Branching models for telomere evolution

Emma Horton
Auckland probability seminar

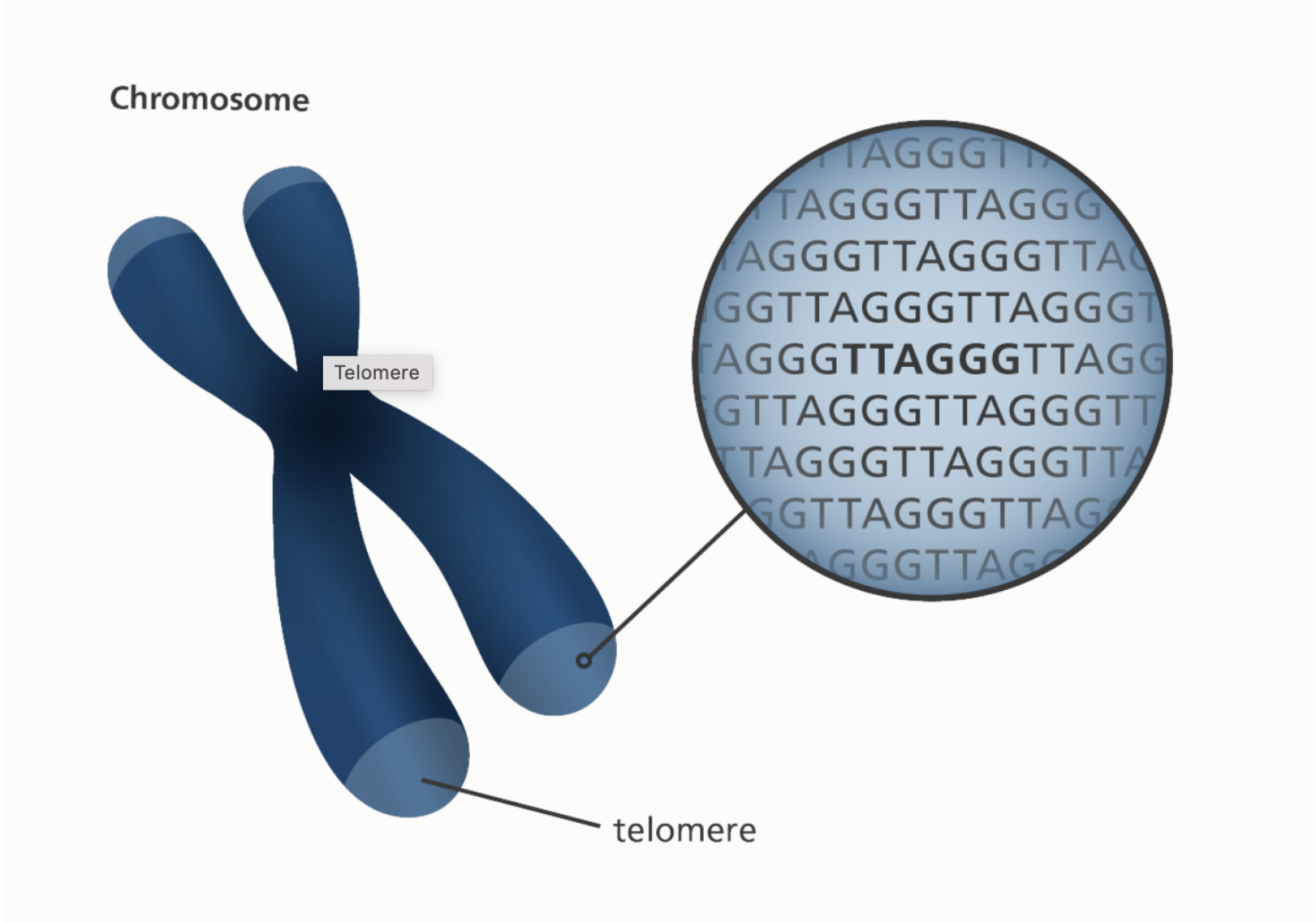
September 2, 2024

Based on joint work with Coralie Fritsch, Anne Gégout-Petit, Simon Toupance, Denis Villemonais, Constance Xhaard.

