Modelling biology from first principles

Wilfred Ndifon

African Institute for Mathematical Sciences, Next Einstein Initiative, Rwanda

January 10, 2023

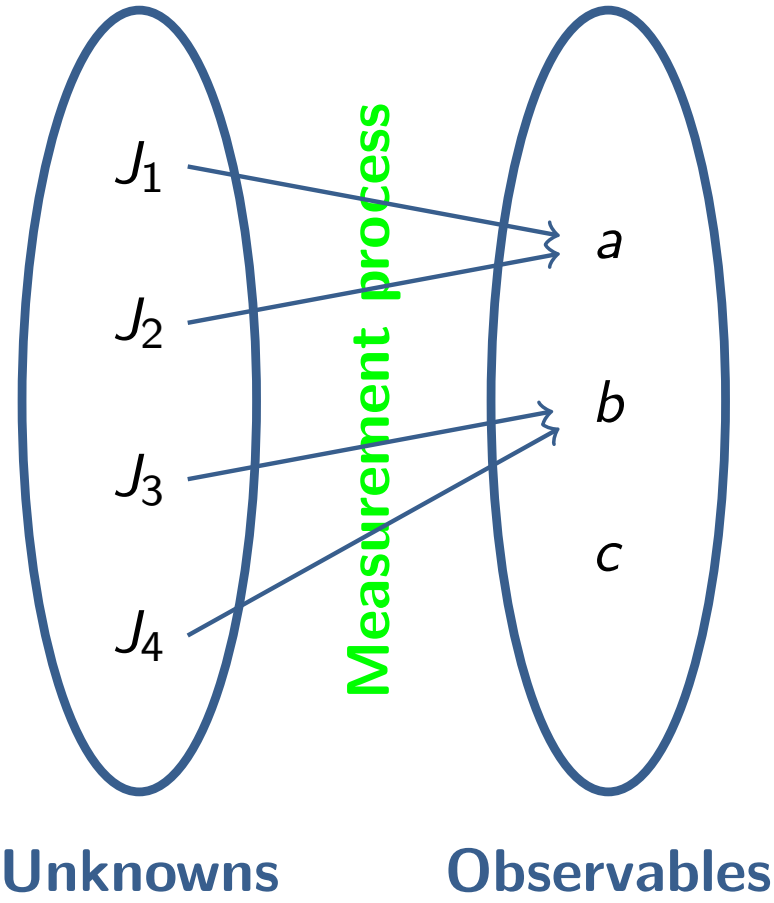
Outline

Setting the scene

Much of our understanding of how Nature works comes from interpreting the outputs of *measurement* processes.

To this end, we seek to understand both

- the *mappings* that underpin measurements (are particular outputs *signals* or *noise*?) and
- the *signals* embedded therein.



AIMS African Institute for Mathematical Sciences NEXT EINSTEIN INITIATIVE

First-principles modelling

In essence, understanding requires reverse engineering:

- Breaking up a system and putting the parts together to reconstruct observations of interest

We look for a parsimonious set of **propositions** that organizes the parts to produce the observations.

This is what I call a first-principles approach to modelling.

First-principles modelling

More specifically,

- We want to show **that** the validity of our propositions implies the observations of interest
- We also want to show **how** the propositions give rise to the observations – i.e. the underlying rules.
- An even more ambitious goal is to explain **why** the propositions give rise to the observations.

To my mind, these are some of the most important goals in all of science.

AIMS | African Institute for Mathematical Sciences NEXT EINSTEIN INITIATIVE

First-principles modelling

A proposition is in general only valid in a particular frame of reference

- for our purposes, the reference frame is the scale of physical organization of interest

It might be possible to deduce a proposition that applies to one scale from propositions that apply at a smaller scale, although this is not always desirable or more informative.

First-principles modelling

Within this "first principles" framework, how do we evaluate our *understanding*?

- Because we know how the rules that organize the parts generate observations, we can tinker with those rules and/or alter the initial conditions to produce new observations and then compare these to reality.
- Conversely, given new observations we can predict the underlying organization of the parts and then check whether this prediction is consistent with reality.

AIMS African Institute for Mathematical Sciences NEXT EINSTEIN INITIATIVE

Example: How J segment biases are generated in T cells

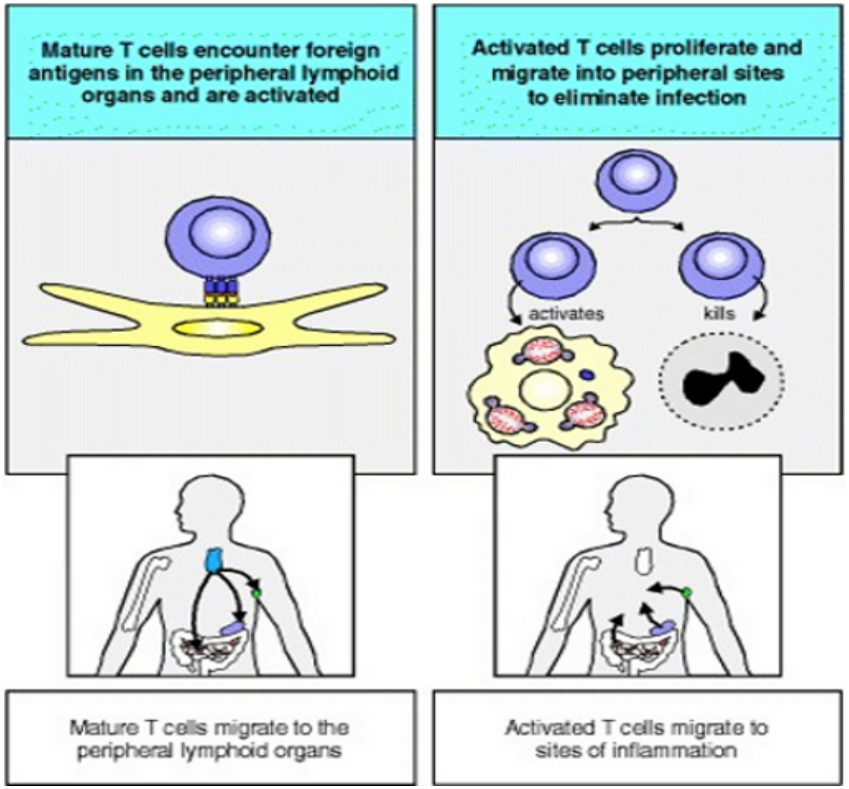


Figure: T cells use a segmented receptor to detect pathogens (Janeway Immunology)

