



## **Behind Enemy Lines: Investigating Crozier's Paradox and Recognition of Nestmates through an Individual-Based Model**

Research project report

by

Lakshya Chauhan

MSc Ecology and Evolution, University of Groningen

MEME Masters

s5381029

[lakshya.chauhan@evobio.eu](mailto:lakshya.chauhan@evobio.eu); [lakshya3141@gmail.com](mailto:lakshya3141@gmail.com)

Supervised by:

Dr. Margaux Bieuville<sup>1</sup>, Dr. Florian Menzel<sup>1</sup>,  
Prof. Hanna Kokko<sup>1</sup> and Prof. G. S. van Doorn<sup>2,\*</sup>

<sup>1</sup> Institute of Organismic and Molecular Evolution (iomE), Johannes Gutenberg University,  
55128, Mainz, Germany

<sup>2</sup> Centrum Levenswetenschappen, Groningen Institute for Evolutionary Life Sciences,  
University of Groningen, Nijenborgh 7, 9747 AG Groningen, The Netherlands

\* Internal supervisor

## Abstract

In social insects, nestmate recognition and cooperation mechanisms are pivotal for colony success. Crozier's Paradox challenges these systems, as the maintenance of high cue diversity necessary for genetic recognition can be undermined by the homogenizing effects of selection. Specifically, the benefits of cooperation or aggression directed towards particular groups may drive increased similarity within populations, potentially leading to the collapse of recognition systems as individual cues converge. To address this paradox, we developed an individual-based simulation model focused on intraspecific parasitism of food between colonies. Our model separates the processes of cue recognition and expression, allowing for a recognition system that can either remain static or evolve alongside the cues. The simulations reveal that high cue diversity can be maintained across a broad spectrum of conditions, suggesting that selection pressures from intraspecific parasitism can preserve distinct recognition cues among colonies. Furthermore, the model explores various biological scenarios of recognition, providing robust evidence that ants are more likely to recognize nestmates rather than non-nestmates. This study offers new insights into the evolution of recognition systems in social insects, demonstrating how selection pressures can resolve Crozier's Paradox and maintain the diversity necessary for effective nestmate recognition.

## Introduction

Recognition of individuals similar/dissimilar to oneself is widespread throughout all clades of life. This recognition is crucial for social interactions, mating behaviors, and survival strategies, as it influences how organisms respond to their environment and each other (Gokcekus et al. 2021). From microbes recognizing highly related individuals for increased survival (Strassmann et al. 2011; Wall 2016) to complex mammals forming intricate social bonds (Penn and Potts 1998), altruism towards group members is often observed. Social insects like ants, bees, and wasps face selection on multiple levels, and recognition is a key player in mediating cooperation and conflict within and between colonies (Korb and Heinze 2004). Discriminating a non-nestmate from a nestmate is essential for maintaining colony integrity, as it helps to prevent the infiltration of outsiders that could threaten resources or disrupt social structure. Often this recognition system is based on chemical cues, although visual and auditory signals can also play a role in identifying familiar individuals (Ratnieks 1991). In social insects like ants, bees, and wasps, Cuticular Hydrocarbons (CHCs) on their exoskeleton serve as important chemical markers that convey information about identity and kinship (Guerrieri et al. 2009; Sturgis and Gordon 2012; Giuseppe Di Mauro et al. 2015).

While the evolutionary origin of CHCs is protection against desiccation, a large variety of CHCs have evolved to be used in intraspecific and interspecific recognition and communication. While as many as 150 compounds can exist on a single individual, the specific profiles of these compounds can vary significantly among castes, colonies, species, and environmental conditions (Sprenger et al. 2021). Often, these compounds exist in homologous groups (such as the sum of 15-MeC25, 15-MeC27, and 15-MeC29) where several chemicals can have similar functions, but groups can vary independently (Holze et al. 2021). The chemical distance between these profiles is perceived by a neuronal template in the individual, prompting further actions (Brandstaetter et al. 2011).

The mapping between recognition cues and the internal neuronal template is a complex process, and much is yet to be deciphered about it (Antoine Couto et al. 2023). Multiple recognition models like Gestalt, Undesirable-absent, and Desirable-present have been proposed (Crozier and Dix 1979; d'Ettorre and Lenoir 2009). In the Gestalt model, the overall similarity between the chemical profiles of the acting individuals is perceived. Hence both individuals would perceive each other identically, subject to low stochasticity in the recognition process. However, this is not a necessity as individuals' perceptions of each other may differ in a natural setting. The Undesirable-absent (U-absent) model depicts such a scenario of asymmetric recognition, where an ant perceives components that are not a

part of its own template, identifying a non-nestmate. Under this system, an individual would reduce their chemical signature to avoid detection and successfully exploit other colonies. The Desirable-present (D-present) model highlights the other scenario, where the presence of familiar chemicals is perceived by an individual as a nestmate. We draw that the prevalence of the U-absent model could lead to organisms losing chemical cues or lowering their values to be recognized more by others. In contrast, the D-present model could lead to an increase in cue values being produced (a biological extension being the evolution of newer chemicals or homologous groups). Some empirical evidence points to the prevalence of the U-absent model (van Zweden and d'Ettorre 2010; Nehring et al. 2016), but more research is yet needed.

A theoretical problem in the maintenance of such recognition systems in populations was highlighted by R. H. Crozier in 1986 (Crozier 1986). Using a single locus, bi-allelic model for recognition in a panmictic population, Crozier demonstrated that altruistic benefits from recognition would lead to an increase in the numbers of the group that is present in larger numbers. Eventually, rarer alleles would perish, and the diversity required for such a recognition system to function would be lost. To resolve this paradox, opposing selection pressures would need to function on the recognition system or the cue diversity. In social insects, high aggression in interaction with a non-nestmate would be detrimental for both individuals. This would drive colonies to be more similar but would eventually lead to the loss of cue diversity required for such a recognition system to sustain. However, cue diversity is seen aplenty at the colony level in social insects (Sprenger and Menzel 2020; Sprenger et al. 2021).

Multiple solutions have been proposed for Crozier's paradox, but none sufficiently explains nestmate recognition in social insects. Allowing multiple encounters between individuals such that they can find similar individuals and pair up before an action is taken seemed to solve the paradox in a recent individual-based study (Scott et al. 2022; Scott 2024). However, this study cannot be extended to social insects in its current state as multiple interactions may be extremely costly from a defense perspective in social insects. High-stakes interactions that require instantaneous decision-making often occur when intruders attempt to infiltrate a colony, or when foraging ants from different colonies or species encounter each other and must defend their resources (Kleineidam et al. 2017). The occurrence of these situations raises doubts about whether cue diversity in social insects can be maintained solely through multiple interactions between individuals. Our study does not require multiple interactions, and individuals make instantaneous decisions. Disassortative mating, where individuals prefer to mate with partners that are genetically or phenotypically dissimilar has been shown to resolve Crozier's paradox and sustain high cue diversity in an individual-based study (Holman et al. 2013). In this scenario, rarer individuals in the population do not go extinct as rarer alleles persist as heterozygotes. While CHCs have been shown to undergo sexual selection (Steiger and Stöckl 2014), the occurrence of disassortative mating in ants has little empirical evidence and often the mating function is reserved for a small proportion within the colony (Baer 2011). A pleiotropic role of recognition cues unrelated to social behaviors has been another theoretical solution to Crozier's paradox (Scott et al. 2023). Host-parasite interactions provide one such avenue, as parasites that evolve fast or have high virulence would evolve to prefer the more common individuals, thus stabilizing genetic kin recognition. However most ant workers spend most of their lives inside the colony, and this limited risk of parasites can reduce the selection pressure that parasitic interactions may provide on cue diversity (Cremer et al. 2007). Theoretical models have demonstrated that genetic kin recognition can remain stable provided certain conditions are met: extreme spatial structure with limited migration, tight linkage between the tag locus and the helping trait locus, and strong selection pressures (Axelrod et al. 2004; Rousset and Roze 2007; Scott et al. 2022). This concept has been referred to as 'beard chromodynamics' due to its reliance on the physical linkage between a tag and a trait, which is a hallmark of greenbeard genes (Madgwick et al. 2019). However previous attempts at explaining Crozier's paradox also utilized multiple social interactions (Scott 2024), an assumption that does not hold for social insects. Frequent intraspecific and interspecific socio-parasitic

interactions in social insects suggest another resolution to the paradox and our model explores an alternate simpler scenario of intraspecific parasitism. In our model, colonies are the reproductive units and act both as the host and the parasite, competing and stealing for food resources.

We used an individual-based model inspired by intraspecific parasitism to investigate Crozier's paradox in social insects. Workers can forage resources from the environment but can also exploit resources from other colonies subject to a recognition process. Recognition cues are expressed as a chemical profile and have a metabolic cost associated with their production. We hypothesize that a pull between opposing selection forces due to resource competition and metabolic costs on CHC production would evolve high cue diversity required to sustain the recognition system. Our model allows the chemical distance between these profiles to be perceived according to either the Gestalt, U-absent, or D-present recognition model detailed above. Perception of these cues is separated from their production, where each colony has a tolerance curve that dictates whether an intruder is accepted or not for a specific chemical distance. Our model allows us to fix the tolerance curve such that only the recognition cues evolve. Additionally, we can also co-evolve the cue profiles and perception together where the population starts with an indiscriminate recognition system. The model is parsimonious and does not assume disassortative mating. The workers act in continuous time, and this is implemented using Gillespie's algorithm. In each simulation, the population undergoes seasonal mortality cycles with a constant period, where a constant fraction of the total number of colonies dies every season. A biological understanding of this can be the winter season in temperate regions, where extreme starvation due to lack of food often causes colonies to die out. Reproduction occurs at the end of every season in our model.

## Materials and Methods

### The Model

The purpose of the model is to study how cue diversity and abundance may evolve in the context of Crozier's paradox focusing on ant colonies that can forage or steal resources from another colony. The simulated population comprises of  $N_{colonies}$  colonies containing  $N_{workers}$  workers each. Each worker ant can choose between foraging from a global resource pool or stealing from another colony based on the relative availability of food in the environment and the number of available colonies available for exploitation. Workers undergo a recognition process whenever they try to enter a colony and entry is permitted if it is recognized as a nestmate. Gillespie algorithm was utilized to schedule foraging and stealing in the model dynamically. Once an action of foraging, stealing, or returning to one's colony is completed, the period before the next action is sampled from an exponential distribution with a rate parameter 1. Our model defines 1 time unit as the average time taken for one action to occur (inverse of rate parameter 1).

