Stochastic model of tumor evolution for cancer etiology and risk

Sophie PÉNISSON

Center for Cancer Prevention and Early Detection, Translational Genomics Research Institute / City of Hope (USA)

University Paris-Est Créteil (France)

Forschungsseminar Warscheinlichkeitstheorie

Institut für Mathematik der Universität Potsdam

10th May 2023

UPEC
Connaissance - Action
UNIVERSITÉ PARIS-EST CRÉTEIL VAL DE MARNE

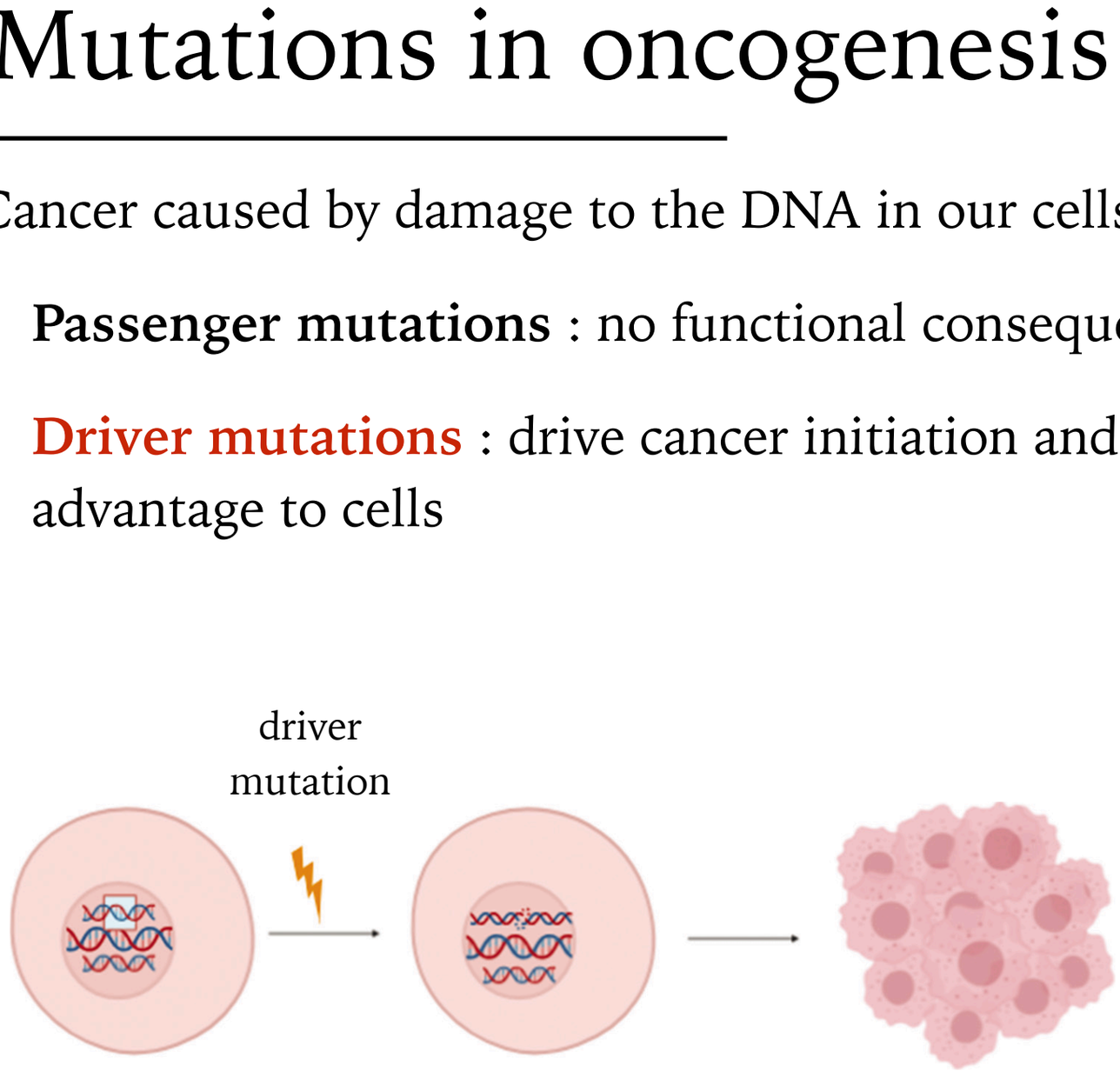
City of Hope®

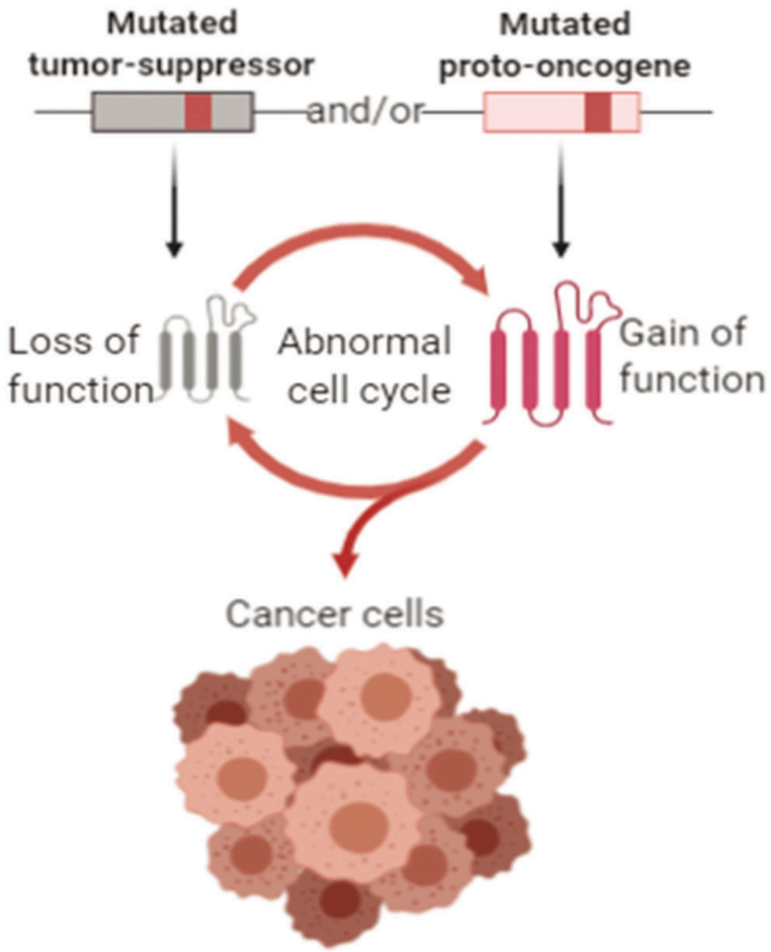
tgen® AN AFFILIATE OF City of Hope.

Mutations in oncogenesis

Cancer caused by damage to the DNA in our cells, called *gene mutations*.

- **Passenger mutations** : no functional consequences
- **Driver mutations** : drive cancer initiation and progression by confering a selective advantage to cells