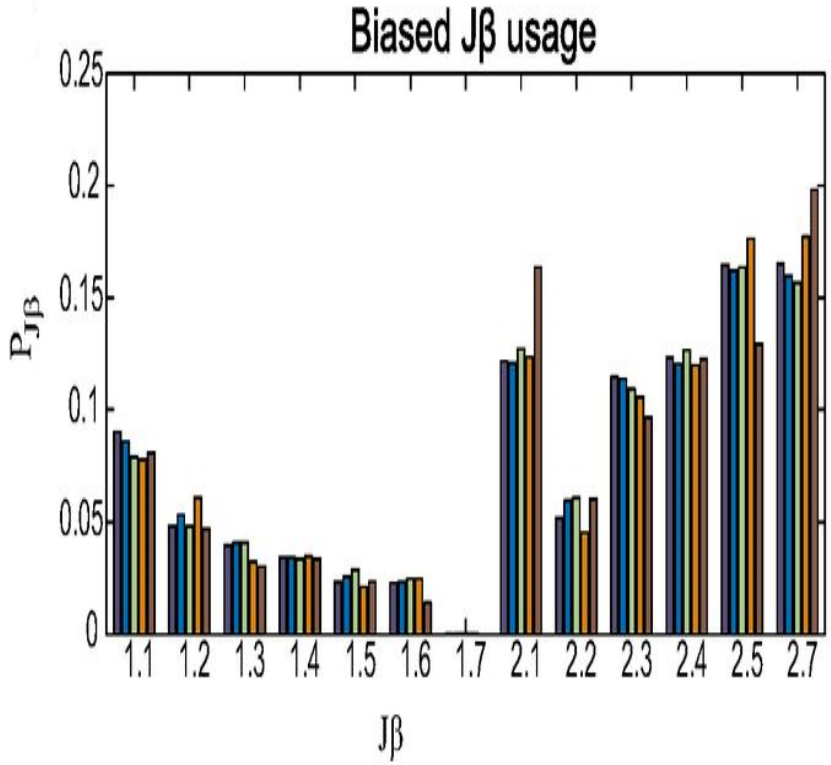


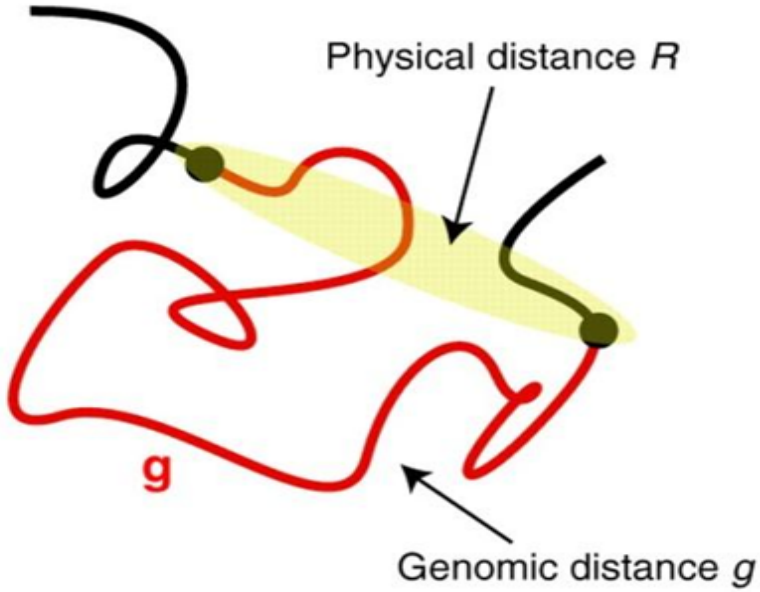
Figure: Different mice exhibit similar biases in the receptor's J segments (Ndifon et al. PNAS 2012)

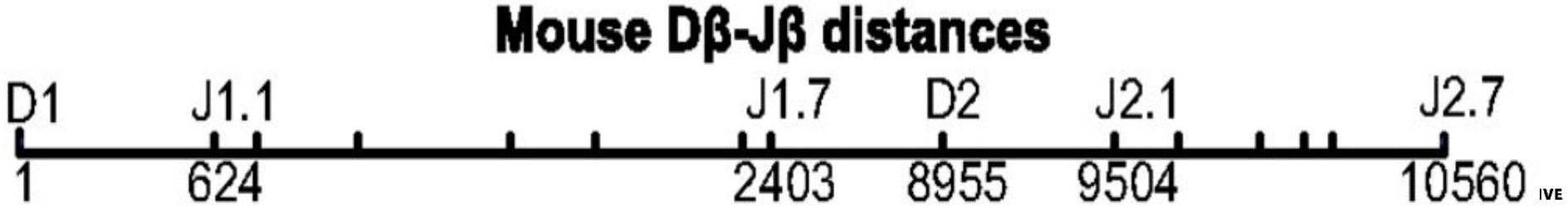
AIMS African Institute for Mathematical Sciences NEXT EINSTEIN INITIATIVE

Example: How J segment biases are generated in T cells

Proposition

The biases are generated by the conformation of the region of chromatin where the J segments are embedded





Example: How J segment biases are generated in T cells

Question

Is our proposition that biases in J segments are generated by chromatin conformation deductively valid?

Answer

It appears so!

A mechanical model instantiating our proposition

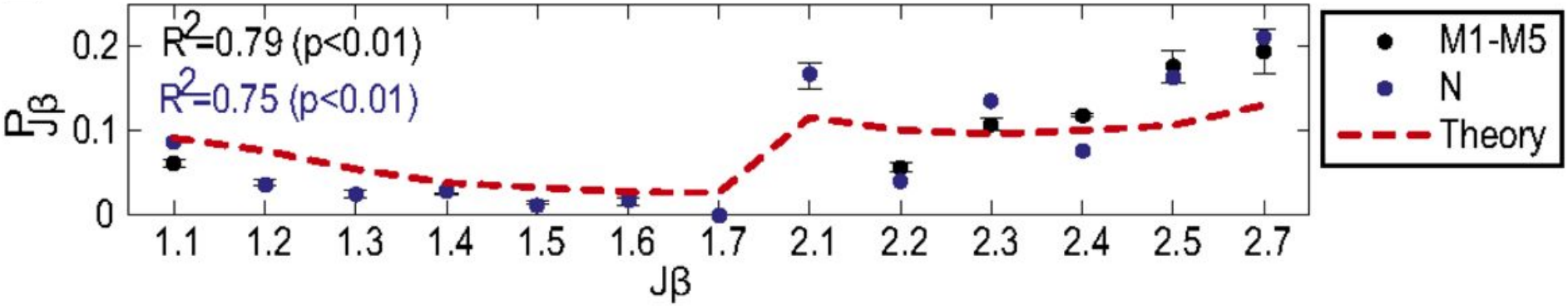
$P(J_i) \propto K \sum_{j=1}^{2} \alpha_{i,j}^{-3/2} e^{-2\alpha_{i,j}^{-2}}$, where $\alpha_{i,j} = (d_{i,j}/b)(1 - d_i/c)$

b : DNA flexibility

c : DNA curvature

$d_{i,j}$: genomic distance (in base pairs) between J_i and D_j

Fit of mechanical model to mouse data



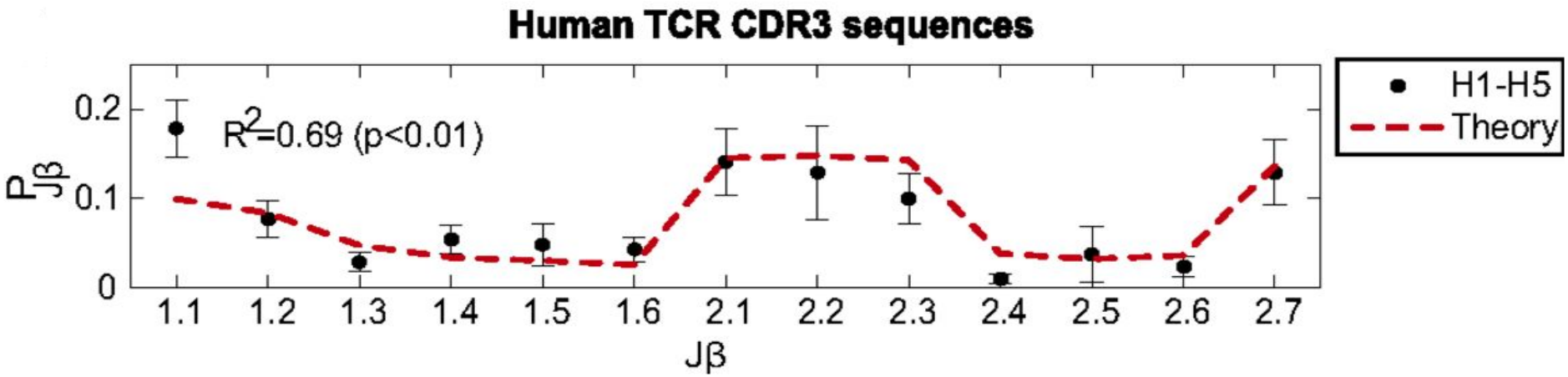
Source: Ndifon et al. PNAS 2012

Testing our understanding of what causes J segment biases

Fact: $D-J$ segment distances differ between mice and humans

Question: If we use the $D-J$ distances from humans, will we reproduce the biases observed in humans?

Answer: We can predict 69% of the human biases!



Source: Ndifon et al. PNAS 2012

This example demonstrates the amazing power of first principles modelling.

