

Bayesian inference methods for ecology

Theory

What is Bayesian inference?

Probability 101

Bayes' theorem in biology

Priors and likelihood

Interpretation of posteriors

Markov Chain Monte Carlo

Applications

Population dynamics in
microscopic and
macroscopic communities

Phylogeny and model
selection

ODE generative models:
Morphogen patterning of
embryonic tissues

Example 1.1: Determining a proportion

Hypothesis testing What about p-values?

(AKA, the dangers of p-values, and how Bayesian inference comes to the rescue)

Question: Is proliferation more probable than cell death?

P-value strategy:

Set null hypothesis: Rate proliferation same as death $p = 1/2$

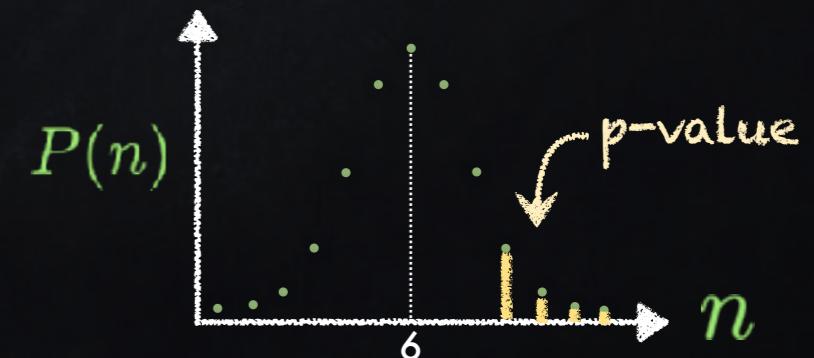
Perform experiment: Measure $N = 12$ events and count number of proliferations
For instance in this case you measured $n = 9$ proliferations

Calculate probability of observation:

Binomial distribution $P(n) = \binom{N}{n} p^n (1-p)^{N-n}$

calculate p-value:
(deviation from null)

Probability the observation or more extreme values



$$P(n = 9) + P(n = 10) + P(n = 11) + P(n = 12) \simeq 0.073$$

Evaluate significance: P-value is > 0.05 therefore we can't reject the null hypothesis

Example 1.1: Determining a proportion

Hypothesis testing What about p-values?

(AKA, the dangers of p-values, and how Bayesian inference comes to the rescue)

Question: Is proliferation more probable than cell death?

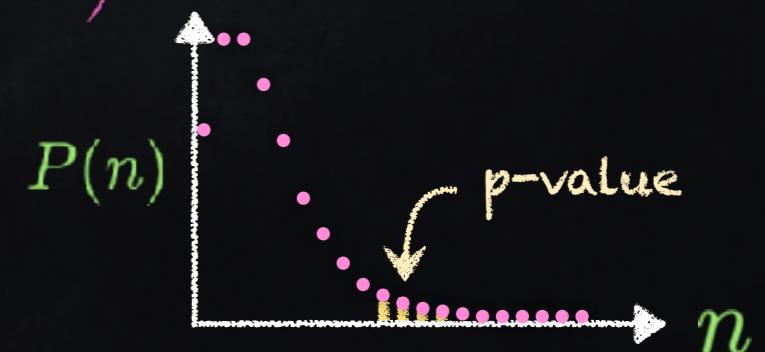
P-value strategy:

Set null hypothesis: Rate proliferation same as death $p = 1/2$

Perform experiment: Observe cells until $d = 3$ deaths are detected,
for instance $n=9$ proliferations were observed before $d = 3$
(different intention but same data as previous example!)

Calculate probability of observation:
Negative binomial $P(n) = \binom{n+d-1}{n} p^n (1-p)^d$

calculate p-value:
(deviation from null)
Probability the observation
or more extreme values



$$P(n = 9) + P(n = 10) + P(n = 11) + \dots = 0.033$$

Evaluate significance: P-value is < 0.05 therefore we can reject the null hypothesis

Example 1.1: Determining a proportion

Hypothesis testing What about p-values?

(AKA, the dangers of p-values, and how Bayesian inference comes to the rescue)

Question: Is proliferation more probable than cell death?

Bayesian strategy:

Set null hypothesis:

Perform experiment:

Measure $N = 12$ events
and count number of
proliferations $n = 9$

Observe cells until $d = 3$ deaths
 $n=9$ proliferations were observed

Calculate likelihood
of observation:

Binomial distribution

$$P(n|p) = \binom{N}{n} p^n (1-p)^{N-n}$$

Negative binomial

$$P(n|p) = \binom{n+d-1}{n} p^n (1-p)^d$$

Calculate posterior
distribution of p

$$P(p|n=9) \propto P(n=9|p) = Cp^9(1-p)^3$$

Uninformative prior

Posterior is independent of the observatory intention (as it should be)

Calculate HDR:

$$\text{HDR}(95\%) = [0.091, 0.537]$$

Contains 0.5 so $p=0.5$ is
a credible value!

Example 1.1: Determining a proportion

Hypothesis testing What about p-values?

(AKA, the dangers of p-values, and how Bayesian inference comes to the rescue)

Question: Is proliferation more probable than cell death?

