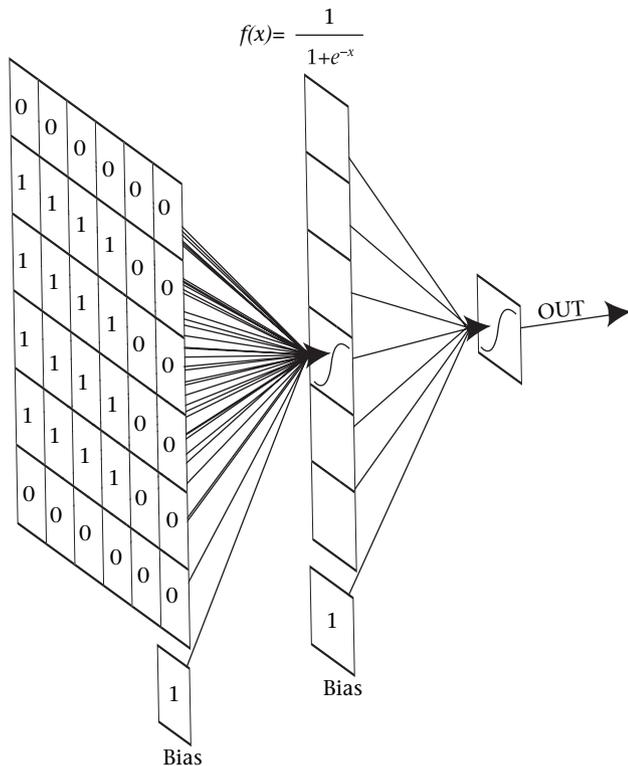


Figure 2. An artificial neural network with an input layer representing a carrion crow (the 4×4 square filled with ones corresponds to a larger black square), one hidden layer and one output layer. All incoming signals to a given neuron are summed up in a sigmoid activation function. See text for details.

a picture consisting of a grid of 6×6 cells the ‘crow’ was a 4×4 square (Fig. 1, bottom row) that could occur in any position in the 6×6 grid (Fig. 1, bottom row). Variation between pictures is necessary to obtain generalization ability in networks and avoid overtraining. An overtrained network may have learnt a specific picture or pattern almost perfectly without any generalization ability (Haykin 1999). Transferred to the crow recognition problem, this corresponds to a crow chick that could recognize its parent only if it were sitting exactly in a specific position at a specific distance.

Cells could be white (empty), grey or black. We assigned no input signal (0) to represent empty cells, a signal of -1 to represent the grey cells and 1 to the black cells. The reason for the choice of these numbers was to ensure that the black and grey cells would be an equal numerical distance from 0 to balance the network. Such a balanced network will have equal ‘learning power’ for grey and black signals. Such balance around a mean of 0 is a desirable property for the input layer (Haykin 1999).

Our network was fully connected which means that each neuron was connected to all neurons in the next layer (Fig. 2). Connections in neural networks are called weights and consist of decimal numbers. A propagated signal $A_i w_i$ coming from neuron i will be the product of the weight w and the signal A_i , with A being a vector of the signals from the previous layer. All incoming signals to a neuron are then summed, making the total incoming signal $x = \sum_{i=0}^n A_i w_i$. Here n is the number of neurons in the previous layer. In the hidden layer, A would thus be a vector of inputs from a crow picture. Note that $i = 0$ symbolizes the occurrence of an extra cell in the input and hidden layers. Such extra cells are called biases and these always fire a signal of 1. Graphically, learning in an ANN means that the network obtains separation capacity, which in our case can be seen as drawing lines or curves between the crows and the background, and biases are necessary to obtain such high capacity (Haykin 1999).

As an activation function for the output signal we used the logistic function $f(x) = 1 / (1 + e^{-ax})$ which produces a sigmoid response. Compared to a simpler step function it makes it easier to determine in which direction weights should be adjusted during training. A step function, on the other hand, would produce identical output signals (for example 0) for many different sets of weights. The logistic function is also the most common activation function in this type of application (e.g. Kamo et al. 1998; Haykin 1999). Before training the network, we assigned random values between -5.0 and 5.0 to all weights and we set a to 1 to adjust the slope of the sigmoid to the range of the values of the weights (Haykin 1999).