Overview

Contents

Telomeres

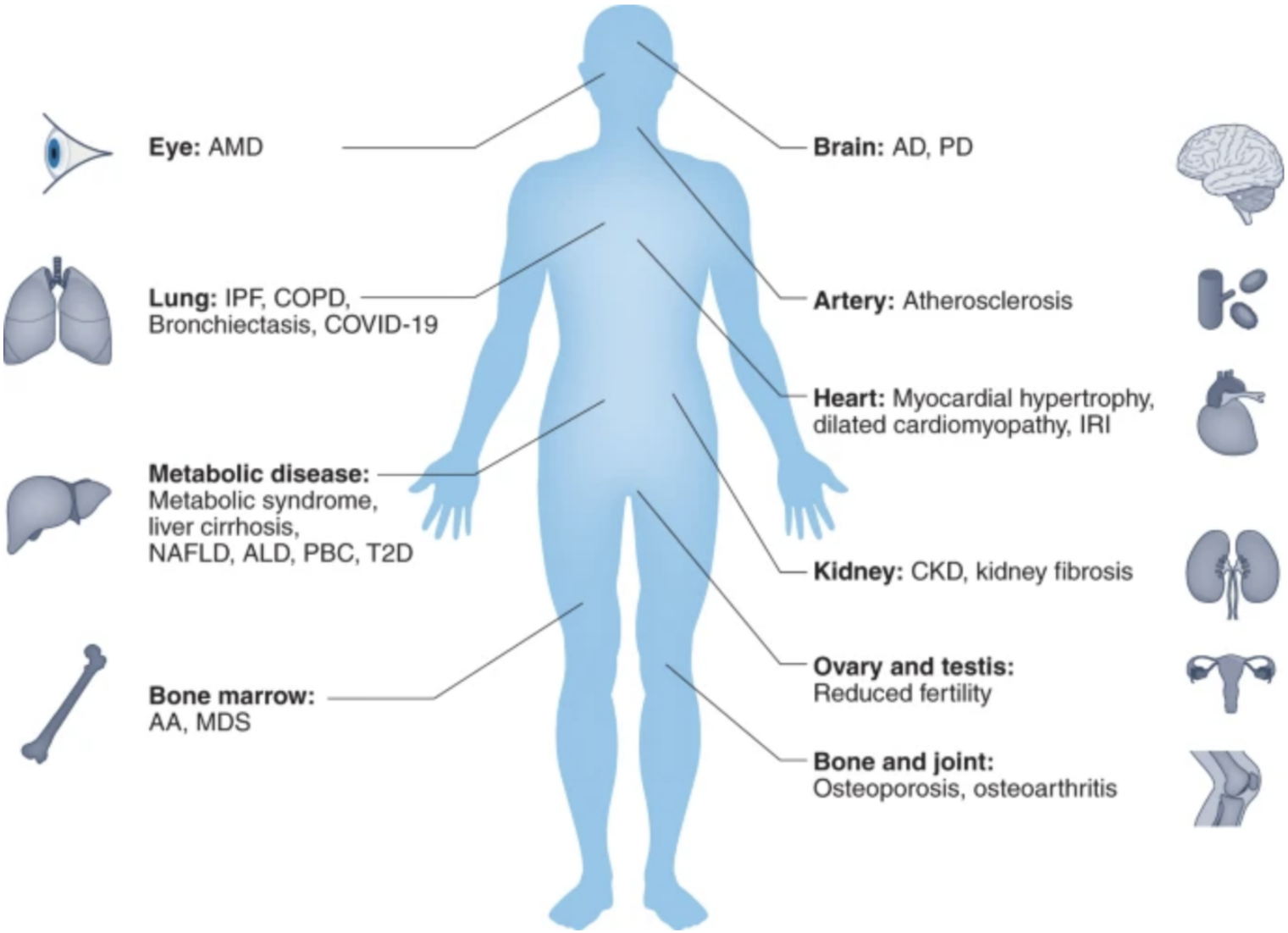

Chromosome
Telomere
TAGGGTTAGGG
AGGGTTAGGGTTAG
GGTTAGGGTTAGGGT
AGGGTTAGGGTTAGG
GTTAGGGTTAGGGTT
TAGGGTTAGGGTTA
GGTTAGGGTTAG
telomere

Telomeres

Many other factors affect the length of telomeres:

- Lifestyle factors: exercise, smoking, drinking
- Stress
- Genetics
- Exposure to pollution