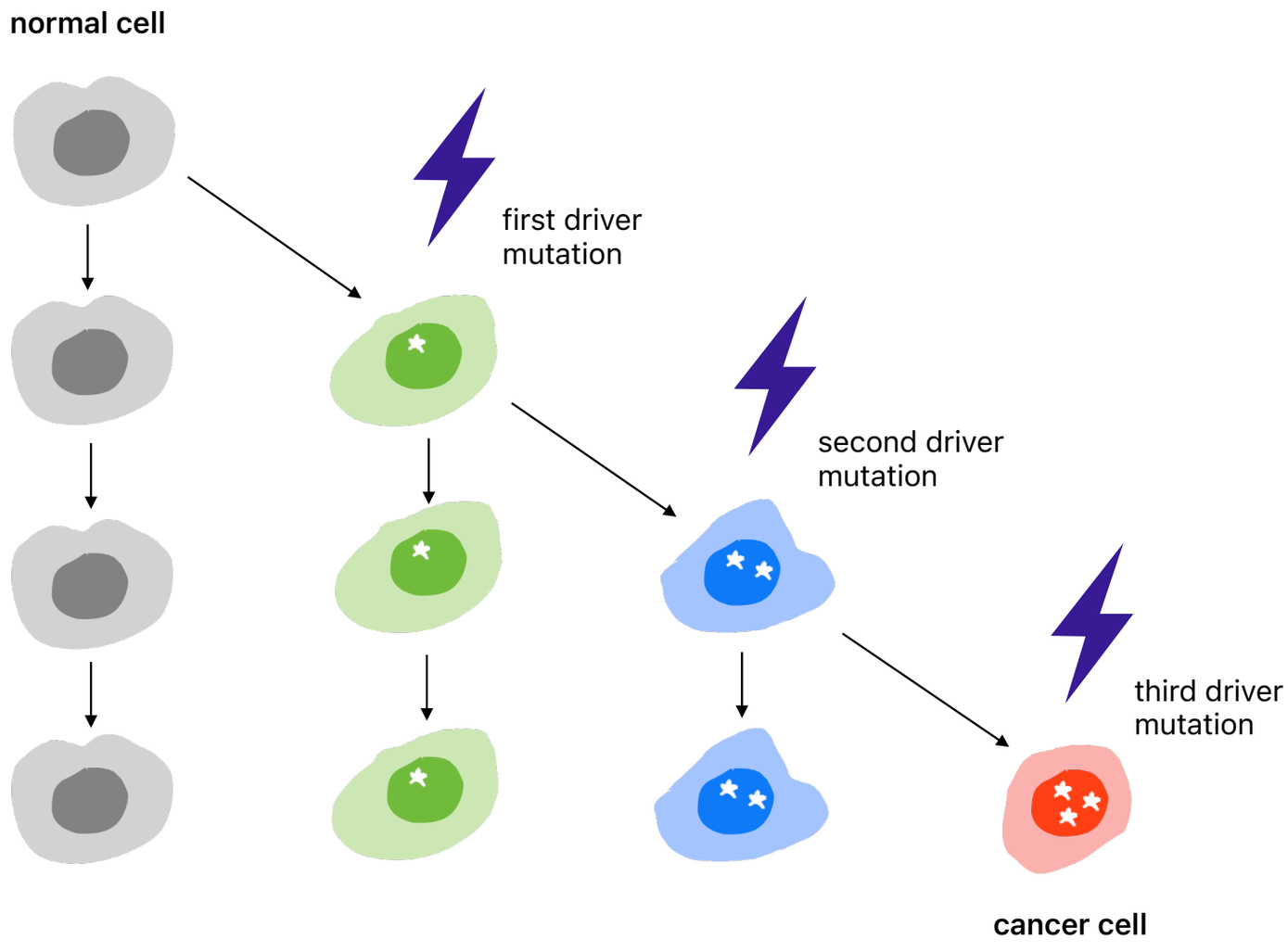




Morjaria *Int J Mol Immuno Oncol* 2021

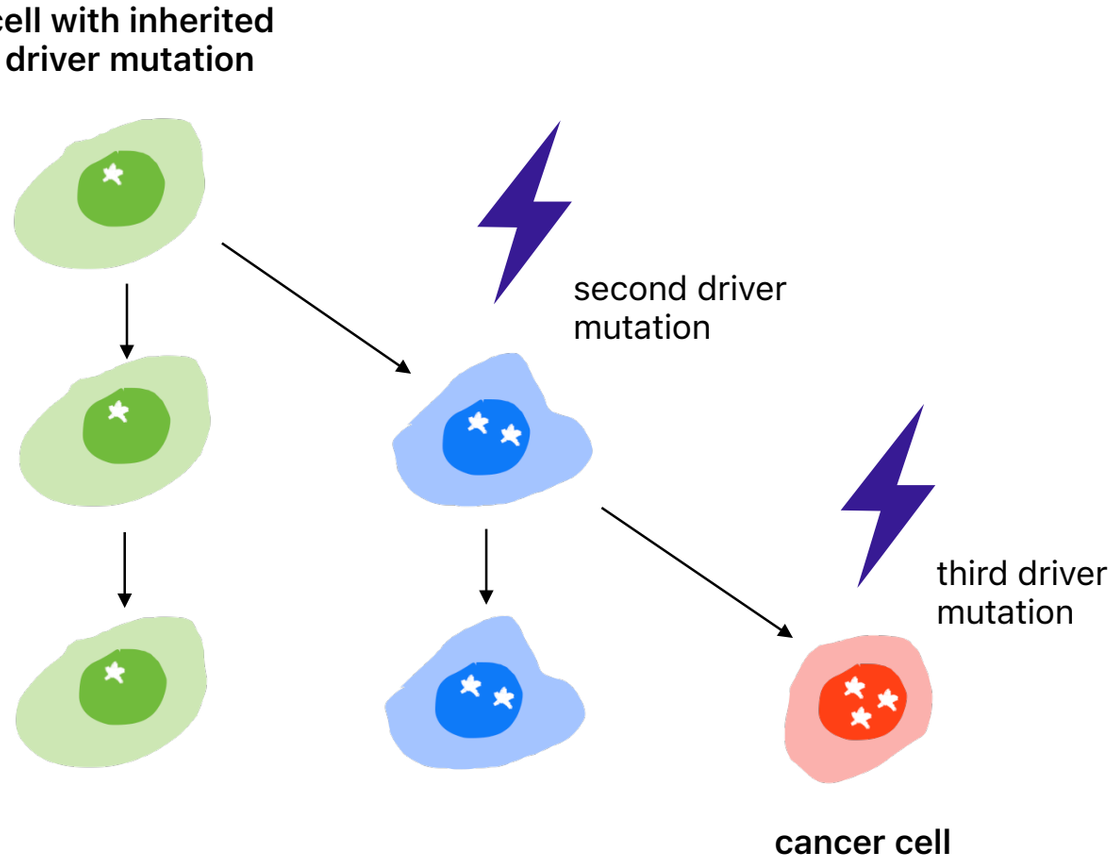
Mutations in oncogenesis

Cancers result from the gradual **accumulation** of *n* driver mutations (*n* between 1 and 5) in at least one stem cell.



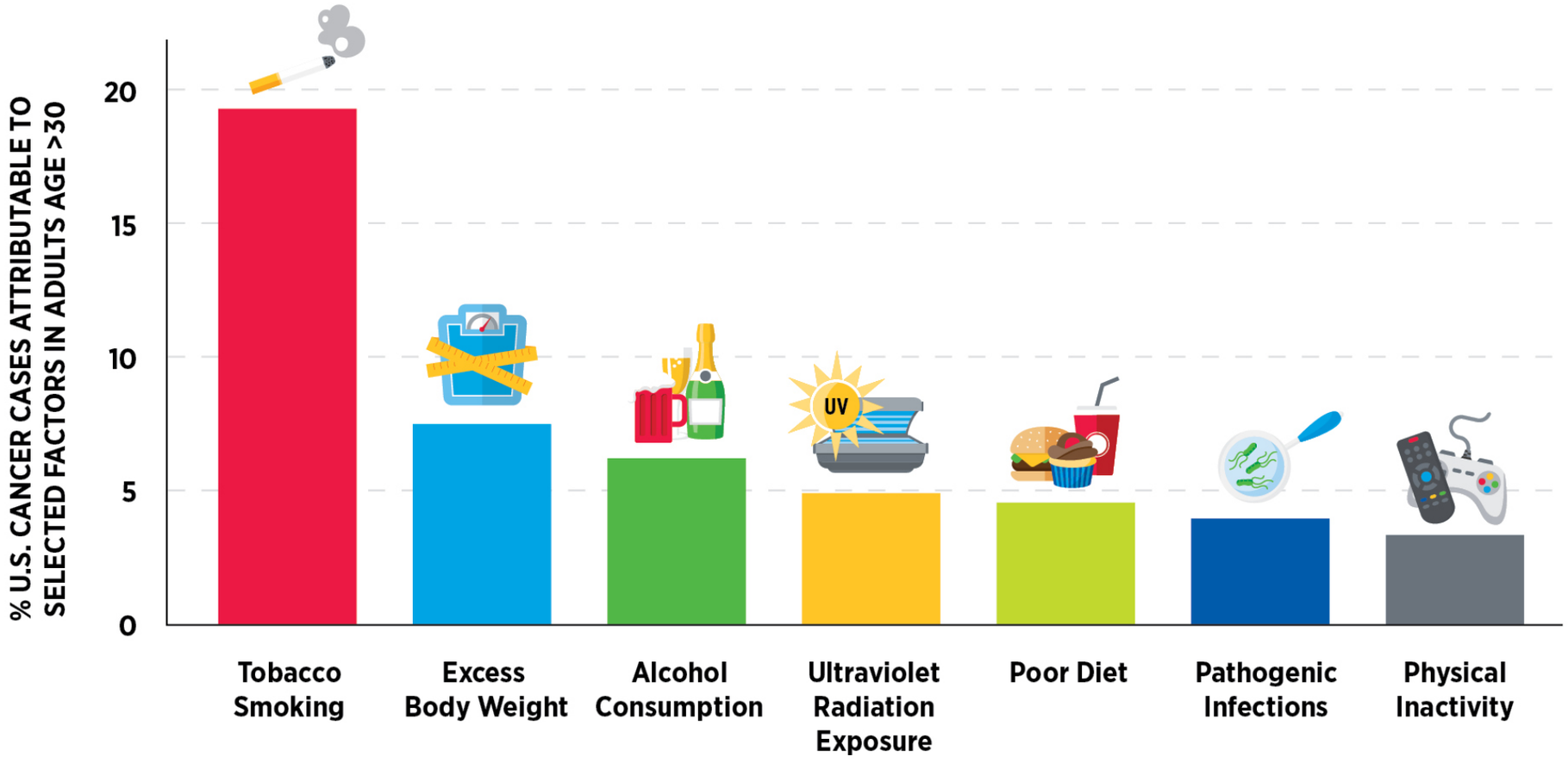
Causes of mutations

- **Hereditary factors**
  - For example, women with a BRCA gene mutation have a higher risk of developing breast and ovarian cancer (4 to 7 times more likely).
  - About 5% of cancers are hereditary.



Causes of mutations

- **Environmental factors**
  - For example, smokers have a higher risk of developing lung cancer (15 to 30 times more likely).
  - About 40% of cancers are considered "preventable".



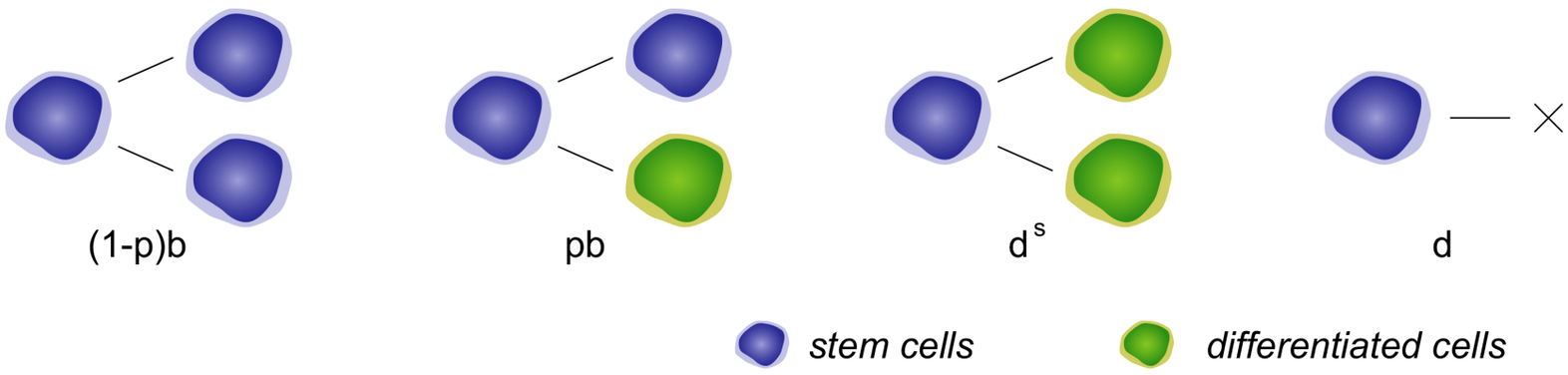
American Association for Cancer Research (AARC) Cancer Progress Report 2022