We trained the network in supervised mode, meaning that the output result was compared to a training criterion. Since the activation function will produce output between 0 and 1, we set the target output to 1 for images of the correct parental type and 0 for the other two types. During one training session we exposed the network to a batch of 700 ‘crows’ of which 350 were of the parental (correct) type and 350 randomly chosen out of the other two (erroneous) types. We updated the values of the weights after each batch.

We estimated deviations from the target output (1 or 0) as mean squared error (MSE). This means that we summed the squared error from a batch of 700 exposures and divided it by 700. The maximum error will then be 1, and an MSE of 0.2 means that the actual mean error is approximately 0.45. Depending on the sigmoid nature of the activation function, the output signals will approach 0 and 1 asymptotically but never become exactly 0 or 1. We measured the performance during training as the change in MSE. Training of a network means that the weights are updated after which the output from the new network is compared with that from the previous set of weights. The weights that produce the lowest MSE are saved, whereafter the weights are updated again according to some learning rule. Random assignments of weights would be inefficient and the role of the learning rule is to show in which direction the weights should be adjusted. As mentioned above, such directional change is facilitated by a sigmoid activation function.

To adjust weights in an optimal way we used a genetic algorithm, GA, to train the ANN. This is a powerful learning rule for a network of this type and less elaborate versions of this technique have been used in earlier ANN simulations in behavioural biology (Arak & Enquist 1993; Enquist & Arak 1993; Johnstone 1994; Kamo et al. 1998). The GA means that the weights of the network are inserted as genes in a 'chromosome'. Since we used a real parameter GA (Wright 1991) we could use the decimal numbers of the weights directly as genes. To provide sufficient variation to find optimal solutions, we created a population of 48 chromosomes that competed in a standard GA evolution procedure.

After the initial assignment of random values as weights, the steps were as follows.

(1) 'Fitness evaluation', that is, we calculated the MSE by comparing the output with the criterion output. We considered the chromosome with the lowest MSE to have the highest fitness.

(2) Sorting of chromosomes according to fitness rank from 1 to 48.

(3) Pairing and recombination in which we allowed the 10 best chromosomes to produce offspring by crossing over. We paired the highest-ranked chromosome with the second best, the third with the fourth, etc. We chose the crossing-over point randomly in the chromosomes. We used a crossing-over procedure called linear blending (Wright 1991, see below for an explanation) that could create new weights outside the starting range (-5.0 to 5.0) to ensure that the range was wide enough.

(4) In this linear blending procedure, each pair of chromosomes would produce three offspring (see below) but, commonly, crossing over between two parental chromosomes will produce two new ones. Thus, we randomly picked two of the three offspring and inserted them in the population.

(5) To keep the number of chromosomes constant at 48 we discarded the 10 lowest-ranking ones before we inserted 10 new offspring.

(6) Another sorting according to fitness made sure that the new individuals would be in correct rank position.

(7) All but the 10 highest-ranked chromosomes were exposed to a random mutation probability (baseline value 2%) for all genes. If a gene was selected for mutation we gave it a new random value between -5.0 and 5.0. Such mutations will increase variation by creating new weights that have no relation to the original set of weights. Crossing over, on the other hand, will create new weights that have some relation to the original set of weights.

Analogous to real mutations, the mutation procedure will in many cases mean that nonoptimal genes are inserted in the population. These genes will not 'invade' the best solutions since these are preserved in the 10 highest-ranked chromosomes that are mutation free. Because of this, we calculated the result as the mean of the 10 top-ranking chromosomes. We ran each simulation (including both training and testing) three times and used the average of these runs to produce Figs 3 and 4. This gave us a total of 3×30 optimal networks. Since there were some small variations between the runs we evaluated the test session with a general linear model (GLM) ANOVA using the three crow categories as groups.

