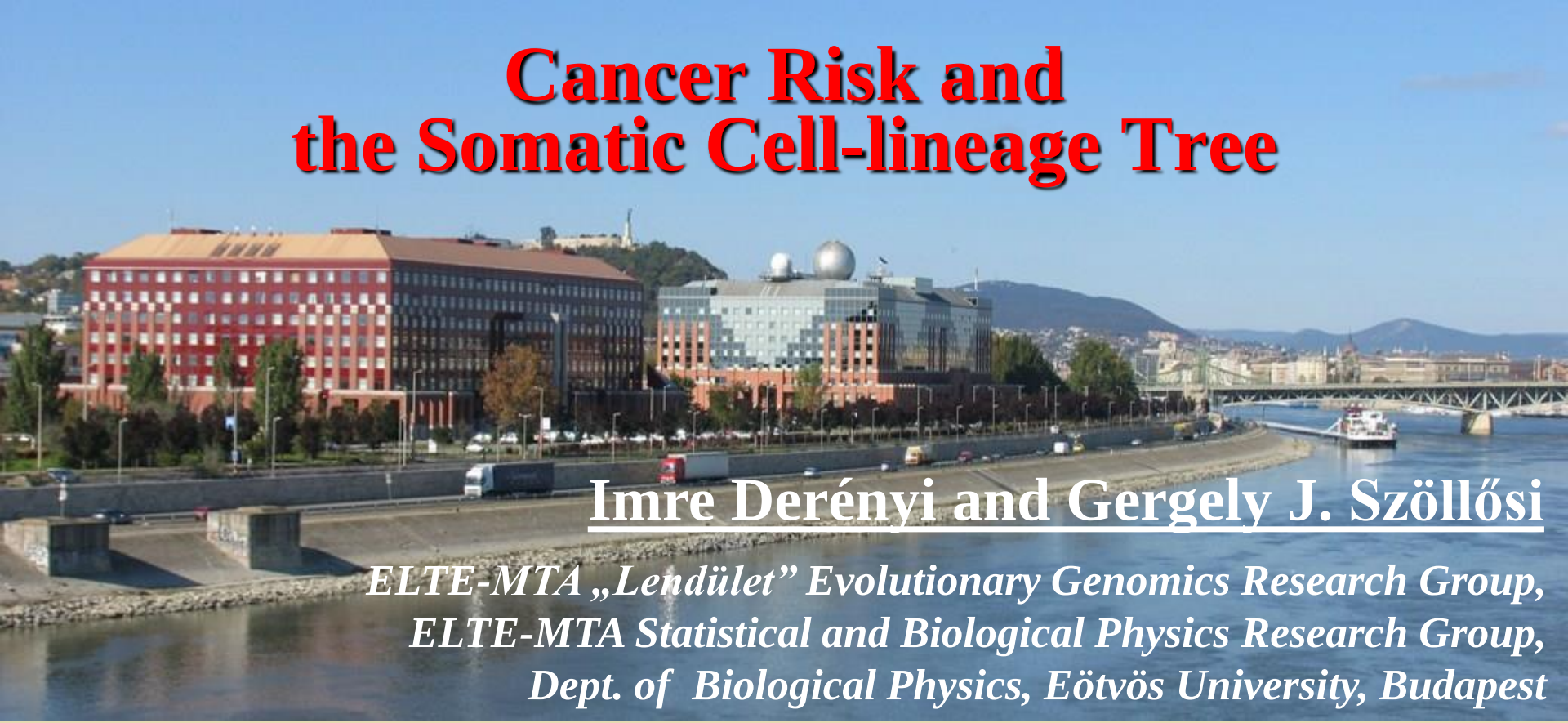
Cancer Risk and the Somatic Cell-lineage Tree



Imre Derényi and Gergely J. Szöllősi

ELTE-MTA „Lendület" Evolutionary Genomics Research Group,
ELTE-MTA Statistical and Biological Physics Research Group,
Dept. of Biological Physics, Eötvös University, Budapest

Postdoc: Gergely Tibély
Students: Márton Demeter, Dániel Grajzel, Máté Kiss

Cancer is the result of somatic evolution

[Nature Communications **8**,14545 (2017)]

RBC, 2018

Somatic Evolution

Reproduction
(cell division)

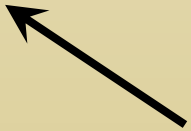
Variation
(mutations)

aging
tissue deterioration
premalignant cells

Gergely Tibély
Máté Kiss

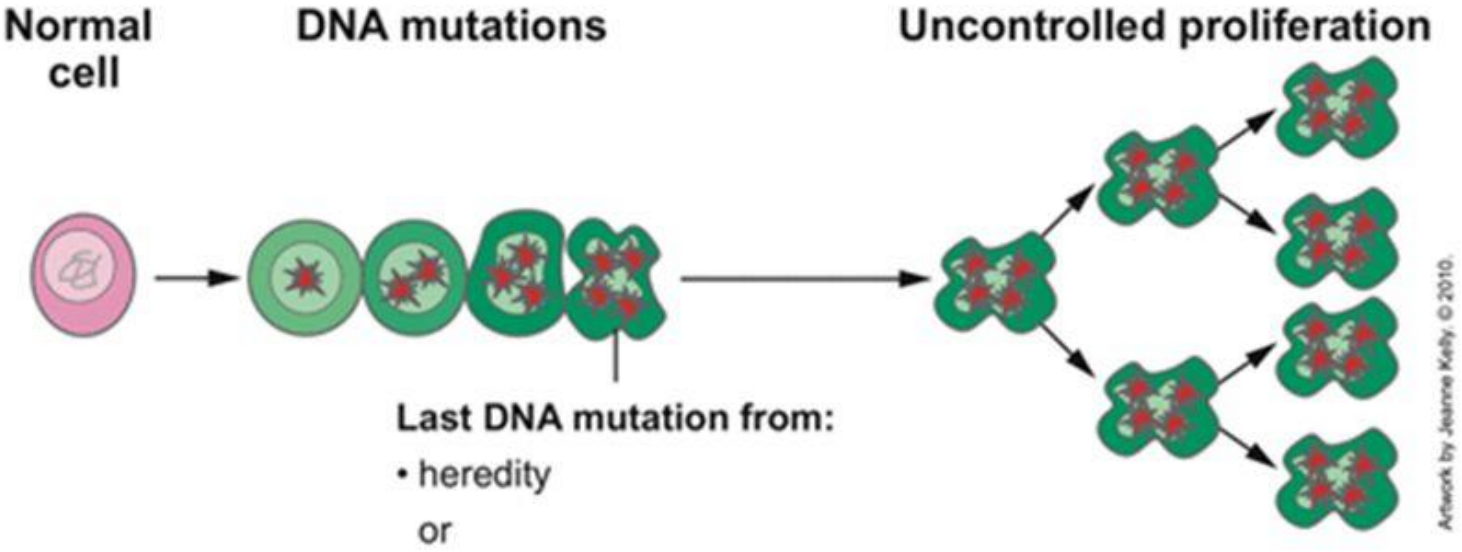
Selection

cancer development



Márton Demeter
Dániel Grajzel

Cancer Arises From DNA Mutations in Cells



Last DNA mutation from:

- heredity

  or

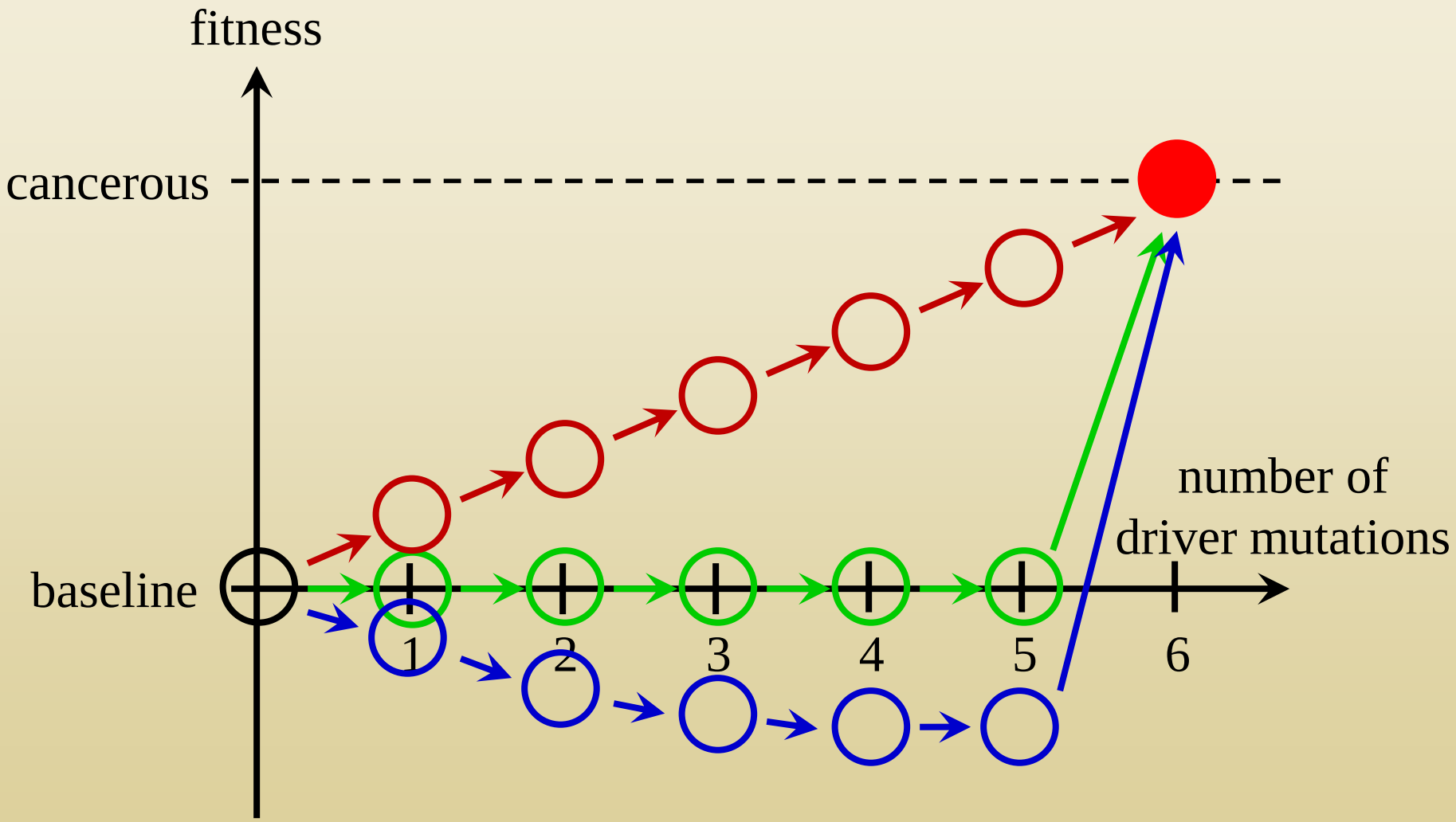
- radiation or chemicals

  or

- spontaneous errors during DNA duplication

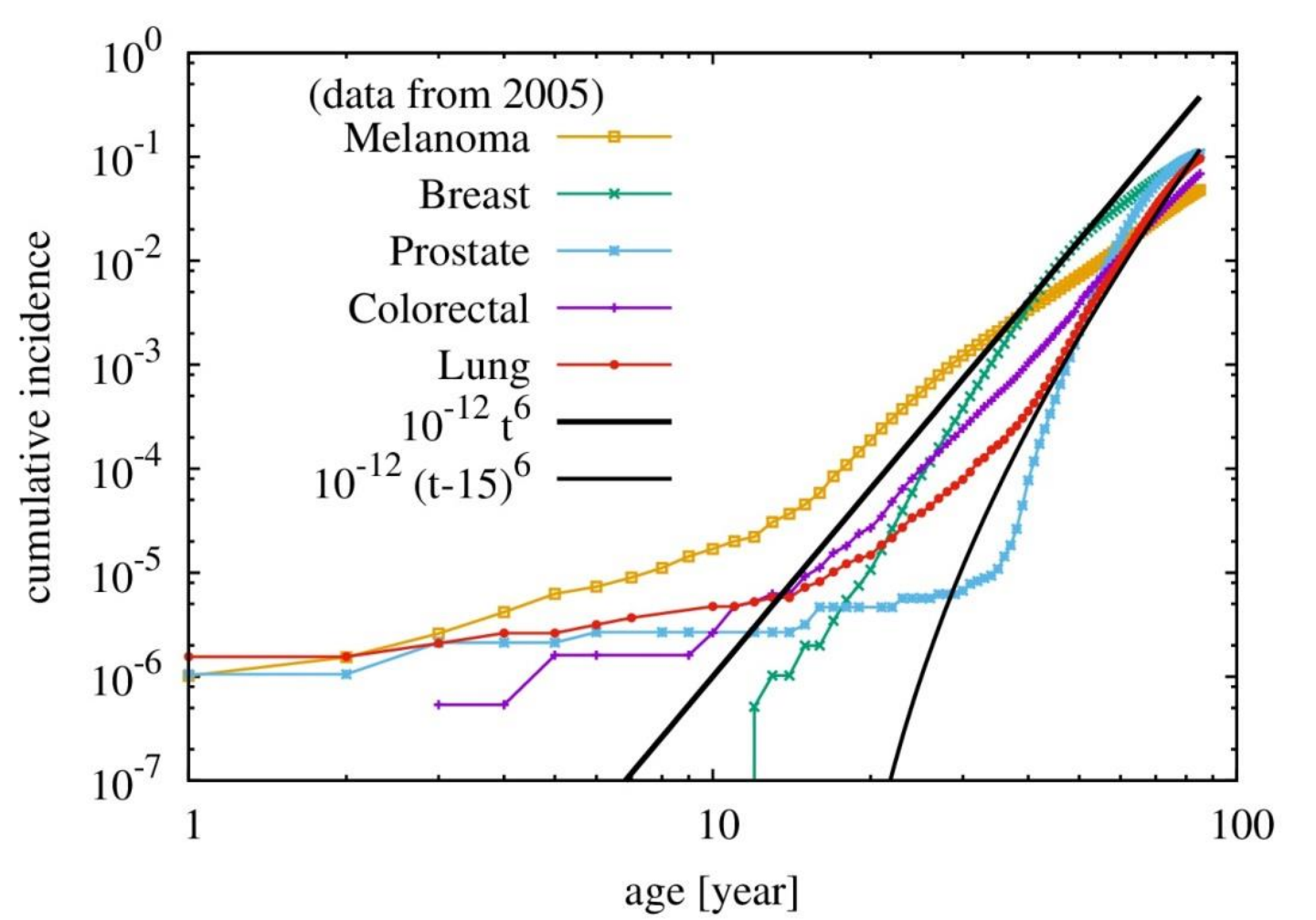
NATIONAL CANCER INSTITUTE

Cancer Development



Epidemiological data

[National Cancer Institute, SEER database, http://seer.cancer.gov/]

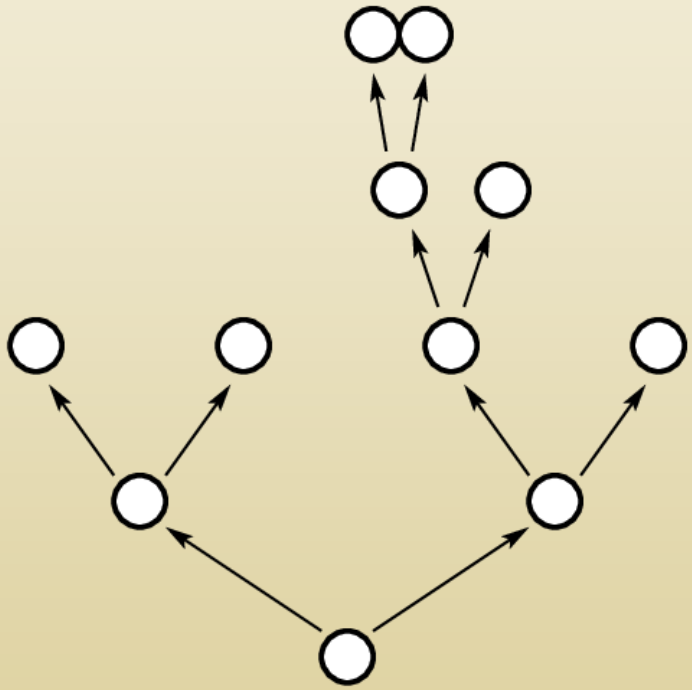