### Recognition Cues

Each colony has a unique chemical profile that serves as the recognition signal against which an intruder is compared and consists of  $N_{cues}$  homologous groups. Said compound groups can also be seen as the principal component along which sets of hydrocarbons vary and are independent of each other by definition.

For a colony  $r$ , the abundance of the  $i^{th}$  compound groups are denoted as  $c_{r,i}$ , and the total cue abundance for a colony denoted by  $C_r$  is the sum of its compound groups' abundances. Each time a colony  $r$  reproduces a daughter colony  $d$ , the  $i^{th}$  compound group abundance  $c_{d,i}$  is chosen from a normal distribution with mean  $c_{r,i}$  and standard deviation  $\mu_{cues}$ . Here,  $\mu_{cues}$  represent the strength of mutation for the chemical cues. Chemical abundance values  $c_{r,i}$  have no upper limit but are lower bound at 0. If the simulated population evolves a chemical profile where all cues have a value of 0, we have essentially lost the cue recognition system, and the simulation is terminated.

Workers have a profile that varies around the parent colony profile. The  $i^{th}$  chemical group abundances for a worker  $w$  from a parent colony  $r$  is sampled from a normal distribution with mean  $c_{r,i}$  and standard deviation  $\mu_{cues}/10$ . The smaller spread of the sampling curve ensures that workers in a colony have sufficiently similar profiles and provides a high relatedness within a colony to the model.

### Population initialization

For each colony initialized at the start of a simulation, cue values  $c_{r,i}$  were sampled independently from an exponential distribution with rate parameter  $\lambda_{cues}$ . The global pool of resources is initialized with 300 units of food and each colony is initialized with 25 units of food in its reserve. In the co-evolving scenario, for a colony  $r$  the tolerance variables  $T_{slope,r}$  and  $T_{int,r}$  are sampled from normal distributions with mean 0 and 0.5 (and common standard deviation  $u_{tolerance}$ ) respectively. Thus, colonies initially have a flat tolerance curve averaged over the population scale.

### Foraging, stealing, and resource consumption

All workers have access to a global pool of resources. When a worker leaves the colony, it probabilistically decides to forage or steal food from another colony. This decision is dependent on the amount of food available in the global pool ( $F$ ) and the number of colonies alive  $N_{alive}$ . The probability of foraging is then defined as:

$$P(\text{foraging}) = \text{bernoulli}\left(\frac{F}{F + N_{alive} - 1}\right)$$

If the worker forages, then one unit of food is acquired from the population food source by the individual. If the worker is stealing resources from another colony, a colony is chosen at random for the worker to invade. The intruder undergoes a recognition process in which a worker ant in the target colony determines whether the intruder is a nestmate or a competitor (see section below). If accepted as a nestmate, the intruder raids one unit of food from the target colony. Since we use the Gillespie algorithm for continuous-time simulation, actions are performed in a random order, and all colonies forage and invade simultaneously. Workers which have successfully acquired food and are returning to their colonies face another recognition process from a nestmate worker. If accepted as a nestmate, a unit of food resource is added to the colony stock. In cases where recognition fails and the worker is misidentified as an invader, the resource is discarded, and the worker is rescheduled for the next task.

Our model also accounts for the metabolic costs needed to produce CHCs. Whenever an individual from the colony  $r$  acts, there is a decrease in the colony's food stock  $F_r$  given by

$$\Delta F_r = \frac{m * C_r}{2000 * N_{workers}}$$

Where  $m$  is the cost of CHC biosynthesis per unit of CHC. This models a biological scenario where CHCs are lost outside the colony, hence average individual CHC production costs are subtracted at action points. This formula was chosen such that costs increase as the total cue abundance in a colony increases. The division by 2 and  $N_{workers}$  in the denominator is to normalize across two actions needed for gaining one food source and to normalize the total number of workers, respectively. If a colony exhausts its food stock, it dies.

Our model allows for discrete regeneration of the global pool of resources. 300 units of food are provided periodically as the global pool to the population after a constant time interval  $t_{gen}$ . Workers can then forage from this global pool but must rely on stealing from other colonies once the pool is exhausted before the regeneration point is reached. If there are any food resource remnants in the global pool right before the regeneration point, they are discarded.

## Recognition Process

Every time an individual attempts to enter a colony, a recognition procedure takes place to determine if the individual is a nestmate or intruder. A worker from the resident colony intercepts the intruder and calculates the chemical distance between the intruder's chemical profile and the resident colony's chemical profile. This chemical distance can be calculated in three ways deriving from the Bray-Curtis distance in our model:

- Overall similarity (Gestalt): The chemical distance is calculated as the Bray-Curtis distance between the profile of the intruder  $i$  and resident colony  $r$ , given by

$$BC_{r,i}^G = 1 - \frac{2 \sum_{k=1}^{N_{cues}} \min(c_{k,r}, c_{k,i})}{\sum_{k=1}^{N_{cues}} (c_{k,r} + c_{k,i})}$$

This is quite similar to the conjectured recognition in the Gestalt model (insert citation). Here, both individuals would either reject each other or accept each other due to the symmetric nature of the formula.

- Undesirable-absent (U-absent): This recognition model focuses on 'unknown' compounds that are not present or less abundant in the resident colony profile compared to the intruder profile. The chemical distance between the profile of the intruder  $i$  and resident colony  $r$  under this model is given by a modified Bray-Curtis distance as

$$BC_{r,i}^{UA} = 1 - \frac{2 \sum_{k=1}^{N_{cues}} \min(c_{k,r}, c_{k,i})}{\sum_{k=1}^{N_{cues}} \begin{cases} 2c_{k,i} & \text{if } c_{k,i} < c_{k,r} \\ (c_{k,r} + c_{k,i}) & \text{otherwise} \end{cases}}$$

Here, the compound groups that are present in the resident but absent or less abundant in the intruder ( $c_{k,i} < c_{k,r}$ ) are ignored, and both the values are set to  $c_{k,i}$  before calculating the distance. Thus, the compounds less abundant in the intruder lower the chemical distance.

- Desirable-present (D-present): In this recognition system, emphasis is placed on whether the intruder has all the compounds that are present in the resident colony. Thus, while calculating distance, if compound groups in the intruder that have a higher abundance than residents are ignored, we arrive at the formula:

$$BC_{r,i}^{DP} = 1 - \frac{2 \sum_{k=1}^{N_{cues}} \min(c_{k,r}, c_{k,i})}{\sum_{k=1}^{N_{cues}} \begin{cases} 2c_{k,r} & \text{if } c_{k,r} < c_{k,i} \\ (c_{k,r} + c_{k,i}) & \text{otherwise} \end{cases}}$$

Here cues that are higher in abundance for the intruder help lower the perceived chemical distance.

Our model separates cue expression from cue perception. Each colony has a tolerance curve, which gives the rejection probability for a specific value of chemical distance. The rejection probability is the binomial probability with which the intruder or worker is rejected entry into the colony. Each colony  $r$  genetically encodes two variables  $T_{int,r}$  and  $T_{slope,r}$ , which provides for a linear tolerance curve. The rejection probability for a given chemical distance  $d$  is given by:

$$P(rejection) = T_{slope,r} * d + T_{int,r}$$

Once the rejection probability is obtained from the chemical distance and tolerance curve of the target colony, a Bernoulli sampling is performed. This recognition process occurs every time an individual attempts to steal food from another colony, or when individuals with food return to their colony.

The tolerance curve itself can either be a fixed property of the model (default) or co-evolve along with recognition cues with mutation strength  $\mu_{tolerance}$ . In the default scenario where tolerance curves are non-evolving, variables  $T_{int,r}$  and  $T_{slope,r}$  are not inherited but are sampled from normal distributions with mean 0 and 1 respectively, with standard deviation  $\mu_{tolerance}$ . This sampling imparts

demographic heterogeneity to our model such that the rejection probability has the same value as the chemical distance between the profiles on average at the population scale.

In the scenario where tolerance co-evolves with recognition cues, the colonies start with a flat tolerance curve such that all entry attempts by workers have an approximately 50% chance of acceptance. The recognition system can evolve adaptively, with the inherited tolerance variables for a daughter colony  $d$  from parent colony  $r$  are given by:

$$T_{int,d} = \text{Normal}(T_{int,r}, \mu_{tolerance})$$

$$T_{slope,d} = \text{Normal}(T_{slope,r}, \mu_{tolerance})$$

Here,  $\mu_{tolerance}$  is the mutation strength for tolerance variables. This is analogous to  $\mu_{cues}$  as mutation strength for compound groups in the recognition profile.

### Reproduction and seasonal mortality

By default, reproduction and mortality are seasonal in our model. Periodically after a time interval  $t_{mort}$  (default value 200 time units), a fraction  $\delta$  of maximum population size  $N_{colonies}$  dies out. This biologically mimics winters, where food availability is low, and ants have a harder time surviving. Colonies can also die due to starvation and the number of colonies to be removed at the end of a season is adjusted to include the colonies that died due to starvation before the seasonal death. The colony food stock acts as a proxy fitness value, where the ones with the lowest values are removed. Producing CHCs is costly in our model as it lowers the food reserves and fitness of colonies. Colonies surviving seasonal death undergo reproduction to maintain population size, where the number of offspring for a colony is proportional to the food stock of that colony relative to the total amount of food available in all living colonies. Thus, a colony can supposedly survive for the entire duration of the simulation. All offspring are initialized with colony food reserve values at 25.

### Controls

Our model allows for three mutually independent controls, and multiple combinations of them can be taken for making various inferences.

- Population Control (Control P): Colonies that die at the end of each season are chosen randomly, minimizing selection pressure in the model. However, this control is not perfect as some colonies may still die of starvation before the mortality point, penalizing colonies producing a high abundance of recognition cues. Hence this control can still experience lower selection due to metabolic costs.
- Model Control (Control M): Individuals have a 50% probability of being rejected irrelevant of their chemical distance, and that removes the influence of recognition systems on fitness. This provides us with a control to test against the three recognition systems highlighted above. Combining this with the Population Control yields Control MP, though it still experiences weak selection like Control P.
- Drift Control (Control D): Each colony also codes for neutral genes that inherit similarly to chemical cues. Since no recognition processes or metabolic costs interact with neutral genes, this control removes the influence of selection and metabolic costs in our model.

### Output and simulations

Various metrics are calculated and reported at periodic intervals:

- The average cue abundance  $C_r$  across colonies is recorded regularly.
- Population-level cue diversity is measured as the average pairwise Bray-Curtis distance (Gestalt formula) between all colonies.
- Average abundance and diversity are also calculated for the neutral genes (Control D).