Causes of mutations

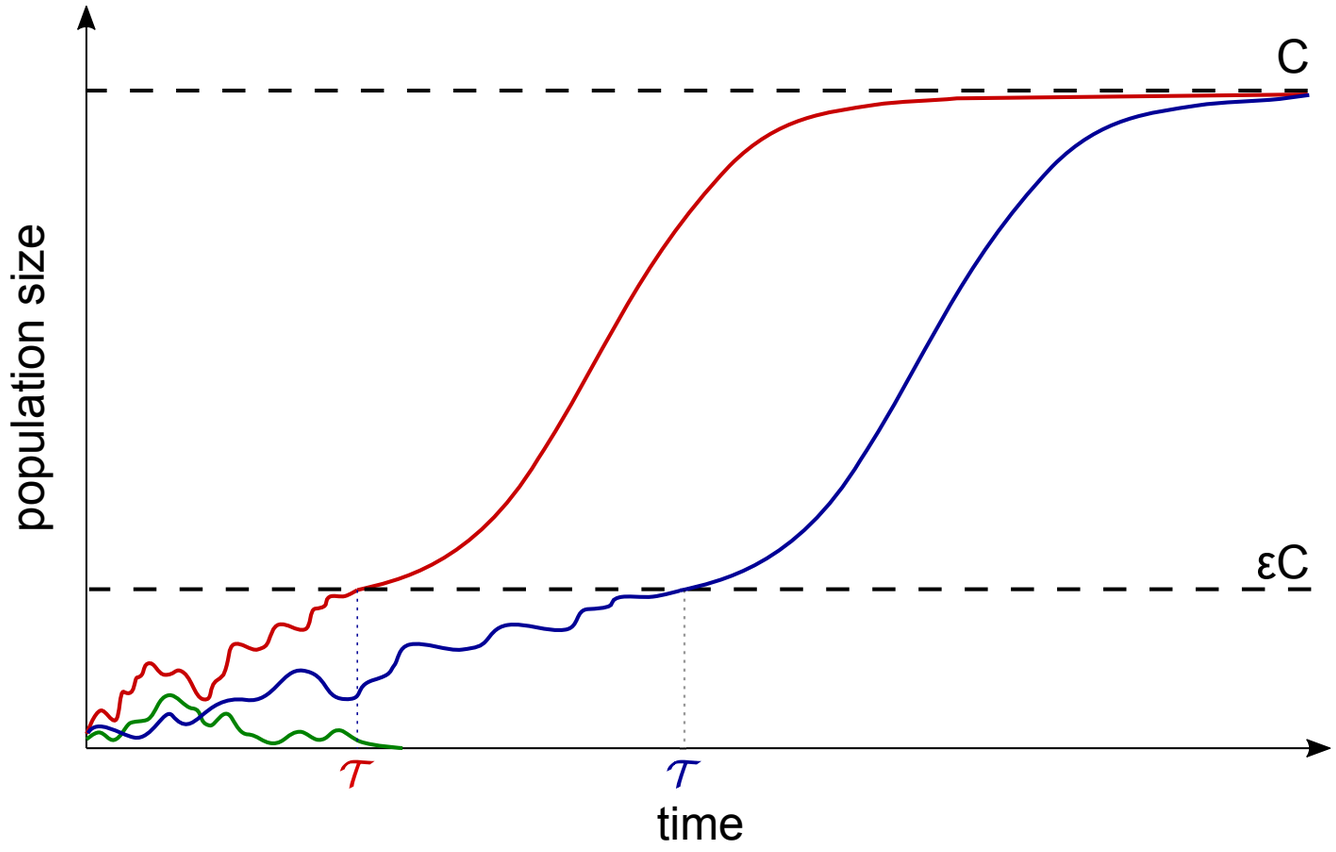
- **Endogenous factors**
    - Mutations occurring naturally during cell divisions (around 3 mutations per cell division), due to the random mistakes made during normal DNA replication ("bad luck").

    - Endogenous factors first suggested and studied in 2015 (Tomasetti & Vogelstein, *Science*).

Model of tumor evolution



Evolution of the stem cell populations :



logistic growth

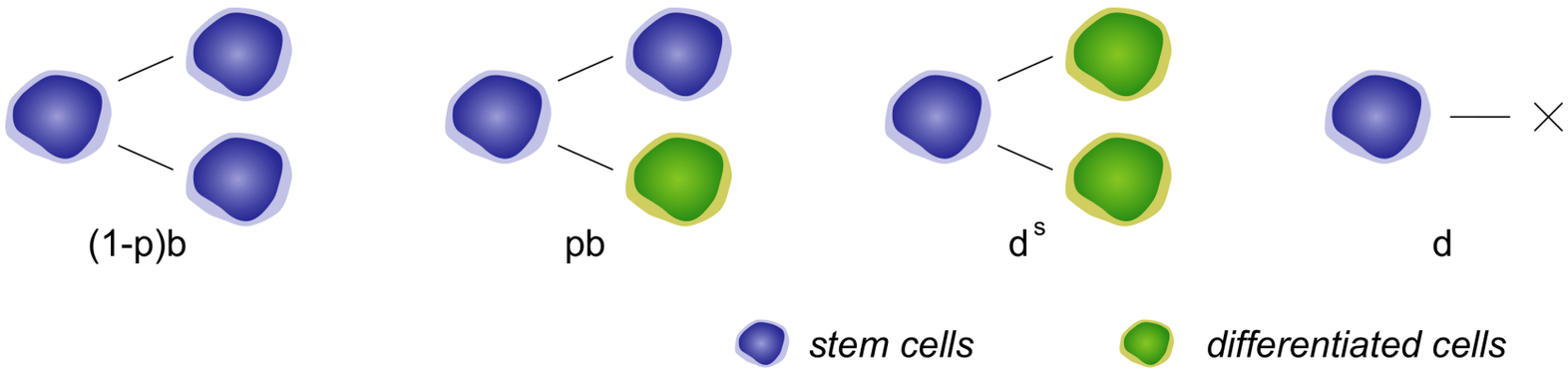
$$\frac{dX_t}{dt} = r\left(1 - \frac{X_t}{C}\right)X_t$$

$$r = (1-p)b - (d^s + d)$$

birth-and-death process

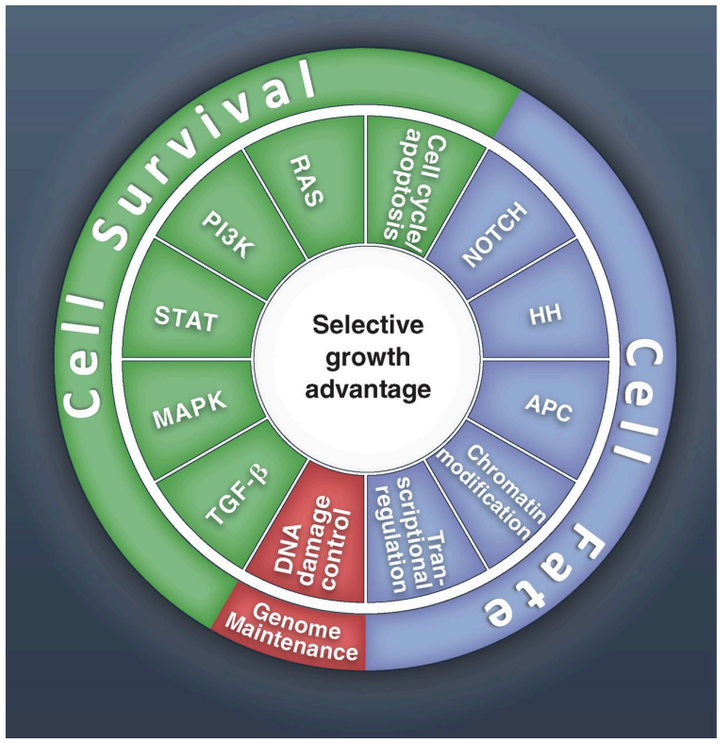
Pénisson, Lambert & Tomasetti *Nat. Commun.* 2022

Model of tumor evolution



Selective growth advantage of the driver mutations :

- **Cell survival** (S)
  proliferation rate $\Delta b > 0$
- **Cell fate** (F)
  asymmetric division probability $\Delta p < 0$
- **Genome maintenance** (M)
  driver mutation probability $\Delta u > 0$

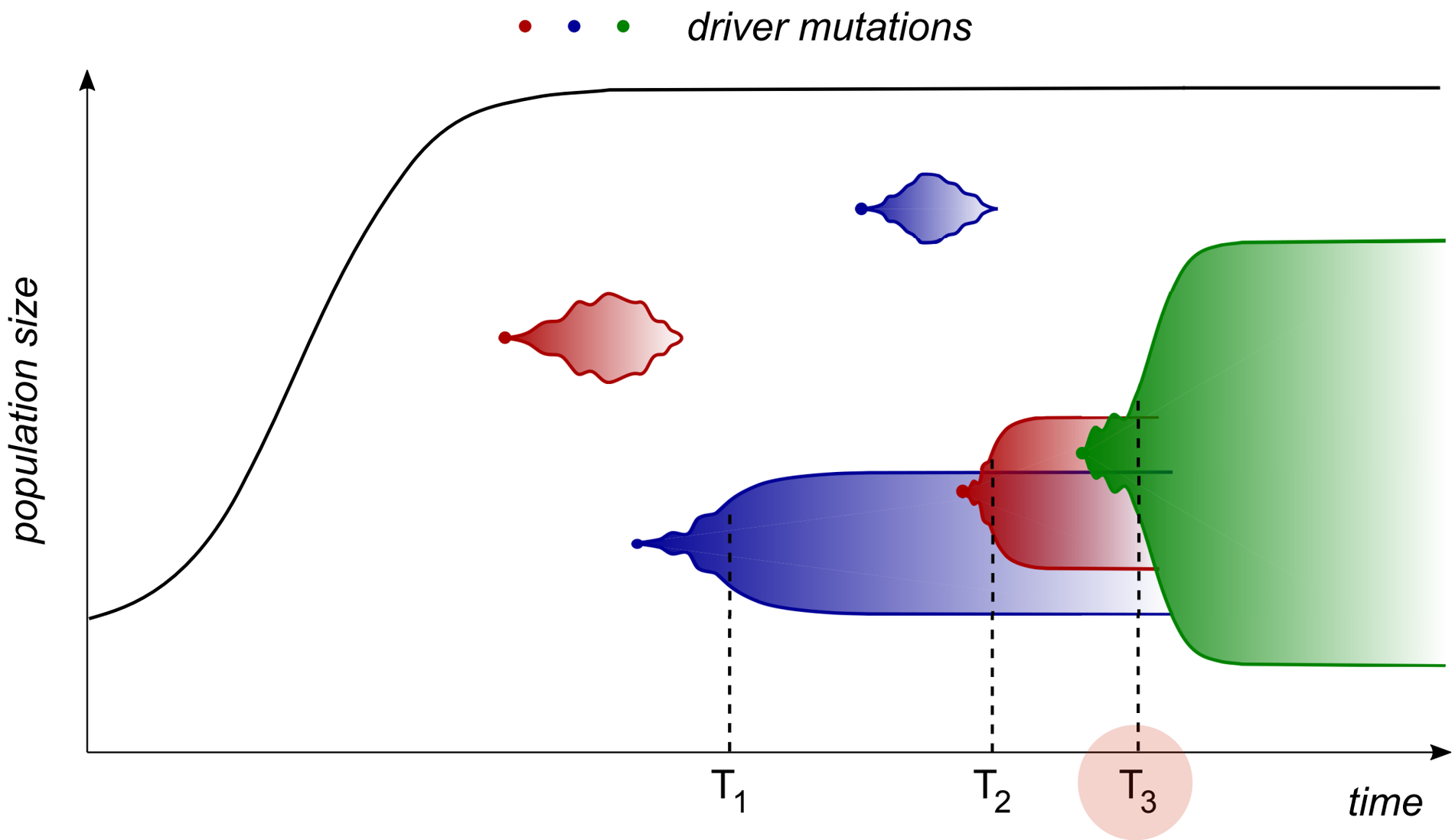


Volgestein et al. *Science* 2013

⟶ Increase of the carrying capacity, logistic growth rate, or mutation rate of the clonal population