Cancer incidence $\sim t^m$

$m \approx 6$ independent driver mutations

[Armitage and Doll, British J. Cancer (1954)]

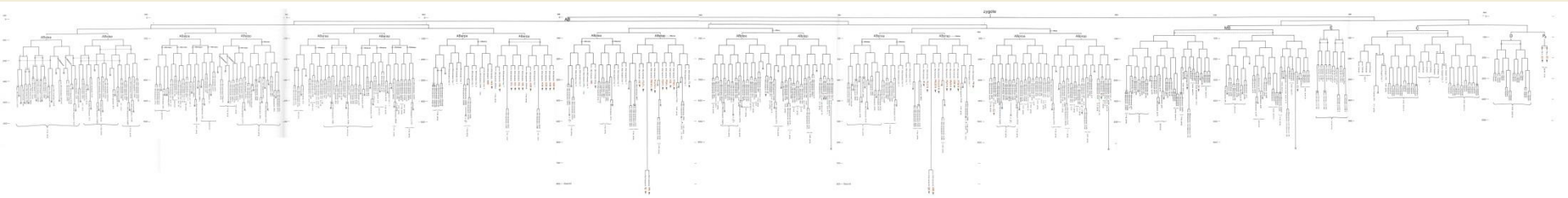
Question: How does the cancer incidence depend on the details of the cell-lineage tree?

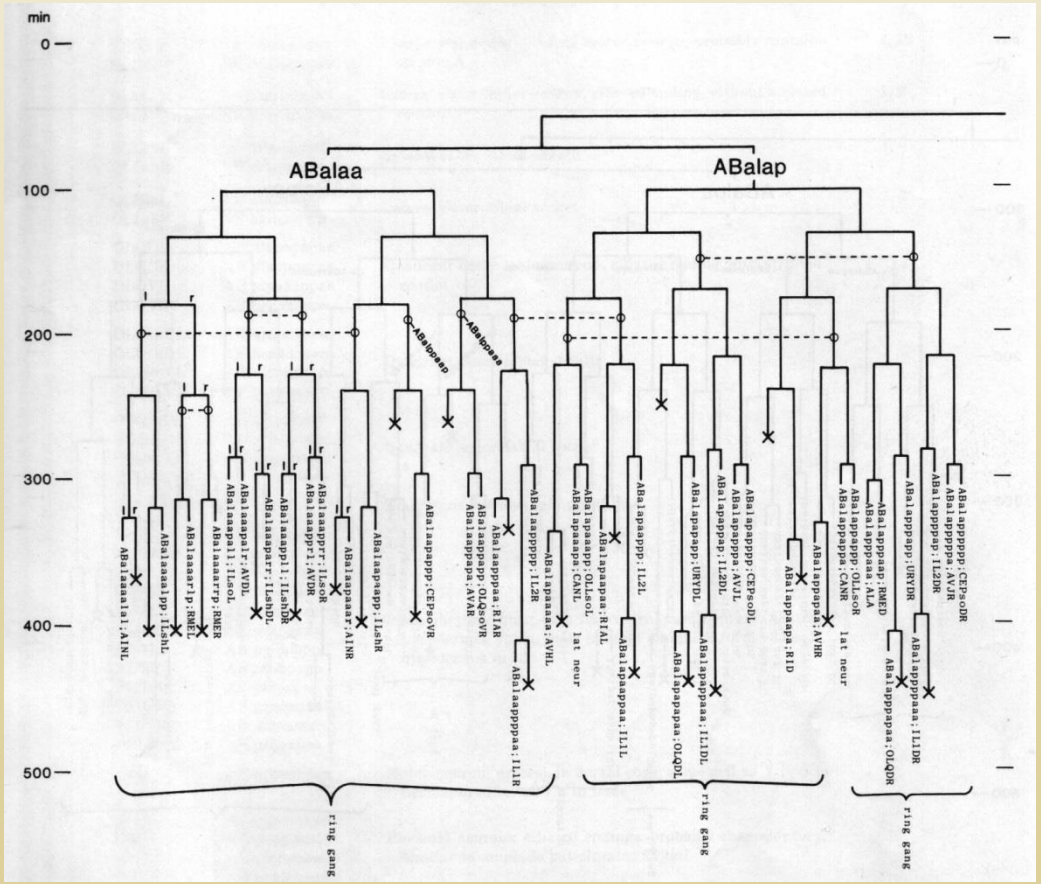


A major, unavoidable source of mutations is the replication error during cell divisions:

rate: $\sim 10^{-9} \frac{1}{\text{bp}} \frac{1}{\text{cell div}}$

Cell-lineage tree of the developing C. elegans



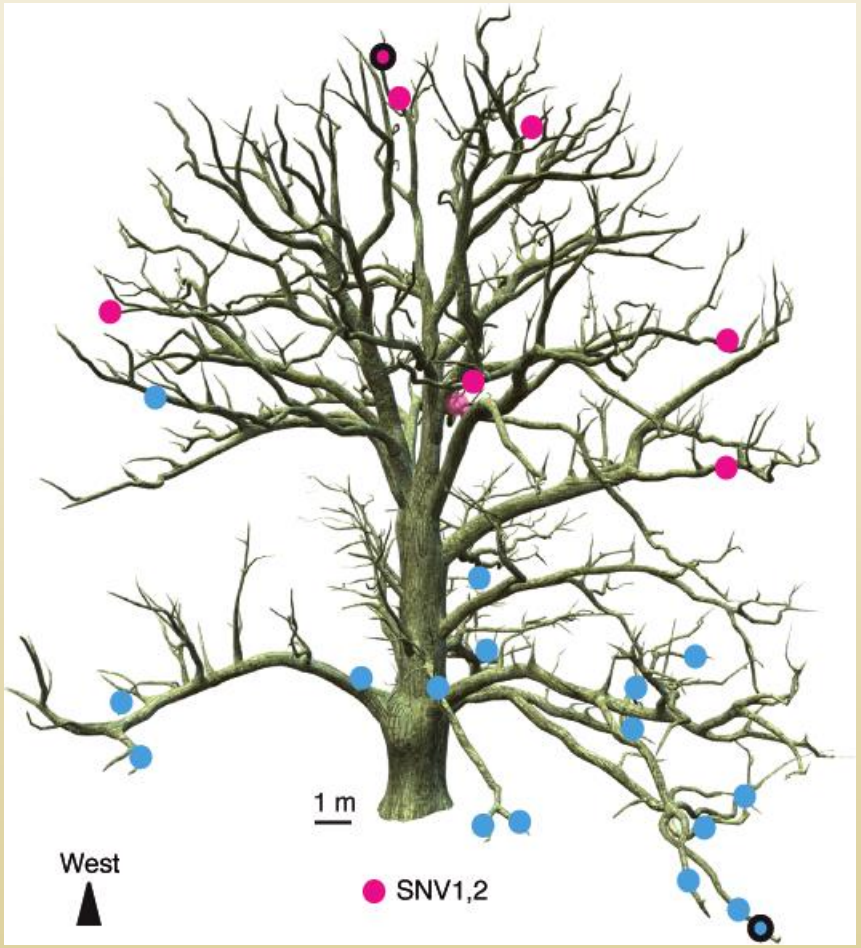


671 cells generated

111 or 113 apoptoses
(depending on the worm's sex)

[Sulston et al., Developmental Biology (1983)]

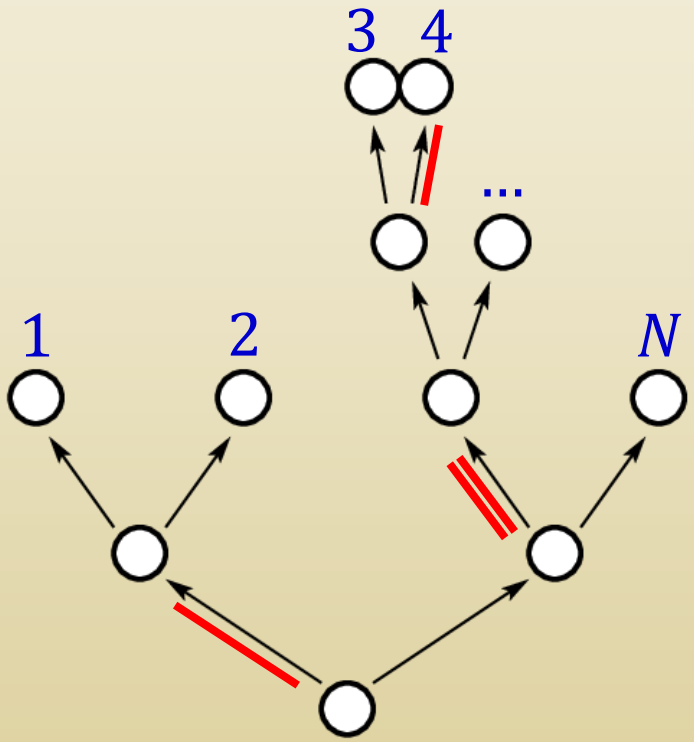
The napoleon oak (a 234-year-old tree)



Only 17 mutations between distant leaves.

[Raymond et al., Low Rate of Somatic Mutations in a Long-Lived Oak Tree, bioRxiv (2017)]

What is the probability that a lineage (path) with m (driver) mutations appear?



μ: driver mutation rate per cell division

$$\sim 10^{-9}\,\frac{1}{\text{bp}}\,\frac{1}{\text{cell div}}\,100\ \text{gene}\ 1000\,\frac{\text{bp}}{\text{gene}}$$

$$= 10^{-4}\,\frac{1}{\text{cell div}}$$

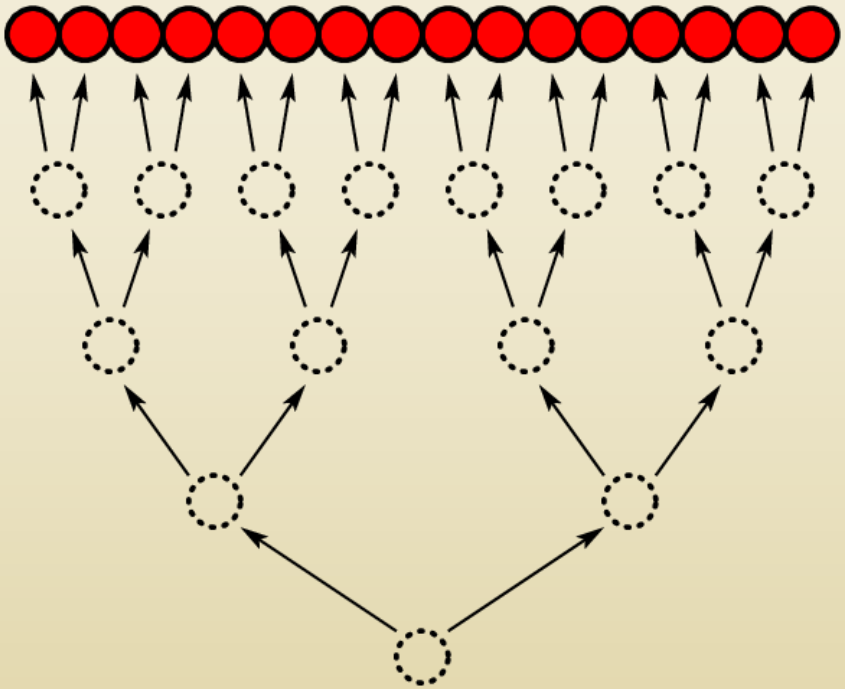
N: "leaves" of the binary tree

D_i : depth (divisional load) of leaf i

$$P(\mu, m) \approx \frac{2\mu^m}{(m-1)!}\sum_{i=1}^{N}(D_i - 1.5)^{m-1}$$

Accurate in leading order of μ and the two highest orders of D_i .