A well-designed genetic algorithm should search a wide 'solution space' at the same time as it preserves the best solutions intact. An algorithm that is too restricted may get stuck on local optima. This means that there must be a balance between production of new genes (crossing over

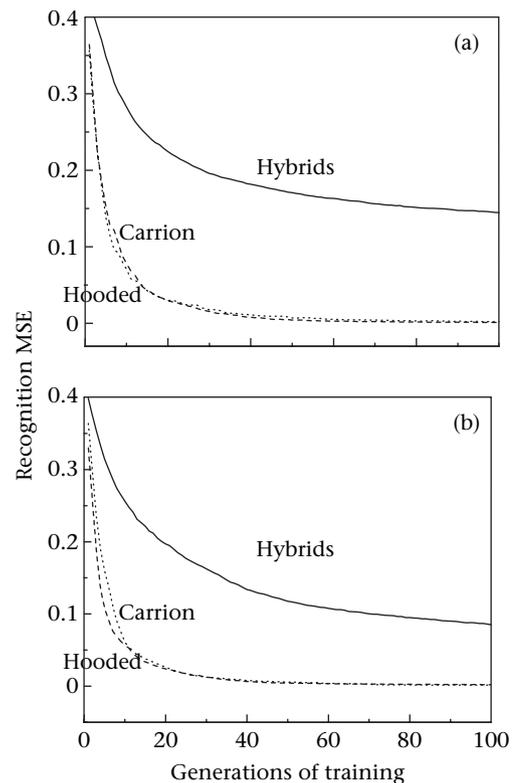


Figure 3. Change in mean square error (MSE) of the artificial neural network over 100 generations of training for the genetic algorithm for carrion (dashed curve), hooded (dotted curve) and hybrid (solid curve) crow patterns. (a) The learning curve for pictures with 14×14 squares. (b) The learning curves for the lower resolution of 6×6 squares.

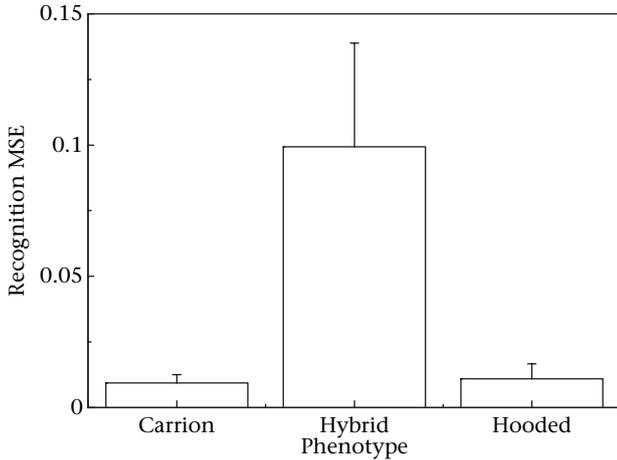


Figure 4. The probability of making an error (mean square error, MSE, + SD) in the separation of own phenotype from the other two when trained networks were exposed to 300 unfamiliar pictures of crows.

and mutation) at the same time as the treatment of the best chromosomes should be conservative. Most previous similar simulations used simpler GAs to train the networks (Arak & Enquist 1993; Enquist & Arak 1993; Johnstone 1994). The one that is most similar to ours was used by Kamo et al. (1998) but their algorithm produced new genes only during mutation (step 7 above), while we also produced new genes during reproduction (step 3 above).

The crossing-over procedure in a continuous GA is not as straightforward as in a binary GA. Simple one-point crossover between paired chromosomes is not a good alternative, since this would mean that the same decimal numbers would be shuffled around instead of new genes being created. Averaging between the values in the parental genes is also not good, since this will reduce the variation over sessions. We used a standard linear blending procedure suggested by Wright (1991) that will maintain variation between the offspring at the same time as it produces new decimal numbers. In this procedure, the two parental genes p_1 and p_2 will produce three offspring, o_i :

$$o_1 = 0.5p_1 + 0.5p_2$$

$$o_2 = 1.5p_1 - 0.5p_2$$

$$o_3 = 1.5p_2 - 0.5p_1.$$

In many cases this procedure will create values outside the range of the weights, -5.0 and 5.0 , so the algorithm's 'search space' will be wide.

We cycled the algorithm for at least 100 generations. After training had finished we exposed the trained networks to 300 new pictures in a test session. Just as during the training sessions, half of these were of the correct parental type and half of them were randomly chosen from the other two types. To verify that our results were solid we tested the sensitivity of the algorithm in

several ways. For example, we increased the number of hidden neurons up to two times the baseline value, we increased the number of generations to 500 and increased mutation rates up to two times the baseline 2%. We also varied the size of the 'crows' between pictures and rotated and turned the 'crows' in Fig. 1 by 180° . These manipulations had very small effects.