Telomere related illness

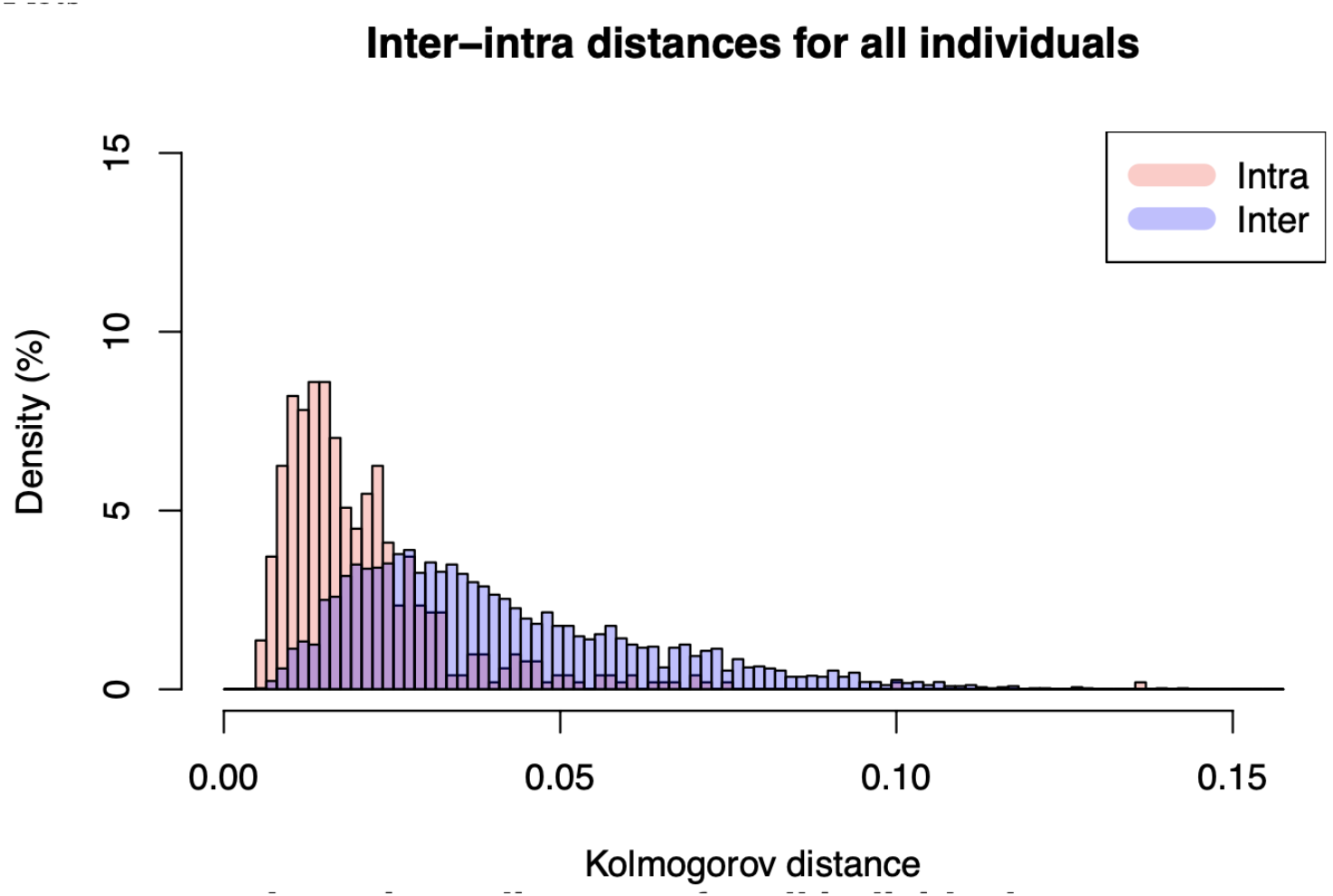


Contents

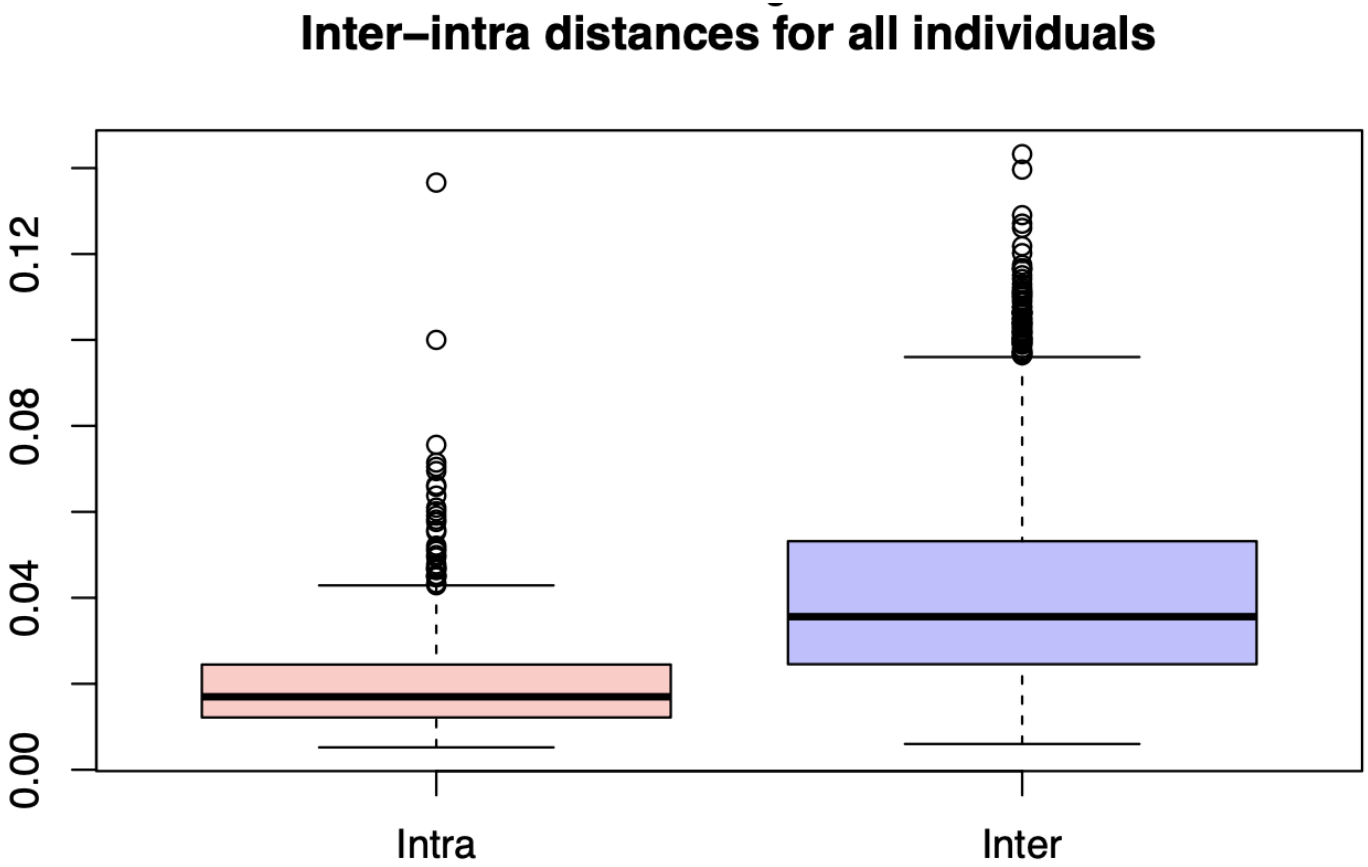
STANISLAS data

- Data on telomere length distribution for ≈ 530 patients.
- For each patient, telomere length distribution (TLD) at times T_1 and $T_2 > T_1$, parent/child, sex, family.
- Aim: to find differences/similarities between TLDs of different groups.