Mutations in oncogenesis

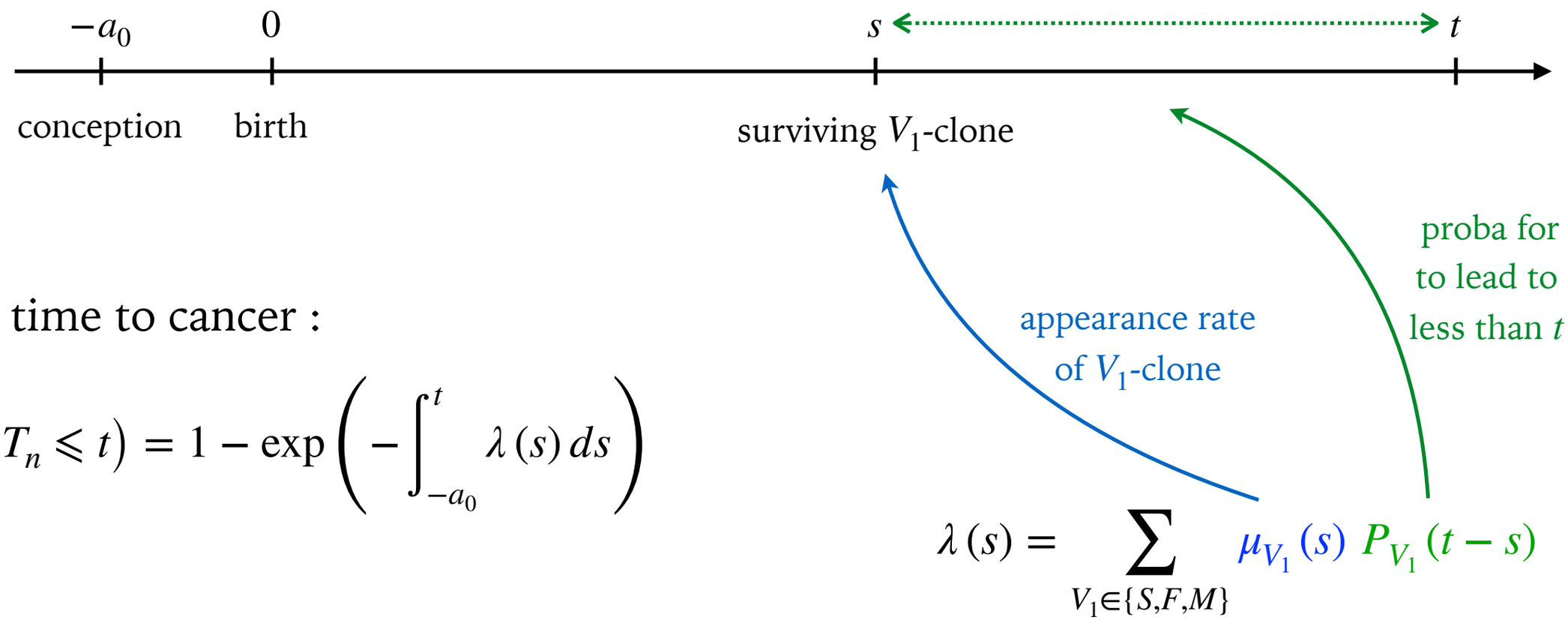
- Gradual accumulation of n driver mutations (here $n=3$)
- T_n time to cancer



Pénisson, Lambert & Tomasetti *Nat. Commun.* 2022

Time to cancer

Cancer occurs when a first "surviving" clone carries n driver mutations $V_1 \dots V_n$, where $V_i \in \{S, F, M\}$.

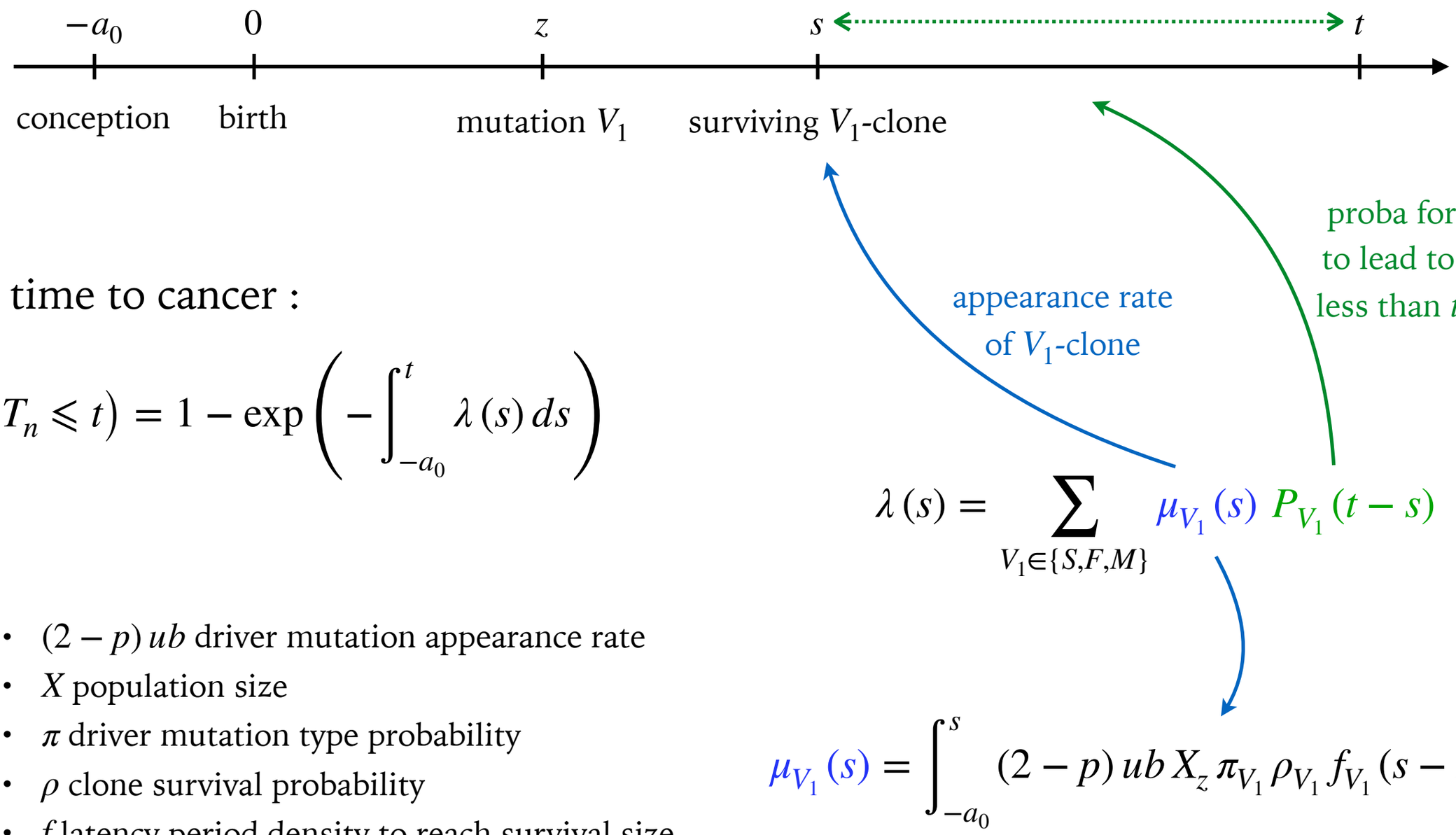


T_n time to cancer :

$$\mathbb{P}\left(T_n \leqslant t\right) = 1 - \exp\left(-\int_{-a_0}^{t} \lambda(s)\, ds\right)$$

Time to cancer

Cancer occurs when a first "surviving" clone carries n driver mutations $V_1 \dots V_n$, where $V_i \in \{S, F, M\}$.



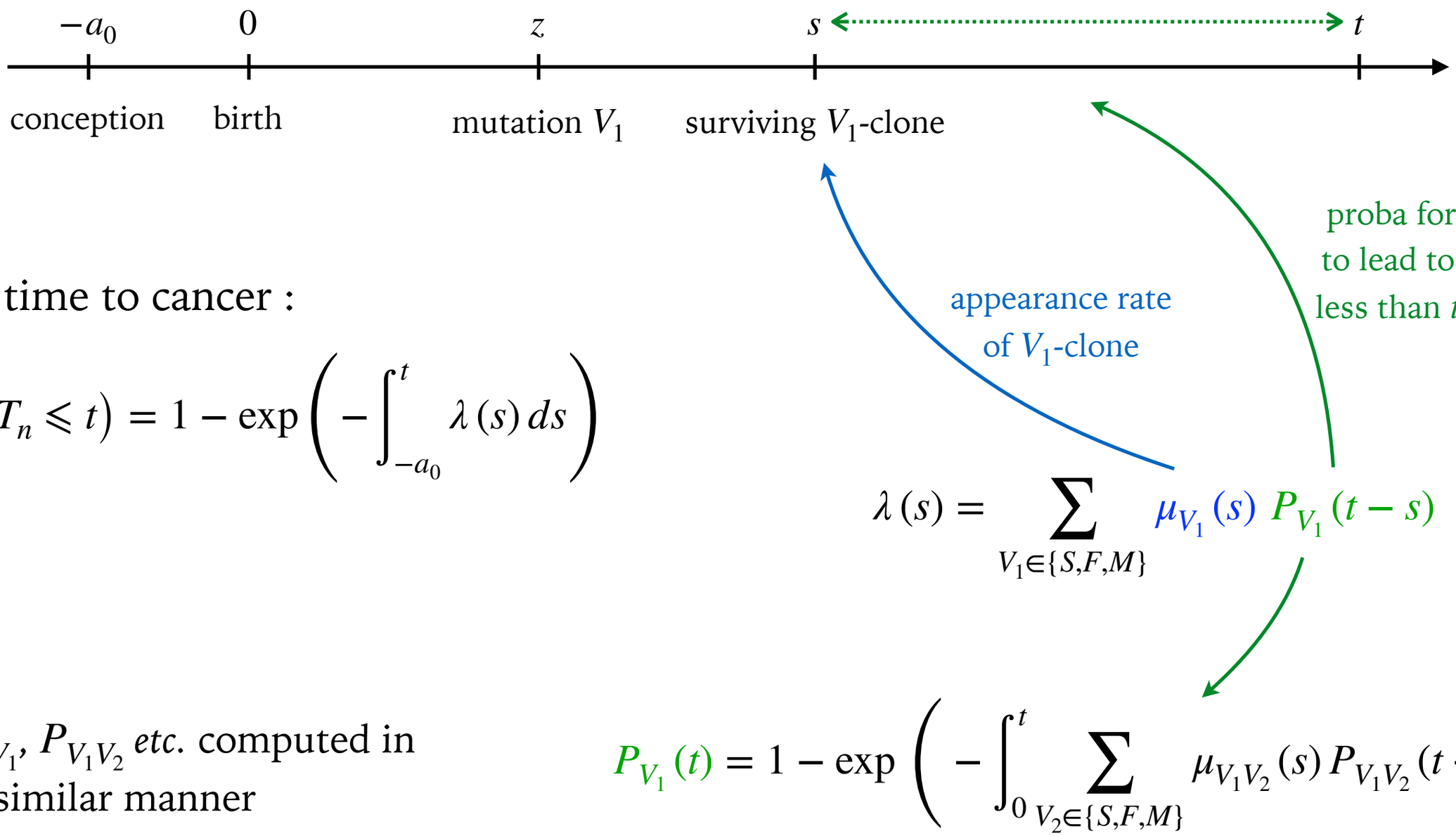
T_n time to cancer :

$$\mathbb{P}\left(T_n \leqslant t\right) = 1 - \exp\left(-\int_{-a_0}^{t} \lambda(s)\, ds\right)$$

- $(2-p)\, ub$ driver mutation appearance rate
- X population size
- π driver mutation type probability
- ρ clone survival probability
- f latency period density to reach survival size

Time to cancer

Cancer occurs when a first "surviving" clone carries n driver mutations $V_1 \dots V_n$, where $V_i \in \{S, F, M\}$.