- When tolerance curves co-evolve with recognition cues, the averages and standard errors of  $T_{int,r}$  and  $T_{slope,r}$  across colonies are reported.
- For the final 1% of the simulation duration, the total number of stealing attempts and successful stealing attempts across the population are recorded. The fractions of successful stealing and re-entry attempts are calculated and reported. The number of original  $N_{colonies}$  lineages surviving is also recorded as a measure of diversity.

To observe how cue diversity evolves when a cue recognition system is present, simulations were performed with default parameters (refer to Table 1). The effect of various parameters on the reported metrics was tested by varying the regeneration time of the global pool ( $t_{gen}$ ), metabolic costs ( $m$ ), mutation strength of cues ( $\mu_{cues}$ ), and seasonal mortality ( $\delta$ ) and keeping other parameters at default values.

Three recognition models (Gestalt, U-absent, and D-present) and the model control (Control M) were also explored for each unique set of parameters. For each unique combination of parameter set and recognition model (including control M), a corresponding population control (Control P) was also performed. In the figures, Control MP is reported alongside the recognition models.

To explore whether a genetic recognition system can co-evolve with a diverse cue system, simulations were conducted in the discrete regeneration scenario across the three recognition models and two controls (M and P), starting with a flat tolerance curve. The effects of food availability on the co-evolving recognition system and cues were also tested by varying the regeneration time of the global pool ( $t_{gen}$ ). All simulations ran for 200,000 time units (or 1000 seasonal cycles). Temporal plots range over the first 100,000 time units for better visualization as simulations achieved equilibrium by then, unless specified otherwise. All metrics plotted in boxplots are from equilibrium values obtained at the end of simulations (200,000 time units).

For each unique combination of parameters and recognition model system, 30 replicates were simulated. However, not all populations survive or sustain non-zero cue values till the end of simulations. 15 replicates that still had colonies at the end of the simulation and retained the cue system (i.e. non-zero cue values) were subsampled from the 30 replicates for the plotting and comparing results.

Temporal trajectories plotted depict the average values of a metric for 15 replicates over simulated time. The standard deviation is demarcated by the shaded region. In the box plots generated, each box represents the interquartile range of the data and shows the middle 50% of the distribution. The horizontal line indicates the median value and the whiskers extend to the maximum and minimum data points within 1.5 times the interquartile range. Points outside this range are considered outliers

### Linear models

The effects of various parameters and recognition models on cue diversity and abundance, as well as their significance, were estimated using linear regressions. Parameters and recognition models served as explanatory variables, with interactions allowed between them. Cue diversity, cue abundance, and evolved linear tolerance slope were set as response variables. Coefficient estimates and significance values were obtained from model summaries for inference. Model summaries with brief explanations are provided in the supplementary section. The emmeans package in R was used to compute estimated marginal means for specific parameter values, and pairwise comparisons were performed to determine significant differences when necessary.

### Code details and availability

The model was coded in the C++ language and compiled with g++ v8.5.0. R v4.3.1 and packages dplyr, readr, ggplot2, emmeans, and tidyR were used to analyse and plot the results.



All scripts are available on <https://github.com/Lakshya3141/CroziersParadox>.

## Model parameters and variables

Table 1: Symbols, definitions, and values of model parameters. Default values are marked in bold.

Symbol	Definition	Values
$N_{colonies}$	Total number of colonies in the population	50
$N_{workers}$	Number of workers in a colony	10
$N_{cues}$	Number of compound groups in a recognition profile	10
$N_{alive}$	Number of colonies alive at an instance in the simulation	Variable
$\lambda_{cues}$	Exponential parameter for the distribution from which initial compound abundances are sampled from	0.1
$m$	Metabolic cost of CHC production per unit	{10; 20; <b>40</b> ; 80}
$\mu_{cues}$	Mutation strength for recognition cues	{0.1; 0.5; 1; <b>5</b> ; 10}
$\mu_{tolerance}$	Mutation strength for tolerance variables (if evolving)	0.1
$t_{mort}$	Generation time for seasonal mortality and reproduction	200
$t_{gen}$	Regeneration time of global pool of food in discrete regeneration scenario	{1; <b>2</b> ; 3; 4}
$F_r$	Food stock available for colony $r$	Variable
$\delta$	Fraction of total colonies that die seasonally	{0.1; 0.3; <b>0.4</b> ; 0.5; 0.7; 0.9}
$T_{int,r}$	Intercept variable for tolerance curve	{0, variable}
$T_{slope,r}$	Slope variable for tolerance curve	{1, variable}
$c_{r,i}$	Cue abundance for compound group $i$ in colony $r$	Variable
$C_r$	Total cue abundance for colony $r$	Variable

## Results

For the default scenario where the recognition system was non-evolving, various values for renewal time of the global resource pool, metabolic costs, mutation strength of cues, and seasonal mortality were explored. The following three subsections and supplementary results contain the findings and inferences from these simulations. The co-evolving scenario was also explored, where the tolerance curves co-evolved with recognition cues. The effect of food availability by varying the renewal time for resources was also obtained. The results from this are highlighted in the fourth subsection and supplementary results.

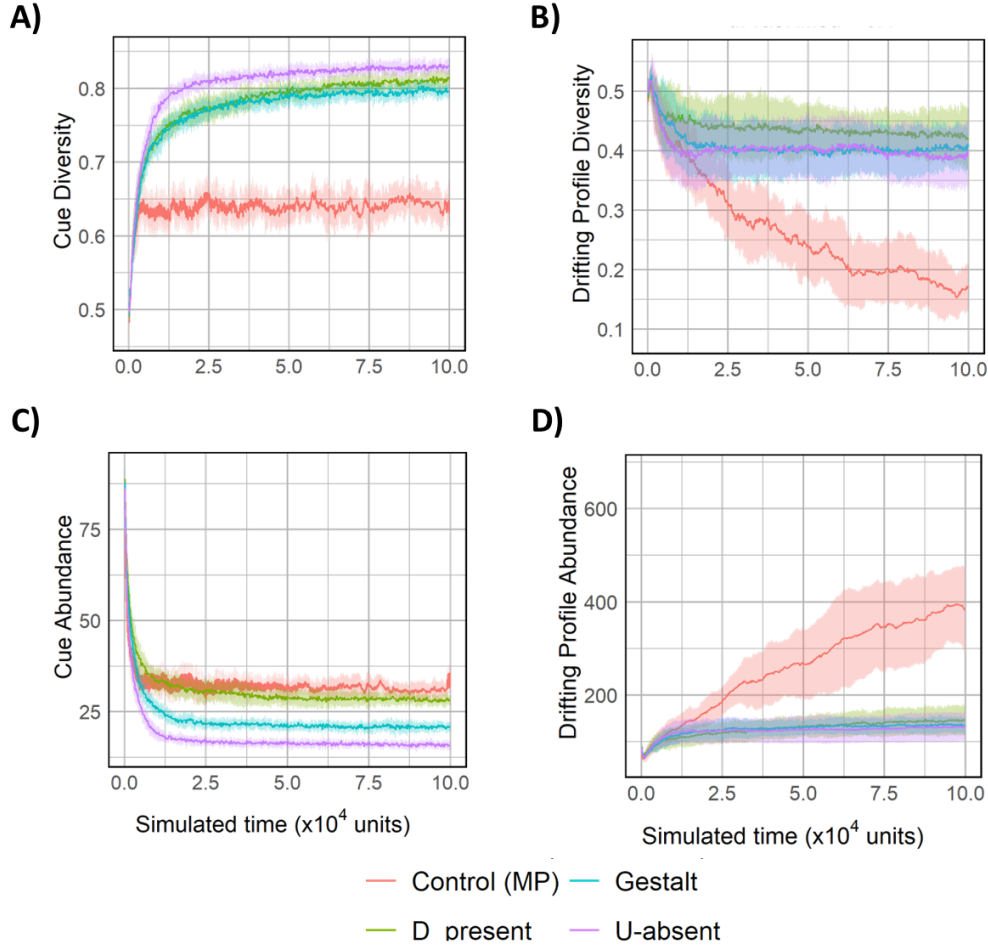


Fig 1: Temporal trajectories for average cue diversity (A), average drifting profile diversity (B), average cue diversity (C), and average drifting profile abundance (D) for various recognition models and control MP. Simulations were run with default parameter values ( $\mu_{cues} = 5$ ;  $m = 40$ ;  $\delta = 0.4$ ;  $t_{gen} = 2$ ). The X-axis is limited to 100,000 time units as equilibrium was observed beyond this point.

### 1) Cue recognition leads to high cue diversity

Cue diversity increased in all recognition models (including control MP) (Fig 1A) compared to drift control (Fig 1B). This result is consistent across all our model explorations, where the cue diversity for drift control declines and has a significantly lower value compared to its recognition model counterparts. U-absent, D-present, and Gestalt models also evolved significantly higher cue diversity compared to control MP (Fig S1A). D-present and U-absent models had significantly higher final diversity compared to the Gestalt model, but no significant difference was observed between the two. Interestingly, while multiple lineages can survive in the population, high cue diversity is also observed in simulations where a single lineage survives (Fig S1B), suggesting high amounts of initial cue diversity are not needed to achieve the high diversity required to sustain a genetic recognition system.

Cue abundance for all recognition models and control MP decreased over time (Fig 1C) compared to drift control (Fig 1D), and this result is also uniform across all subsequent parameter explorations of our model. All four models differed significantly, with the largest abundance observed in Control MP, followed by D-present, Gestalt, and U-absent models (Fig S1C). Thus provided a cue recognition system exists; high cue diversity can evolve under intraspecific parasitism with metabolic costs to cue production.

U-absent and D-present recognition models show higher success in stealing resources from other colonies compared to Gestalt (Fig S1D). This could be driven by the possibility of asymmetric recognition between two interacting workers which would facilitate incorrect recognition of an intruder as a nestmate. This could also be a potential driver for the higher cue diversity observed in U-absent and D-present models compared to the Gestalt model, which would reduce exploitation.