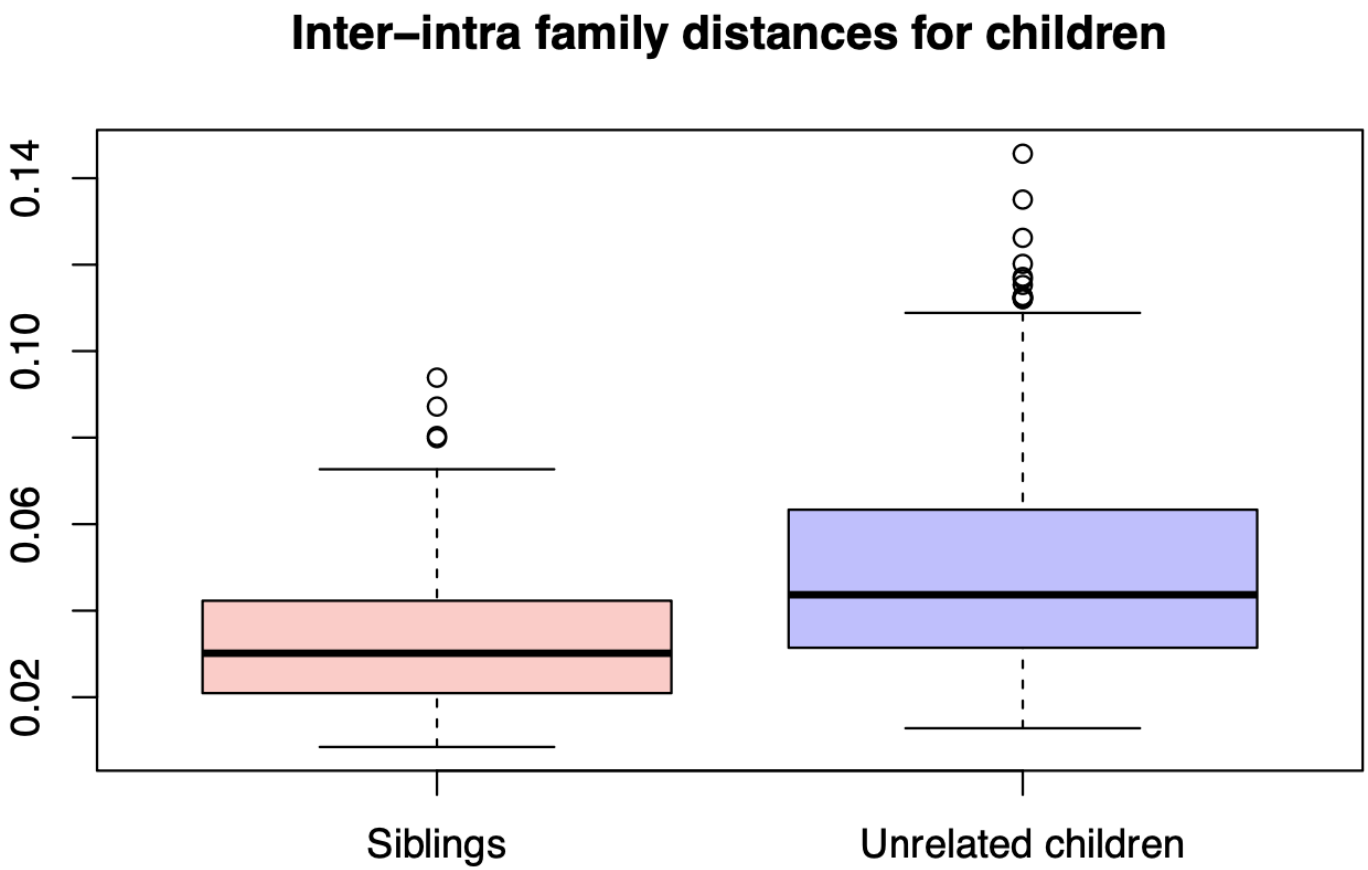
STANISLAS data



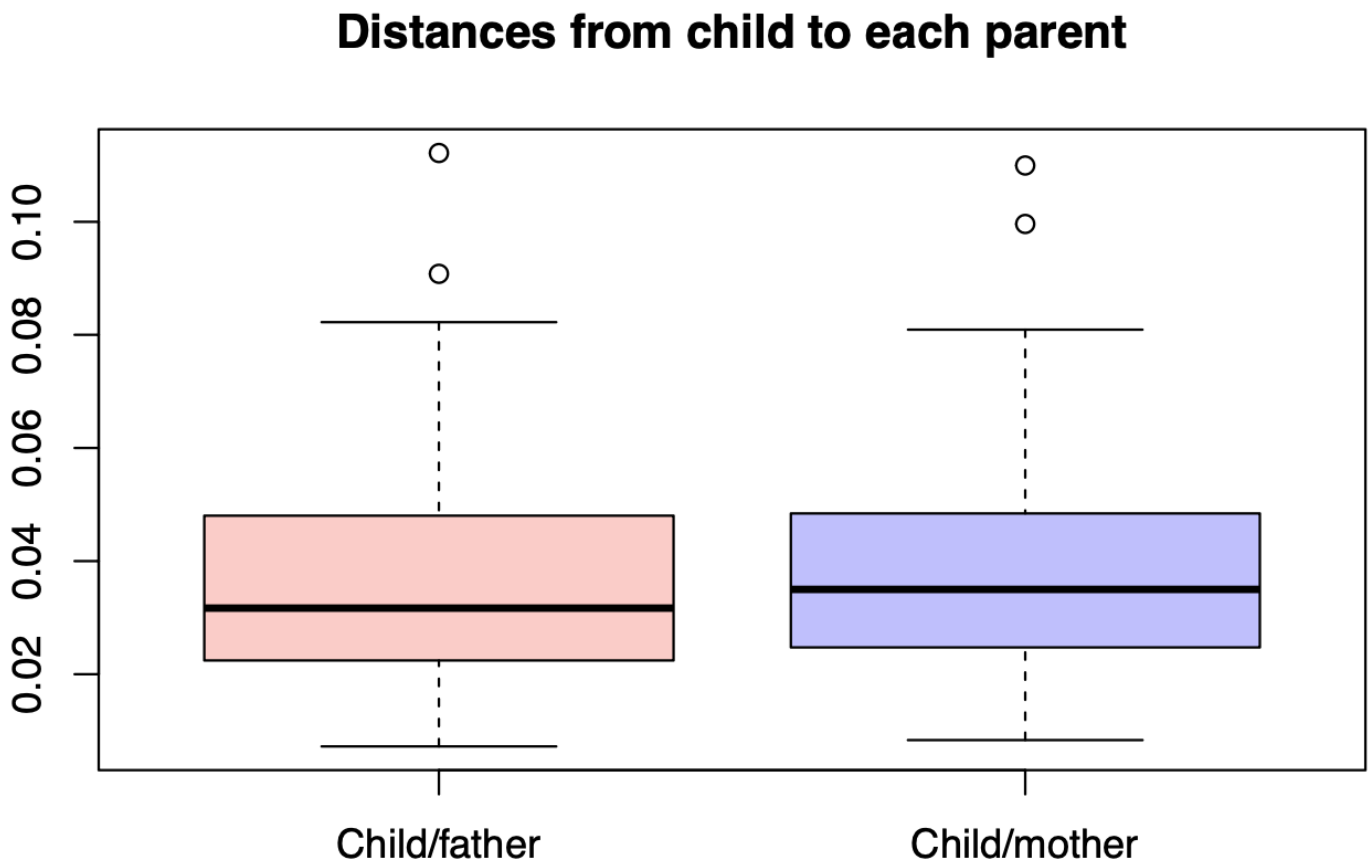
STANISLAS data



STANISLAS data



STANISLAS data



Objectives

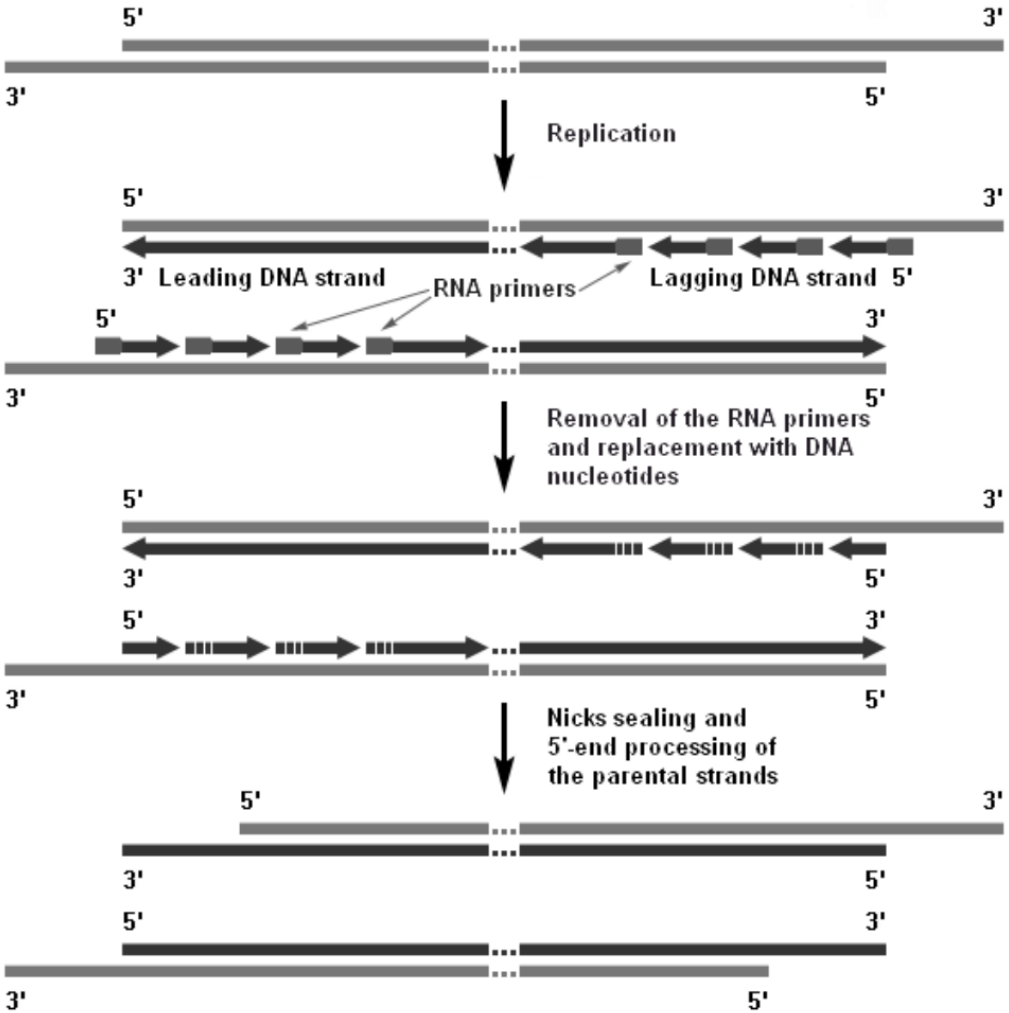
- Develop stochastic models for telomere dynamics
- Model the influence of different factors on telomere dynamics
- Understand the time evolution of these models
- Simulate model and recover characteristic properties

Contents

Telomere dynamics

- There are two types of cells: senescent and non-senescent.
- Senescent cells can no longer divide: they continue to function until "death".
- Non-senescent cells can divide, become senescent and die.
- Cells containing a telomere that is shorter than some value L_{min} die.

