Evolutionary Dynamics: Homework 04

Guido Putignano, Lorenzo Tarricone, Gavriel Hannuna,
Athanasia Sapountzi

February 22, 2024

1 Problem 1: TSG inactivation in small populations

Consider the approximate model of TSG inactivation for $N \ll \frac{1}{\sqrt{u^2}}$:

$$\begin{aligned}\dot{X}_0 &= -u_1 X_0 \\ \dot{X}_1 &= u_1 X_0 - N u_2 X_1 \\ \dot{X}_2 &= N u_2 X_1\end{aligned}$$

Solve this set of ODEs analytically for the initial conditions $X_0(0) = 1$, $X_1(0) = X_2(0) = 0$, to obtain the probability of finding both TSG alleles inactive at time t.

Hint: For the analytical approach, solve the system iteratively for X_0, then for X_1 and X_2. A linear ODE $\frac{dx}{dt} + bx(t) = f(t)$ can be integrated by multiplying both sides with the integrating factor e^{bt}, because $[e^{bt}x]\frac{d}{dt} = e^{bt}\frac{dx}{dt} + be^{bt}x$. Thus, $x(t) - x(t_0)e^{-b(t-t_0)} = e^{-bt}\int_{t_0}^{t} e^{b\tau} f(\tau)d\tau$. To compute X_2, use the conservation of probability.

Solution: Let us start by solving the first differential equation. And since $X_0(0) = 1$ the integral of dX will be from 1 to x.

$$\frac{dX_0}{dt} = -u_1 X_0$$

$$\int_1^x \frac{dX_0}{X_0} dX = \int_0^t -u_1 d\tau$$

$$ln(X_0) = -u_1 t$$

$$X_0 = e^{-u_1 t}$$

Once we solved X_0 we can continue by solving the second ODE, which is of the form $\frac{dx}{dt} + bx(t) = f(t)$ that can be integrated by multiplying both sides with the integrating factor e^{bt}, because $[e^{bt}x]\frac{d}{dt} = e^{bt}\frac{dx}{dt} + be^{bt}x$. Thus, $x(t) - x(t_0)e^{-b(t-t_0)} =$

$e^{-bt}\int_{t_0}^{t} e^{b\tau}f(\tau)d\tau$. So, taking into account that $X_1(0) = 0$, the equation will be

$$X_1 = e^{-Nu_2 t}\int_0^t e^{Nu_2\tau}u_1 e^{-u_1\tau}d\tau$$

$$X_1 = u_1 e^{-Nu_2 t}[\frac{e^{t(Nu_2-u_1)}}{(Nu_2-u_1)} - \frac{1}{(Nu_2-u_1)}]$$

$$X_1 = u_1[\frac{e^{-tu_1} - e^{-Nu_2 t}}{(Nu_2-u_1)}]$$

Finally, we can solve the last ODE, inserting X_1 and integrating.

$$\frac{dX_2}{dt} = Nu_2u_1[\frac{e^{-tu_1} - e^{-Nu_2 t}}{(Nu_2-u_1)}]$$

Knowing that $X_2(0) = 0$ we can write

$$X_2 = \frac{Nu_2u_1}{(Nu_2-u_1)}\int_0^t e^{-\tau u_1} - e^{-Nu_2\tau}d\tau$$

$$X_2 = \frac{Nu_2u_1}{(Nu_2-u_1)}([\frac{1-e^{-tu_1}}{u_1}] - [\frac{1-e^{-tNu_2}}{Nu_2}])$$

$$X_2 = \frac{Nu_2 - u_1 + u_1e^{-tNu_2} - Nu_2e^{-tu_1}}{(Nu_2-u_1)}$$

$$X_2 = 1 + \frac{u_1e^{-tNu_2} - Nu_2e^{-tu_1}}{(Nu_2-u_1)}$$

X_1 and X_2 represent the probabilities at time t of having 1 and 2 TSG allele inactivated.