T_n time to cancer :

$$\mathbb{P}\left(T_n \leqslant t\right) = 1 - \exp\left(-\int_{-a_0}^{t} \lambda(s)\, ds\right)$$

P_{V_1}, $P_{V_1V_2}$ *etc.* computed in a similar manner

$$P_{V_1}(t) = 1 - \exp\left(-\int_0^t \sum_{V_2 \in \{S,F,M\}} \mu_{V_1V_2}(s)\, P_{V_1V_2}(t-s)\, ds\right)$$

Time to cancer

T_n time to cancer :

$$\mathbb{P}\left(T_n \leqslant t\right) = 1 - e^{-\Lambda(t)}$$

Under some simplifying assumptions :

$$\Lambda\left(t\right) = Nu^n \left(b_0^n \sum_{i=1}^{n} c_i \, t^{i-1} + \kappa \, b^n t^n \right)$$

- number of stem cells
- driver mutation probability
- proliferation rate (before and after birth)
- effects of the driver mutations on the clone's carrying capacities and mutation appearance rates

If the probability of mutations occurring before birth is negligible : $\Lambda\left(t\right) = \kappa N u^n b^n t^n$

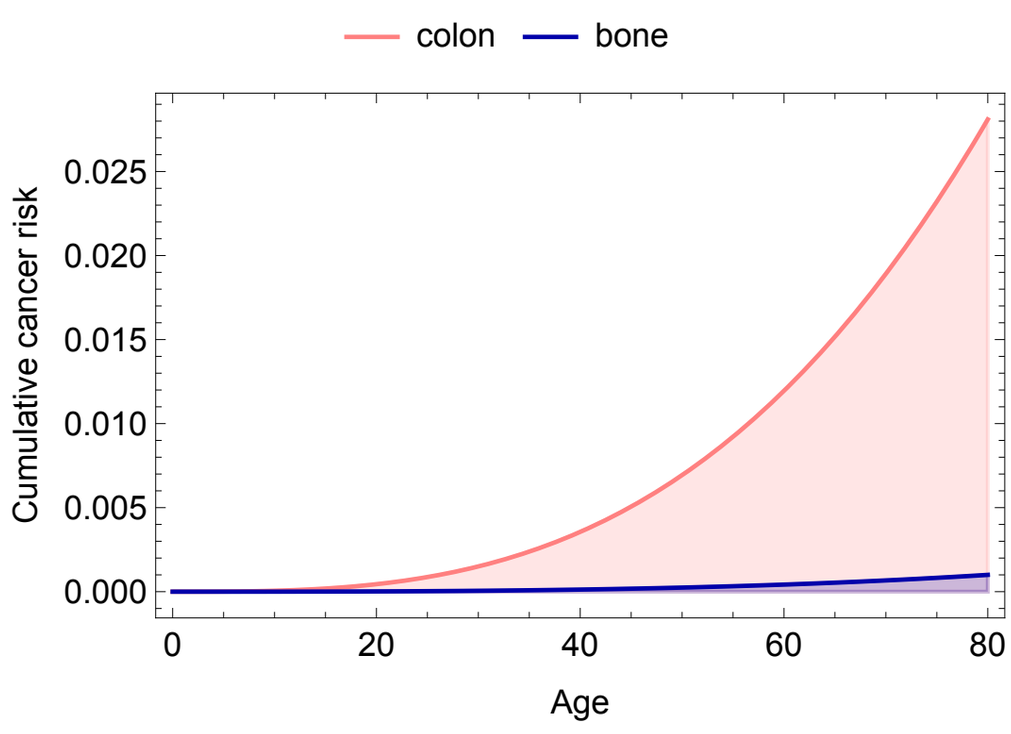
$T_n \sim W\left(n, \kappa N u^n b^n\right)$ Weibull distribution

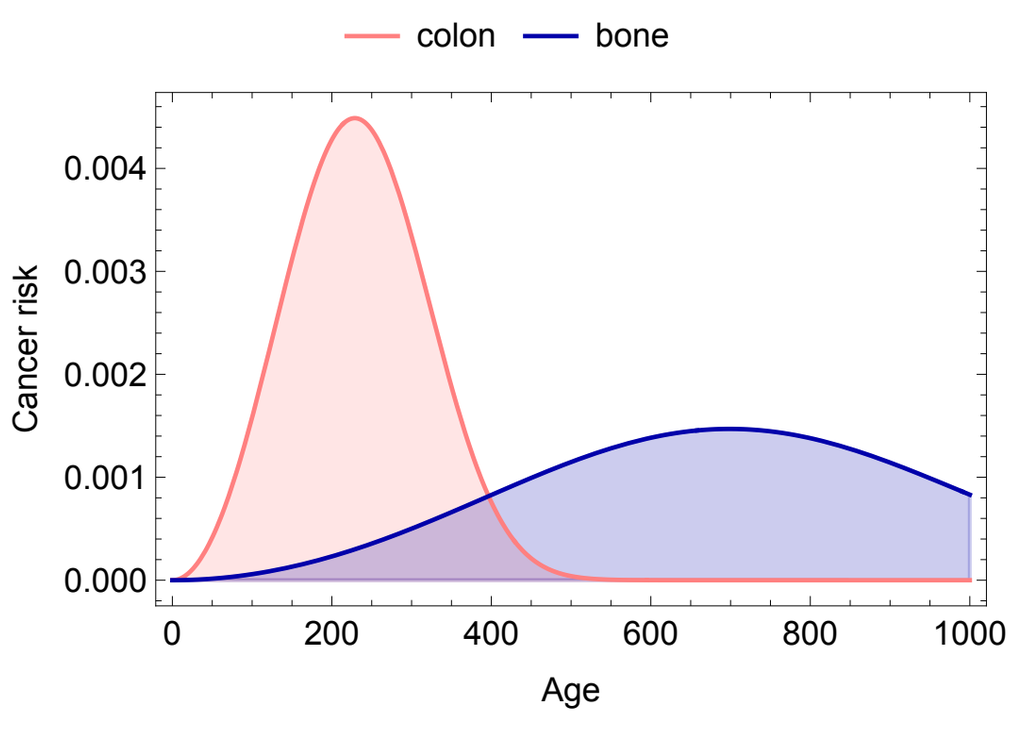
Time to cancer

N number of stem cells
b proliferation rate
u driver mutation proba
n number of drivers