AIMS African Institute for Mathematical Sciences NEXT EINSTEIN INITIATIVE

Motivation for modelling DNA quantification

Article

A pooled testing strategy for identifying SARS-CoV-2 at low prevalence

https://doi.org/10.1038/s41586-020-2885-5

Received: 13 May 2020

Accepted: 12 October 2020

Published online: 21 October 2020

Check for updates

Leon Mutesa[1,2], Pacifique Ndishimye[2,3], Yvan Butera[1,2], Jacob Souopgui[1,2,4], Annette Uwineza[1,2], Robert Rutayisire[1,2], Ella Larissa Ndoricimpaye[2], Emile Musoni[2], Nadine Rujeni[2], Thierry Nyatanyi[2], Edouard Ntagwabira[2], Muhammed Semakula[2], Clarisse Musanabaganwa[2], Daniel Nyamwasa[2], Maurice Ndashimye[2,3], Eva Ujeneza[3], Ivan Emile Mwikarago[2], Claude Mambo Muvunyi[2], Jean Baptiste Mazarati[2], Sabin Nsanzimana[2], Neil Turok[3,5,6]✉ & Wilfred Ndifon[3]✉

Suppressing infections of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) will probably require the rapid identification and isolation of individuals infected with the virus on an ongoing basis. Reverse-transcription polymerase chain reaction (RT–PCR) tests are accurate but costly, which makes the

AIMS African Institute for Mathematical Sciences NEXT EINSTEIN INITIATIVE

Polymerase chain reaction (PCR)

Kary Mullis (Nobel Prize '93) invented the polymerase chain reaction (PCR) in the 1980s to solve the DNA quantification problem.

Figure: Karry Mullis (wikipedia)

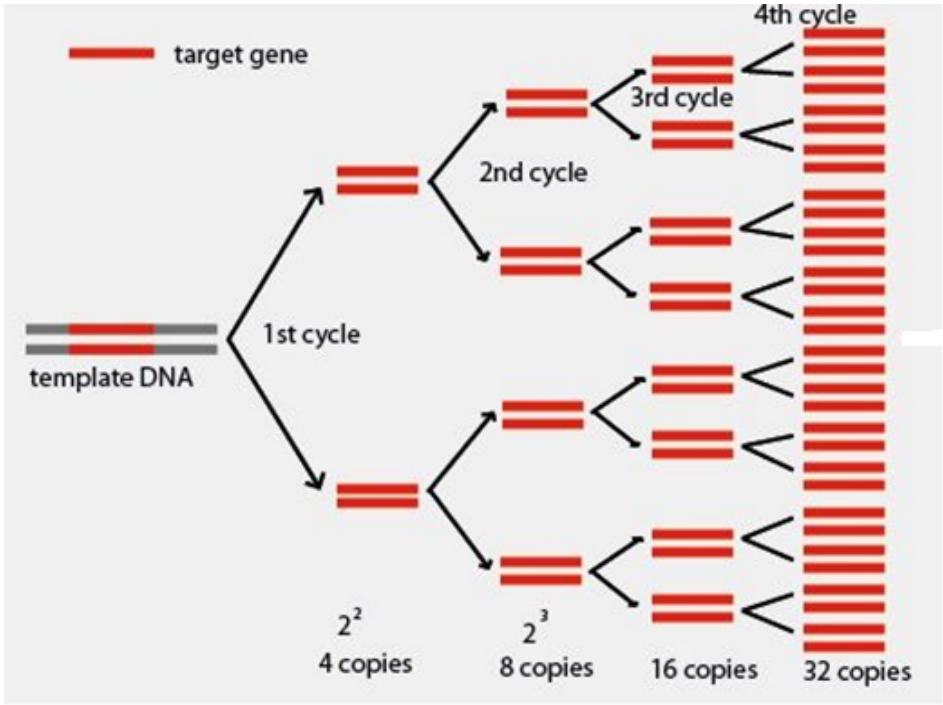


Figure: Basic principle of PCR

AIMS African Institute for Mathematical Sciences NEXT EINSTEIN INITIATIVE

PCR

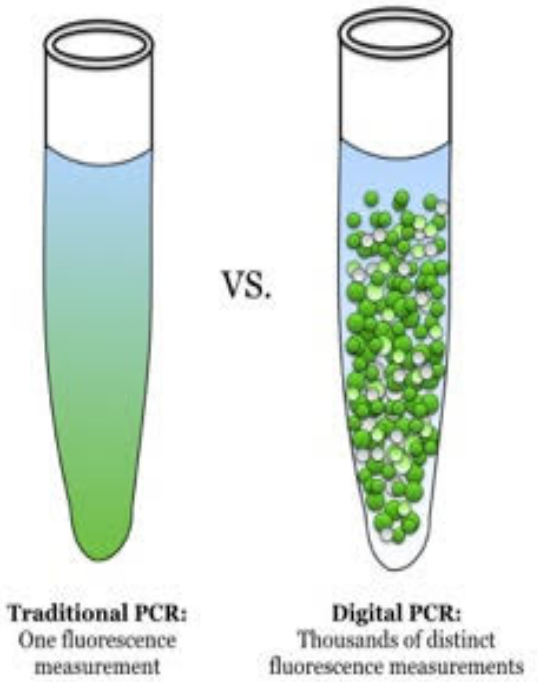
Despite the valiant efforts of many biologists and technologists, most reported PCR data are quantitative only in a relative, rather than an absolute, sense

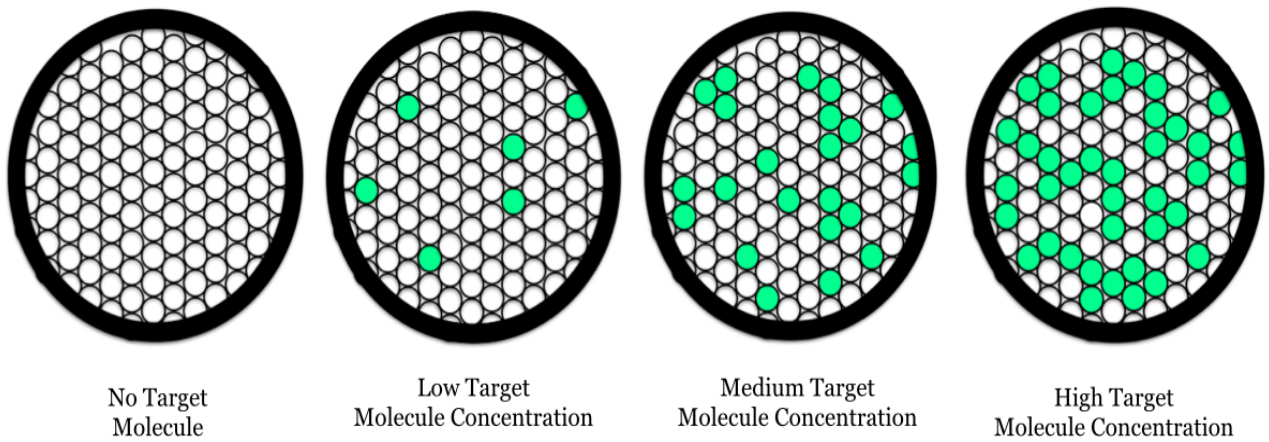
- The ratio of the amount of target DNA to that of a reference DNA is frequently measured
- Mostly (eg. during the response to COVID-19), only the PCR quantification threshold (Ct), an indirect readout of the number of DNA molecules, is measured

Existing mathematical solutions to this problem suffer from several limitations, which I will discuss in the following

Digital PCR

A recent modification of traditional PCR, called digital PCR, was developed to facilitate the absolute quantification of DNA.



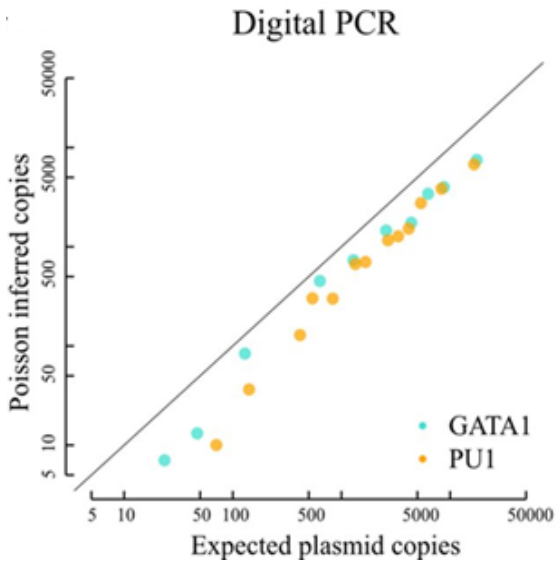


Source: wikipedia

AIMS African Institute for Mathematical Sciences NEXT EINSTEIN INITIATIVE

Digital PCR

However, the standard method used to interpret digital PCR data tends to underestimate the number of input DNA molecules by several fold!