Non-renewing tissues: divisions along a perfect binary tree gives the theoretical minimum number of divisions



Number of cell divisions for generating N cells: $D = \log_2(N)$

At any moment the number of terminally differentiated cells on the skin surface per founder cell: $\sim \frac{(1\ \text{mm})^2}{(10\ \mu\text{m})^2} \approx 10^4$

Total number of skin cells produced by a founder cell over a lifetime (30,000 days; 1 layer/day): $N \approx 3 \times 10^8$

Number of cell divisions: $D = \log_2(N) \approx 28$

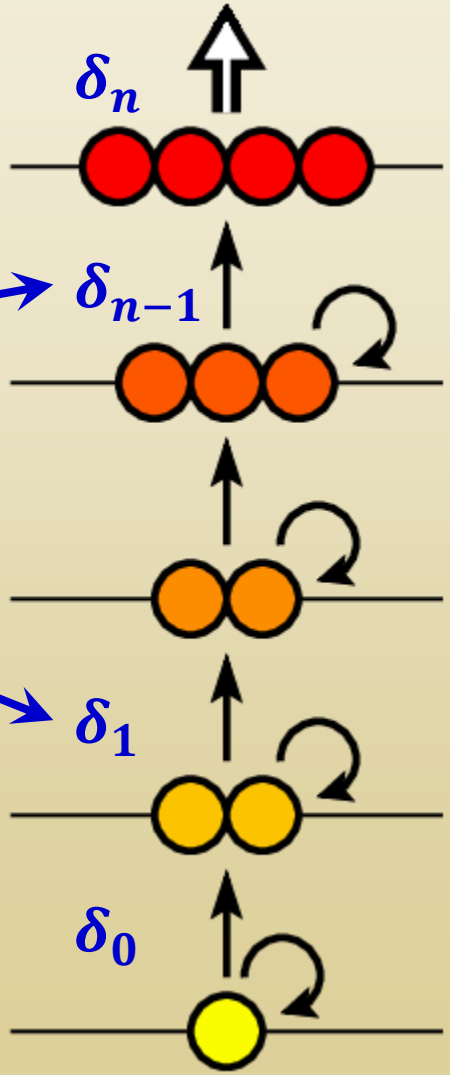
Renewing tissues: hierarchical tissue organization

terminally differentiated cells
(***N*** cells are generated throughout the lifetime of the organism) → N_n

differentiation rates

partially differentiated cells
hierarchically organized into
***n*-1** intermediate levels