Bayesian strategy:

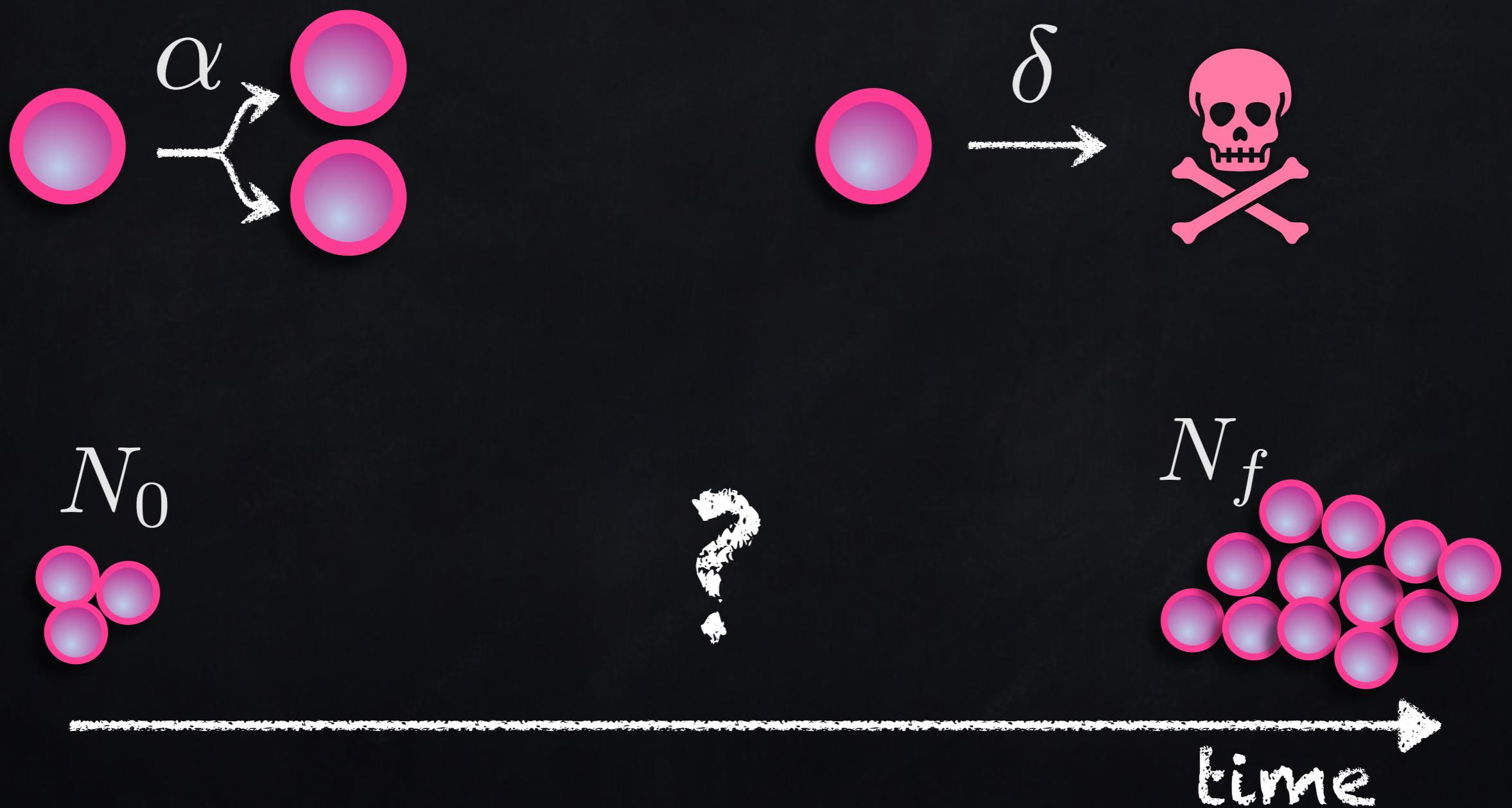


Bayesian inference only depends on the data, not on the observer intention



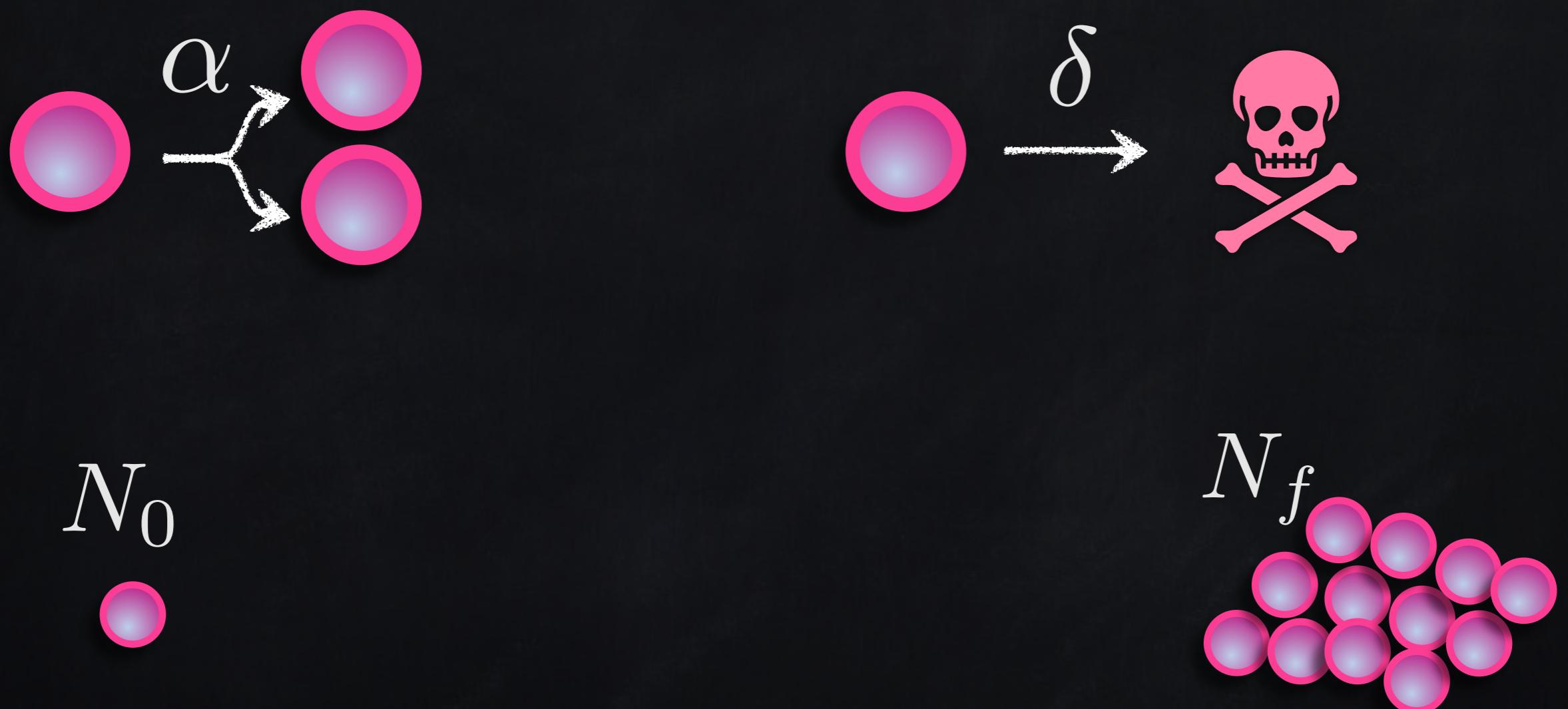
Hypothesis testing does not affect the inference process, this allows for multiple hypothesis testing