Source: Mojtahedi et al. Nucleic Acids Research 2014

In the following slides, we will use a new principled model of PCR to explain why.

A parsimonious mathematical model of PCR

Proposition

The PCR process is essentially a discrete-state, continuous-time Markov process.

Rationale

- Products of PCR reactions, DNA molecules, are countable
- What happens in the next PCR cycle is conditionally independent of what happened in the past given the present state of the reaction
- In experiments, *time* is reported as a positive real number

The following results are a few logical consequences of this proposition.

Ref: Degoot & Ndifon, "Stochastics of DNA quantification", arxiv.org/pdf/2301.02149.pdf

AIMS African Institute for Mathematical Sciences NEXT EINSTEIN INITIATIVE

pgf for the number of molecules (*pgf theorem*)

Theorem

Let $\{X(t), t \in R\}$ *be a continuous-time Markov process with* p *phases* $I_i, i = 1, 2, ..., p$, *a countable state space* $S \subset \mathbb{N}^+$, *phase-specific transition rates* r_i, $i \in \{1, 2, \ldots, p\}$, *and state transition probability given by*

$$P\left(X(t' + \Delta t) = x | X(t') = x'\right) = \delta(x' - x + 1) \sum_{i=1}^{p} r_i 1_{I_i}(t'), \qquad (1)$$

where $1(.)$ *is the indicator function and* $\delta(.)$ *is the Kronecker delta function. If the process starts with* n *molecules, the probability generating function (pgf) for the number of molecules found at time* $t \in I_k, k \leq p$, *is*

$$G(s, t) = \left[\frac{s e^{-\left(r_k t + \sum_{i=1}^{k-1} (r_i - r_k)\tau_i\right)}}{1 - s\left(1 - e^{-\left(r_k t + \sum_{i=1}^{k-1} (r_i - r_k)\tau_i\right)}\right)} \right]^n, \ \tau_i = |I_i|. \qquad (2)$$

Proof of the pgf theorem

We will prove the pgf theorem by mathematical induction on k.

Case 1: $k = 1$:

In this case, the Chapman-Kolmogorov forward equation corresponding to our process is given by:

$$\frac{\partial P(X = x, t|X = x', t')}{\partial t} = r_1(x-1)P(X = x-1, t|X = x', t') - r_1 x P(X = x, t|X = x', t'), \tag{3}$$

where we have set $t = t' + \Delta t$, and r_1 is the replication rate associated with the first phase of the PCR process.

To simplify our notation, we abbreviate $P(X = x, t|X = x', t')$ by $P(x, t)$.

Proof of the pgf theorem

Recall that the pgf of $P(x,t)$ is defined as:

$$G(s,t) = \sum_{x=0}^{\infty} s^x P(x,t).$$

Multiplying both sides of (3) by s^x and summing over all x yields:

$$\begin{aligned}
\sum_{x=0}^{\infty} s^x \frac{\partial P(x,t)}{\partial t} &= r_1 \sum_{x=0}^{\infty} (x-1) s^x P(x-1,t) - r_1 \sum_{x=0}^{\infty} x s^x P(x,t) \\
&= r_1 s^2 \sum_{x=0}^{\infty} (x-1) s^{x-2} P(x-1,t) - r_1 s \sum_{x=0}^{\infty} x s^{x-1} P(x,t) \\
&= r_1 s \left[s \sum_{x=0}^{\infty} (x-1) s^{x-2} P(x-1,t) - \sum_{x=0}^{\infty} x s^{x-1} P(x,t) \right]. \qquad (4)
\end{aligned}$$

AIMS African Institute for Mathematical Sciences NEXT EINSTEIN INITIATIVE

Proof of the pgf theorem

Using

$$\begin{aligned}\frac{\partial G(s,t)}{\partial s} &= \sum_{x=0}^{\infty} x s^{x-1} P(x,t) \text{ and} \\ \frac{\partial G(s,t)}{\partial t} &= \sum_{x=0}^{\infty} s^x \frac{\partial P(x,t)}{\partial t}, \end{aligned} \tag{5}$$

we simplify (4) to obtain

$$\frac{\partial G(s,t)}{\partial t} = r_1 s(s-1)\frac{\partial G(s,t)}{\partial s}, \tag{6}$$

which is a partial differential equation (pde) in $G(s,t)$.

Proof of the pgf theorem

By defining new variables $u = u(s,t) = t$ and $v = v(s,t) = c$, with c an arbitrary constant, using the method of characteristics, we transform

$$\frac{\partial G(s,t)}{\partial t} = r_1 s(s-1)\frac{\partial G(s,t)}{\partial s}$$

into an easy to solve ode, yielding the general solution

$$G(s,t) = \Psi\left(\frac{s-1}{s}e^{r_1 t}\right). \tag{7}$$

Proof of the pgf theorem

Because there are n molecules at the start of the PCR process ($t = 0$), $p(x, 0) = 1$ if $x = n$ and 0 otherwise, so

$$G(s, 0) = \Psi\left(\frac{s-1}{s}\right) = \sum_{x=0}^{\infty} s^x P(x, 0) = s^n. \tag{8}$$

Observe that the argument y of $\Psi(y)$ maps onto $(\frac{1}{1-y})^n$, implying that

$$\begin{aligned} G(s, t) = \Psi\left(\frac{s-1}{s} e^{r_1 t}\right) &= \left(\frac{1}{1 - \frac{s-1}{s} e^{r_1 t}}\right)^n \\ &= \left[\frac{s e^{-r_1 t}}{1 - s\left(1 - e^{-r_1 t}\right)}\right]^n. \end{aligned} \tag{9}$$

Equation (9) matches the pgf when $k = 1$.

AIMS African Institute for Mathematical Sciences NEXT EINSTEIN INITIATIVE

Proof of the pgf theorem

Note: It can be readily shown that

$$G(s,t) = \left[\frac{se^{-r_1 t}}{1 - s\left(1 - e^{-r_1 t}\right)}\right]^n$$

solves

$$\frac{\partial G(s,t)}{\partial t} = r_1 s(s-1)\frac{\partial G(s,t)}{\partial s}$$

by differentiating the latter equation with respect to s and t.

Proof of the pgf theorem

Case 2: $k = 2$:

There are two amplification phases with rates r_1 and r_2:

- The first one runs from $t = 0$ to $t = \tau_1$.
- The second runs from $t = t_1$ to $t = \tau_1 + \tau_2$.

In phase two, the pgf has the same general functional form as in phase one, albeit with a different initial condition, that is

$$G(s,t) = \Psi\left(\frac{s-1}{s}e^{r_2(t-\tau_1)}\right),$$

with the initial condition (at time $t = \tau_1$)

$$G(s,t_1) = \Psi\left(\frac{s-1}{s}\right) = \left[\frac{se^{-r_1 t_1}}{1-s\left(1-e^{-r_1\tau_1}\right)}\right]^n.$$

AIMS | African Institute for Mathematical Sciences NEXT EINSTEIN INITIATIVE

Proof of the pgf theorem

Using the same procedure as before, we obtain

$$\begin{aligned} G(s,t) &= \Psi\left(\frac{s-1}{s}e^{r_2(t-\tau_1)}\right) \\ &= \frac{\left(\frac{1}{1-\frac{s-1}{s}e^{r_2(t-\tau_1)}}\right)^n e^{-nr_1t_1}}{\left[1-\left(\frac{1}{1-\frac{s-1}{s}e^{r_2(t-\tau_1)}}\right)\left(1-e^{-r_1\tau_1}\right)\right]^n} \\ &= \frac{s^n e^{-n[r_2t+(r_1-r_2)\tau_1]}}{\left[1-s\left(1-e^{-[r_2t+(r_1-r_2)\tau_1]}\right)\right]^n}. \end{aligned} \tag{10}$$

The right hand side of (10) matches the pgf when $k = 2$, as expected.

Proof of the pgf theorem

Case 3: We assume the statement is true for an arbitrary k, that is

$$G(s,t) = \left[\frac{se^{-z}}{1-s(1-e^{-z})}\right]^n$$

where $z = r_k t + \sum_{i=1}^{k-1}(r_i - r_k)\tau_i$, and we prove it for $k+1$.