Telomere dyanamics: without telomerase



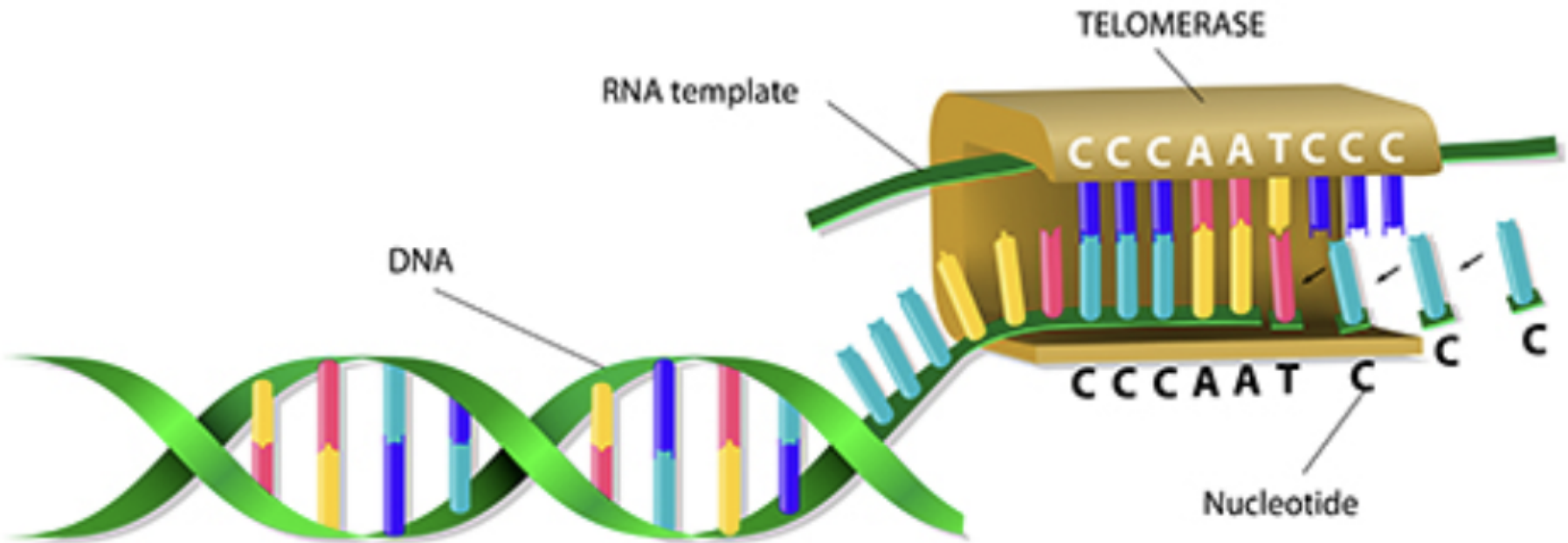
Telomere dynamics: without telomerase

- Let ω denote the length of the overhang at the end of a DNA strand.
- A chromosome with telomeres of length m, $m-\omega$, n and $n-\omega$ is denoted (m, n).
- After replication,

$$(m, n) \mapsto (m, n-\omega) + (m-\omega, n).$$

- If the length of the shortest telomere is less than $L_{\min}$ the cell dies.

Telomere dyanamics: with telomerase



Telomere dynamics: with telomerase

- Let $q(\ell)$ denote the probability that telomerase 'acts' on a telomere of length ℓ and let $\chi_\ell \sim \mathrm{Ber}(q(\ell))$.

- Given $\chi_\ell = 1$, let L denote the random quantity added to the telomere, whose law we denote μ.

- After replication,

$$(m, n) \to (m + \Lambda_m, n - \omega + \Lambda'_{n-\omega}) + (m - \omega + \Lambda''_{m-\omega}, n + \Lambda'''_n),$$

  where Λ_ℓ, Λ'_ℓ, Λ''_ℓ and Λ'''_ℓ are independent copies of $\chi_\ell L$.

- If the length of the shortest telomere is less than $L_{\min}$, say, the cell dies.

The branching process

A cell is represented by a pair (c, x) where $c = (m_j, n_j)_{j=1}^{K} \in (\mathbb{N} \times \mathbb{N})^{2K}$ and $x \in \{\mathtt{s}, \mathtt{ns}\}$.

If $x = \mathtt{ns}$,

- at rate $d_{\mathtt{ns}}(c)$, the cell dies and is removed from the system;
- at rate $s_{\mathtt{ns}}(c)$, the cell becomes senescent and we set $x = \mathtt{s}$;
- at rate $b_{\mathtt{ns}}(c)$, the cell divides into two daugher cells, $(c, \mathtt{ns}) \mapsto (c_1, \mathtt{ns}) + (c_2, \mathtt{ns})$;
- if a cell has a telomere whose length is less than $L_{\min}$, the cell dies.

If $x = \mathtt{s}$,

- at rate $d_{\mathtt{s}}(c)$, the cell dies and is removed from the system.

The branching process

- Let N_t denote the number of cells alive at time t.
- Let $E = \{c \in (\mathbb{N} \times \mathbb{N})^K : \min c > L_{\min}\}$.
- For $i \in \{1, \ldots, N_t\}$, the i-th cell is denoted $(c_i(t), x_i(t))$, where $c_i(t) = (m_{i,j}(t), n_{i,j}(t))_{j=1}^K$ and $x_i(t) \in \{\mathtt{s}, \mathtt{ns}\}$.
- The branching process is defined as

$$X_t := \sum_{i=1}^{N_t} \delta_{(c_i(t), x_i(t))}$$

- For $g : E \times \{\mathtt{s}, \mathtt{ns}\} \to [0, \infty)$ bounded and measurable, we define

$$\psi_t[g](c, x) := \mathbb{E}_{(c,x)} \left[\sum_{i=1}^{N_t} g(c_i(t), x_i(t)) \right], \quad t \geq 0, c \in E, x \in \{\mathtt{s}, \mathtt{ns}\}.$$

The branching process

- Let N_t denote the number of cells alive at time t.
- Let $E = \{c \in (\mathbb{N} \times \mathbb{N})^K : \min c > L_{\min}\}$.
- For $i \in \{1, \ldots, N_t\}$, the i-th cell is denoted $(c_i(t), x_i(t))$, where $c_i(t) = (m_{i,j}(t), n_{i,j}(t))_{j=1}^K$ and $x_i(t) \in \{\mathtt{s}, \mathtt{ns}\}$.
- The branching process is defined as

$$X_t := \sum_{i=1}^{N_t} \delta_{(c_i(t), x_i(t))}$$

- For $g : E \times \{\mathtt{s}, \mathtt{ns}\} \to [0, \infty)$ bounded and measurable, we define

$$\psi_t[g](c, x) := \mathbb{E}_{(c,x)} \left[\sum_{i=1}^{N_t} g(c_i(t), x_i(t)) \right], \quad t \geq 0, c \in E, x \in \{\mathtt{s}, \mathtt{ns}\}.$$