RESULTS AND DISCUSSION

The networks learned to recognize the pure patterns faster and more accurately than they did hybrid patterns (Fig. 3). The learning curves for networks trained on the two pure patterns were almost identical suggesting that they were equally easy to learn. This is important, since beforehand it was possible that the all black types would be easier to learn than the ones representing hooded crows. The difference in learning ability between the pure forms and the hybrids increased as the resolution of the pictures increased, that is, if we used more realistic representations of crows (Fig. 3a, b). This suggests that our simplifications of the crow representations were conservative since they decreased the difference between the crow types. Overall, however, the results were similar regardless of resolution. The effect of varying size and orientation of the crow pictures within training sessions was hardly discernible (not shown). The robustness of the results suggests that the learning principle underlying the results is strong and general.

Networks that had been trained on the pure patterns made fewer mistakes than the ones trained on hybrids when they were tested on unfamiliar 'individuals' (GLM, ANOVA: $F_{2,87} = 148.3$, $P < 0.001$; Fig. 4). This difference shows how accurate the networks were in identifying their own parental type. This finding confirms that the difference in learning ability after training (Fig. 3) will be manifested also when new, unfamiliar individuals are encountered. More importantly, it suggests that this difference may have considerable effect on mate choice preference and that neural receptors will prefer clean uniform patterns to hybrid ones. Since the errors are squared and the maximum error is 1.0, the probability that 'hybrid networks' will choose 'nonself' types may be considerable.

In our specific crow example it seems that individuals of the pure phenotypes will be more prone than hybrid offspring to choose partners of their own kind. Imagine, for example, a hybrid zone where the population consist of equal proportions of all three phenotypes. Then hybrids would be at a disadvantage since they would be preferred as mates by a smaller proportion of the population than the other types.

There is empirical evidence suggesting that sexual imprinting on visual patterns may produce nonadaptive, or even maladaptive, partner choice. There are three distinct colour morphs of the European buzzard, *Buteo buteo*: light, dark and intermediate. Both the light and the dark morphs would maximize fitness by choosing partners of the opposite type but, instead, they prefer their own type even though such matings result in homozygotic offspring with low fitness (Krüger et al. 2001).

Krüger et al explained this partner choice by sexual imprinting on the mother morph. This suggests that 'non-evolved' sexual imprinting may have strong effects.

A common assumption is that assortative mating is evolved, that is, it originates from some genetically based process, for example selection on mate preferences (e.g. Price 1998; Polechkova & Barton 2005) or selection for recognition of conspecifics (Irwin & Price 1999). The nature of the genetic basis of assortative mate choice is important since speciation under parapatric conditions may be impossible if mate preference is genetically coded in many loci (Gavrilets 2004). It is even more difficult to imagine that sympatric speciation could occur under such conditions. In fact, the assumption that a complex trait such as mate choice ought to be coded by multiple loci was Mayr's (1947) main argument against sympatric speciation. In this simulation we have shown that assortative mating may depend on a nongenetic learning process that is so simple that it may be a self-emerging property of signal perception.

There is an increasing awareness that not only imprinting, but also other forms of learning, may be important for assortative mating (Kondrashov & Mina 1986; Beltman et al. 2004). Several forms of learning have been discussed in the context of assortativeness and speciation, for example learning of conspecific signalling in adult brown-headed cowbirds, *Molothrus ater* (Freeberg 1996) and fruit flies, *Drosophila melanogaster* (Dukas 2004) and learning of characteristics of newly colonized niches (Beltman et al. 2004). In birds, however, but maybe also in mammals, sexual imprinting in young individuals must be the most important form of learning in this context (Price 1998; Irwin & Price 1999; Owens et al. 1999). Here we suggest that sexual imprinting may depend on the nervous system being prone to learn some patterns better than others.