As before, in phase $k+1$, the generating function has the functional form

$$G(s,t) = \Psi\left(\frac{s-1}{s}e^{r_{k+1}(t-\sum_{i=1}^{k}\tau_i)}\right).$$

Proof of the pgf theorem

At time $t = \sum_{i=1}^{k-1} \tau_i$, by the induction step, we have

$$G(s,t) = \Psi\left(\frac{s-1}{s}\right) = \left[\frac{se^{-z}}{1-s(1-e^{-z})}\right]^n.$$

Using the same arguments as before, we find that, for t_{k+1},

$$\begin{aligned} G(s,t) &= \Psi\left(\frac{s-1}{s} e^{r_{k+1}(t-\sum_{i=1}^{k}\tau_i)}\right) \\ &= \left[\frac{\frac{1}{1-\frac{s-1}{s}e^{r_{k+1}(t-\sum_{i=1}^{k}\tau_i)}}e^{-z}}{1-\frac{1}{1-\frac{s-1}{s}e^{r_{k+1}(t-\sum_{i=1}^{k}\tau_i)}}(1-e^{-z})}\right]^n \\ &= \frac{s^n e^{-n\left[r_{k+1}t+\sum_{i=1}^{k}(r_i-r_k)\tau_i\right]}}{\left[1-s\left(1-e^{-\left[r_{k+1}t+\sum_{i=1}^{k}(r_i-r_k)\tau_i\right]}\right)\right]^n}, \end{aligned} \tag{11}$$

and this ends the proof of the pgf theorem.

pgf for PCR process with Poisson-distributed initial state

Theorem

Let $\{X(t), t \in R\}$ be the discrete-state, continuous-time Markov process described in the pgf theorem. If the initial state of the process is Poisson-distributed with mean λ, then the pgf for the state of the process at time $t \in I_k, k \leq p$, is given by

$$G(s,t) = e^{\left[\frac{\lambda(s-1)}{1-s\left(1-e^{-\left(r_k t+\sum_{i=1}^{k-1}(r_i-r_k)\tau_i\right)}\right)}\right]}. \tag{12}$$

Probability distribution of the number of molecules

Corollary

Let $\{X(t), t \in R\}$ be the discrete-state, continuous-time Markov process described in the pgf theorem. If the initial state of the process is Poisson-distributed with mean λ, then the probability that there are x molecules at time $t \in I_k, k \leq p$, is given by

$$P(x|\lambda, \vec{r}, t, \vec{\tau}) = e^{-\lambda}\left(1 - e^{-z}\right)^x \sum_{i=1}^{x} \frac{\binom{x-1}{i-1}}{i} \left(\frac{\lambda e^{-z}}{1 - e^{-z}}\right)^i, \qquad (13)$$

where

$$z = r_k t + \sum_{i=1}^{k-1}(r_i - r_k)\tau_i \text{ and } \vec{\tau} = (\tau_1, \tau_2, ..., \tau_{k-1}).$$

AIMS African Institute for Mathematical Sciences NEXT EINSTEIN INITIATIVE

Probability density function (pdf) of the Ct value

Let t be the Ct value of a PCR process with up to p phases with lengths $\vec{\tau} = (\tau_1, \tau_2, \ldots, \tau_k)$ and replication rates $\vec{r} = (r_1, r_2, \ldots, r_k)$.

By definition, the Ct value t is the time at which the number of molecules reaches the quantification threshold, denoted x.

By Bayes' theorem, the probability density of t is given by

$$P(t|\lambda, \vec{r}, \vec{\tau}, x) = \frac{P(\lambda, \vec{r}, \vec{\tau}, x|t)P(t)}{P(\lambda, \vec{r}, \vec{\tau}, x)}. \quad (14)$$

pdf of the Ct value

Because λ is independent of $\vec{r}$, t, and $\vec{\tau}$, and $\vec{r}$ is also independent of both s and the entries of $\vec{\tau}$, we simplify $P(t|x,\vec{r},\vec{\tau},\lambda)$ as follows:

$$\begin{aligned} P(t|\lambda,\vec{r},\vec{t},x) &= \frac{P(x|\lambda,\vec{r},t,\vec{\tau})P(\lambda)P(\vec{r})P(\vec{\tau})P(t|\vec{\tau})}{P(x|\lambda,\vec{r},\vec{\tau})P(\lambda)P(\vec{r})P(\vec{\tau})} \\ &= \frac{P(x|\lambda,\vec{r},t,\vec{\tau})P(t|\vec{\tau})}{P(x|\lambda,\vec{r},\vec{\tau})} \\ &= \frac{P(x|\lambda,\vec{r},t,\vec{\tau})P(t|\vec{\tau})}{\int_{\sum_{i=1}^{k-1}(r_i-r_k)\tau_i}^{\infty} P(x|\lambda,\vec{r},s,\vec{\tau})P(s|\vec{\tau})ds}. \end{aligned} \tag{15}$$

In the next slides, we will use Eqn. (15) to derive the pdf for a single-phase PCR process, other key statistical features of which we will also compute.

pdf of the Ct value for a single-phase PCR process

By assuming a uniform prior for t, we obtain the following pdf for t:

$$P(t|x,\lambda,r_1) = \frac{r_1\lambda x e^{-r_1 t}(1-e^{-r_1 t})^{x-1}\ {}_1F_1\left(1-x,2,\frac{-\lambda e^{-r_1 t}}{1-e^{-r_1 t}}\right)}{e^{\lambda}-1}, \tag{16}$$

where ${}_1F_1$ is the hypergeometric function, i.e.

$${}_1F_1\big(a;b;c\big) = \sum_{k=0}^{\infty}\frac{(a)_k}{(b)_k}\frac{c^k}{k!},$$

and $(a)_k$ is the rising factorial, i.e. $(a)_k = a(a+1)(a+2)\ldots(a+k-1)$ with $(a)_0 = 1$.

Mean and variance of the Ct value

- **Mean**

$$\mathbb{E}(t) = \frac{\psi(x+1)}{r_1} - \frac{\sum_{i=1}^{\infty} \frac{\lambda^i}{i!}\psi(i)}{r_1\left(e^{\lambda}-1\right)}, \tag{17}$$

where $\psi(\cdot)$ is the first polygamma function.

- **Variance**

$$\mathbf{Var}(t) = \frac{\left(e^{\lambda}-1\right)\sum_{j=1}^{x}\frac{\lambda^j}{j!}\left[\psi_1(j)+\psi(j)^2\right] - \left[\sum_{j=1}^{x}\frac{\lambda^j}{j!}\psi(j)\right]^2}{\left(r_1(e^{\lambda}-1)\right)^2} - \frac{\psi_1(x+1)}{r_1^2}, \tag{18}$$

where $\psi_1(\cdot)$ is the second polygamma function.

AIMS | African Institute for Mathematical Sciences NEXT EINSTEIN INITIATIVE

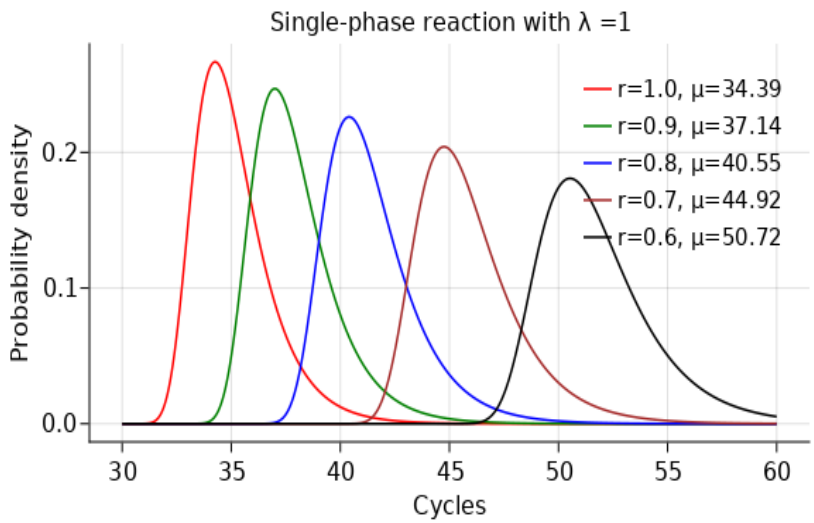
Cumulative density function (cdf) of the Ct value

