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

Credentials

# Abstract

# Introduction

# Materials and Methods

## The Model

The purpose of the model is to study how cue diversity and abundance may evolve despite Crozier’s paradox in a situation where ant colonies can forage or steal resources from another nest. The model was developed in C++ (enter citation), and an instance of a simulated population comprises of ant colonies containing workers. Each worker ant can adaptively choose between foraging or stealing resources from another colony. Workers stealing from another colony or returning with resources to their home colony undergo a recognition check by a random worker, where entry is permitted if the intruder is recognized as a nestmate. Gillespie algorithm was utilized for dynamically scheduling events like foraging and stealing in the model. Once an action is completed, the period before the next action is sampled from an exponential distribution with expected value 1.

## Recognition Cues

Each colony has a chemical profile which serves as the recognition signal against which an intruder is compared. Ant CHCs don’t all vary independently, and this can be due to physical as well as biosynthetic constraints. Thus, many CHC dynamics and related recognition procedures are quite difficult to encode in a model. In our model, a colony’s recognition profile consists of compound groups. These compound groups can vary independently from each other and can be viewed as homologous series of CHCs (such as sum of 15-MeC25, 15-MeC27 and 15-MeC29) for a clearer biological understanding. 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 , the abundance of the compound groups are denoted as , and the total cue abundance for a colony denoted by is the sum of its compound groups’ abundances. Initial abundance for each group were sampled from an exponential distribution with the rate parameter . Each time a colony reproduces a daughter colony , the compound group abundance is chosen from a normal distribution with mean and standard deviation . Here, denotes the mutation step size for the chemical cues. Chemical abundance values have no upper limit but are lower bound at 0.

Workers have a profile that varies around the parent colony profile. The chemical group abundances for a worker from a parent colony is sampled from a normal distribution with mean and standard deviation . The smaller spread of the sampling curve ensures that workers in a colony have sufficiently similar profiles.

## Foraging and resource consumption

When a worker leaves the nest, it stochastically makes the decision to forage or steal food from another colony. This decision is dependent on the amount of food available and number of colonies alive . The probability of foraging is then defined as:

If the worker forages, then one unit of food is acquired from the population food source by the individual. In the event of colony invasion, 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. If accepted as a nestmate, the intruder raids one unit of food from the target colony. Since we are utilising Gillespie algorithm for scheduling time, individuals perform their actions in a random order and hence all colonies are foraging and invading simultaneously. Workers which have successfully acquired food and are returning to their colonies face another recognition process from a nestmate worker. If they are determined as nestmates, the colony acquires one unit of food. In the scenario where faulty recognition as invader happens, the resource is discarded, and the incoming worker is scheduled for the next task.

Our model also accounts for the metabolic costs needed for the production of CHCs. Whenever an individual from colony performs an action, there is a decrease in the colony food stock given by

Where is the cost of CHC biosynthesis per unit CHC. This formula was chosen such that costs increase as the total number of cues in a colony increases. The division by 2 and in the denominator is to normalise across two actions needed for gaining one food source and to normalise the total number of workers, respectively. A colony without any resources (non-positive value of colony food stock) dies.

Our model allows for multiple forms for food source regeneration, which signifies different biological scenarios (default?):

* Constant population food source: Under this condition, the amount of population food source stays constant across the simulation.
* Pulsing population food source: In this scenario, 300 units of food are provided to the population at a constant time interval . If there are any food sources 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 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 and resident colony , given by

* This is quite similar to the conjectured recognition in the Gestalt model (insert citation).

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

* Here, the compound groups that are present in the resident but absent or less abundant in the intruder are ignored, and both the values are set to before calculating the distance. Thus, the compounds less abundant in the intruder do not contribute to calculating the chemical distance.

Desirable-present (D-present): In this recognition system, emphasis is placed on whether the intruder ant 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

Each colony also has a tolerance curve, which gives the rejection propensity for a specific value of chemical distance. The rejection propensity is the binomial probability with which the intruder is rejected entry into the colony. Each colony has variables and , which code for a linear or logistic tolerance curve:

* Linear: Here, the rejection propensity for chemical distance is given by
* Logistic: The rejection propensity for chemical distance is given by

The tolerance curve itself can either be a fixed property of the model or co-evolve along with recognition cues with mutation strength . In the scenario where tolerance curves are non-evolving, linear curves are the default and variables and are sampled from normal distributions with mean 0 and 1 respectively, with standard deviation . This sampling imparts demographic heterogeneity to our model.

In the scenario where tolerance co-evolves with recognition cues, the tolerance variables for a daughter colony from parent colony are given by

Here, is the mutation step size for tolerance variables. This is analogous to as mutation step size for compound groups in the recognition profile. Once the rejection propensity has been obtained from the chemical distance and tolerance curve of the target colony, a Bernoulli sampling is performed with probability of rejection as the rejection sensitivity.

This recognition process occurs every time an individual attempts to steal food from another colony, or when individuals with food return to their own colony. Our model also defines a control scenario where the acceptance or rejection of an intruder is a coin-toss, and not dependent on any chemical profiles or tolerance curves.

## Reproduction and selection

By default, reproduction and selection are seasonal in our model. After a generation time , a fraction of maximum population size dies out. The number of colonies to be removed is adjusted to include the colonies that died due to starvation in this generation. The colony stocks act as proxy fitness values, where the ones with the lowest values are removed. The remaining colonies undergo reproduction, where the number of offspring for a colony is proportional to the food collected by 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 default colony food stock value at 25.

Our model also includes another control for the seasonal cycles: a tgen, X% of the colonies are randomly sampled, independently of their respective food stocks, and die out. This control acts as a relative measure for judging performance of Gestalt vs U-absent vs D-present recognition models.

The model also encodes for non-seasonal death and reproduction, where the death of a colony is followed by an offspring colony taking its place. Similar to seasonal death and reproduction, the probability of a colony reproducing (once another colony dies) is proportional to the food stock of the colony relative to total food available in all colonies.

## Data recorded

The average cue abundance across colonies was recorded at regular time intervals. Population level cue diversity was measured as the average of all pairwise Bray-Curtis distances between all colonies. In cases where tolerance curves co-evolve with recognition cues, the average of and were also measured. The average and standard deviation of number of offsprings per colony was also recorded.

## Code details and availability

<<need to write>>

## Model parameters and variables

| **Symbol** | **Definition** | **Default value** |
| --- | --- | --- |
|  | Total number of colonies in the population | 50 |
|  | Number of workers in a colony | 10 |
|  | Number of compound groups in a recognition profile | 10 |
|  | Number of colonies alive at an instance in the simulation | Variable |
|  | Exponential parameter for the distribution from which initial compound abundances are sampled from | 0.1 |
|  | Metabolic cost of CHC production per unit | 40 |
|  | Mutation step size for recognition cues | 5 |
|  | Mutation step size for tolerance variables (if evolving) | (don’t know yet) |
|  | Generation time for seasonal death and reproduction | 200 |
|  | Regeneration time of population food source in pulsing food source scenario | 2 |
|  | Food stock available for colony | Variable |
|  | Food available at population level in constant food source scenario | (don’t know yet) |
|  | Fraction of total colonies that die seasonally | 0.4 |
|  | Intercept variable for tolerance curve | \* |
|  | Slope variable for tolerance curve | \* |
|  | Cue abundance for compound group in colony | Variable |
|  | Total cue abundance for colony | Variable |