The branching process

- Let N_t denote the number of cells alive at time t.
- Let $E = \{c \in (\mathbb{N} \times \mathbb{N})^K : \min c > L_{\min}\}$.
- For $i \in \{1, \ldots, N_t\}$, the i-th cell is denoted $(c_i(t), x_i(t))$, where $c_i(t) = (m_{i,j}(t), n_{i,j}(t))_{j=1}^K$ and $x_i(t) \in \{\mathtt{s}, \mathtt{ns}\}$.
- The branching process is defined as

$$X_t := \sum_{i=1}^{N_t} \delta_{(c_i(t), x_i(t))}$$

- For $g : E \times \{\mathtt{s}, \mathtt{ns}\} \to [0, \infty)$ bounded and measurable, we define

$$\psi_t[g](c, x) := \mathbb{E}_{(c,x)} \left[\sum_{i=1}^{N_t} g(c_i(t), x_i(t)) \right], \quad t \geq 0, c \in E, x \in \{\mathtt{s}, \mathtt{ns}\}.$$

Perron-Frobenius result

Theorem

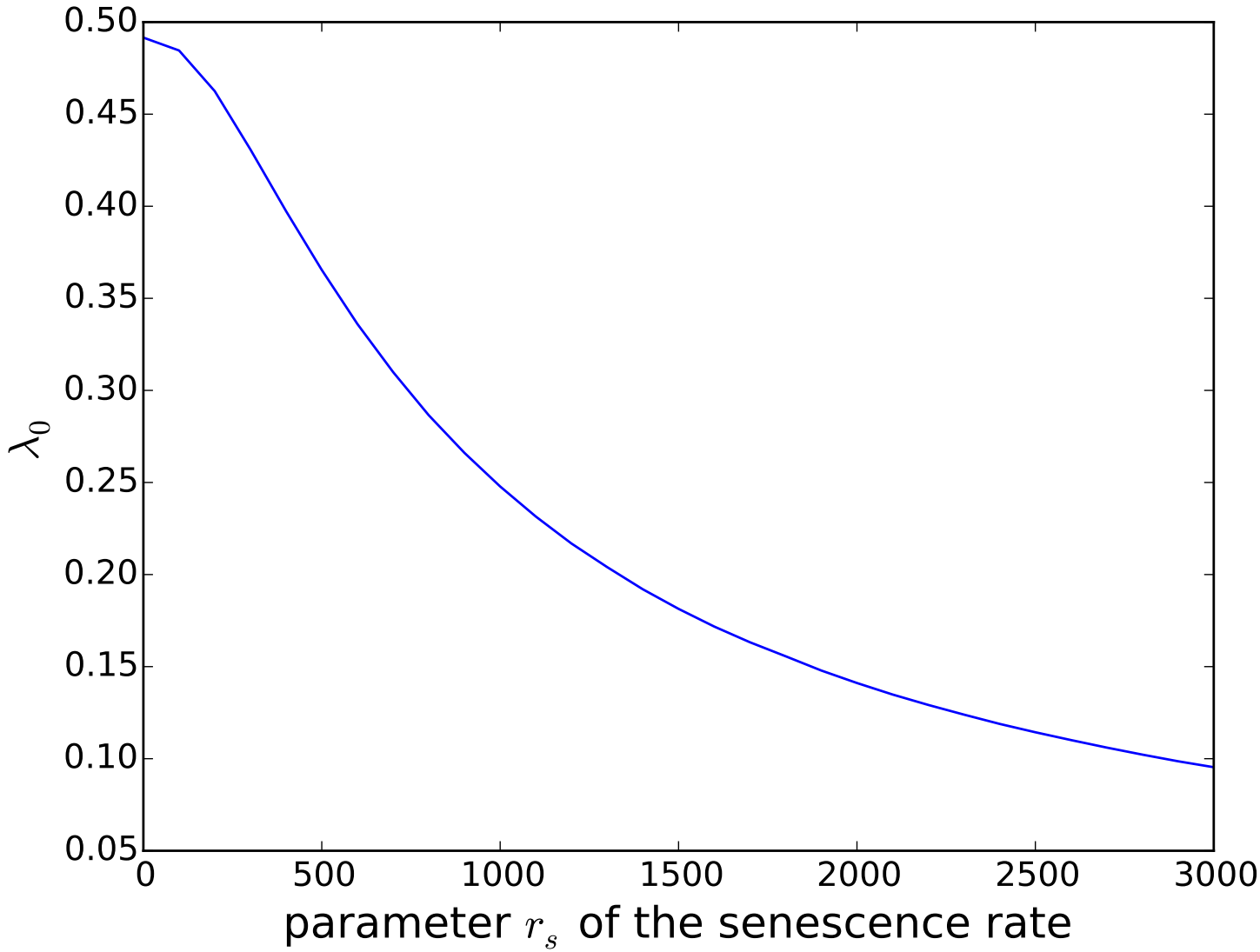
There exists a constant $\lambda_0 \in \mathbb{R}$, a function $\eta : E \times \{\mathtt{s}, \mathtt{ns}\} \to \mathbb{R}$, positive on $E \times \{\mathtt{ns}\}$ and vanishing on $E \times \{\mathtt{s}\}$, a probability measure ν on $E \times \{\mathtt{s}, \mathtt{ns}\}$, and a function $V : E \times \{\mathtt{s}, \mathtt{ns}\} \to [1, \infty)$ such that, for all $g : E \times \{\mathtt{s}, \mathtt{ns}\} \to \mathbb{R}$ with $|g| \leq V$, all $t \geq 0$, and for all $(c, x) \in E \times \{\mathtt{s}, \mathtt{ns}\}$,

$$\left| e^{-\lambda_0 t} \psi_t[g](c, x) - \eta(c, x)\nu[g] \right| \leq C e^{-\gamma t} V(c, x),$$