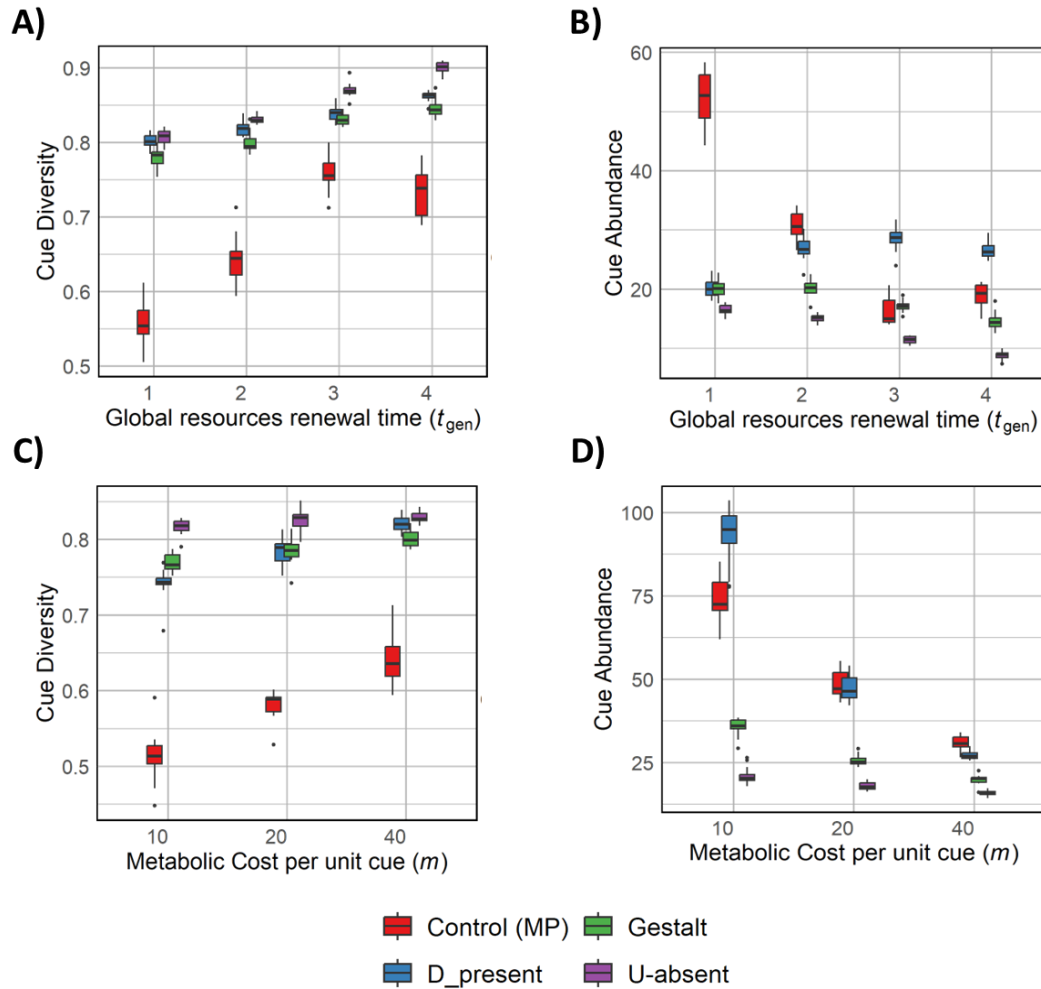


Fig 2: Box plots for cue diversity (A, C) and cue abundance (B, D) for varying  $t_{gen}$  (A, B) and varying metabolic costs (C, D). Apart from the parameter being varied, default values were used ( $\mu_{cues} = 5$ ;  $m = 40$ ;  $\delta = 0.4$ ;  $t_{gen} = 2$ ).

## 2) Decreasing food availability causes an increase in cue diversity

Increasing  $t_{gen}$  indicates a decreasing availability of food, as colonies must rely on stealing resources for longer once the global pool is depleted. Therefore, we observe that the cue diversity increases across all recognition models as the time required for regenerating the global pool of resources ( $t_{gen}$ ) increases (Fig 2A). Cue diversity increased the most for U-absent model with an increase in  $t_{gen}$ , followed by Gestalt, followed closely by D-present (SLM1). Irrespective of the availability of food, U-absent, and D-present models had higher cue diversity and stealing success compared to Gestalt model (Fig S2A). This suggests asymmetric recognition models where two individuals might not perceive each other identically are better at maintaining diversity and raiding other colonies. Additionally, higher diversity under starvation conditions leads to better colony defense, as reflected in the reduced stealing success. Stealing success also reduced as the availability of food decreased, suggesting a need for better defense to safeguard resources (Fig S2A).

Decreasing food availability (increasing  $t_{gen}$ ) led to lower cue abundance for U-absent and Gestalt models (Fig 2B, SLM2). Under low food availability, selection favors lower cue values to reduce the metabolic cost of CHC production (Control MP; Fig 2B). Interestingly, the opposite trend was observed in D-present model, where higher cue abundance is seen with increasing  $t_{gen}$ . This could be explained by the need for high cue diversity to better

Consistent with previous results, the U-absent model had the lowest cue abundance of all models across all  $t_{gen}$  values. This reflects the nature of the recognition models. Under the U-absent model, lower cue abundances compared to other colony profiles are selected as they are often accepted as nestmates. Contrary to this, in the D-present model, higher cue abundances are selected for as producing more cues could help in being similar to multiple other colonies. It should be noted though that these values are still a magnitude of order smaller than the total abundance of drift control cues (data not attached).

### **3) Increasing metabolic costs lead to higher cue diversity and lower cue abundance**

Cue diversity increased for all models with increasing metabolic costs (Fig 2C). The increase is largest for D-present, followed by Gestalt and then U-absent (SLM3). U-absent model consistently shows the highest cue diversity, followed closely by D-present. Increasing metabolic cost is expected to select for lower cue values, which is reflected in cue abundance decreasing for all models with increasing metabolic cost (Fig 2D). This effect is strongest for D-present, followed by Control MP, Gestalt, and U-absent (SLM4). Consistent with previous results, asymmetric recognition models (U-absent and D-present) had higher success in raiding other colonies compared to the Gestalt model (Fig S2B) while showcasing higher cue diversity. Since not all cues are perceived in D-present and U-absent models, higher exploitation may occur due to faulty recognition which could drive higher cue diversity. Raiding success did not vary considerably with varying metabolic costs.

Since the D-present model selects for higher cue values, metabolic costs provide an opposing selection force. This opposing selection explains the sensitivity of the D-present model to metabolic costs. Many D-present replicates also lose their recognition systems by evolving all cue values to be zero under extremely high metabolic costs ( $m = 80$ ; only 1 replicate retains the recognition system; data not shown).

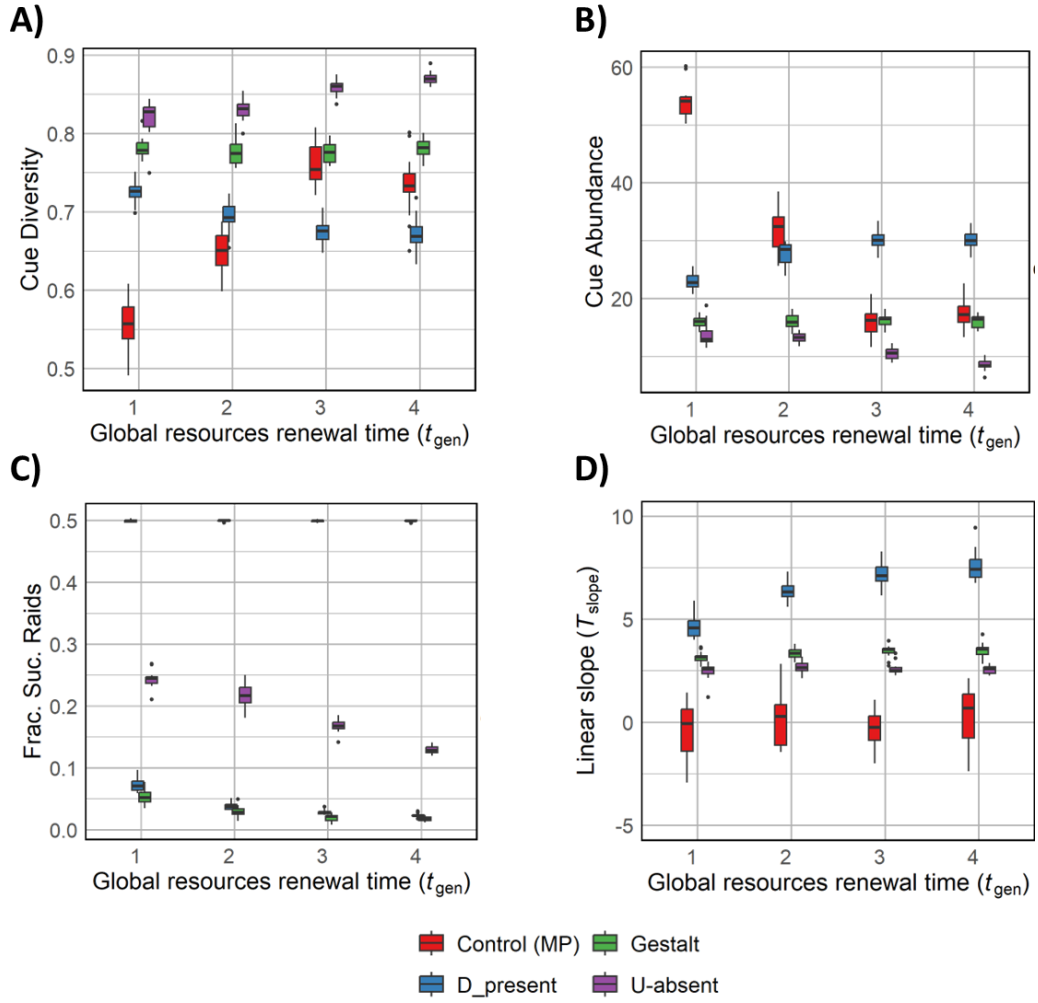


Fig 3: Box plots for the linear slope of cue diversity (A), cue abundance (B), fraction of successful stealing attempts (C), and evolved tolerance slope (D) across varying  $t_{gen}$  in the scenario where the tolerance curve and recognition cues are both subject to evolution. Default values were used for other parameters ( $\mu_{cues} = 5$ ;  $m = 40$ ;  $\delta = 0.4$ ).

#### 4) Stringent recognition systems can co-evolve alongside high cue diversity

Under default parameter conditions, simulations that started with an indiscriminate adaptive recognition system (characterized by a flat tolerance curve) evolved higher cue diversity and lower cue abundance for all recognition models compared to Control MP (Fig 3A, 3B;  $t_{gen} = 2$ ) and drift control (data not shown). The recognition models also evolved positive values for the tolerance curve slope compared to control MP (Fig 3C;  $t_{gen} = 2$ ). This indicates that our model can evolve highly stringent recognition systems ( $T_{slope} > 1$ ) while maintaining the high diversity required to support such systems. This is further supported by lower cue diversity and lower raiding success (Fig 3A, 3D;  $t_{gen} = 2$ ) in the co-evolving scenario compared to earlier simulations where the tolerance curve is non-evolving (Fig 1A, 1D). When the recognition system is fixed and comparatively lenient ( $T_{slope} \approx 1$ ), colonies have no choice but to evolve high cue diversity to avoid exploitation. However, an adaptive recognition system can evolve higher stringency, which could reduce the selection force on recognition cues and result in lower diversity. From this, we can infer that recognition systems and cues can evolve separately but may exert strong selection onto each other.