$$T_n \sim W\left(n, \kappa u^n b^n N\right)$$

Probability of getting cancer by age a : $1 - e^{-\kappa u^n b^n a^n N}$

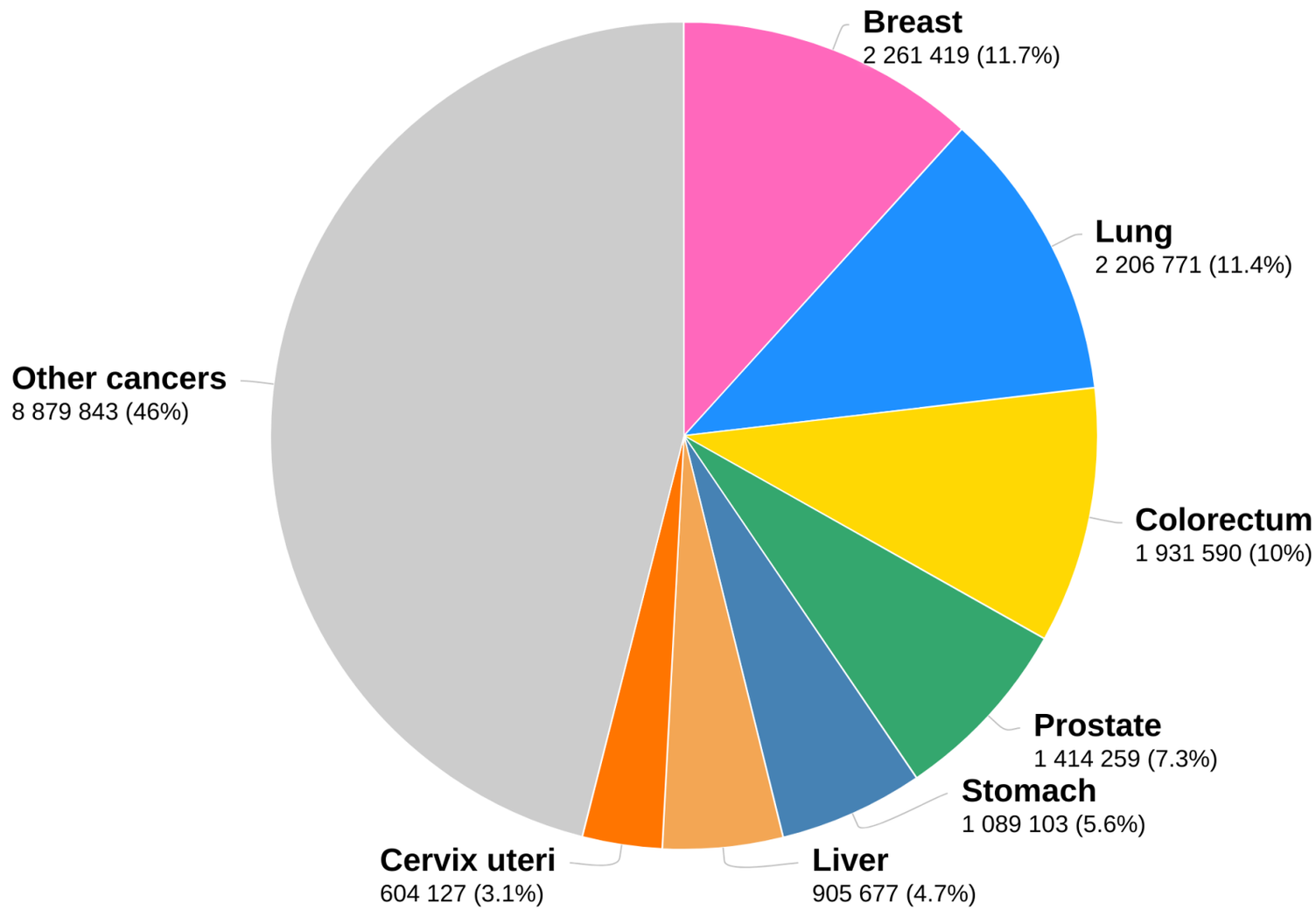




→ Lifetime cancer risk approximately $\kappa u^n D^n N$

$D = bT$ lifetime number of divisions per stem cell

Variation in cancer incidence



Total : 19 292 789

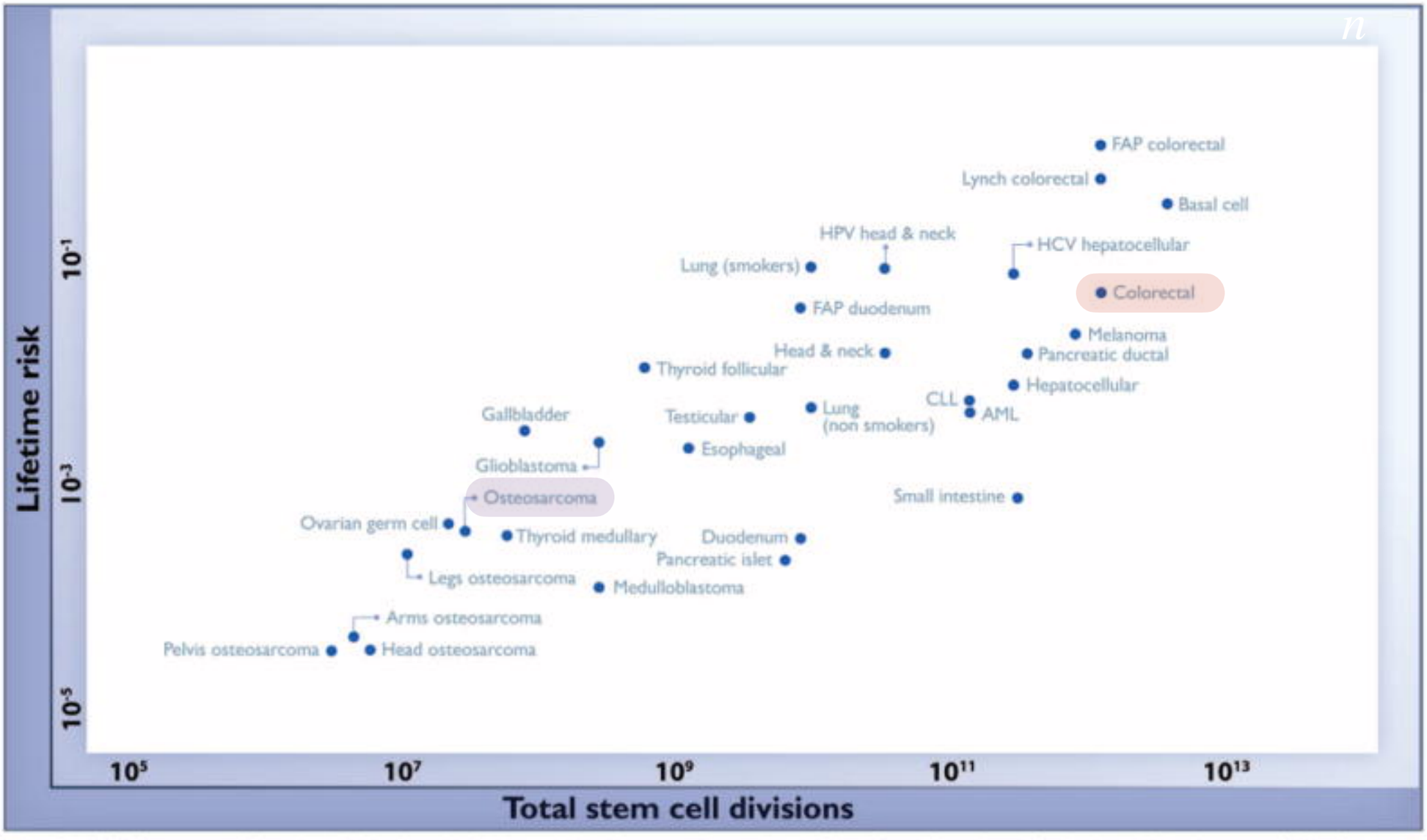
Estimated number of new cases in the world in 2020.

Globocan 2020

Variation in cancer incidence

N number of stem cells

D number of divisions



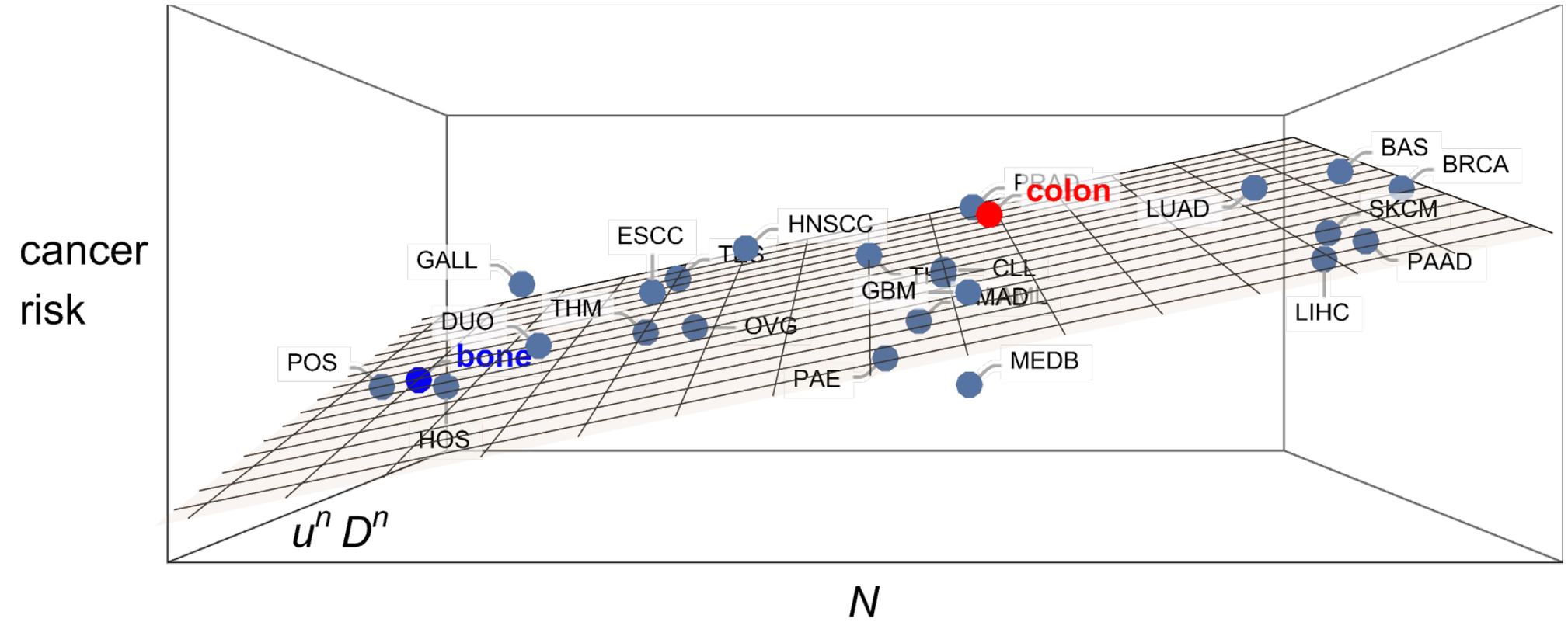
$$P \propto DN$$

Tomasetti & Volgestein *Science* 2015

→ 2/3 of variation in cancer risk explained by endogenous mutational processes

Variation in cancer incidence

N number of stem cells
D number of divisions
u driver mutation proba
n number of drivers

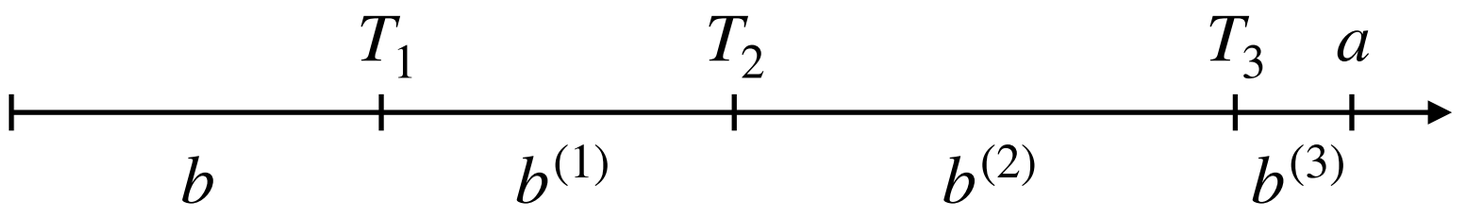


$$P \propto u^n D^n N$$

Pénisson, Lambert & Tomasetti *Nat. Commun.* 2022

→ 4/5 of variation in cancer risk explained by endogenous mutational processes

Number of somatic mutations



n number of drivers

μ background mutation rate

D_0 number of divisions during development phase

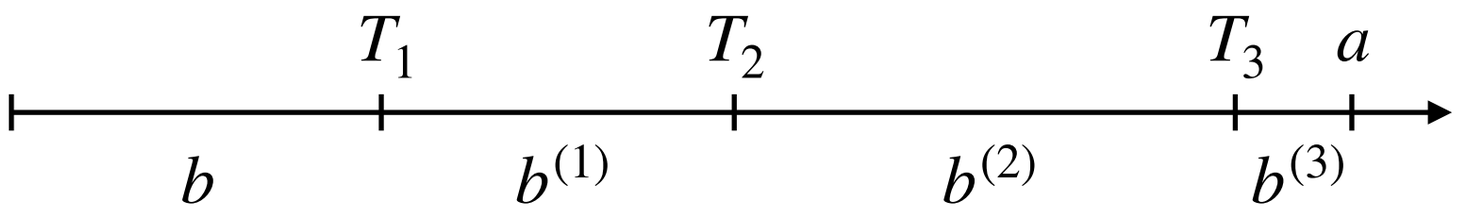
T_i hitting time of i^{th} driver

$b^{(i)}$ proliferation rate after T_i

In a cell lineage of a cancer patient of age a, in the absence of inherited or environmental factors :

$$\eta(a) = \mu\left(D_0 + bT_1 + b^{(1)}\left(T_2 - T_1\right) + \ldots + b^{(n)}\left(a - T_n\right)\right)$$