Example 1.2: Population growth



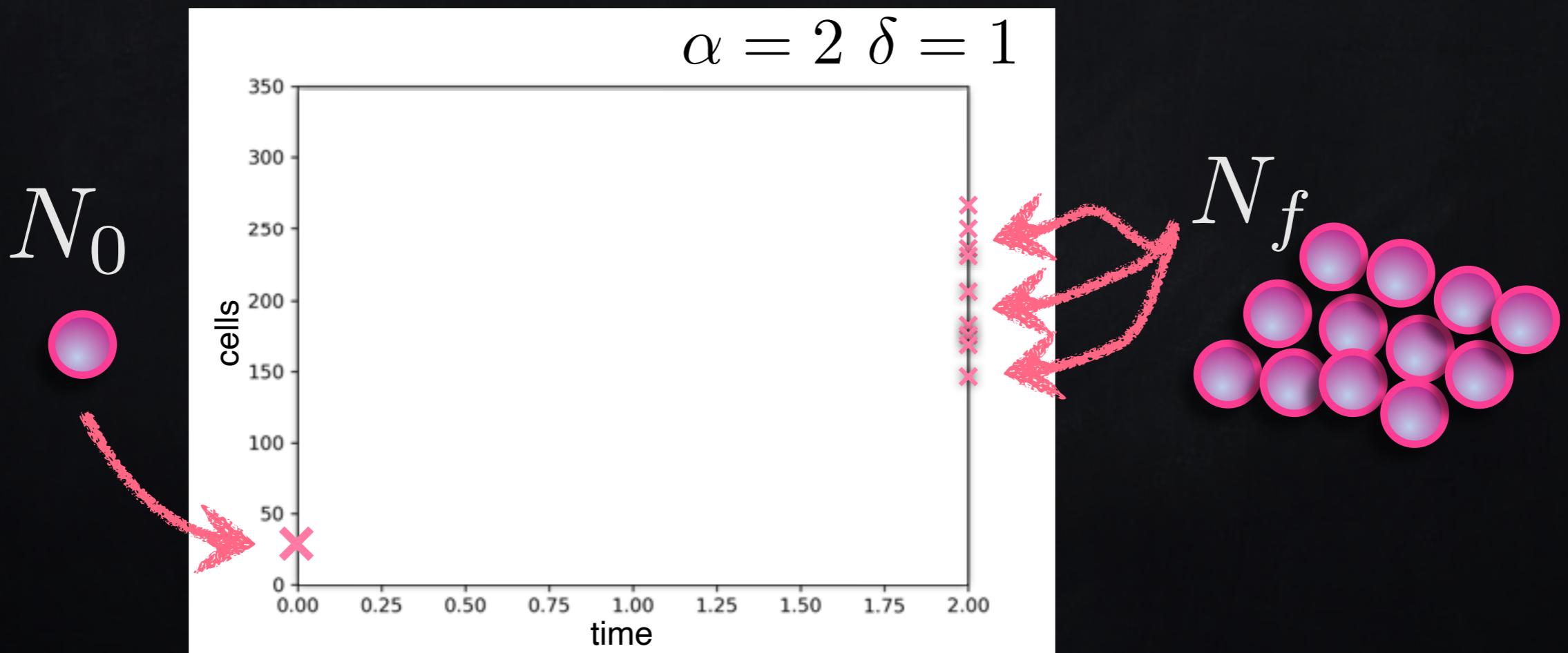
How do I infer the values of α and δ ?

Example 1.2: Population growth



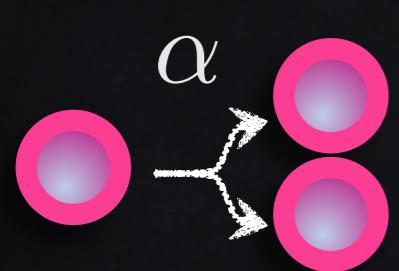
How do I infer the values of α and δ ?