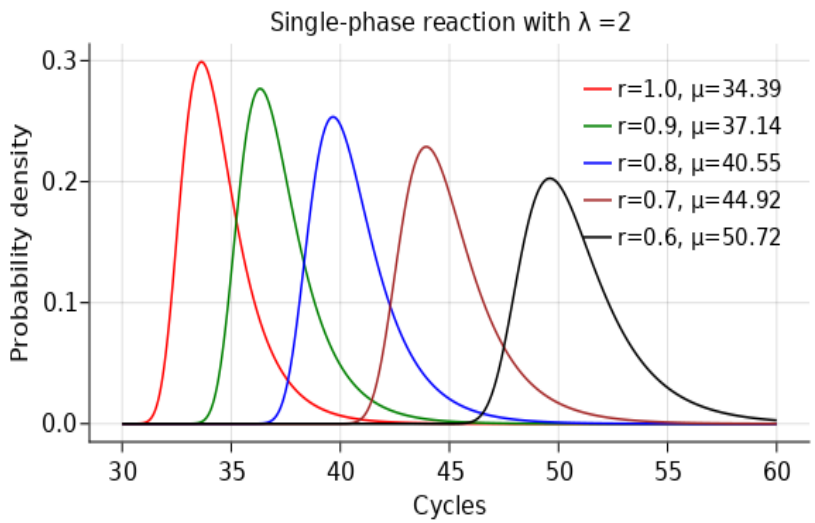
The cdf is given by

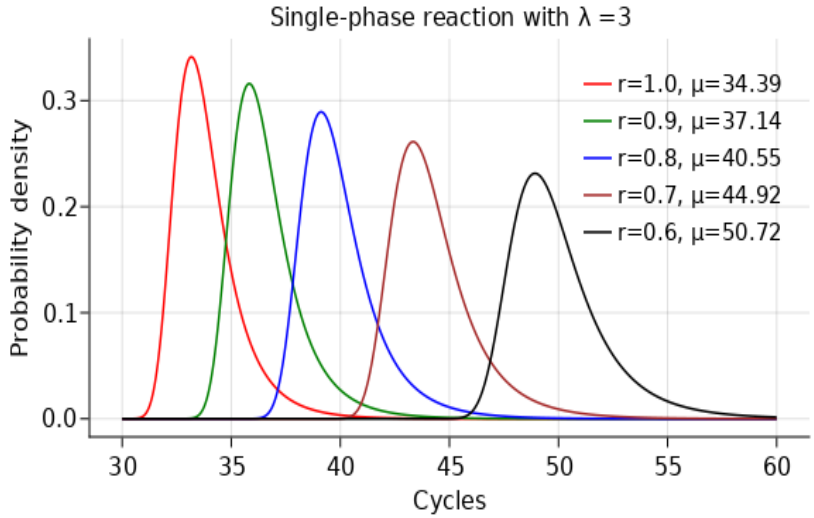
$$F(t) = 1 - \frac{\sum_{i=1}^{x} \frac{\binom{x}{i}}{(i-1)!} \lambda^i B_{e^{-r_1 t}}(i, x-i+1)}{e^{\lambda} - 1}, \tag{19}$$

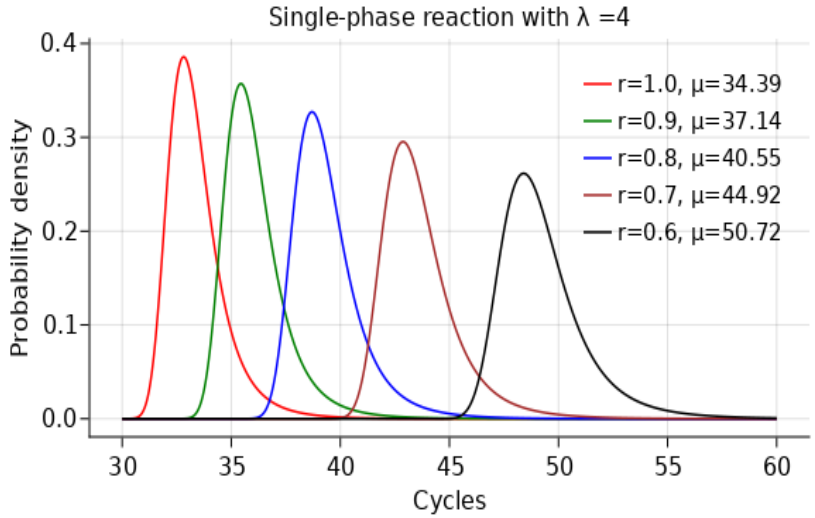
where $B_{e^{-r_1 t}}(i, x-i+1)$ is the incomplete Beta function.

Shape of the pdf

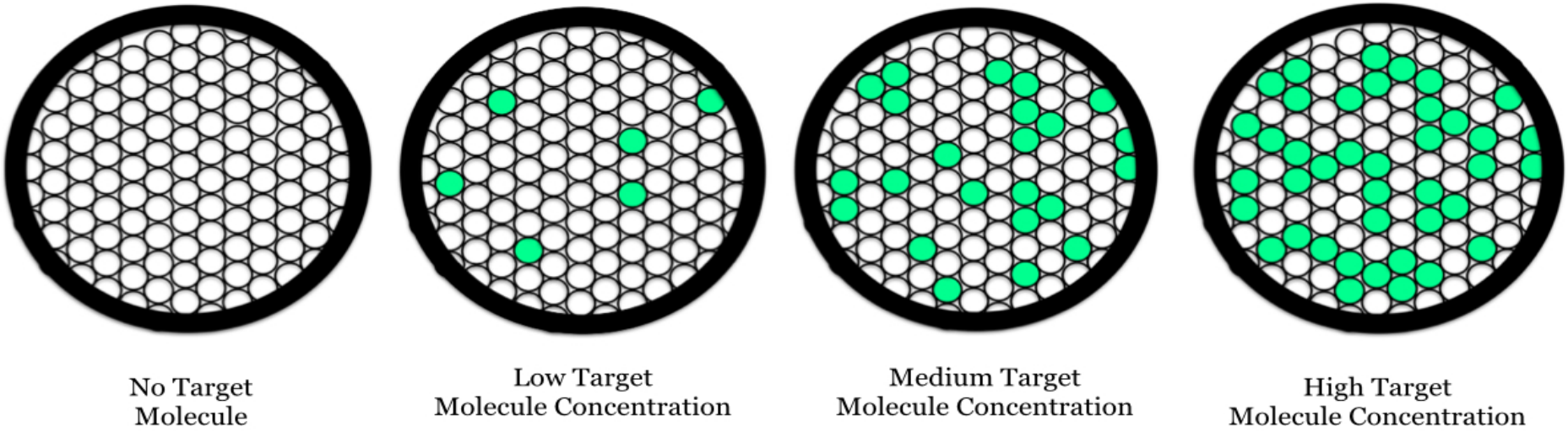








Revisiting digital PCR: standard way of estimating fraction of positive droplets



The standard method of interpreting digital PCR data calculates the expected fraction of positive droplets as