This interplay between the recognition system and cues becomes more apparent when the amount of food provided to the population is varied. As food availability decreased (increasing  $t_{gen}$ ), The U-

absent model evolved higher cue diversity (Fig 3A, SLM5) and lower cue abundance (Fig 3B, SLM6; consistent with previous results Fig 2A, 2B). While cue abundance increased with increasing  $t_{gen}$  for the D-present model like in the non-adaptive tolerance curve scenario, cue diversity shows a decreasing trend surprisingly.

As anticipated, raiding success decreased with decreasing food availability across all models (Fig 3C), which is consistent with previous results for a non-coevolving tolerance curve (Fig S2A). This suggests the high importance of colony defense during low food availability. We also observed lower raiding success and lower cue diversity in general for the co-evolving scenario compared to the non-adaptive tolerance slope scenario (Fig 3C vs Fig S2A; Fig 3A vs 2A). Decreasing food availability resulted in an increase in stringency for D-present, while U-absent and Gestalt were not significantly affected (Fig 3D, SLM7).

Highly stringent recognition would differentiate between nestmates and non-nestmates even when the chemical distance between interacting individuals is low and would mitigate the need for maintaining high cue diversity in the population. This effect is highlighted in the D-present model, where the impact of food scarcity is especially pronounced due to its inherent tendency to select for higher cue values which is costly. Instead, the D-present population evolves a highly stringent recognition system when food is scarce and reduces the load on cue production.

It was observed that cue diversity could increase (Fig 2A) and decrease (Fig 3A) as total cue abundance increases. It highlights the need for a holistic diversity metric that looks at chemical profiles and their distances. It should once again be noted that all systems evolved higher cue diversity and lower cue abundance compared to the drift control.

## Discussion

In this study, we introduced an evolutionary individual-based simulation model to find a possible solution to Crozier's paradox for nestmate recognition in ants. Population-level diversity is essential for maintaining recognition systems, as it allows for the differentiation between nestmates and non-nestmates. However, higher aggression towards non-nestmates would lead to colonies evolving similar profiles to avoid competition. This would render the recognition system moot as population diversity is lost, which suggests opposing selection forces are at play to maintain the necessary diversity for effective recognition (Crozier 1986). Our model is inspired by an intraspecific parasitic setting where colonies can exploit each other for resources, and it contains several novel mechanisms centered around the life histories and recognition systems of social insects. In the scenario where the recognition curve doesn't evolve, we found that high cue diversity evolved for all recognition models and increased with increasing starvation. Higher metabolic costs of CHC production also led to an increase in cue diversity and raiding success. For the coevolving scenario, we see the evolution of stringent recognition curves and high cue diversity. We observed that the D-present model is the most affected by metabolic costs and low food availability.

In the original study where Crozier highlighted the dilemma of maintaining population diversity for functional recognition systems, the model assumed that cue expression and recognition both occur using the same genetic architecture (Crozier 1986), while the CHC production and recognition are separate processes in social insects. Previous attempts at resolving Crozier's paradox also make this unifying assumption, which is not biologically realistic, as cue production and perception probably evolved separately (Leonhardt et al. 2016). CHCs initially evolved to safeguard against desiccation in dry environments (Wang et al. 2022) and were probably later co-opted for communication and recognition purposes within social insects. This decoupling of recognition and production systems of cues is a key aspect of our model, as tolerance curves and chemical cues are independent modalities. Higher cue diversity evolved in our model under the condition that individuals have a pre-existing

recognition system and reject individuals chemically distant from them. This high cue diversity is consistently sustained across a range of different parameter explorations, suggesting that opposing selection forces due to resource competition and metabolic costs on cue production can sustain the high diversity required for recognition systems to be functional. While multiple initial lineages often persisted until the end of our simulations, the equilibrium population often resulted from a single lineage. This indicates that the high cue diversity required to resolve Crozier's paradox can also be obtained from low initial genetic diversity in our framework.

A co-evolution scenario was also tested in our study, where the population starts with no recognition between nestmates and non-nestmates, and the tolerance curves and cue production are both adaptive and can exert selection on each other. We consistently observed high cue diversity and the evolution of a stringent recognition system across multiple parameter combinations. Compared to the non-evolving case, more stringent tolerance curves and lower cue diversities evolved. The evolution of a stringent recognition system alleviates the pressure on diversity. We also observe that the effect of model system choice and metabolic costs on CHCs are linked, as starvation conditions in the D-present model see higher stringency and decreasing cue diversity, while in the U-absent model, cue diversity increases with no effect on the recognition system. Since the raiding success decreased with increasing starvation for all recognition models, our model shows that recognition cues and perception systems exert selection pressure on each other, and further studies could benefit from keeping this consideration in mind. Chemical receptors and production in social insects can evolve quickly and chemical signals can even vary over an individual's lifetime (Zhou et al. 2015; Menzel et al. 2018), further highlighting the need to study them adaptively in tandem.

In our models, colonies are both hosts and parasites and can steal food from each other. Our framework reflects a realistic biological scenario, as such intraspecific food parasitism has empirical evidence in ants across a variety of climate conditions (*Messor aciculatus*: temperate, *Solenopsis molesta*: subtropical, *Ectatomma ruidum*: tropical and many others) (Grasso et al. 2010; Guénard and McGlynn 2013). While this study considers food as a resource to justify metabolic costs for CHC production, it can be extended to other forms of resources that increase the survival and reproductive success of colonies. One such example is parasitism of brood or workers, instances of which are widely reported in ants across a range of climates, especially temporarily during colony foundation (Davies et al. 1989). This form of parasitism is referred to as social parasitism and well-established colonies have been reported to increase their workforce by raising brood from younger colonies and has been shown to lead to higher cue diversity (Jongepier and Foitzik 2016). The high relatedness between socially parasitic species and their hosts suggests that interspecific parasitism could have evolved from intraspecific parasitism, which further supports our model. Pleiotropic effects due to host-parasite interactions have been proposed as a possible mechanism for sustaining cue diversity by exerting negative frequency-based selection (Scott et al. 2023). Our study provides a specific intraspecific condition of this in the context of social insects, but the framework can further be extended to study the evolution of recognition perception and traits in other organisms, perhaps even in interspecific parasitism.

Various recognition models that highlight the complexities of these interactions were explored in our model for the first time (Crozier and Dix 1979; d'Ettoire and Lenoir 2009). Individuals possess the capacity to either comprehend the comprehensive chemical profile of other individuals (Gestalt) or discern solely the amicable (D-present) or antagonistic (U-absent) signals in other individuals. While all recognition models show high cue diversity, our results indicate that U-absent is more optimal than D-present and Gestalt as a recognition scheme. The U-absent recognition model generally evolved higher cue diversity and raiding success than other recognition schemes and this behavior is consistent across a larger parameter space than other models. The D-present model showed the highest sensitivity to various parameters (Fig S4, S5), and population extinction was observed disproportionately for extreme parameter values. This can be explained by the biological basis for this

recognition model, as well as its implementation in our study. Since individuals exhibiting the D-present model rely on the presence of friendly cues, colonies that can exploit other colonies successfully must express high abundances of cues. However, an opposing selection pressure exists on cue values due to metabolic costs, and this increases the susceptibility of the D-present model to metabolic costs and food availability. Increasing mutation strength resulted in higher diversity for all recognition models, presumably due to a much larger numeric space available for the chemical profiles to vary over. The U-absent model may benefit the most from this effect, as even small mutations in cue values can now explore a broader space of potential profiles that might better exploit the population.

Empirical evidence suggests that ants recognize foes and not friends (Guerrieri et al. 2009), providing biological significance to the U-absent recognition model and empirical support to our study. While our study does not address the nuances of different CHC and cue types, one can theoretically argue that losing an existing chemical pathway is considerably easier than gaining new ones. Hence, a recognition scheme like U-absent can potentially adapt faster than D-present or Gestalt to exploit others. Some social parasites that lack all their chemical cues represent a limiting condition of our study. Though such chemical insignificance did evolve in our results, further analyses are required to determine whether this is due to the prevalence of Crozier's paradox or a response to metabolic costs.

In conclusion, we developed a robust model that reported high cue diversity across colonies while maintaining or evolving a recognition system. High cue diversity is maintained by the pull of opposing selection forces due to metabolic costs and resource competition, which provides a solution to Crozier's paradox for nestmate recognition in ants. Our results provide deeper insights into how various recognition models may work and highlight the importance of integrating ecological perspectives (like resource availability and the absence of disassortative mating) and evolutionary perspectives (separation of cue production and perception) to fully understand the complexities of social behavior and cue diversity in organisms.

## References

- Antoine Couto, Simon Marty, Erika H. Dawson, Patrizia d'Ettorre, Jean-Christophe Sandoz, and Stephen H. Montgomery. 2023. Evolution of the neuronal substrate for kin recognition in social Hymenoptera. 98:2226–2242.
- Axelrod, R., R. A. Hammond, and A. Grafen. 2004. ALTRUISM VIA KIN-SELECTION STRATEGIES THAT RELY ON ARBITRARY TAGS WITH WHICH THEY COEVOLVE. *Evolution* 58:1833–1838.
- Baer, B. 2011. The copulation biology of ants (Hymenoptera: formicidae). *Myrmecol. News* 14:55–68.
- Brandstaetter, A. S., W. Rössler, and C. J. Kleineidam. 2011. Friends and Foes from an Ant Brain's Point of View – Neuronal Correlates of Colony Odors in a Social Insect. *PLOS ONE* 6:e21383. Public Library of Science.