Previous speciation models that incorporate assortative mating (reviewed in Polechkova & Barton 2005) require stricter assumptions than ours. Felsenstein (1981) suggested that genes that could promote premating isolation must either be linked to genes for postmating isolation or to the traits that already differ between the populations. Some models, for example Felsenstein's, include evolution of 'preference genes', whereas others suggest that assortative mating may be a by-product of habitat preferences (Rice 1987), bimodal resource distribution (Doebeli 1996) or, in accordance with our model, sexual imprinting (Irwin & Price 1999; Slabbekoorn & Smith 2002; Servodio & Noor 2003).

It has been questioned whether prezygotic isolation barriers can evolve in hybrid zones. The argument against this has been that genes coding for selectivity only will be advantageous inside the zone, but not outside (Moore 1957). Genes from areas outside the zone would then swamp the effects of selection within the zone since only small fractions of the populations occur in it (Bige-low 1965). Finally, genes selected to avoid the cost of hybridizing could be disadvantageous outside the zone (Barton & Hewitt 1981). All these arguments lose strength if assortative mating is based on a simple nongenetic process of the type we present here.

Our simulation suggests that assortative mating can develop relatively easily. In that case, not only parapatric but also sympatric speciation may be more common than generally believed. Sympatric and parapatric speciation models in an adaptive dynamics setting (Dieckmann & Doebeli 1999; Doebeli & Dieckmann 2003) have recently been extensively criticized (e.g. Gourbiere & Mallet 2005; Polechkova & Barton 2005; Waxman & Gavrilets 2005). One of the main objections has been that the assortativeness in these models will be transient and hence disruptive selection will not be maintained (Polechkova & Barton 2005). Clearly, a mechanism of the type we propose here will be stable over time. Given the existence of a polymorphic character prominent enough to be perceived, imprinting may provide an automatic mechanism promoting assortativeness. The tendency for neural receptors to learn uniform signals more easily than 'shaggy' ones is thus a simple mechanism that may reduce gene flow between diverging groups.

We do not suggest that mate choice, habitat preference, selection against hybrid unfitness or reinforcement may not be genetically based. On the contrary, we consider it likely that imprinting of the type we suggest here may be the starting point for a gradual development of disruption by selection. After an initial effect of imprinting, the strength of disruptive selection could increase through reinforcement and selection against hybrid unfitness (Dobzhansky 1940; Felsenstein 1981). In addition, nothing restricts the imprinting process we suggest here from interacting with other factors that facilitate the evolution of assortative mating.

Our simulation does not say anything about the strength of selection. For European crows, assortative mating appears to be strong enough to maintain the zone, but not to obtain full isolation between the two types. With no assortative mating, the hybrid zone would gradually widen and some sort of hybrid phenotype could take over. Our results suggest that this may be prevented by the fact that assortative mating is not as strong in hybrids as in the pure phenotypes.

In conclusion, our results show that already at the signal reception level hybrids may be at a disadvantage even if natural selection does not reduce fitness directly. The example of hooded and carrion crows is easy to envisage since it is graphically illustrative, but the same principle could apply to many cases when mate choice depends on signalling between the sexes.

Acknowledgments

We thank Lars Hård and Mattias Ohlsson for helping us with the neural networks. A.B. was financed with a grant from the Swedish Research Council, VR.

References

- Arak, A. & Enquist, M. 1993. Hidden preferences and the evolution of signals. *Philosophical Transactions of the Royal Society of London, Series B*, **340**, 207–213.
- Barton, N. H. & Hewitt, G. M. 1981. Hybrid zones and speciation. In: *Evolution and Speciation: Essays in Honor of M.J.D. White* (Ed.