Example 1.2: Population growth

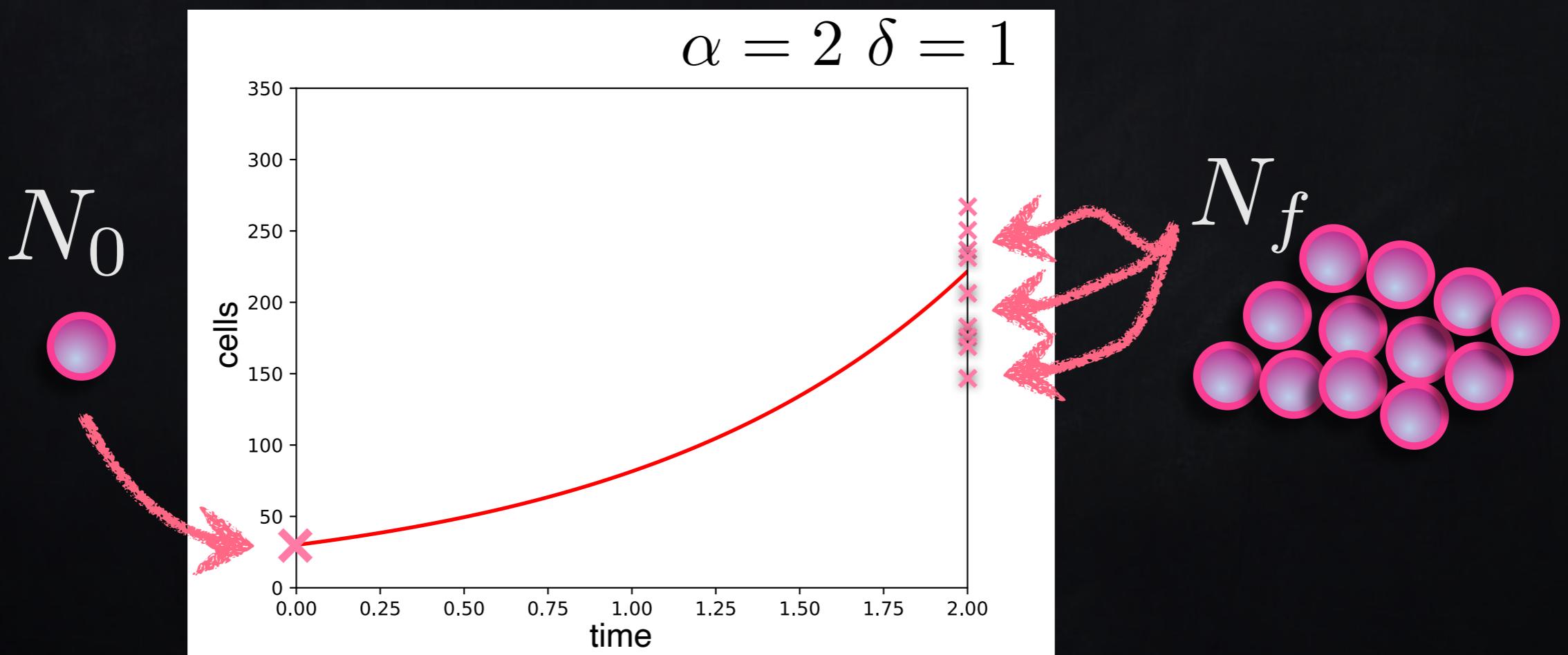


How do I infer the values of α and δ ?

Example 1.2: Population growth



$$\frac{dN(t)}{dt} = \alpha N - \delta N$$



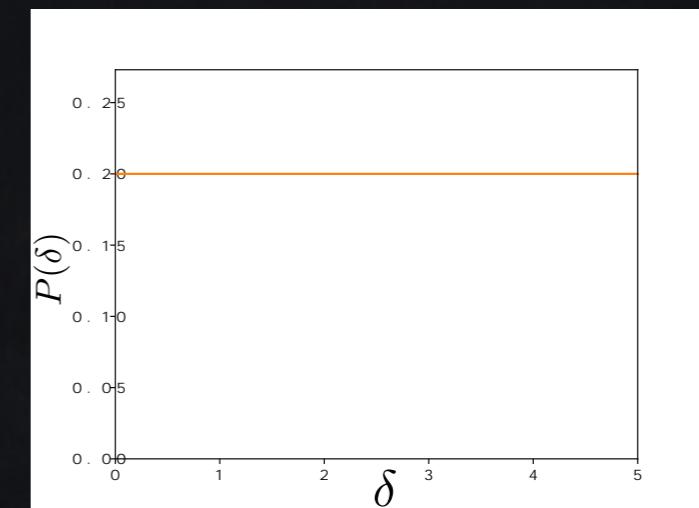
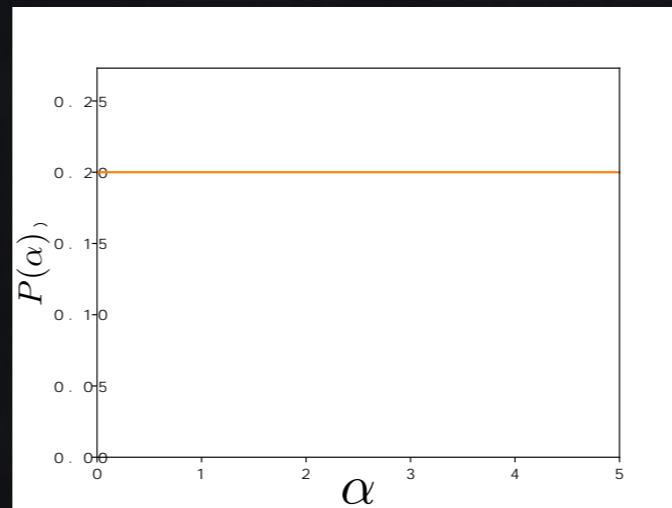
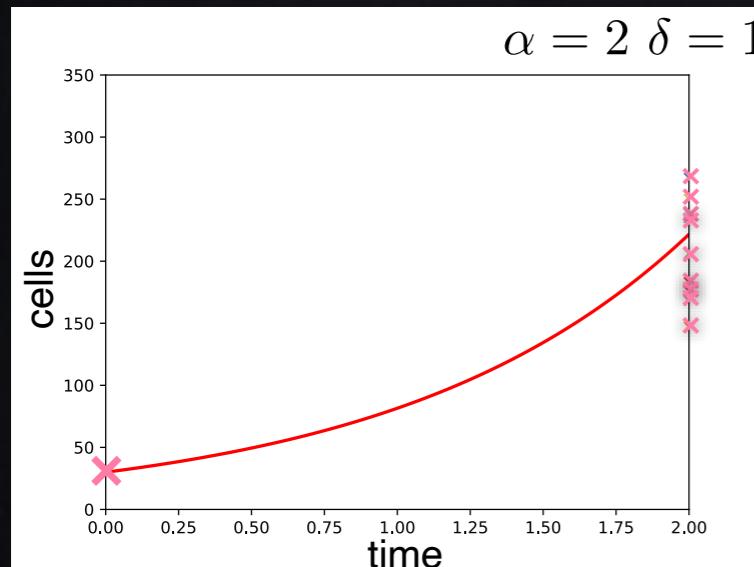
How do I infer the values of α and δ ?