stem cell → $N_0 = 1$



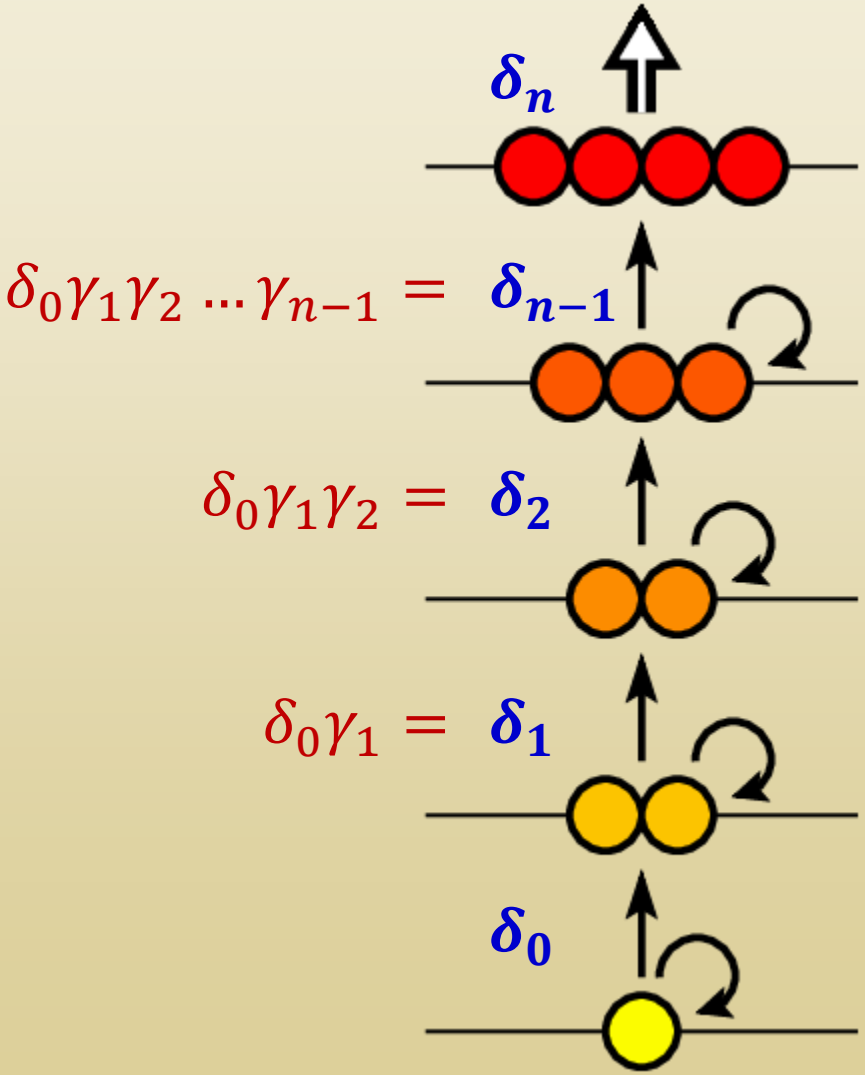
N_{n-1}

N_1

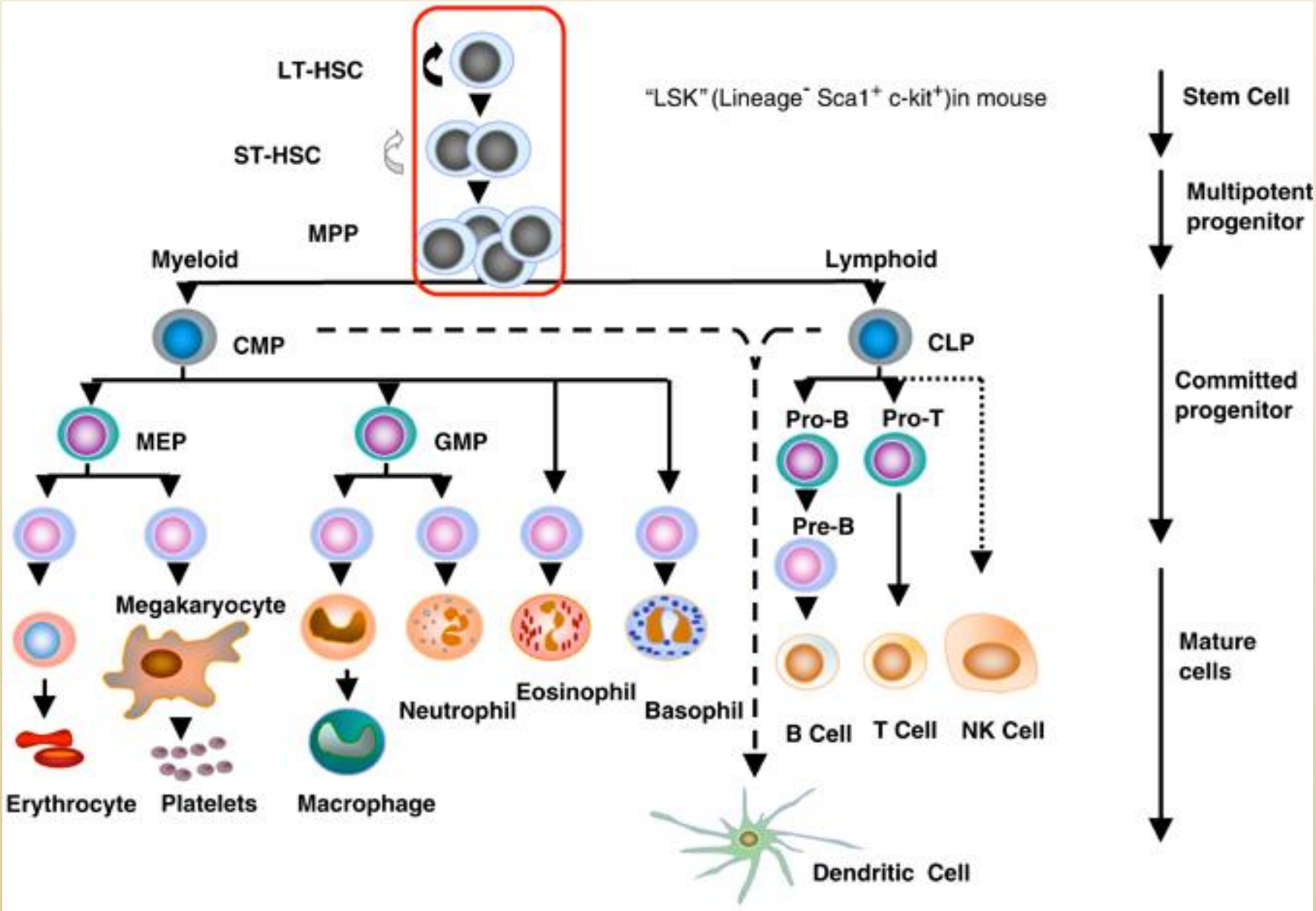
Renewing tissues: hierarchical tissue organization

amplification factor:

$$\gamma_k = \frac{\delta_k}{\delta_{k-1}}$$

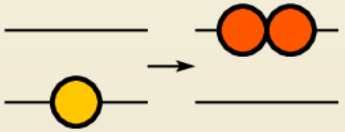


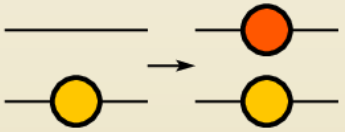
Hematopoiesis



Microscopic events in the optimal tissue dynamics

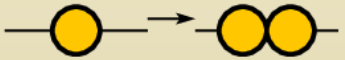
Differentiation

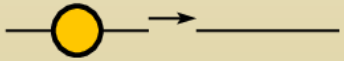
p_k:  Symmetric cell division with differentiation

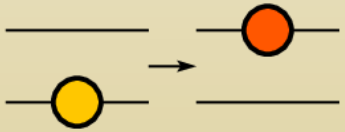
$1 - p_k$:  Asymmetric cell division

Cell replacement (fill empty site)

q_k: Differentiation from a lower level

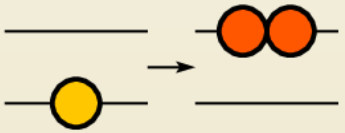
$1 - q_k$:  Symmetric cell division

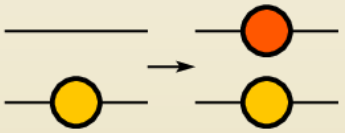
No cell death 

No single cell differentiation 

Microscopic events in the optimal tissue dynamics

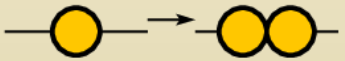
Differentiation

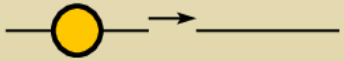
p_k:  $\frac{\delta_k}{2} p_k$

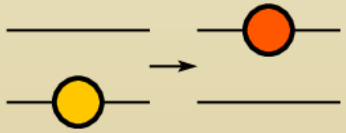
$1 - p_k$:  $\delta_k(1 - p_k)$

Cell replacement (fill empty site)

q_k: Differentiation from a lower level $\frac{\delta_k}{2} p_k q_k = \delta_{k-1}$

$1 - q_k$:  $\frac{\delta_k}{2} p_k (1 - q_k)$

No cell death 

No single cell differentiation 

Time evolution of the mean number of cell divisions (the divisional load) of cells at level k:

$$\dot{D}_k N_k = -\frac{\delta_k}{2} p_k D_k + \delta_k (1 - p_k) + \frac{\delta_k}{2} p_k [q_k (D_{k-1} + 1) + (1 - q_k)(D_k + 2)]$$

Average length of the longest cell-lineage

At the end of the lifetime (i.e. after generating N terminally differentiated cells) the mean number of cell divisions D of the last terminally differentiated cell is:

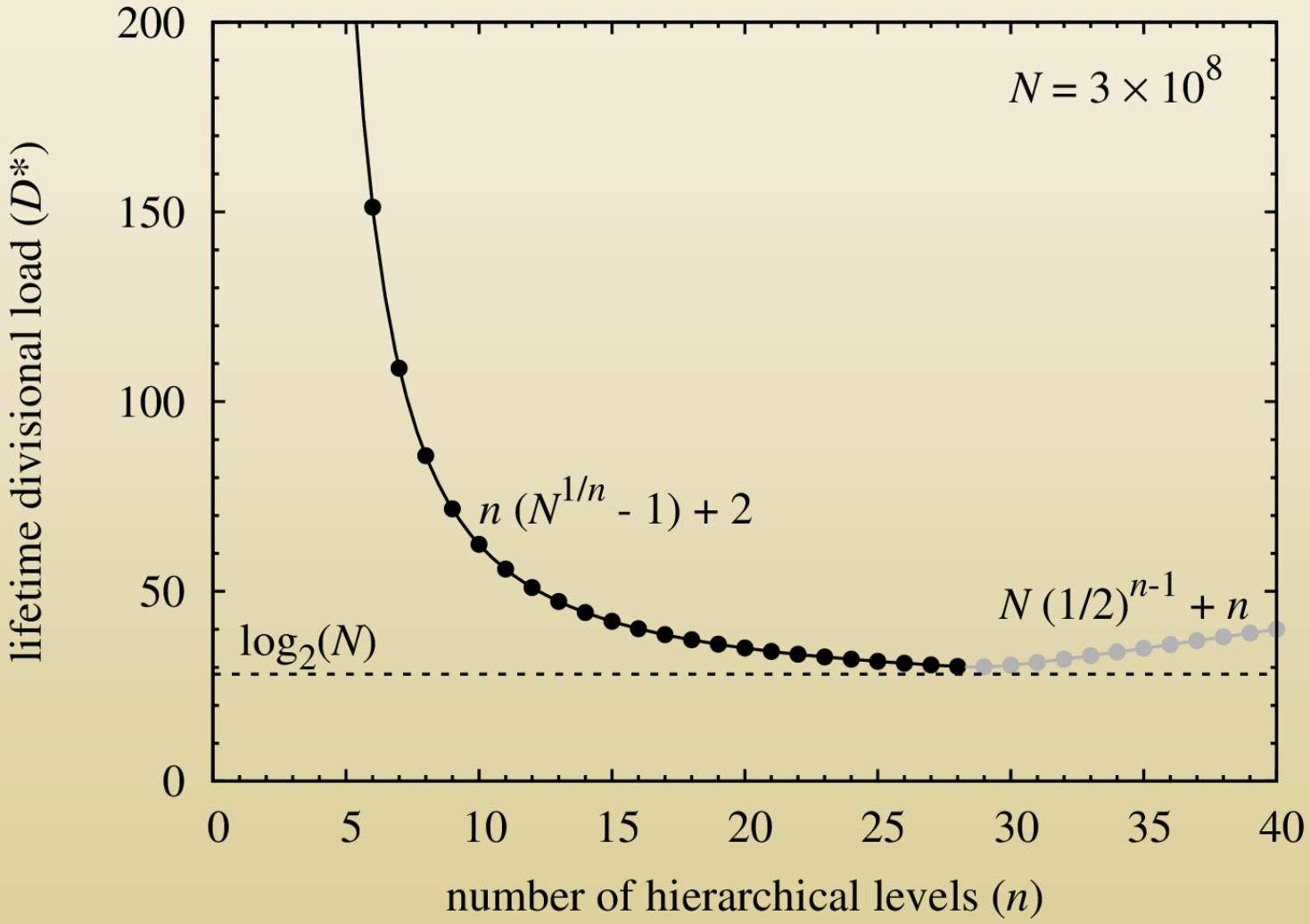
$$D = N\prod_{l=1}^{n-1}\gamma_l^{-1} + \sum_{l=1}^{n-1}(\gamma_l - 1) + 1$$

For given N and n, the divisional load D has a minimum at $\gamma_l^* = N^{1/n}$ (as long as $\gamma_l^* \geq 2$ holds):

$$D^* = n\left(N^{1/n} - 1\right) + 2$$

[Nature Communications **8**,14545 (2017)]

Optimal hierarchy



D^* reaches its minimum, $D^* \approx \log_2(N) + 2$ at $n \approx \log_2(N)$.

Conclusions

- The theoretical minimum of the number of cell divisions during tissue maintenance can almost be reached with a hierarchical tissue architecture.

- Thus, one of the main reasons of the existence of the hierarchical architecture in multicellular organisms is to minimize somatic evolution (e.g., to prevent cancer, or to reduce germline mutations).

- We predict that tissue-specific stem cells may in general correspond to a diverse set of slower dividing cell types.

- This hierarchical tissue architecture is vital in understanding and modelling cancer development.

Acknowledgments

Postdocs and Students:

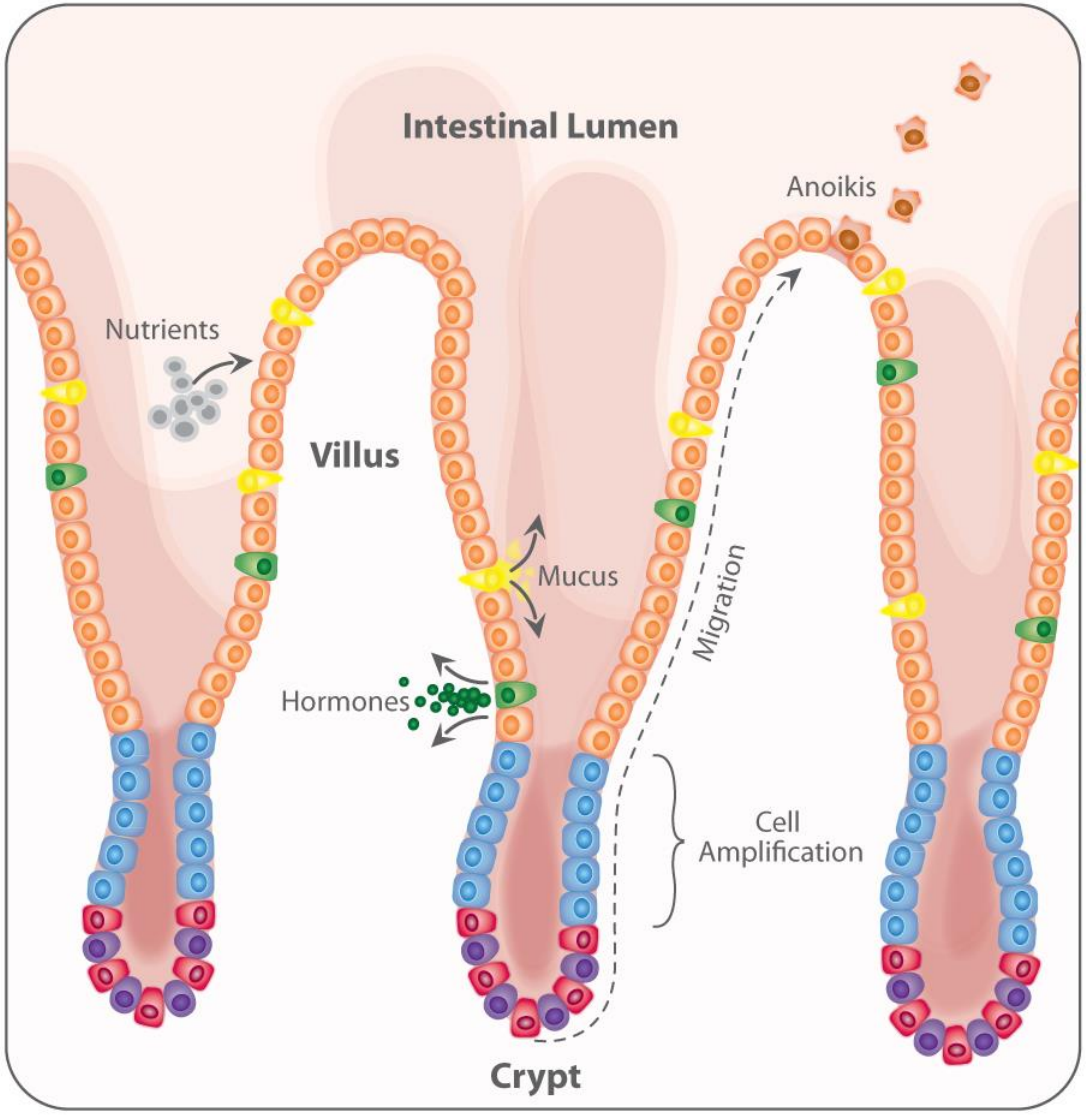
- Gergely Tibély
- Márton Demeter
- Dániel Grajzel
- Máté Kiss