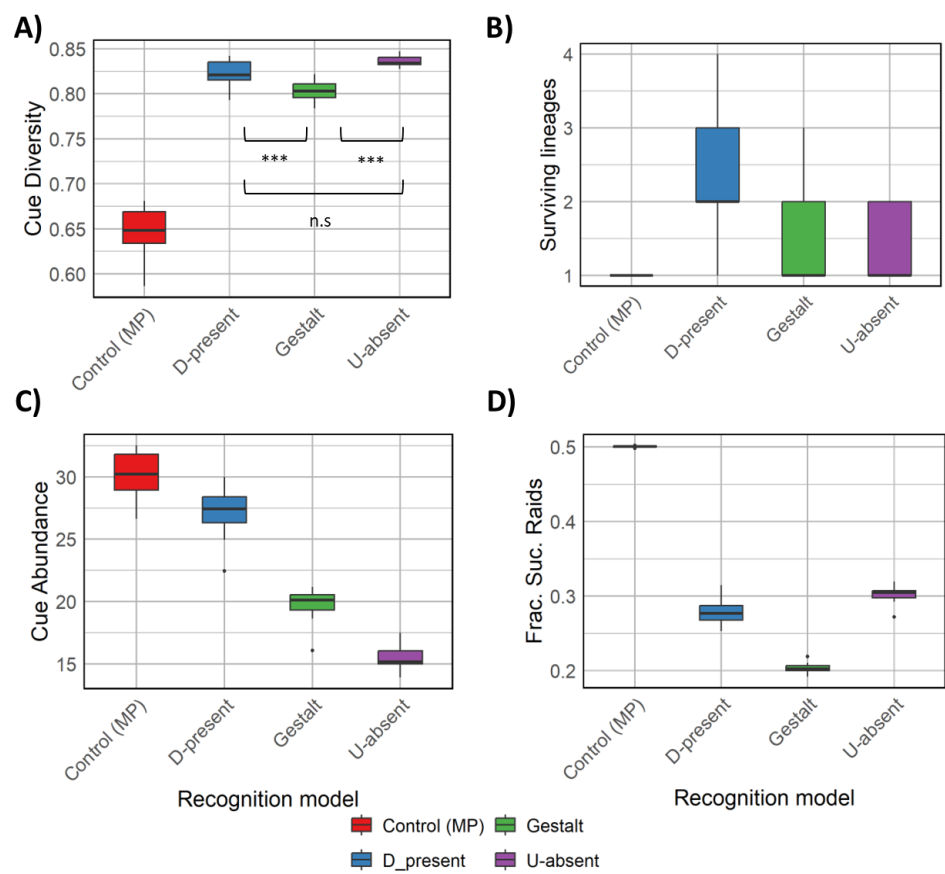


- Cremer, S., S. A. O. Armitage, and P. Schmid-Hempel. 2007. Social immunity. *Curr. Biol.* CB 17:R693-702.
- Crozier, R. H. 1986. GENETIC CLONAL RECOGNITION ABILITIES IN MARINE INVERTEBRATES MUST BE MAINTAINED BY SELECTION FOR SOMETHING ELSE. *Evol. Int. J. Org. Evol.* 40:1100–1101.
- Crozier, R. H., and M. W. Dix. 1979. Analysis of Two Genetic Models for the Innate Components of Colony Odor in Social Hymenoptera. *Behav. Ecol. Sociobiol.* 4:217–224. Springer.
- d’Ettorre, P., and A. Lenoir. 2009. Chapter 11 Nestmate Recognition. P. 0 *in* L. Lach, C. Parr, and K. Abbott, eds. *Ant Ecology*. Oxford University Press.
- Davies, N. B., A. F. G. Bourke, and M. de L. Brooke. 1989. Cuckoos and parasitic ants: Interspecific brood parasitism as an evolutionary arms race. *Trends Ecol. Evol.* 4:274–278.
- Giuseppe Di Mauro, G. Di Mauro, Giuseppe di Mauro, Margot Perez, M. Perez, María Cristina Lorenzi, M. C. Lorenzi, Fernando J. Guerrieri, F. J. Guerrieri, Jocelyn G. Millar, J. G. Millar, Patrizia d’Ettorre, and P. d’Ettorre. 2015. Ants Discriminate Between Different Hydrocarbon Concentrations. *Front. Ecol. Evol.* 3:133.
- Gokcekus, S., J. A. Firth, C. Regan, and B. C. Sheldon. 2021. Recognising the key role of individual recognition in social networks. *Trends Ecol. Evol.* 36:1024–1035.
- Grasso, D., A. Mori, M. Giovannotti, and F. Moli. 2010. Interspecific interference behaviours by workers of the harvesting ant *Messor capitatus* (Hymenoptera Formicidae). *Ethol. Ecol. Evol.* July 2004:197–207.
- Guénard, B., and T. P. McGlynn. 2013. Intraspecific Thievery in the Ant *Ectatomma ruidum* is Mediated by Food Availability. *Biotropica* 45:497–502.
- Guerrieri, F. J., V. Nehring, C. G. Jørgensen, J. Nielsen, C. G. Galizia, and P. d’Ettorre. 2009. Ants recognize foes and not friends. *Proc. Biol. Sci.* 276:2461–2468.
- Holman, L., J. S. van Zweden, T. A. Linksvayer, and P. d’Ettorre. 2013. Crozier’s paradox revisited: maintenance of genetic recognition systems by disassortative mating. *BMC Evol. Biol.* 13:211.

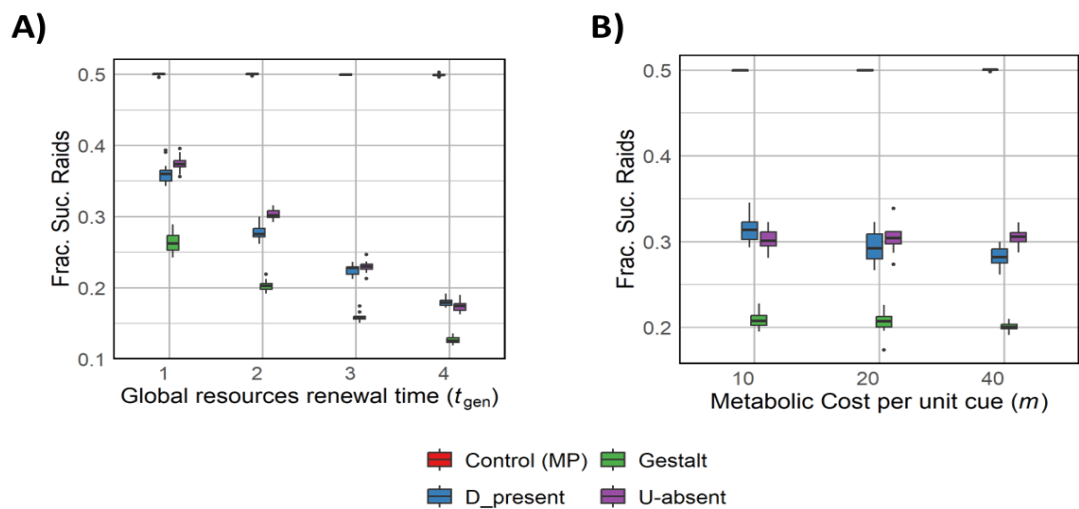
- Holze, H., L. Schrader, and J. Buellesbach. 2021. Advances in deciphering the genetic basis of insect cuticular hydrocarbon biosynthesis and variation. *Heredity* 126:219–234. Nature Publishing Group.
- Jongepier, E., and S. Foitzik. 2016. Ant recognition cue diversity is higher in the presence of slavemaker ants. *Behav. Ecol.* 27:304–311.
- Kleineidam, C. J., E. L. Heeb, and S. Neupert. 2017. Social interactions promote adaptive resource defense in ants. *PLoS ONE* 12:e0183872.
- Korb, J., and J. Heinze. 2004. Multilevel selection and social evolution of insect societies. *Naturwissenschaften* 91:291–304.
- Leonhardt, S. D., F. Menzel, V. Nehring, and T. Schmitt. 2016. Ecology and Evolution of Communication in Social Insects. *Cell* 164:1277–1287.
- Madgwick, P. G., L. J. Belcher, and J. B. Wolf. 2019. Greenbeard Genes: Theory and Reality. *Trends Ecol. Evol.* 34:1092–1103. Elsevier.
- Menzel, F., M. Zumbusch, and B. Feldmeyer. 2018. How ants acclimate: Impact of climatic conditions on the cuticular hydrocarbon profile. *Funct. Ecol.* 32:657–666.
- Nehring, V., F. R. Dani, L. Calamai, S. Turillazzi, H. Bohn, K.-D. Klass, and P. d’Ettorre. 2016. Chemical disguise of myrmecophilous cockroaches and its implications for understanding nestmate recognition mechanisms in leaf-cutting ants. *BMC Ecol.* 16:35.
- Penn, D., and W. Potts. 1998. How do major histocompatibility complex genes influence odor and mating preferences? *Adv. Immunol.* 69:411–436.
- Ratnieks, F. 1991. The Evolution of Genetic Odor-Cue Diversity in Social Hymenoptera. *Am. Nat.*
- Rousset, F., and D. Roze. 2007. CONSTRAINTS ON THE ORIGIN AND MAINTENANCE OF GENETIC KIN RECOGNITION. *Evolution* 61:2320–2330.
- Scott, T. W. 2024. Crozier’s paradox and kin recognition: Insights from simplified models. *J. Theor. Biol.* 581:111735.
- Scott, T. W., A. Grafen, and S. A. West. 2023. Host–parasite coevolution and the stability of genetic kin recognition. *Proc. Natl. Acad. Sci.* 120:e2220761120. Proceedings of the National Academy of Sciences.

- Scott, T. W., A. Grafen, and S. A. West. 2022. Multiple social encounters can eliminate Crozier's paradox and stabilise genetic kin recognition. *Nat. Commun.* 13:3902. Nature Publishing Group.
- Sprenger, P. P., L. J. Gerbes, J. Sahm, and F. Menzel. 2021. Cuticular hydrocarbon profiles differ between ant body parts: implications for communication and our understanding of CHC diffusion. *Curr. Zool.* 67:531–540.
- Sprenger, P. P., and F. Menzel. 2020. Cuticular hydrocarbons in ants (Hymenoptera: Formicidae) and other insects: how and why they differ among individuals, colonies, and species. , doi: 10.25849/MYRMECOL.NEWS\_030:001. The Austrian Society of Entomofaunistics.
- Steiger, S., and J. Stökl. 2014. The Role of Sexual Selection in the Evolution of Chemical Signals in Insects. *Insects* 5:423–438.
- Strassmann, J., O. Gilbert, and D. Queller. 2011. Kin discrimination and cooperation in microbes. *Annu. Rev. Microbiol.*
- Sturgis, S. J., and D. M. Gordon. 2012. Nestmate recognition in ants (Hymenoptera: Formicidae): a review.
- van Zweden, J. S., and P. d'Ettorre. 2010. Nestmate recognition in social insects and the role of hydrocarbons. Pp. 222–243 in A.-G. Bagnères and G. J. Blomquist, eds. *Insect Hydrocarbons: Biology, Biochemistry, and Chemical Ecology*. Cambridge University Press, Cambridge.
- Wall, D. 2016. Kin Recognition in Bacteria. *Annu. Rev. Microbiol.* 70:143–160.
- Wang, Z., J. P. Receveur, J. Pu, H. Cong, C. Richards, M. Liang, and H. Chung. 2022. Desiccation resistance differences in *Drosophila* species can be largely explained by variations in cuticular hydrocarbons. *eLife* 11:e80859.
- Zhou, X., A. Rokas, S. L. Berger, J. Liebig, A. Ray, and L. J. Zwiebel. 2015. Chemoreceptor Evolution in Hymenoptera and Its Implications for the Evolution of Eusociality. *Genome Biol. Evol.* 7:2407–2416.