Example 1.2: Population growth



$$\frac{dN(t)}{dt} = \alpha N - \delta N$$

Prior distributions

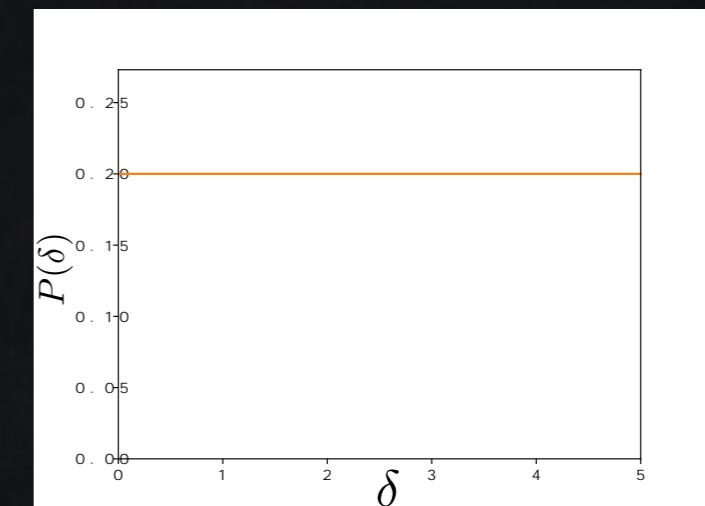
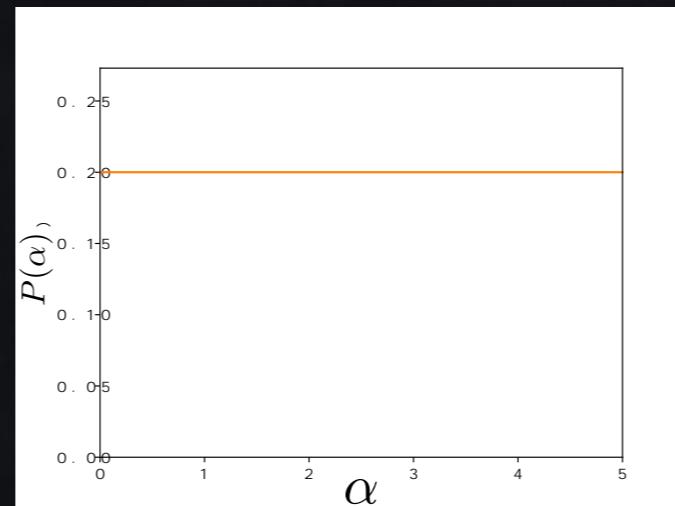
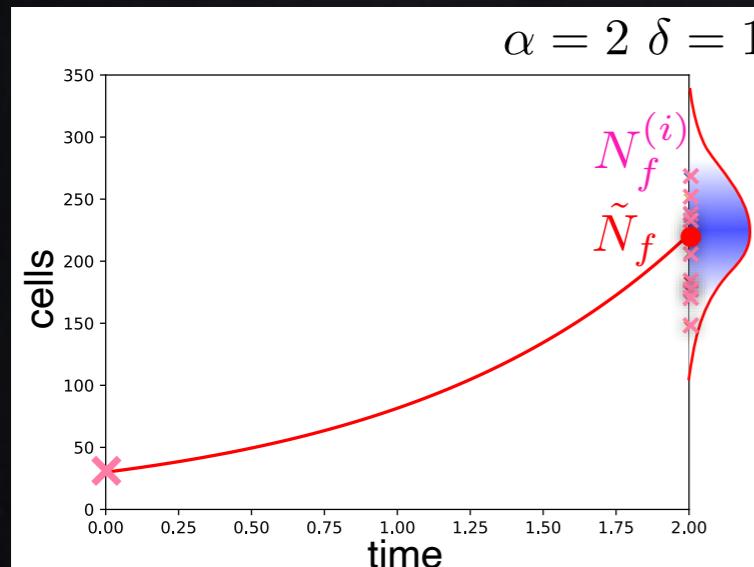


Example 1.2: Population growth



$$\frac{dN(t)}{dt} = \alpha N - \delta N$$

Prior distributions



Likelihood



$$P(\text{data}|\alpha, \delta) = \prod_{i=1}^N P(\text{datapoint } i|\alpha, \delta)$$