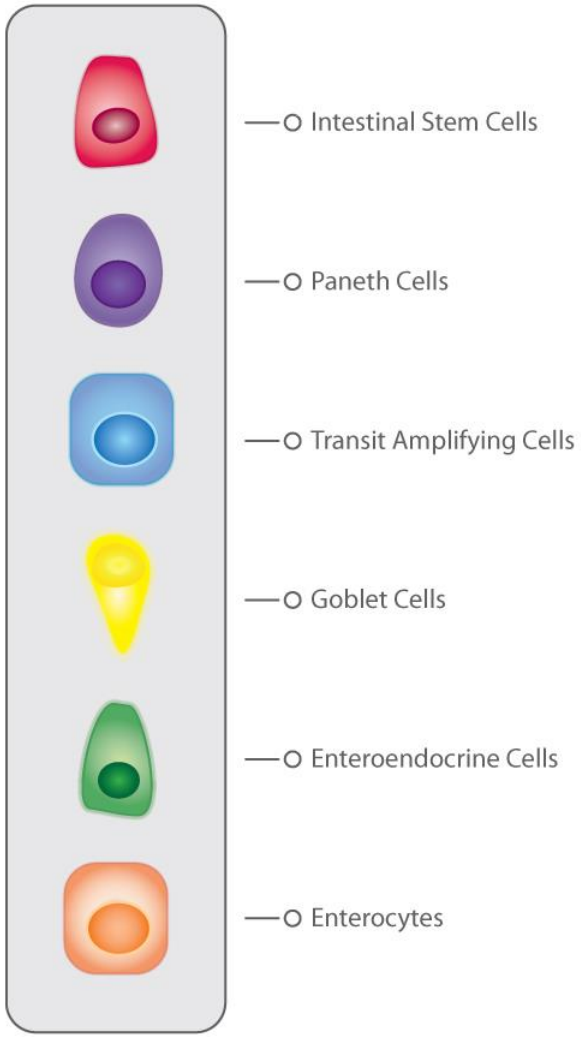
Funding sources:

- European Research Council
- Momentum Program of the Hungarian Academy of Sciences
- Hungarian Scientific Research Fund

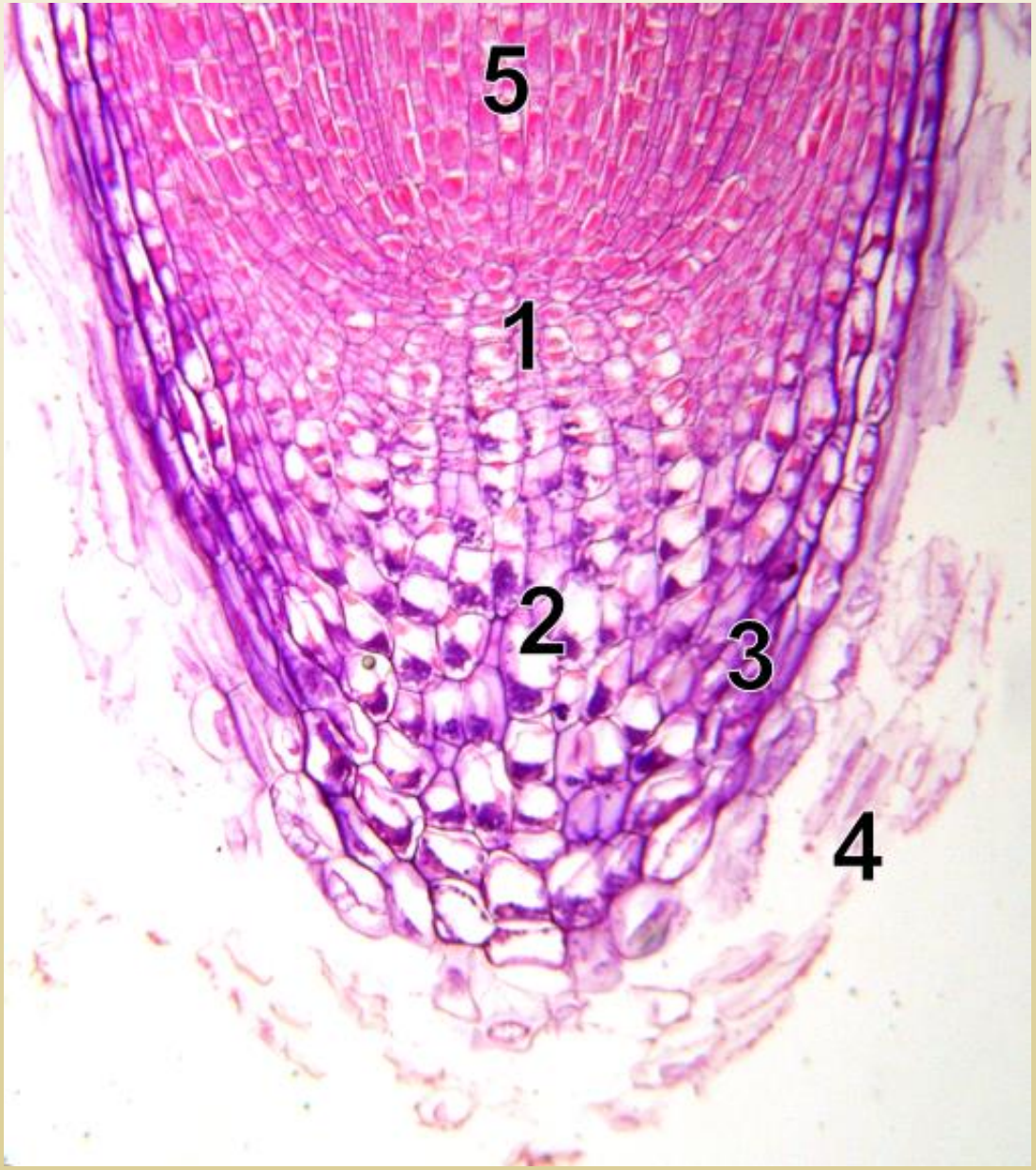
We are looking for postdocs & PhD students!

The Intestinal Epithelium





Root tip



1 – stem cells
2 – live rootcap cells
3 – rootcap
4 – dead rootcap cells
5 – elongation zone