Supplementary Figures and Results



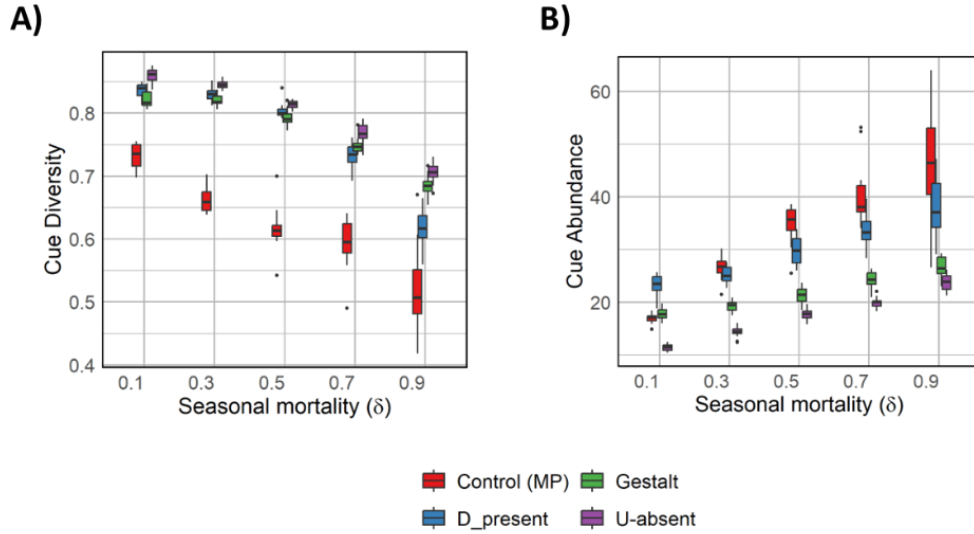
**Fig S1:** Boxplots for cue diversity (A), surviving lineages (B), cue abundance (C), and the fraction of successful stealing attempts (D) for default parameter values ( $\mu_{cues} = 5$ ;  $m = 40$ ;  $\delta = 0.4$ ;  $t_{gen} = 2$ ) across various recognition models and Control MP.



**Fig S2:** Box plots for fraction of successful raids over varying  $t_{gen}$  (A) and metabolic costs (B) under

the non-evolving tolerance curves scenario. Default parameter values are used except for the parameter being varied ( $\mu_{cues} = 5$ ;  $m = 40$ ;  $\delta = 0.4$ ;  $t_{gen} = 2$ ).

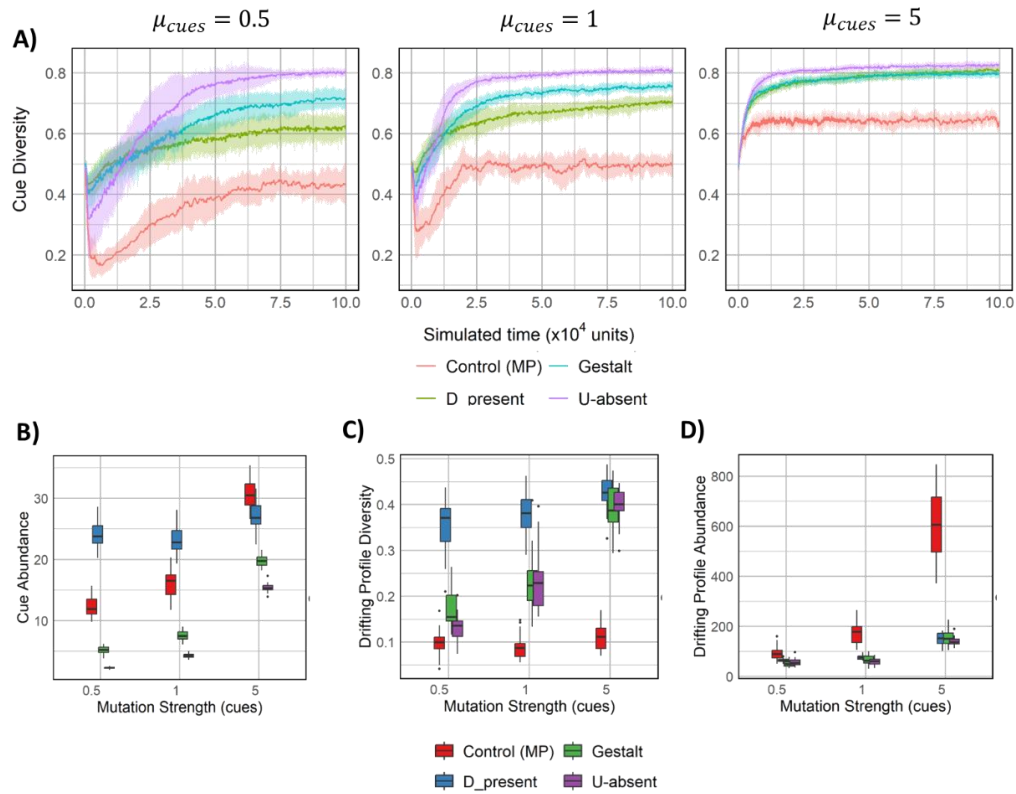
### Seasonal mortality acts as a bottleneck for diversity



**Fig S3:** Box plot for cue diversity (A) and cue abundance (B) across varying seasonal mortality under the non-evolving tolerance curves scenario. Default parameter values are used except for the parameter being varied ( $\mu_{cues} = 5$ ;  $m = 40$ ;  $\delta = 0.4$ ;  $t_{gen} = 2$ ).

Since seasonal mortality ( $\delta$ ) defines the fraction of colonies that die at the end of every season, it can biologically also be understood as the width of bottleneck that the population goes through, albeit based on a proxy fitness measure (food stock of the colony at that instant). This is reflected in the steep decline in cue diversity once more than half the population dies at the end of every season ( $\delta \geq 0.5$ ), (Fig S3A). Cue abundances increase with an increase in seasonal mortality for all recognition models (Fig S3B). The reasoning behind this increase is still unclear.

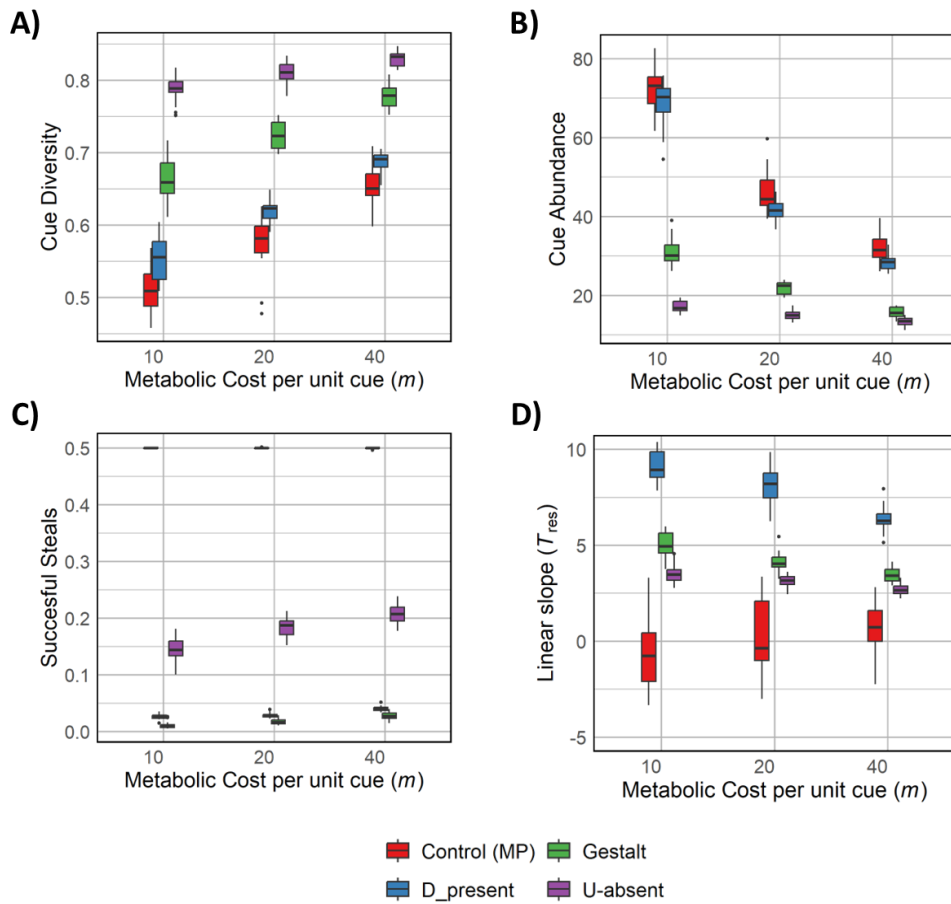
### Increasing mutation strength of cues increases cue diversity and abundance



**Fig S4:** Temporal trajectories (A) for cue diversity for various values of mutation strength of cues ( $\mu_{cues}$ ); Boxplots for cue abundance (B), drifting profile diversity (C), and drifting profile abundance (D) for varying values of  $\mu_{cues}$ . Apart from the mutation strength of cues, all parameters have default values ( $m = 40$ ;  $\delta = 0.4$ ;  $t_{gen} = 2$ ). The X-axis is limited to 100,000 time units as equilibrium was observed beyond this point.

Cue diversity and abundance increased as mutation strength increased (Fig S4A, S4B). This is expected as the cue profile values would explore a larger numeric space due to increased drift under higher mutation strength (Fig S4C, S4D). Interestingly, an initial dip in cue diversity occurs for low mutation strength values. This initial loss in diversity could be due to a bunch of fit lineages that benefit from a low mutation strength and monopolize the population with their offspring at the start of the simulation. However, a subsequent rise in cue diversity is seen, indicating our model resolves the Crozier paradox across a range of mutation strengths. This is subject to limits, as higher mutation strength ( $\mu_{cues} = 10$ ) resulted in population extinction for D-present model. This suggests sensitivity of D-present model to high mutation strengths, presumably due to too much stochasticity in the numeric space that cue values exist over. This would make staying at the equilibrium point for the population harder, possibly resulting in extinction.