approximate likelihood distance



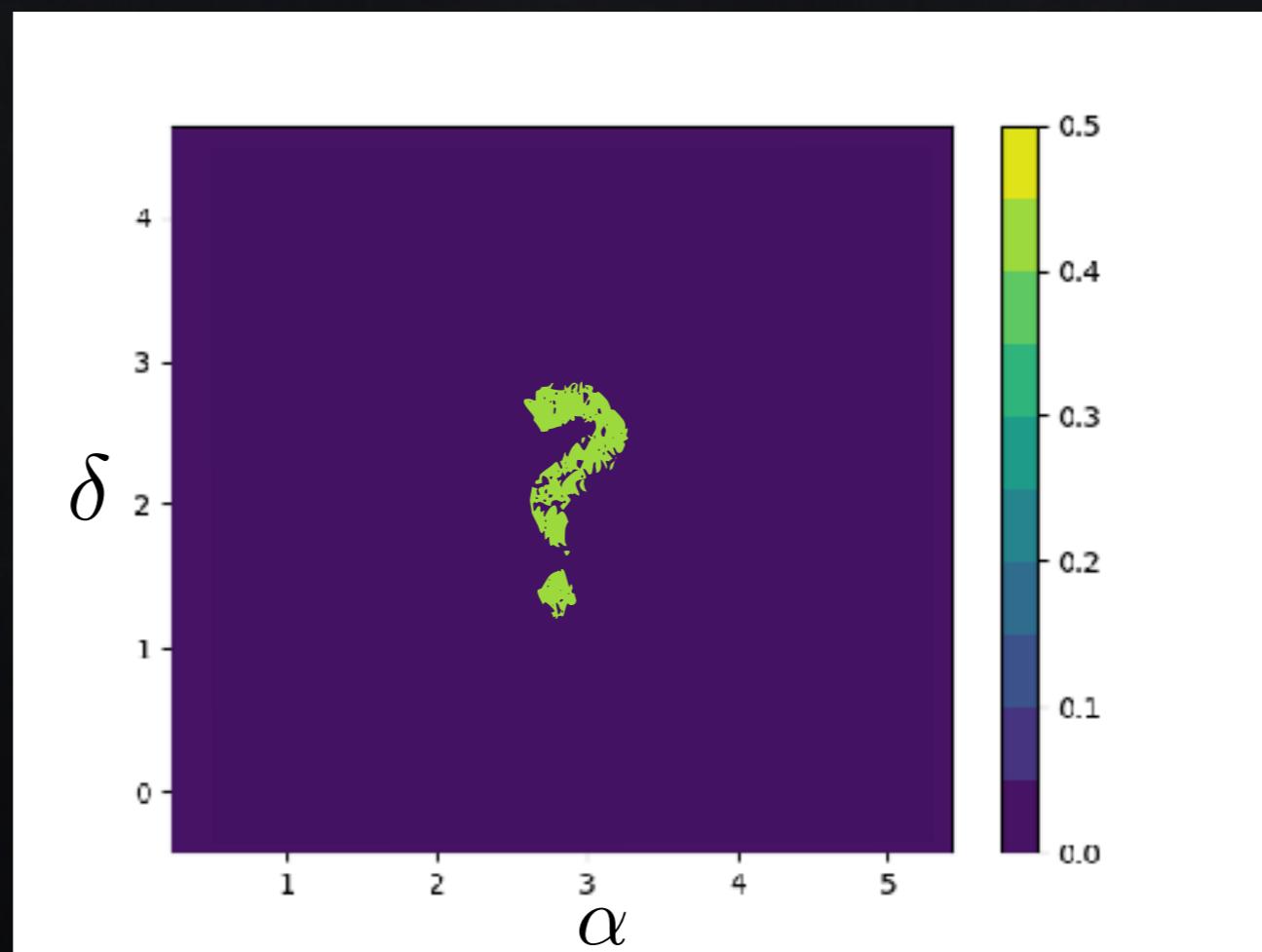
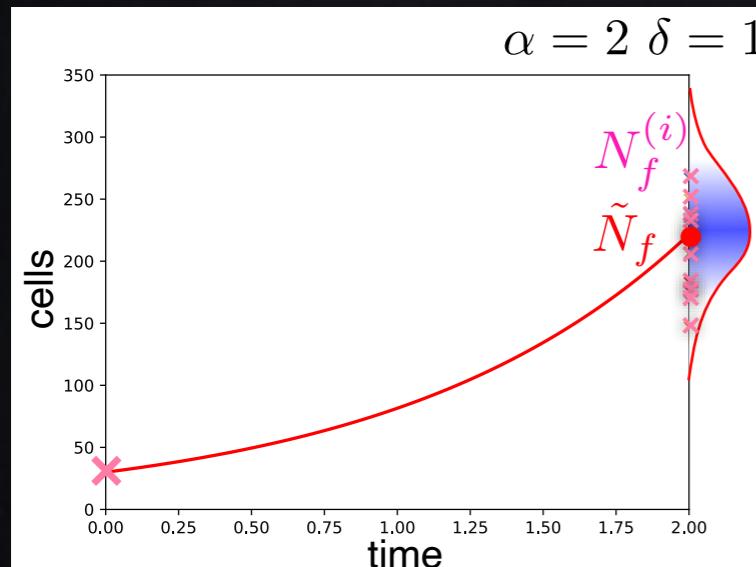
$$\prod_{i=1}^N \mathcal{N}(\tilde{N}_f - \tilde{N}_f^{(i)}, \sigma)$$