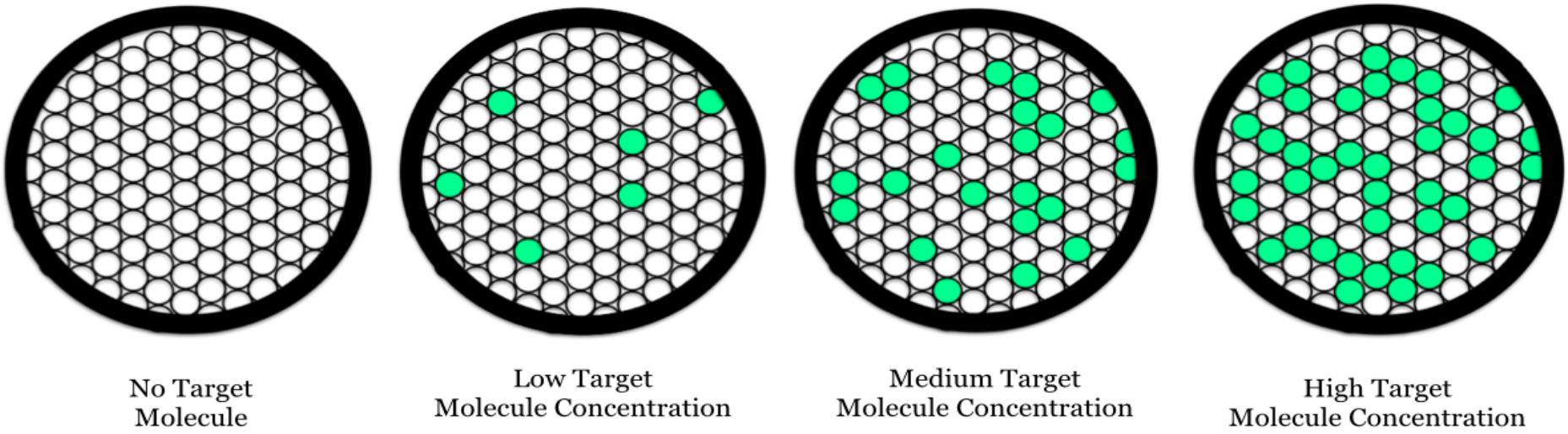
$$\hat{f} = 1 - e^{-\lambda}, \tag{20}$$

from which λ is estimated as

$$\hat{\lambda} = -\ln(1 - \hat{f}).$$

AIMS African Institute for Mathematical Sciences NEXT EINSTEIN INITIATIVE

Revisiting digital PCR: a new way of estimating fraction of positive droplets

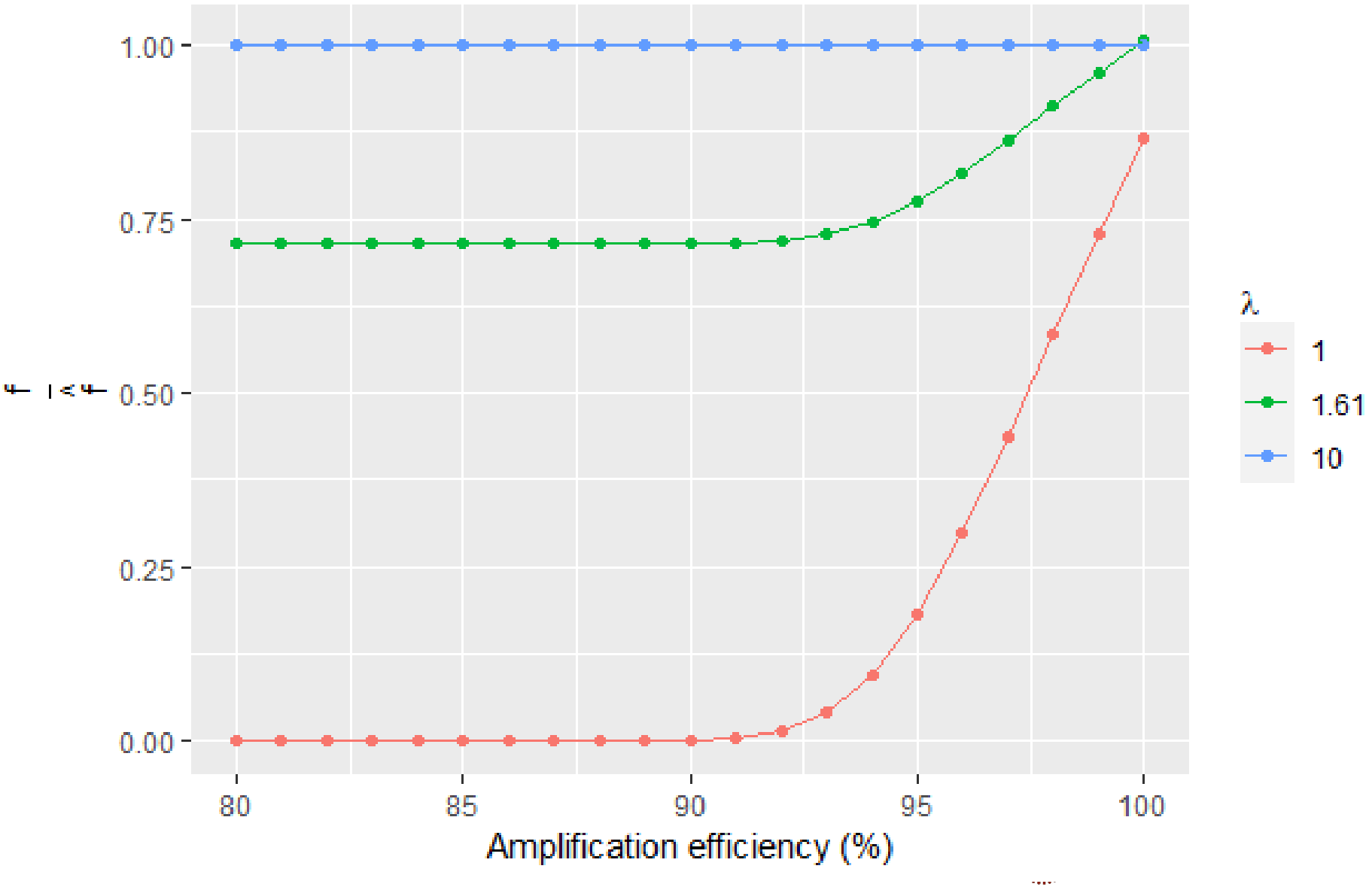


However, the model described here indicates that a much more accurate expression for the expected fraction of positive droplets is:

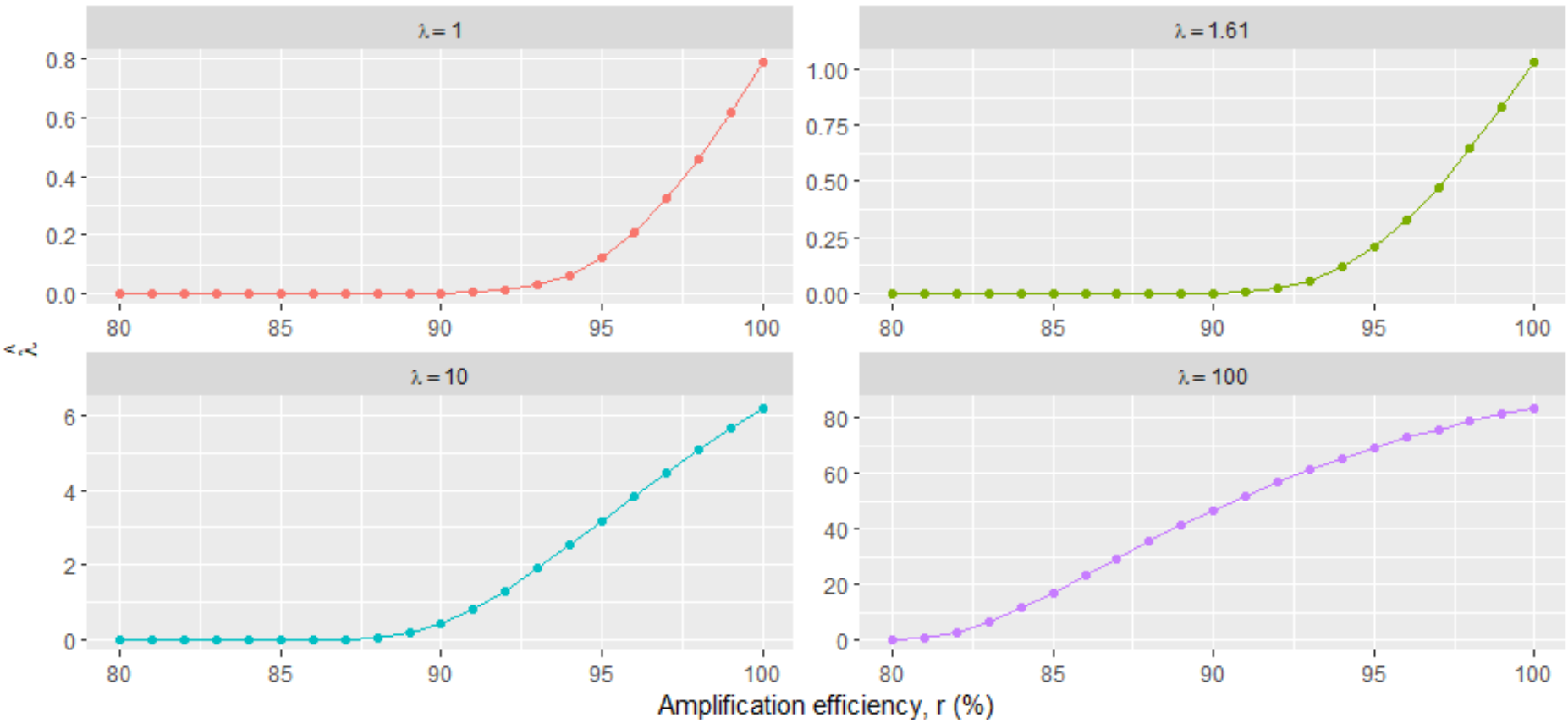
$$\hat{f} = F(T) = 1 - \frac{\sum_{i=1}^{x} \frac{\binom{x}{i}}{(i-1)!} \lambda^i B_{e^{-r_1 T}}(i, x-i+1)}{e^{\lambda} - 1}, \tag{21}$$

where T is the maximum practical duration of PCR.