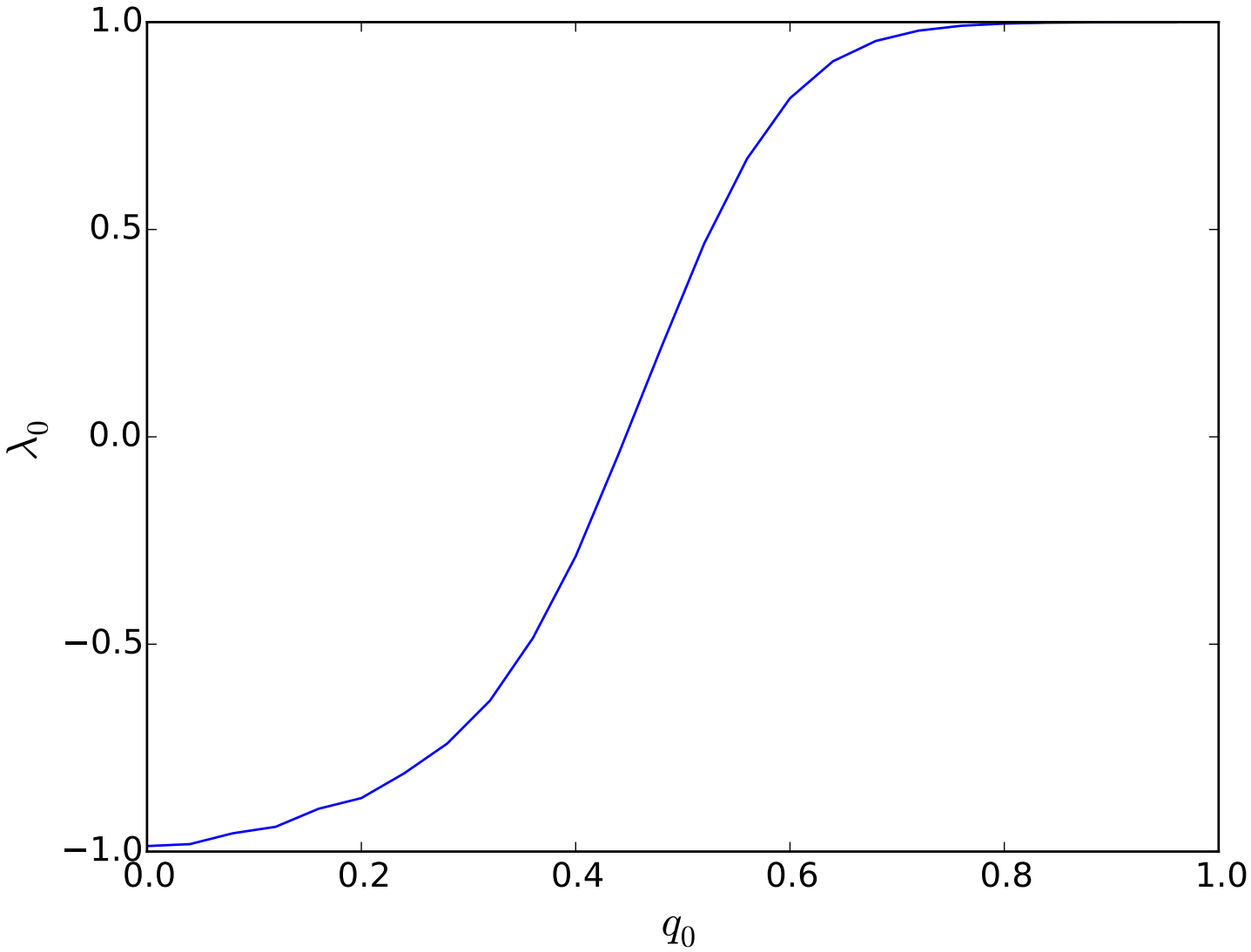
where C, γ are positive constants.