Example 1.2: Population growth



$$\frac{dN(t)}{dt} = \alpha N - \delta N$$

Posterior distributions

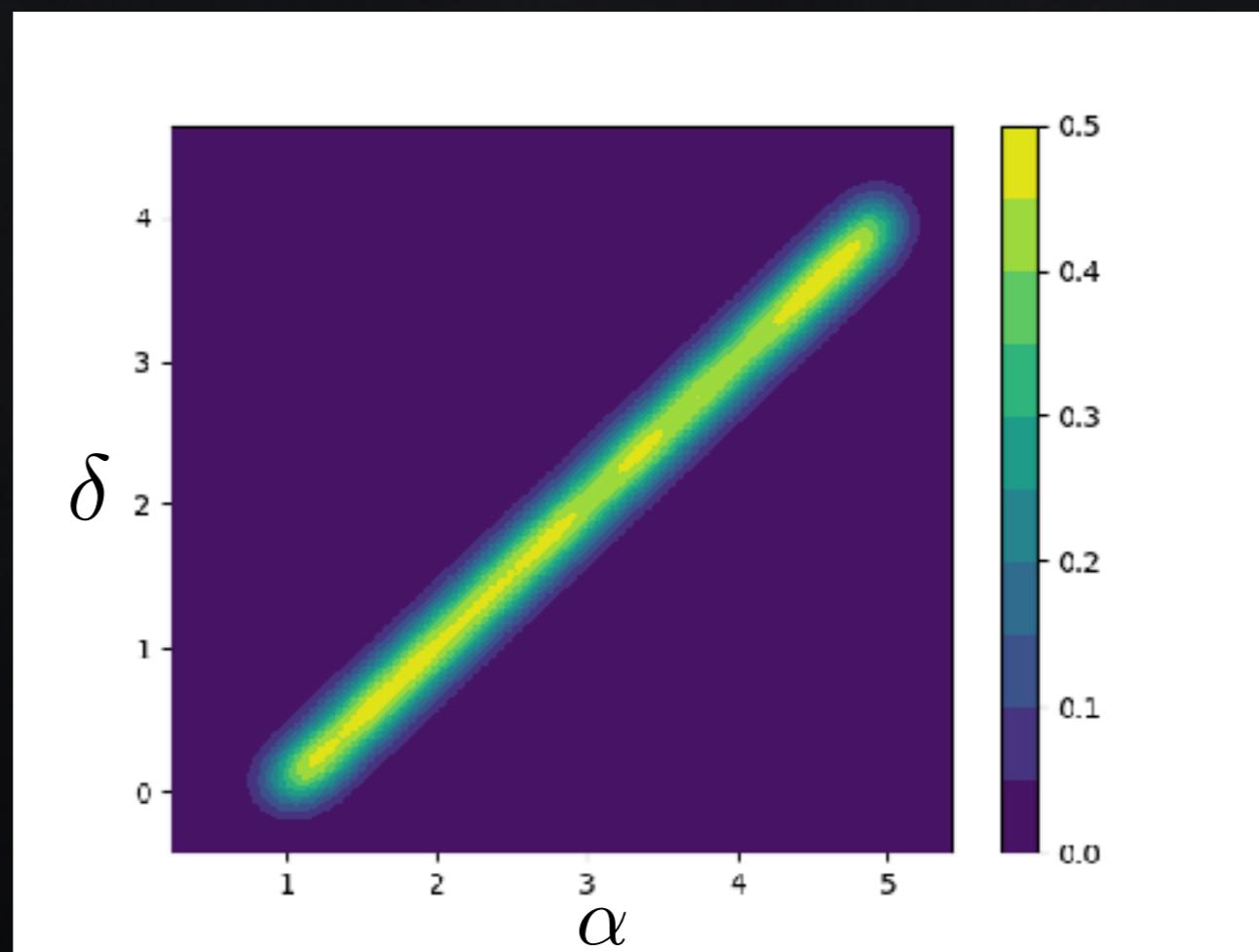
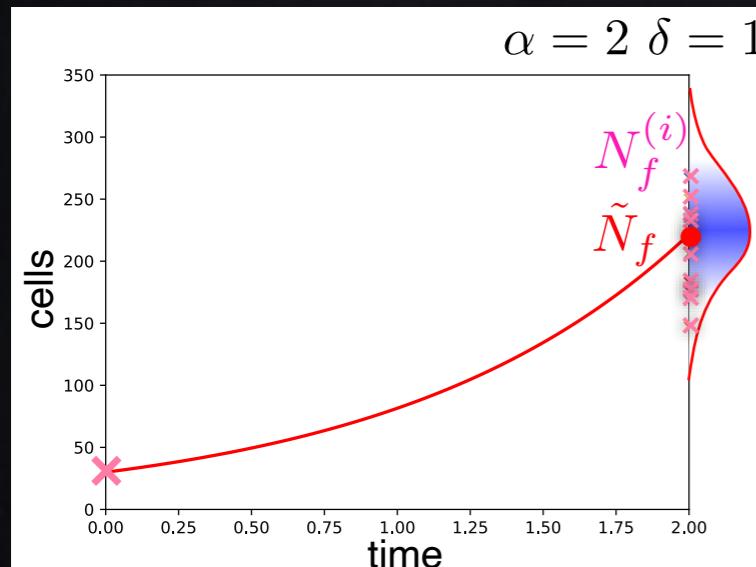


Example 1.2: Population growth



$$\frac{dN(t)}{dt} = \alpha N - \delta N$$

Posterior distributions

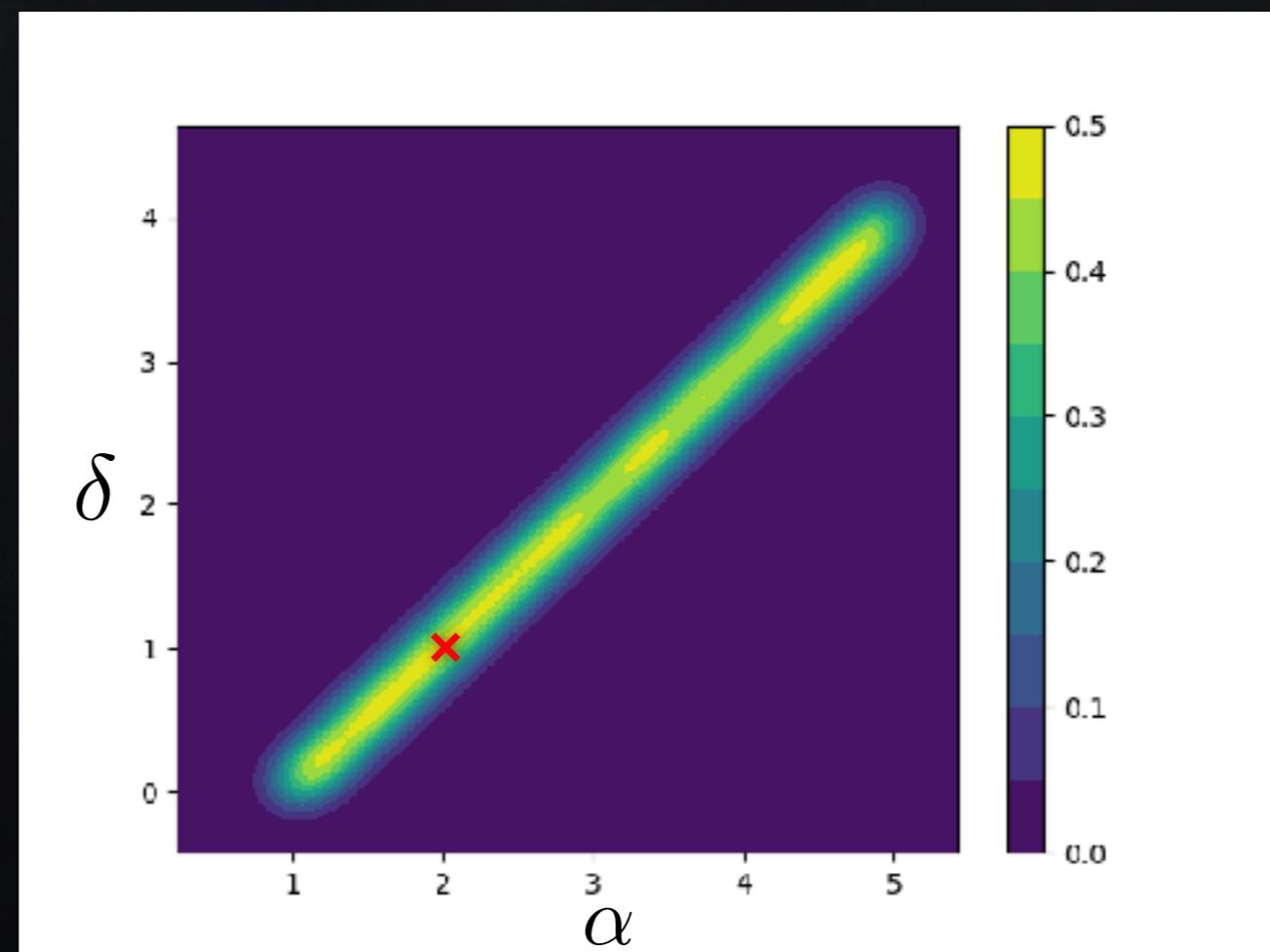
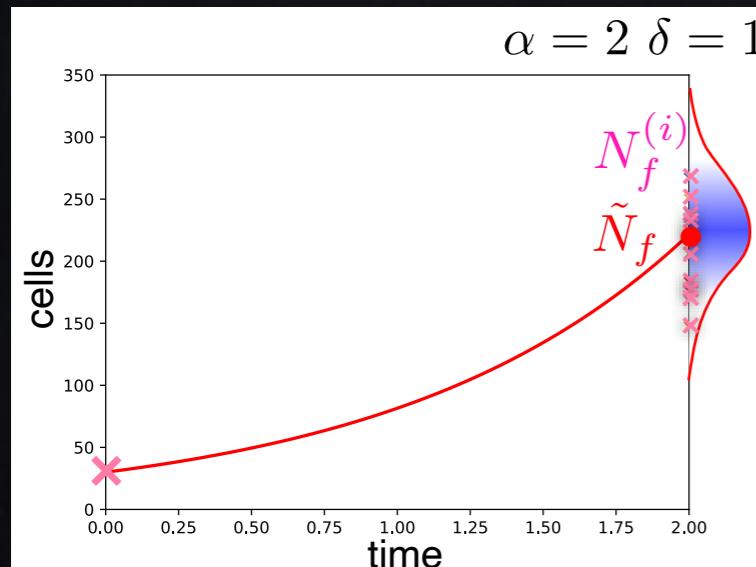


