Miniprojects — Computational Biology 2023-2024

General information

The idea of the miniprojects is to work on a small computational/modeling project in groups using the (spatial) modeling skills you have gained in this course. The expected output is a short research report, with an abstract, introduction, results, discussion, conclusions, and methods. You need to discuss your findings in the light of at least two papers. The focus should be on your own modeling efforts, with *some* discussion in light of recent literature so you don't work in a vacuum. We will also discuss the miniproject outcomes in a plenary session, either by every group giving short pitches about what they found, or by Paulien discussing each project and asking groups who worked on it for explanations or comments.

Start simple: it is often good to think about the minimal implementation you can make that still satisfies what you're trying to model well enough. You can add more complexity as you see fit once you have something basic up and running. As for the grading of the projects, since they are on wildly different topics, in different frameworks, and you have different coding experience, we will just be grading on a scale of insufficient to exceptional. If you are unsure on how to write a research report, following the parts of this GSLS rubric (https://rubric.gsls-uu.nl/rubrics/rubric-research-report) that are applicable (most points except for the layman's summary and appointment scheduling) should yield a good report.

As a final tip, don't forget to check your results with a few seeds to make sure they are not flukes. Good luck and have fun!

Project I — Computational Biology 2023-2024

Does high cost minimize the production of public good?

The potential for cooperators to establish themselves in a non-cooperative populations, or, conversely, for selfish individuals to invade and possibly destroy a cooperative population has mostly been formulated in terms of relative costs and benefits of the 2 populations. This is true both in the context of 'group selection' and in the context of 'kin selection' as well as the well-known game theoretical models of cooperation (prisoners dilemma, hawk and dove). The general model result is that cooperation should not be too costly, or it will not persist.

You will investigate these results in the context of public good production. Many micro-organisms secrete products in the medium to digest food sources and/or to enable the uptake of essential resources. Sidophores are a pre-eminent example. The production is generally costly, whereas the products and therewith benefits are shared by all in the environment.

For your model, make sure the public good decays fast, and assume the public good is *essential for growth*. Also assume that both the cost of production and the benefits are proportional to the amount produced/used. First assume the cost of production of a unit of public good is low relative to its benefit. (e.g. a ratio 1:10) and study invasion of bacteria who produce more or less public good. and then study what increasing costs does to these invasion properties. After these ecological experiments, let production of public good evolve.

Compare your model and your results to the article by Ohtsuki *et. al* [1] Can you check if, under certain parameter settings there is quantitative agreement?

Discuss your finding in the light of at least one relevant, recent paper.

[1] Ohtsuki, Hisashi, et al. "A simple rule for the evolution of cooperation on graphs and social networks" Nature 441.7092 (2006): 502-505.

Project II — Computational Biology 2023-2024

Why do animals/cells die?
Evolution of programmed death

Many organisms die because they are eaten or because of some accident. However even without such mishaps, they have a finite lifetime. The lifetime of closely related species can vary by an order of magnitude. Moreover it it has been experimentally shown that simple (knockout) mutations can increase the lifespan (e.g. *Extended longevity in mice lacking the insulin receptor in adipose tissue* M Blüher, BB Kahn, CR Kahn - Science, 2003). In other words it is apparently possible to live longer, and we should therefore conclude that the extant lifetime of a species is an "evolutionary choice". Explanations have been sought in terms of continued evolutionary adaptability to changing environments, or in terms of trade-offs (e.g. longer lifespan implies less fecundity) but data do not support this notion.

Here we will ask if, and how, higher death-rate could give a long-term competitive advantage in a constant environment due to emergent pattern formation.

You can for example explore this problem in a system of predators and preys.

First study ecological dynamics between species differing in death rates, and look for example at the following:

- Who out competes whom?
- Can otherwise identical lineages with different death rates coexist? Why?
- If you find coexistence, test whether these different lineages can exist without the other? (i.o.w., can high or low death rate lineages be 'rescued' by lineages with other death rates?)

Next, study how death rates evolve (focus on the interesting cases you found from the questions above). Note: the *interesting cases* you should find exist in a rather narrow parameter regime. If you don't find them, then feel free to skip to the evolutionary simulations: the system might more easily evolve it than you can find it by manually trying, and you can then do ecological experiments with information from the evolved values.

Compare your results with results/discussions found in the literature, e.g. in similar models (discuss both differences in the model/experiments and in the general discussions).

Project III — Computational Biology 2023-2024

Influence of HGT on the information threshold

The Error Threshold describes the phenomenon that, given that replication isn't 100% accurate, survival of the fittest may not hold. As described in the lectures and in one of the practical questions, this error threshold depends on how much faster the "fittest" replicator replicates compared to its mutants. Survival of the fittest holds if:

$$\text{quality of replication} > \frac{a_2}{a_1} = \frac{1}{\sigma} \tag{1}$$

a1 and a2 are replication rates of master and mutants, respectively

σ is the selection coefficient a1/a2

In the simple ODEs from the practicals we have studied this problem in well-mixed conditions with an infinite population size. We have also studied the problem in a CA with finite populations and stochastic dynamics. The latter problem is often described as Mullers Ratchet: the unavoidable accumulation of deleterious mutations.

In this mini-project, study the influence of horizontal gene transfer (HGT) on the Error Threshold. Focus on the following questions:

- Does HGT alleviate or hamper this Error Threshold? What are the conditions?
- In a spatial model, study whether HGT can evolve and study which features of genes and HGT can help to strengthen your results
- Consider the work by Takeuchi *et al.*[2]. How does your model compare to this study? What is similar/different?

Discuss your results in the light of other (more recent) literature.

[2]Horizontal gene transfer can rescue prokaryotes from Muller's ratchet: benefit of DNA from dead cells and population subdivision, Takeuchi et al., 2014

Project IV — Computational Biology 2023-2024

RNA Recombination & Spatial Structure: effects on the evolution of neutrality

Neutrality is a key property of RNA adaptive landscapes that makes them efficient. The one-mutational neighborhood of RNA molecules is largely neutral, meaning that a change in sequence does not change the structure of the molecule. The sequence space is percolated with neutral networks or neutral paths, interconnected networks of sequence mutations that are neutral with respect to the structure and span the whole sequence space. Evolution on this landscape leads the population to flatter, more connected, parts of the neutral network. The measure of neutrality of a sequence, that reflects its connectivity in the neutral network is λ, the fraction of neutral mutations in the one-mutational neighborhood. In an evolving RNA population, the master's value of λ tends to increase.

Here we will study the effect (if any) of recombination and spatial structure on the evolution of λ and the one-mutational neighborhood. A model of an RNA population evolving on a 2D grid towards a target structure will be given (https://tbb.bio.uu.nl/BINF/software/RNA2D.py). You will be able to edit the model, to investigate the following questions:

- Does space affect the evolution of the master's λ?(e.g. mixing vs non-mixing)
- Does space affect the evolution of the master's mutational neighborhood?
- Does recombination with neighbors or with non-neighbors affect the evolution of λ?

You will run multiple simulations and compare the results for different settings and parameters (e.g. try different mutation/ recombination rates). Compare your results to the literature.