## Effect of metabolic costs on the model under co-evolution scenario



**Fig S5:** Box plot for cue diversity (A), cue abundance (B), fraction of successful raids (C), and evolved linear slope (D) across varying metabolic costs under the co-evolving scenario. Default parameter values are used except for the parameter being varied ( $\mu_{cues} = 5$ ;  $\delta = 0.4$ ;  $t_{gen} = 2$ ).

Cue diversity increased (Fig S5A) and cue abundance decreased (Fig S5B) with increasing metabolic costs in the co-evolving scenario. While trends are similar to the scenario where tolerance slopes are fixed and non-evolving (Fig 2C; 2D), the effect of metabolic costs on cue diversity is much higher in the co-evolving scenario. Interestingly, lower success in raiding is observed for all metabolic costs in the co-evolving scenario (Fig S5C) compared to earlier results where the recognition system couldn't evolve (Fig S2B).

An increase in stealing success with increasing metabolic costs is seen for U-absent (Fig S5C), even though cue diversity increases (Fig S5A). Reduced stringency with increasing metabolic costs may explain this, as smaller chemical distances between profiles may still be recognized as nestmates, increasing stealing success.

This suggests that our way of measuring cue diversity may be incomplete for the coevolution scenario as increasing diversity does not necessarily correlate to a decrease in raiding success. The adaptive nature of the recognition system might explain this dilemma, as higher evolved stringency can help differentiate between nestmates and non-nestmates even when diversity (and thus chemical distance) is low. Another possibility is that higher diversity in such a scenario may in fact increase raiding success in asymmetric models as diverse profiles have a higher likelihood of deceiving other colonies. Thus, average cue diversity calculated using pairwise distances is not an optimal metric for measuring diversity as raiding success is also determined by an adaptive tolerance curve.

## Supplementary Linear Models

The tables given below give the marginal mean estimates, standard errors, t-values, and significance values from the linear models. To determine the effect of a parameter on a model system, the estimate of its interaction term (for example, D\_present:t\_gen in SLM1) is summed with the estimate for the control MP (t\_gen in SLM1). These inferences can also be made from plots generated with the *predict()* function in R. All reported values except significance values (p-val) have been rounded to 3 decimal points unless precision calls for otherwise. \* in R formula used for linear models refers to independent and interaction effects.

### SLM1: Effect of recognition model and $t_{gen}$ on cue diversity

	Estimate	Std. Error	t value	p-val
(Intercept)	0.507	0.007	66.098	1.72E-152
D_present	0.275	0.011	25.403	5.95E-69
Gestalt	0.251	0.019	23.198	2.24E-62
U-absent	0.268	0.011	24.722	5.99E-67
t_gen	0.066	0.002	23.667	8.38E-64
D_present:t_gen	-0.046	0.004	-11.86	1.12E-25
Gestalt:t_gen	-0.043	0.004	-11.052	4.35E-23
U-absent:t_gen	-0.034	0.004	-8.72	5.44E-16

Residual standard error: 0.023 on 232 degrees of freedom

Adjusted R-squared: 0.9307

F-statistic: 459.7 on 7 and 232 DF

Reference: Control MP

Formula:  $\text{lm}(\text{Avg. Cue Diversity} \sim \text{recog\_model} * t_{gen})$

A positive coefficient for the interaction terms indicates an increase in cue diversity for all recognition models with increasing  $t_{gen}$ . This increase is the highest for U-absent followed by Gestalt and D-present.

### SLM2: Effect of recognition model and $t_{gen}$ on cue abundance

	Estimate	Std. Error	t value	p-val
(Intercept)	61.265	1.282	47.782	4.69E-122
D_present	-40.523	1.813	-22.350	8.81E-60
Gestalt	-38.508	1.813	-21.236	2.58E-56
U-absent	-41.101	1.813	-22.666	9.38E-61
t_gen	-12.507	0.468	-26.713	9.86E-73
D_present:t_gen	14.574	0.662	22.0122	9.82E-59
Gestalt:t_gen	10.673	0.662	16.12	9.99E-40
U-absent:t_gen	9.715	0.662	14.672	6.39E-35

Residual standard error: 4.055 on 232 degrees of freedom

Adjusted R-squared: 0.855

F-statistic: 201.4 on 7 and 232

Reference: Control MP

Formula:  $\text{lm}(\text{Avg. Cue Abundance} \sim \text{recog\_model} * t_{gen})$



A positive coefficient for the interaction term of D-present with  $t_{gen}$  indicates an increase in cue abundance as time of resource renewal increases. A negative coefficient for Gestalt and U-absent indicates decreasing cue abundance with increasing  $t_{gen}$ . This decreasing effect is stronger for U-absent than Gestalt.

### SLM3: Effect of recognition model and metabolic costs on cue diversity

	Estimate	Std. Error	t value	p-val
(Intercept)	0.519049	0.004413	117.6176	3.34E-199
D_present	0.211388	0.007182	29.43216	6.97E-78
Gestalt	0.253262	0.006241	40.58053	1.42E-103
U-absent	0.294483	0.006241	47.18547	2.15E-116
metabolic_cost	0.002469	9.57E-05	25.79207	3.65E-68
D_present:metabolic_cost	-0.00032	0.000219	-1.46122	0.145396
Gestalt:metabolic_cost	-0.00176	0.000135	-13.0185	4.81E-29
U-absent:metabolic_cost	-0.002	0.000135	-14.7887	9.81E-35

Residual standard error: 0.02139 on 218 degrees of freedom

Adjusted R-squared: 0.9522

F-statistic: 641.9 on 7 and 218 DF

Reference: Control MP

Formula:  $\text{lm}(\text{Avg. Cue Diversity} \sim \text{recog\_model} * \text{metabolic\_cost})$

An increase in cue diversity is observed with an increase in metabolic costs for all recognition models. This effect is the strongest for D-present followed by Gestalt and then U-absent.

### SLM4: Effect of recognition model and metabolic costs on cue abundance

	Estimate	Std. Error	t value	p-val
(Intercept)	67.929	1.8669	36.3857	1.44E-94
D_present	24.615	3.0384	8.1011	3.85E-14
Gestalt	-34.795	2.6402	-13.1791	1.47E-29
U-absent	-46.872	2.6402	-17.7532	3.22E-44
metabolic_cost	-0.6706	0.0405	-16.5581	2.04E-40
D_present:metabolic_cost	-0.88	0.0928	-9.4853	4.35E-18
Gestalt:metabolic_cost	0.428	0.0573	7.4853	1.74E-12
U-absent:metabolic_cost	0.558	0.0573	9.7368	7.83E-19

Residual standard error: 8.419 on 218 degrees of freedom

Adjusted R-squared: 0.8524

F-statistic: 186.7 on 7 and 218 DF

Reference: Control MP

Formula:  $\text{lm}(\text{Avg. Cue Abundance} \sim \text{recog\_model} * \text{metabolic\_cost})$

Cue abundance decreases for increasing metabolic costs for all recognition systems. This effect is the strongest for D-present, followed by Gestalt and U-absent respectively.

### SLM5: Effect of recognition model and $t_{gen}$ on cue diversity in co-evolution scenario

	Estimate	Std. Error	t value	p-val
--	----------	------------	---------	-------

(Intercept)	0.5302	0.0079	66.8028	4.66E-153
D_present	0.1883	0.0112	16.7805	7.37E-42
Gestalt	0.2469	0.0112	21.9952	1.39E-58
U-absent	0.2698	0.0114	23.761	5.79E-64
t_gen	0.0575	0.0029	19.8412	8.25E-52
D_present:t_gen	-0.0693	0.0041	-16.9112	2.74E-42
Gestalt:t_gen	-0.0565	0.0041	-13.794	5.62E-32
U-absent:t_gen	-0.0406	0.0041	-9.8394	2.64E-19

Residual standard error: 0.02461 on 232 degrees of freedom

Adjusted R-squared: 0.9081

F-statistic: 338.2 on 7 and 232 DF

Reference: Control MP

Formula:  $\text{lm}(\text{Avg. Cue Diversity} \sim \text{recog\_model} * t_{\text{gen}})$

A decrease in cue diversity is observed with increasing  $t_{\text{gen}}$  for D-present model. An opposite trend is seen in U-absent. The coefficient for Gestalt is an order of magnitude less than the reference estimate, suggesting low effect of increasing  $t_{\text{gen}}$  on the recognition model.

#### SLM6: Effect of recognition model and $t_{\text{gen}}$ on cue abundance in co-evolution scenario

	Estimate	Std. Error	t value	p-val
(Intercept)	60.0405	1.165459	51.5166	5.49E-129
D_present	-38.4705	1.648207	-23.3408	8.24E-63
Gestalt	-43.589	1.648207	-26.4463	5.71E-72
U-absent	-43.9453	1.648207	-26.6625	1.38E-72
t_gen	-12.2031	0.425565	-28.6752	3.21E-78
D_present:t_gen	14.7344	0.60184	24.48224	3.08E-66
Gestalt:t_gen	12.0229	0.60184	19.97693	2.54E-52
U-absent:t_gen	10.3478	0.60184	17.19359	2.82E-43

Residual standard error: 3.686 on 232 degrees of freedom

Adjusted R-squared: 0.889

F-statistic: 274.3 on 7 and 232 DF

Reference: Control MP

Formula:  $\text{lm}(\text{Avg. Cue Abundance} \sim \text{recog\_model} * t_{\text{gen}})$

Cue abundance shows an increasing trend for D-present and a decreasing trend for U-absent as  $t_{\text{gen}}$  is increased. Similar to cue diversity (SLM6), the coefficient for Gestalt model is quite close to zero, albeit statistically significant.

#### SLM7: Effect of recognition model and $t_{\text{gen}}$ on tolerance slope in co-evolution scenario

	Estimate	Std. Error	t value	p-val
(Intercept)	0.247509	0.277595	0.891618	0.373521
D_present	4.175933	0.392579	10.63718	8.81E-22
Gestalt	2.687111	0.392579	6.844766	6.77E-11
U-absent	2.075587	0.392579	5.287056	2.87E-07
t_gen	-0.01079	0.101363	-0.10644	0.915327

D_present:t_gen	0.819278	0.14335	5.715242	3.35E-08
Gestalt:t_gen	0.203363	0.14335	1.41865	0.157343
U-absent:t_gen	0.112421	0.14335	0.784246	0.433696

Residual standard error: 0.8439 on 232 degrees of freedom

Adjusted R-squared: 0.888

F-statistic: 271.7 on 7 and 232 DF

Reference: Control MP

Formula:  $\text{lm}(\text{Avg. Evolved Slope} \sim \text{recog\_model} * t_{\text{gen}})$

Increasing  $t_{\text{gen}}$  causes a significant increase in evolved tolerance slope for D-present model. Other models are not significantly affected by varying  $t_{\text{gen}}$ .