Number of somatic mutations



n number of drivers

μ background mutation rate

D_0 number of divisions during development phase

T_i hitting time of i^{th} driver

$b^{(i)}$ proliferation rate after T_i

In a cell lineage of a cancer patient of age a, in the absence of inherited or environmental factors :

$$\eta(a) = \mu\left(D_0 + bT_1 + b^{(1)}\left(T_2 - T_1\right) + \ldots + b^{(n)}\left(a - T_n\right)\right)$$

→ comparison of *observed* number in a cancer patient of age a :

$\eta^{obs}(a)$ (sequencing data)

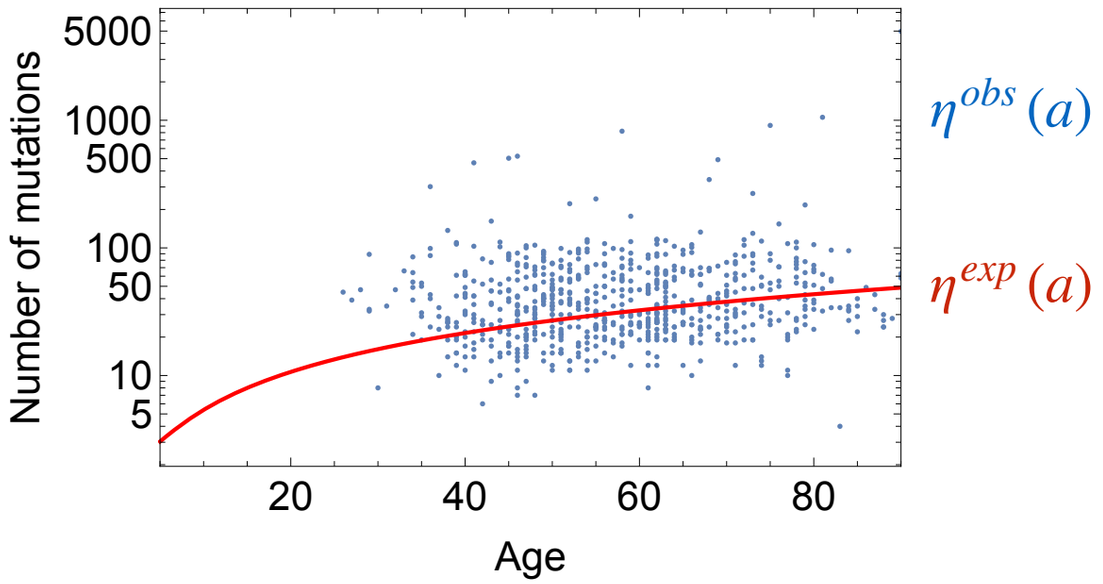
vs. *expected* number in the absence of inherited or environmental factors :

$$\eta^{exp}(a) = \mu\left(D_0 + \left(b + b^{(1)} + \cdots + b^{(n-1)}\right)\frac{1}{n}\mathbb{E}\left(T_n \mid a - 5 \leqslant T_n \leqslant a\right)\right)$$

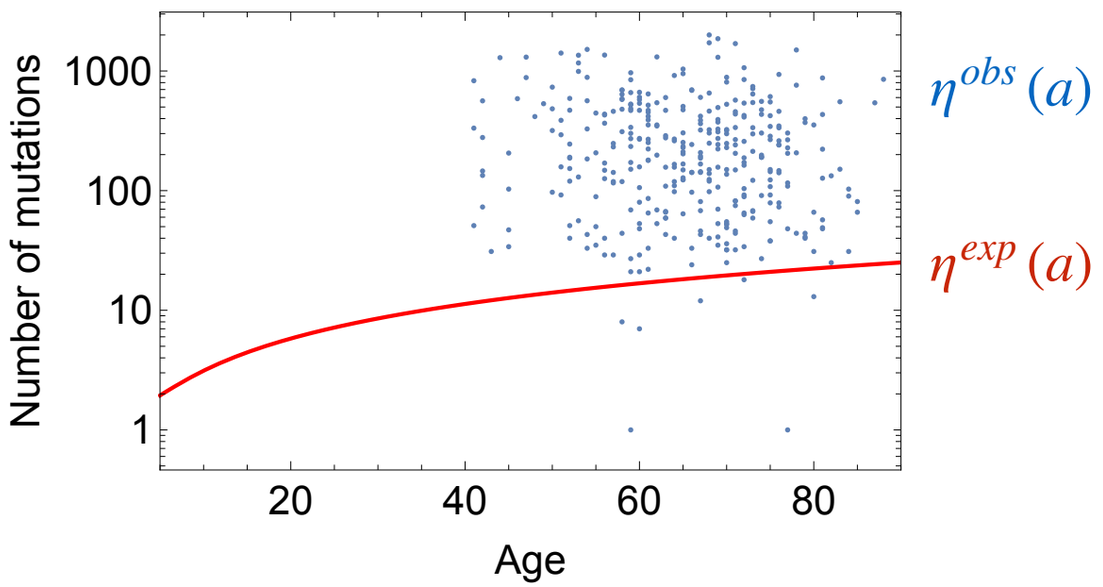
Number of somatic mutations

Observed *vs.* expected number of somatic mutations in a cancer cell lineage of patient of age a in the absence of inherited or environmental factors