2 Problem 2: Chromosomal instability

For TSG inactivation with chromosomal instability (CIN), we have distinguished three cases: neutral CIN, costly CIN in small compartments, and costly CIN in large compartments. Cancer can thus either arise from two subsequent "normal" mutations, or by a normal mutation followed by CIN. Without CIN, the rate of the second mutation can be expressed as $u_2 = u + p_0 \approx 2u$, where $u \approx 10^{-7}$ is the normal point mutation rate per cell division and p_0 is the rate of LOH. With CIN, the rate of the second mutation is approximately $u_3 = 0.01$. Assuming there are n_1 class I and n_2 class II CIN genes, the rate u_c of producing a CIN mutant can be expressed as:

$$u_c = 2n_1(u + p_0) + 2n_2u \approx 2(2n_1 + n_2)u.$$

(a) Calculate the ratio C of cancers that are initiated with CIN to those initiated without CIN for the three CIN scenarios and show that C is independent of time.

Solution: We are asked to calculate ratios between "quantities of two types of cells", but instead in what follows we look at the ratio between "probabilities of finding a cell in that state at time t. (With of course the standard initial condition of all the cells starting being non-CIN and with no mutations). This is justified by the fact that the number of cells in a given state will be given in expectation by $N \cdot P(A_{\bullet}^{\#,\#})$, where $P(A_{\bullet}^{\#,\#})$ is an element of the set $\{P(A_{\bullet}^{\#,\#}) | \# \in \{+,-\}, \bullet \in \{CIN, NORMAL\}\}$. We shall call these probabilities respectively $X_{0,1,2}$ and $Y_{0,1,2}$, in order to be consistent with the notation introduced during the lecture. We will thus always calculate ratios of the kind $\frac{Y_2(t}{X_2(t)}$

- **Neutral CIN**: In this case with the correct approximation we have $X_2(t) = \frac{Nu_1u_2t^2}{2} = \frac{Nu2ut^2}{2} = Nu^2t^2$, $Y_2(t) = u_1u_ct^2 = 2(2n_1 + n_2)u^2t$ leading us to $\frac{Y_2(t)}{X_2(t)} = \frac{2(2n_1+n_2)}{N}$
- **Costly CIN in small comparments**: In this case we have $X_2(t) = Nu_1u_2t^2/2 = \frac{Nu^22t^2}{2} = Nu^2t^2/2$, $Y_2(t) = N\rho u_1u_2t^2 = N\rho u^2 2(2n_1 + n_2)t^2$ leading us to the following ratio $\frac{Y_2(t)}{X_2(t)} = \rho 2(2n_1 + n_2)$
- **Costly CIN in large comparments**: In this last case we have $X_2(t) = Nu_1u_2t^2/2 = \frac{Nu^22t^2}{2} = Nu^2t^2/2$, $Y_2(t) = Rut^2/2$. Now the ratio will be $\frac{R}{2Nu} = \frac{1}{2Nu}\frac{Nu_cru_3}{1-r} = \frac{2(2n_1+n_2)u_3r}{2(1-r)}$

(b)Explicitly compute C for $n_1 = 6$, $n_2 = 3$, and $N = 10$ cells, in the case of neutral CIN and costly CIN in small compartments, and $N = 100$ for large compartments. Assume a relative fitness disadvantage of $r = 0.9$ in the costly cases. Compare the ratios in a few sentences.

Solution:

- **Neutral CIN**: Substituting here gives a value of $\frac{2(2\cdot 6+3}{10} = 3$
- **Costly CIN in small comparments**: In this case we have $\frac{1-10/9}{1-(10/9)^{10}} 2(2 \cdot 6 + 3) \approx 1,784466$
- **Costly CIN in large comparments**: In this last case we have $\frac{2(2\cdot 6+3)\cdot 0.9\cdot 0.01}{2(1-0.9)} \approx 1.35$

When comparing the ratio C of the Neutral CIN case with respect to the costly CIN in small compartments we see a decrease. This is due to the reduced number of cells that pass to the CIN state in response to a decrease in the rate between these two states ($\frac{1}{N}Nu_3 \longrightarrow \rho Nu_3$).