- by W. R. Atchley & D. S. Woodruff), pp. 109–145. Cambridge: Cambridge University Press.
- Barton, N. H. & Hewitt, G. M.** 1985. Analysis of hybrid zones. *Annual Review of Ecology and Systematics*, **16**, 113–148.
- Bateson, P. P. G.** 1966. The characteristics and context of imprinting. *Biological Reviews*, **41**, 177–220.
- Beltman, J. B., Haccou, P. & ten Cate, C.** 2004. Learning and colonization of new niches: a first step toward speciation. *Evolution*, **58**, 35–46.
- Bigelow, R. S.** 1965. Hybrid zones and reproductive isolation. *Evolution*, **19**, 449–458.
- Cain, M. L., Andreassen, V. & Howard, D. J.** 1999. Reinforcing selection is effective under a relatively broad set of conditions in a mosaic hybrid zone. *Evolution*, **53**, 1343–1353.
- Clayton, N. S.** 1990. Assortative mating in zebra finch subspecies, *Taeniopygia guttata guttata* and *T. g. castanotis*. *Philosophical Transactions of the Royal Society of London, Series B*, **330**, 351–370.
- Coyne, J. A. & Orr, H. A.** 2004. *Speciation*. Sunderland, Massachusetts: Sinauer.
- Dieckmann, U. M. & Doebeli, M.** 1999. On the origin of species by sympatric speciation. *Nature*, **400**, 354–357.
- Dobzhansky, T.** 1940. Speciation as a stage in evolutionary divergence. *American Naturalist*, **74**, 312–321.
- Doebeli, M.** 1996. A quantitative genetic competition model for sympatric speciation. *Journal of Evolutionary Biology*, **9**, 893–909.
- Doebeli, M. & Dieckmann, U. M.** 2003. Speciation along environmental gradients. *Nature*, **421**, 259–264.
- Dukas, R.** 2004. Male fruit flies learn to avoid interspecific courtship. *Behavioral Ecology*, **15**, 695–698.
- Endler, J. A.** 1977. *Geographic Variation, Species and Clines*. Princeton, New Jersey: Princeton University Press.
- Enquist, M. & Arak, A.** 1993. Selection of exaggerated male traits by female aesthetic senses. *Nature*, **361**, 446–448.
- Felsenstein, J.** 1981. Skepticism towards Santa Rosalia, or why are there so few kinds of animals? *Evolution*, **35**, 124–138.
- Freeberg, T.** 1996. Assortative mating in captive cowbirds is predicted by social experience. *Animal Behaviour*, **52**, 1129–1142.
- Gavrilets, S.** 2004. *Fitness Landscapes and the Origin of Species*. Princeton, New Jersey: Princeton University Press.
- Gourbiere, S. & Mallet, J.** 2005. Has adaptive dynamics contributed to the understanding of adaptive and sympatric speciation? *Journal of Evolutionary Biology*, **18**, 1201–1204.
- Haas, F. & Brodin, A.** 2005. The crow *Corvus corone* hybrid zone in southern Denmark and northern Germany. *Ibis*, **147**, 649–656.
- Haykin, S.** 1999. *Neural Networks: A Comprehensive Foundation*. Englewood Cliffs, New Jersey: Prentice-Hall.
- Howard, D. J.** 1993. Reinforcement: origin, dynamics, and fate of an evolutionary hypothesis. In: *Hybrid Zones and the Evolutionary Process* (Ed. by R. G. Harrison), pp. 46–69. New York: Oxford University Press.
- Immelmann, K.** 1972. Sexual and other long-term aspects of imprinting in birds and other species. *Advances in the Study of Behaviour*, **4**, 147–174.
- Immelmann, K.** 1975. Ecological significance of imprinting and early learning. *Annual Review of Ecology and Systematics*, **6**, 15–37.
- Irwin, D. E. & Price, T.** 1999. Sexual imprinting, learning and speciation. *Heredity*, **82**, 347–354.
- Johnstone, R. A.** 1994. Female preference for symmetrical males as a by-product of selection for mate recognition. *Nature*, **372**, 172–175.
- Kamo, M., Kubo, T. & Iwasa, Y.** 1998. Neural network for female mate preference, trained by a genetic algorithm. *Philosophical Transactions of the Royal Society of London, Series B*, **353**, 399–406.
- Kendrick, K. M., Hinton, M. R., Atkins, K., Haupt, M. A. & Skinner, J. D.** 1998. Mothers determine sexual preferences. *Nature*, **395**, 229–230.