Breast



Lung

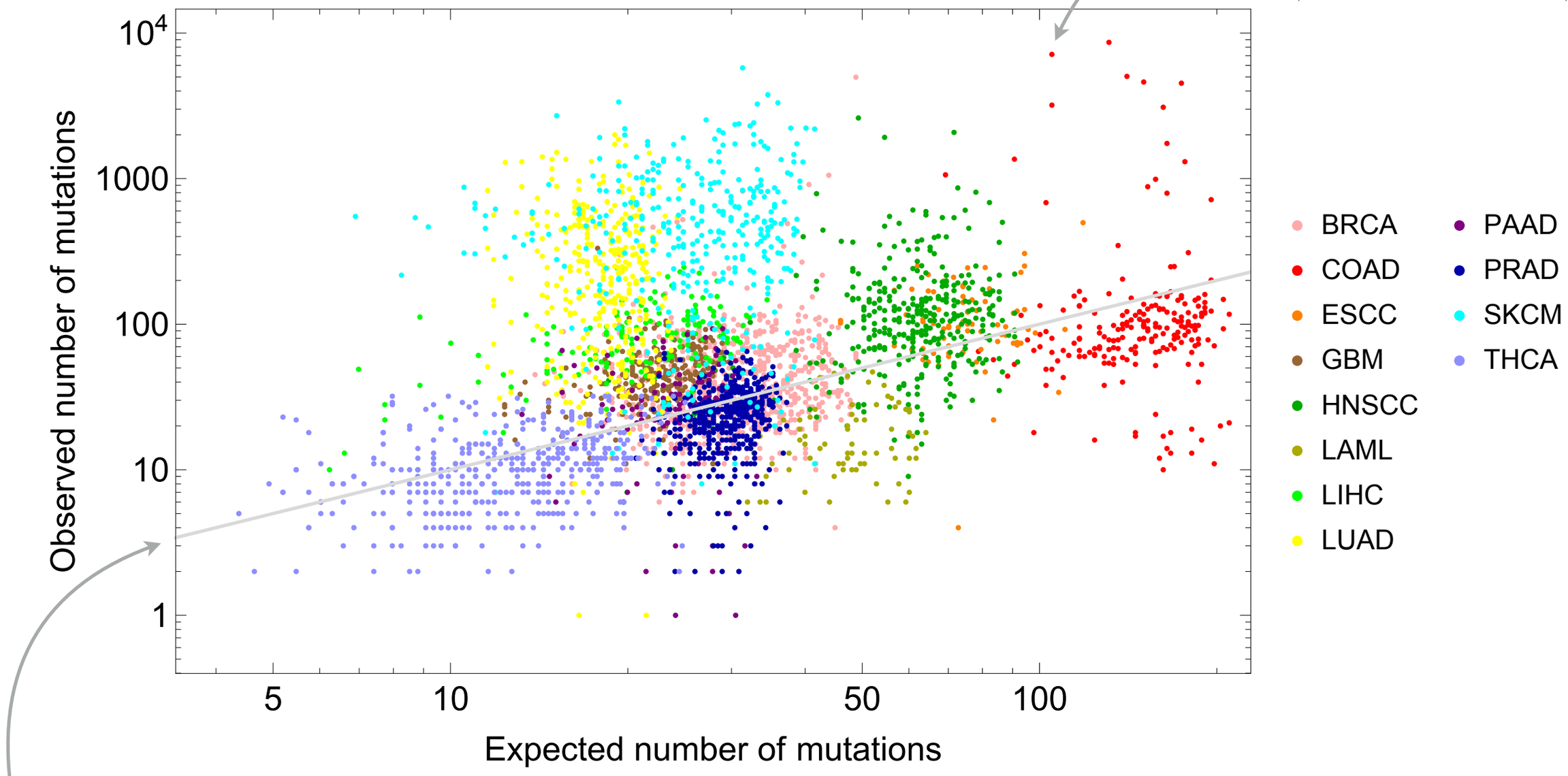


Pénisson, Lambert & Tomasetti *Nat. Commun.* 2022

Number of somatic mutations

patient i of age a_i

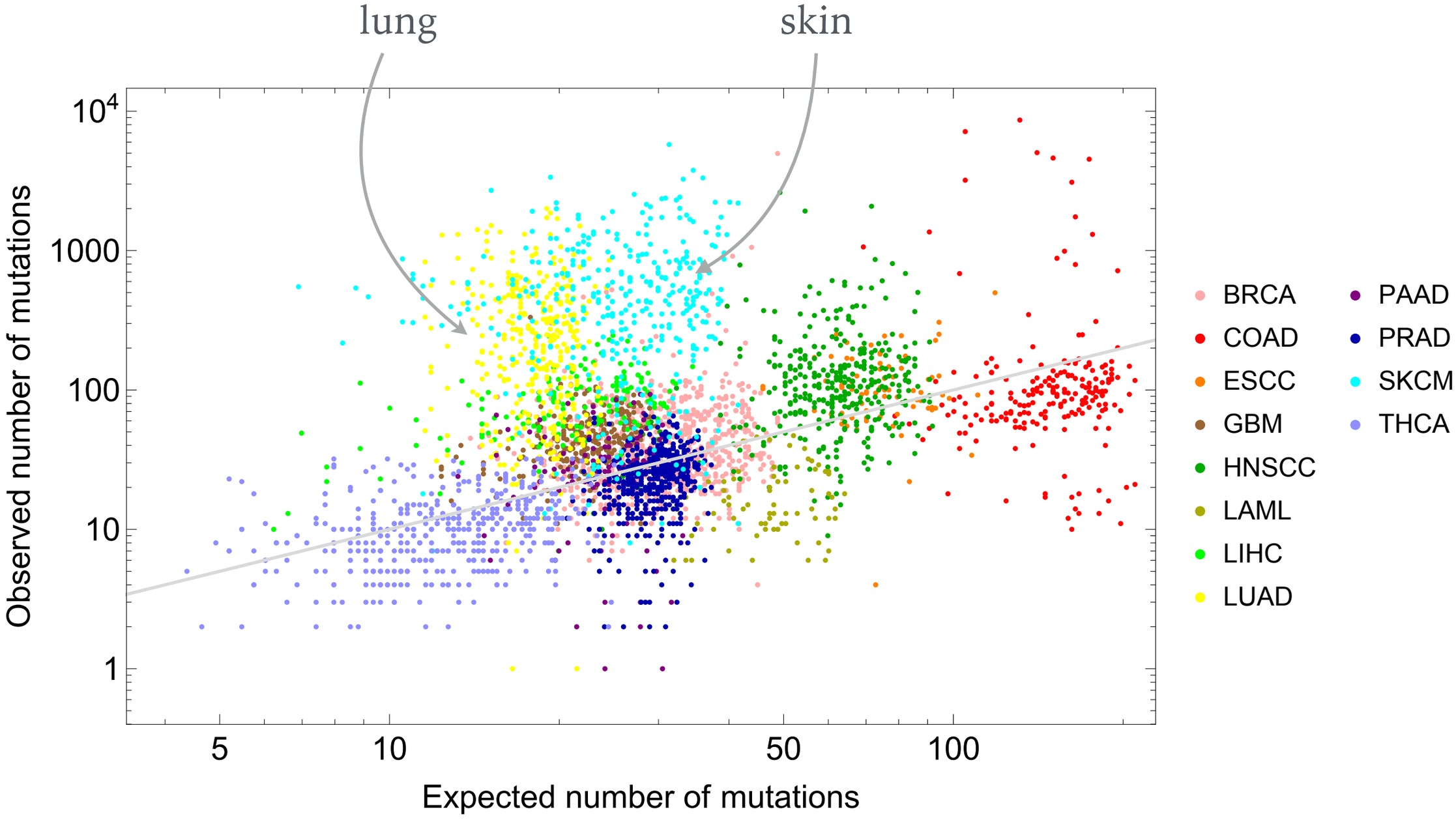
$\left(\eta^{exp}(a_i), \eta^{obs}(a_i)\right)$



theoretical case « observed number of mutations » = « expected number of mutations »

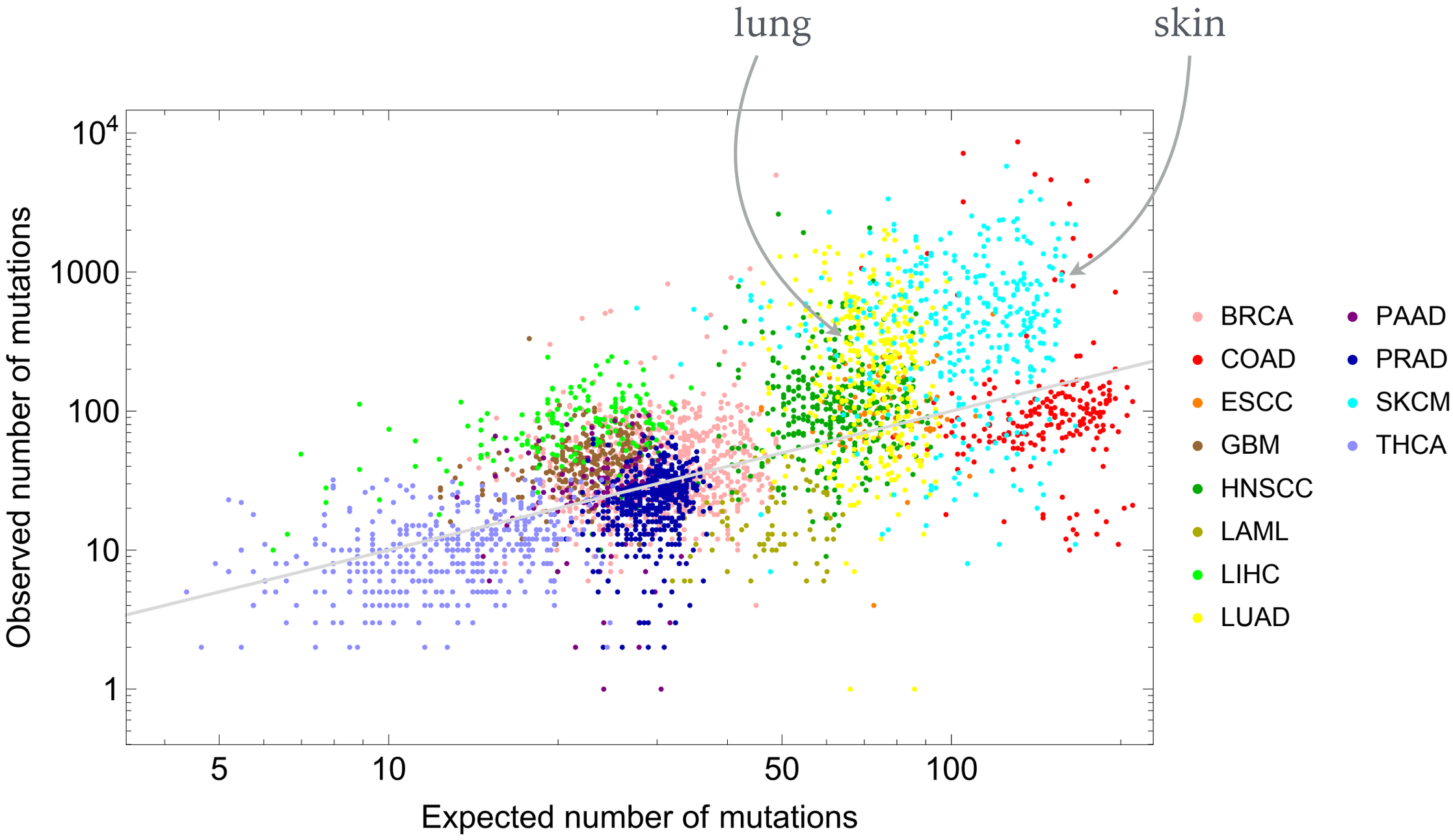
Pénisson, Lambert & Tomasetti *Nat. Commun.* 2022

Number of somatic mutations



Pénisson, Lambert & Tomasetti *Nat. Commun.* 2022

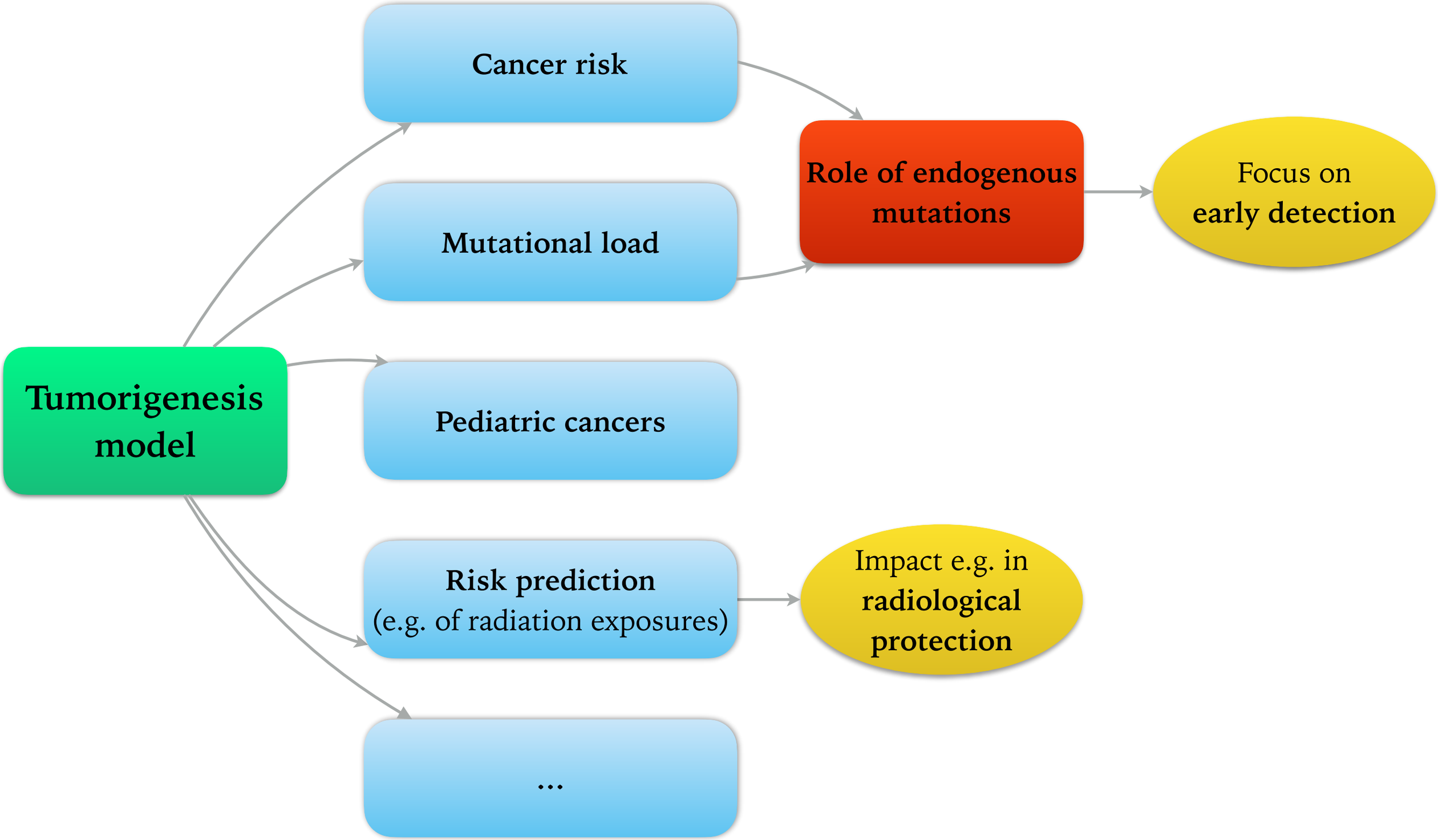
Number of somatic mutations



Including the effect of tobacco smoking and UV light exposure

Pénisson, Lambert & Tomasetti *Nat. Commun.* 2022

Overview


Cancer risk
Mutational load
Role of endogenous mutations
Focus on early detection
Tumorigenesis model
Pediatric cancers
Risk prediction (e.g. of radiation exposures)
Impact e.g. in radiological protection
…