The ratio when comparing C for small and large compartments is even smaller for large compartments. This might be due to the fact that the first hitting time is reached very close to the beginning and therefore there will be a lot of cells in the population with already one mutation ready to initialize a normal (i.e. non-CIN) type of cancer. This is reflected by an increase in probability $X_2(t)$

3 Problem 3: Linear Process of Colonic Crypt Transformation

The colon consists of approximately $M = 10^7$ crypts, each consisting of $N = 10^3$ cells. The mutation rate per gene is of the order of $u = 10^{-8}$ per cell division. We assume a generation time of 1 day. Consider an oncogenic mutation causing a selective advantage $s = 5\%$ in the following three scenarios:

(a) Homogeneous Tissue: Use the Moran model to calculate the expected number of neoplastic crypts (i.e., crypts in which all cells are mutated) at age 50. (1 point)

Solution:

The selective advantage is $s = 5\%$, so the relative fitness r of the mutant cells compared to normal cells is $r = 1 + 0.05 = 1.05$.

The fixation of probability of a single mutant with relative fitness r is:

$P_\rho = \frac{1-\frac{1}{r}}{1-\frac{1}{r^N}} = \frac{1-\frac{1}{1.05}}{1-\frac{1}{1.05^{10^3}}} = 0.048.$

Since we assume a generation time of 1 day:

$t = 50 * 365$ `days` $= 18250$ `days`.

The probability that mutated cells have taken over a crypt after 50 years is:

$P(t) = 1 - \exp(-Nu\rho t) = 1 - \exp(-10^3 * 10^{-8} * 0.048 * 18250) = 0.0087.$

The expected number of neoplastic crypts is $MP(t) = 10^7 * 0.0087 = 87000$.

(b) Single Stem Cell: By a similar calculation, compute the number of transformed crypts at age 50 if the cells in each crypt originate from a single stem cell. Stem cells replicate more slowly with a generation time of approximately 10 days. (1 point)

Solution:

The selective advantage is $s = 5\%$, so the relative fitness r of the mutant cells compared to normal cells is $r = 1 + 0.05 = 1.05$.

Since the cells in each crypt originate from a single stem cell the fixation of probability of a single mutant with relative fitness r is:

$P_\rho = \frac{1-\frac{1}{r}}{1-\frac{1}{r^N}} = \frac{1-\frac{1}{1.05}}{1-\frac{1}{1.05^1}} = 1.$

Since stem cells replicate more slowly with a generation time of approximately 10 days:

$t = 50 * 365/10$ `days` $= 1825$ `days`.

The probability that mutated cells have taken over a crypt after 50 years is:

$P(t) = 1 - \exp(-Nu\rho t) = 1 - \exp(-1 * 10^{-8} * 1 * 1825) = 0.000018 = 1.8 * 10^{-5}$.

The expected number of neoplastic crypts is $MP(t) = 10^7 * 1.8 * 10^{-5} = 1.8 * 10^2 = 180$.

(c) Multiple Stem Cells: Use the Moran model to compute the number of neoplastic crypts after 50 years, assuming that each crypt is maintained by a pool of $N_S = 5$ stem cells. Assume the same fitness advantage s = 5%.

Solution:

The selective advantage is $s = 5\%$, so the relative fitness r of the mutant cells compared to normal cells is $r = 1 + 0.05 = 1.05$.

Assuming that each crypt is maintained by a pool of $N_S = 5$ stem cells the fixation of probability of a single mutant with relative fitness r is:

$P_\rho = \frac{1-\frac{1}{r}}{1-\frac{1}{r^N}} = \frac{1-\frac{1}{1.05}}{1-\frac{1}{1.05^5}} = 0.22.$

Since stem cells replicate more slowly with a generation time of approximately 10 days:

$t = 50 * 365/10$ `days` $= 1825$ `days`.

The probability that mutated cells have taken over a crypt after 50 years is:

$P(t) = 1 - \exp(-Nu\rho t) = 1 - \exp(-5 * 10^{-8} * 0.22 * 1825) = 0.00002 = 2 * 10^{-5}$.

The expected number of neoplastic crypts is $MP(t) = 10^7 * 2 * 10^{-5} = 2 * 10^2 = 200$.

(d) Discussion: Discuss which tissue architecture prevents cancer initiation best.

Solution:

The Single Stem Cell architecture best prevents cancer initiation since it reduces the expected number of neoplastic crypts after 50 years down to 180. This number is smaller than the expected number of neoplastic crypts (200) using the Multiple Stem Cells architecture and far smaller than the expected number of neoplastic crypts (87000) using the Homogeneous Tissue architecture.