- Kirkpatrick, M. & Ravigné, V.** 2002. Speciation by natural and sexual selection: models and experiments. *American Naturalist*, **159**, S22–S35.
- Kondrashov, A. S. & Mina, M. V.** 1986. Sympatric speciation: when is it possible? *Biological Journal of the Linnean Society of London*, **27**, 201–223.
- Kondrashov, A. S. & Shpak, M.** 1998. On the origin of species by means of assortative mating. *Proceedings of the Royal Society of London, Series B*, **265**, 2273–2278.
- Krüger, O., Lindström, J. & Amos, W.** 2001. Maladaptive mate choice maintained by heterozygotic advantage. *Evolution*, **55**, 1207–1214.
- Laland, K. N.** 1994. On the evolutionary consequences of sexual imprinting. *Evolution*, **48**, 477–489.
- Mayr, E.** 1947. Ecological factors in speciation. *Evolution*, **1**, 263–288.
- Mayr, E.** 1963. *Animal Species and Evolution*. Cambridge, Massachusetts: Harvard University Press.
- Meise, W.** 1928. Die Verbreitung der Aaskrähe (Formenkreis *Corvus corone* L.). *Journal für Ornithologie*, **76**, 1–203.
- Moore, J. A.** 1957. An embryologist's view of the species concept. In: *The Species Problem* (Ed. by E. Mayr), pp. 325–338. Washington, DC: American Association for the Advancement of Science.
- Owens, I. P. F., Rowe, C. & Thomas, A. L. R.** 1999. Sexual selection, speciation and imprinting: separating the sheep from the goats. *Trends in Ecology and Evolution*, **14**, 131–132.
- Penn, D. & Potts, W.** 1998. MHC-disassortative mating preferences reversed by cross-fostering. *Proceedings of the Royal Society of London, Series B*, **265**, 1299–1306.
- Polechkova, J. & Barton, N. H.** 2005. Speciation through competition: a critical review. *Evolution*, **59**, 1194–1210.
- Price, T.** 1998. Sexual selection and natural selection in bird speciation. *Philosophical Transactions of the Royal Society of London, Series B*, **353**, 251–260.
- Rice, W. R.** 1987. Speciation via habitat specialization: the evolution of reproductive isolation as a correlated character. *Evolutionary Ecology*, **1**, 301–314.
- Saino, N. & Villa, S.** 1992. Pair composition and reproductive success across a hybrid zone of carrion crows and hooded crows. *Auk*, **109**, 543–555.
- Seiger, M. B.** 1967. A computer simulation study of the influence of imprinting on population structure. *American Naturalist*, **101**, 47–57.
- Servedio, M. R.** 2000. Reinforcement and the genetics of nonrandom mating. *Evolution*, **54**, 21–29.
- Servedio, M. R. & Noor, M. A. F.** 2003. The role of reinforcement in speciation: theory and data. *Annual Review of Ecology, Evolution, and Systematics*, **34**, 339–364.
- Siefke, A.** 1994. Wanderungen ostdeutscher Raben- und Nebelkrähen *Corvus corone* nach Beringungsergebnissen. *Vogelwelt*, **115**, 83–89.
- Slabbekoorn, H. & Smith, T. B.** 2002. Bird song, ecology and speciation. *Philosophical Transactions of the Royal Society of London, Series B*, **357**, 493–503.
- Slagsvold, T., Hansen, B. T., Johannessen, L. E. & Lifjeld, J. T.** 2002. Mate choice and imprinting in birds studied by cross-fostering in the wild. *Proceedings of the Royal Society of London, Series B*, **269**, 1449–1455.
- ten Cate, C. & Vos, D. R.** 1999. Sexual imprinting and evolutionary processes in birds: a reassessment. *Advances in the Study of Behavior*, **28**, 1–31.

- ten Cate, C., Vos, D. R. & Mann, N.** 1993. Sexual imprinting and song learning; two of one kind? *Netherlands Journal of Zoology*, **43**, 34–45.
- Turelli, M., Barton, N. H. & Coyne, J. A.** 2001. Theory and speciation. *Trends in Ecology and Evolution*, **16**, 330–343.
- Waxman, D. & Gavrillets, S.** 2005. Issues of terminology, gradient dynamics and the ease of sympatric speciation in Adaptive Dynamics. *Journal of Evolutionary Biology*, **18**, 1214–1219.
- Wright, A. H.** 1991. Genetic algorithms for real parameter optimization. In: *Foundations of Genetic Algorithms* (Ed. by G. J. E. Rawlings), pp. 205–218. San Francisco: Morgan Kaufman.