Standard way over-estimates fraction of positive droplets



Over-estimation of fraction of positive droplets leads to under-estimation of λ



Limit of detection (LoD) and limit of quantification (LoQ)

LoD and LoQ are two of the most important operating characteristics of a PCR process.

Both LoD and LoQ are frequently estimated by using *ad-hoc* mathematical techniques, eg.

- LoD estimated based on receiver-operator-characteristic curves
- LoQ estimated as value of λ for which the coefficient of variation of the Ct value exceeds an arbitrary threshold

Our model permits the development and execution of mathematically precise statements of the estimation problem.

LoD

Definition

The LoD is the smallest number of molecules that can be detected with a failure rate not exceeding a threshold α. Specifically,

$$\begin{aligned} \text{LoD} = \quad & \textbf{min } \lambda \\ & \text{s.t. } F(T|\lambda, \vec{r}, \vec{s}, x) > 1 - \alpha, \end{aligned} \tag{22}$$

where T is the maximum practical duration of PCR.

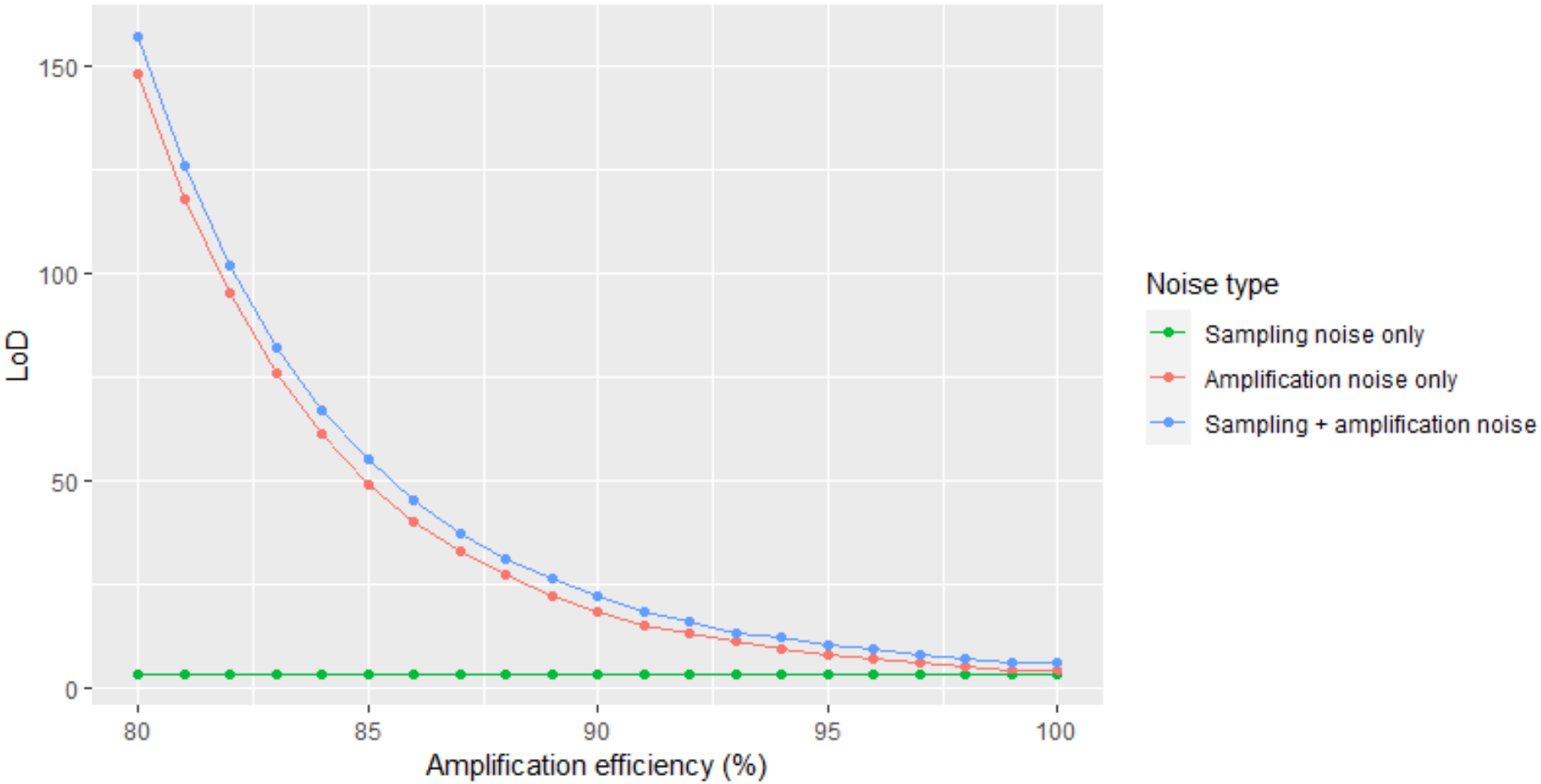
LoQ

Definition

Suppose that a Ct value t, generated by some PCR process, is used to obtain an estimate, denoted $\hat{\lambda}$, of λ. Let $P\left(\lambda/\beta \leq \hat{\lambda} \leq \beta\lambda \mid \lambda, \vec{r}, \vec{\tau}, x\right)$ denote the probability that, for any data t generated by the same process, $\hat{\lambda}$ will not differ from λ by more than a factor $\beta, \beta \geq 1$. We define the LoQ as

$$\begin{aligned} \text{LoQ} = & \quad \textbf{min } \lambda \\ & \text{s.t. } P\left(\lambda/\beta \leq \hat{\lambda} \leq \beta\lambda \mid \lambda, \vec{r}, \vec{\tau}, x\right) > 1 - \alpha. \end{aligned} \tag{23}$$

Calculating the LoD: examples



AIMS African Institute for Mathematical Sciences NEXT EINSTEIN INITIATIVE

Summary

- I have highlighted the importance of first principles modelling
  - The starting point is a parsimonious set of propositions that is postulated to organize a system's parts to generate observations of interest
  - The propositions are instantiated in a model that permits assessment of their deductive validity
  - Further testing is done by using out-of-sample data
- I described how a first-principles model we developed for reverse-engineering PCR data allows to correct estimation errors produced by a purely statistical model

AIMS | African Institute for Mathematical Sciences NEXT EINSTEIN INITIATIVE

Next steps

- Use existing data to compare the accuracy and precision of the new model vs. existing models, which are mostly phenomenological
- Investigate the new model's power to inform the design of more efficient experiments
- Apply the new model to improve on the state of the art in pool testing

AIMS African Institute for Mathematical Sciences NEXT EINSTEIN INITIATIVE

Collaborators and funders

- Dr Abdoelnaser Degoot (co-author of work presented)
- Prof Neil Turok (collaborator)
- International Development Research Centre (funder)
- Carnegie Corporation of New York (funder)

AIMS | African Institute for Mathematical Sciences NEXT EINSTEIN INITIATIVE

Thank you!

wndifon@aims.ac.za

AIMS African Institute for Mathematical Sciences NEXT EINSTEIN INITIATIVE