Contents

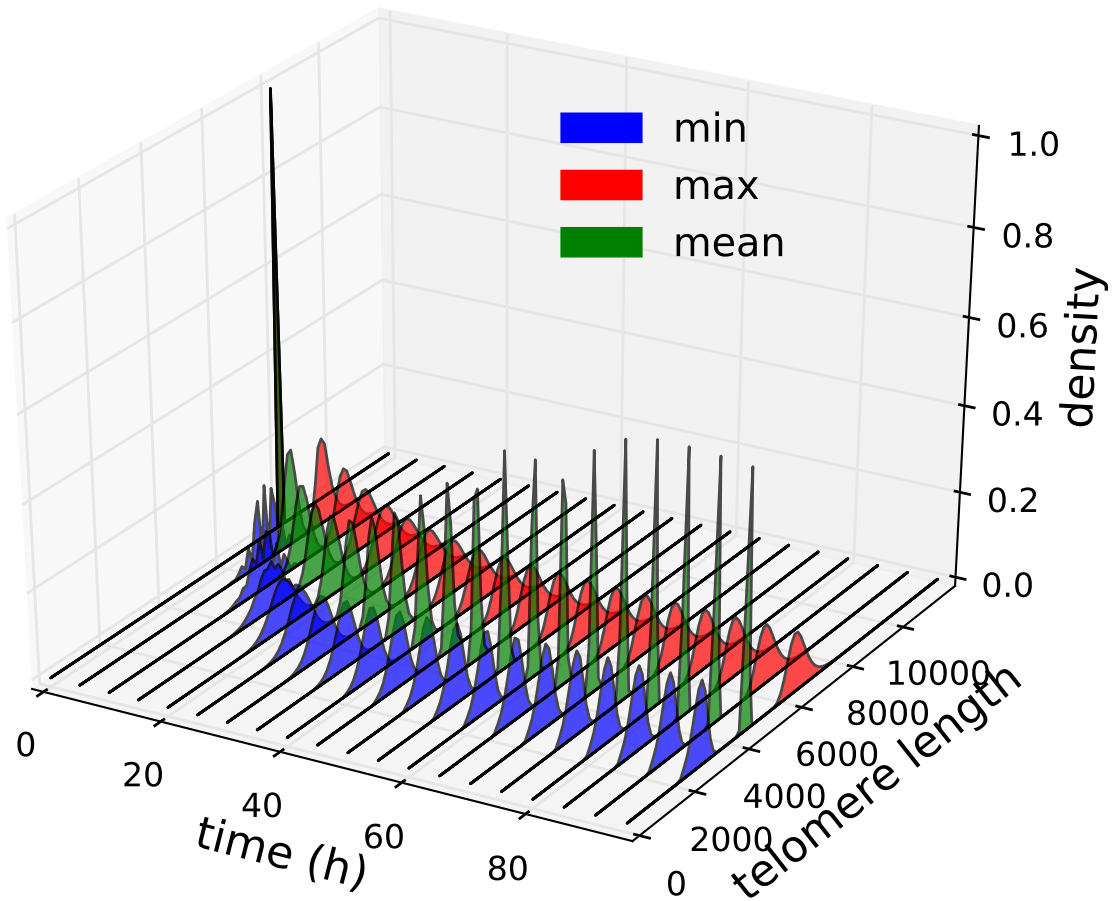
Evolution of λ_0 with respect to s_{ns}



Evolution of λ_0 with respect to q



Evolution of telomere length distribution



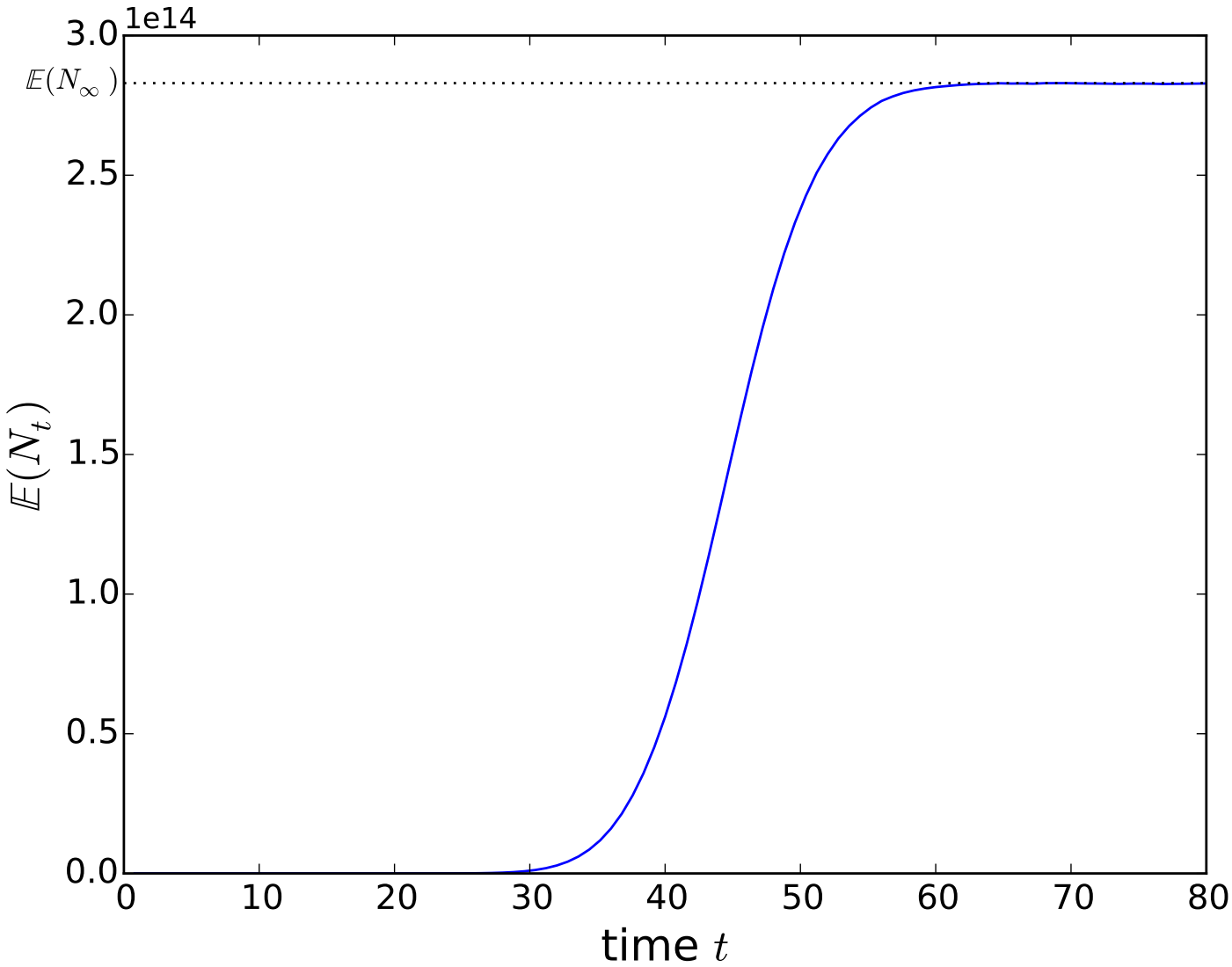
Hayflick limit

- When telomerase isn't present or the rate of senescence if sufficiently high, the population stops growing.
- The Hayflick limit is the total number of doublings of the population:

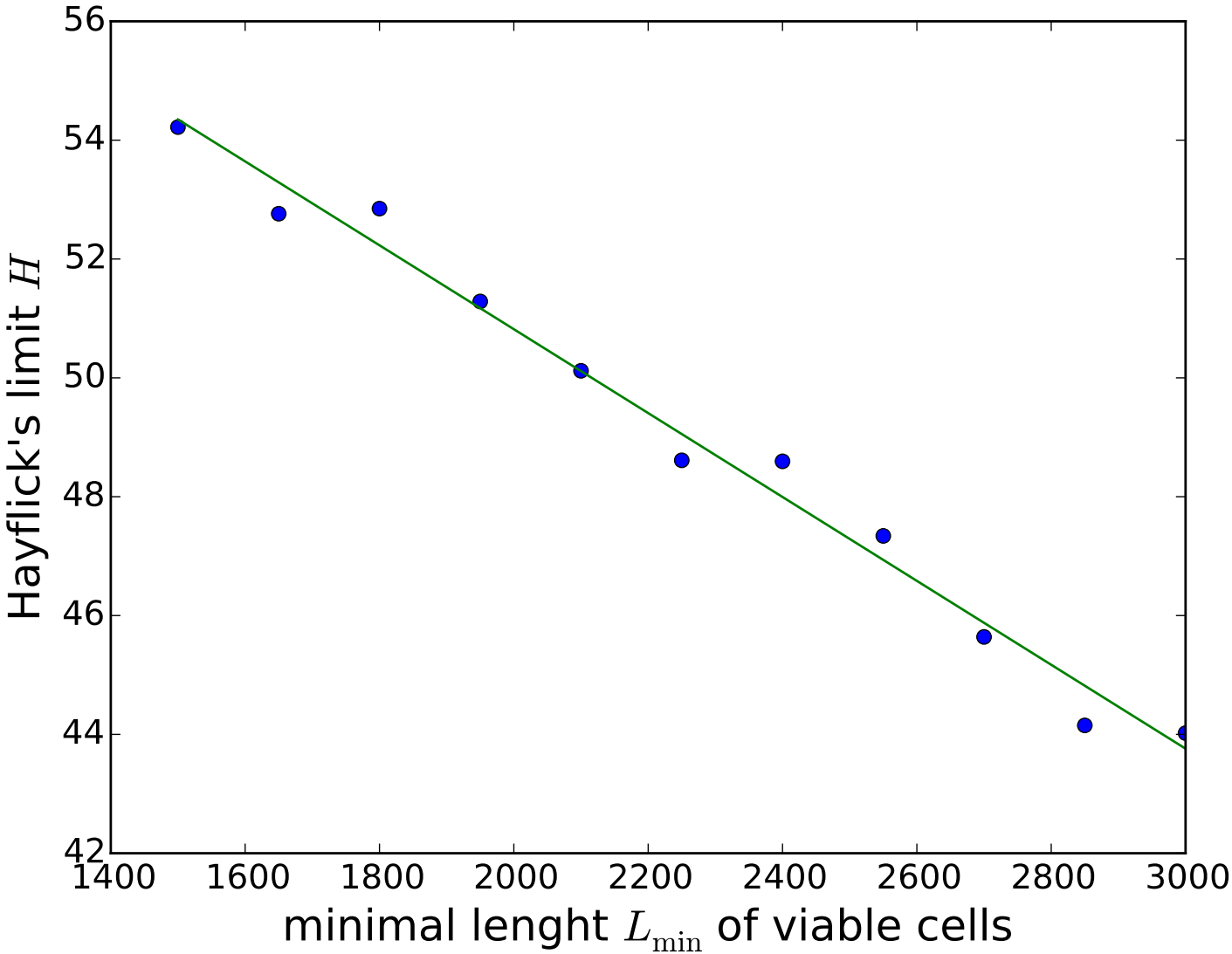
$$H := \log_2 \mathbb{E}_{(c,\mathtt{ns})}[N_\infty].$$

- Aim is to see how the value of H varies as the model parameters vary.

Evolution of number of particles



Evolution of H with $L_{\min}$



Future work

- Understand the distribution of the time of senescence
- Develop statistical models for the influence of external factors on telomere evolution
- Understand the relationship between telomere length and COVID-19
- Parameter estimation