Example 1.2: Population growth



$$\frac{dN(t)}{dt} = \alpha N - \delta N = (\alpha - \delta)N$$

Posterior distributions

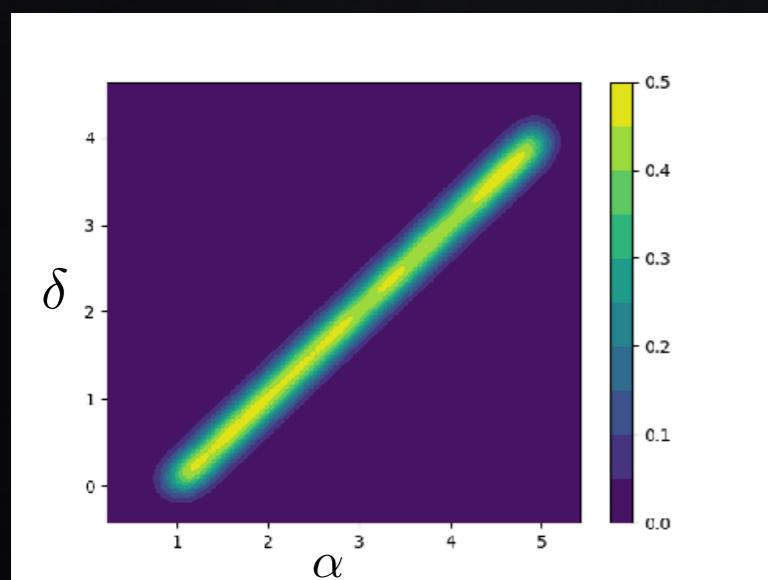
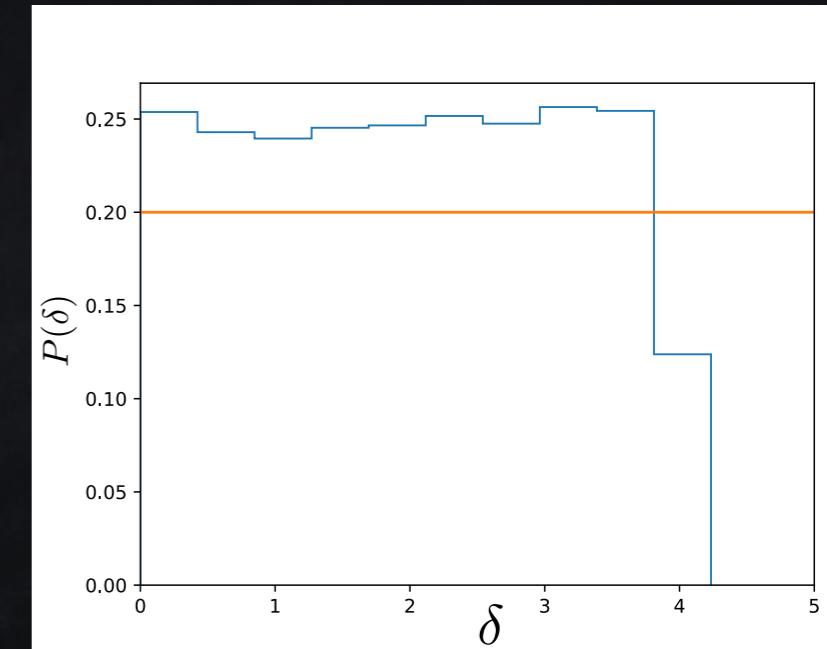
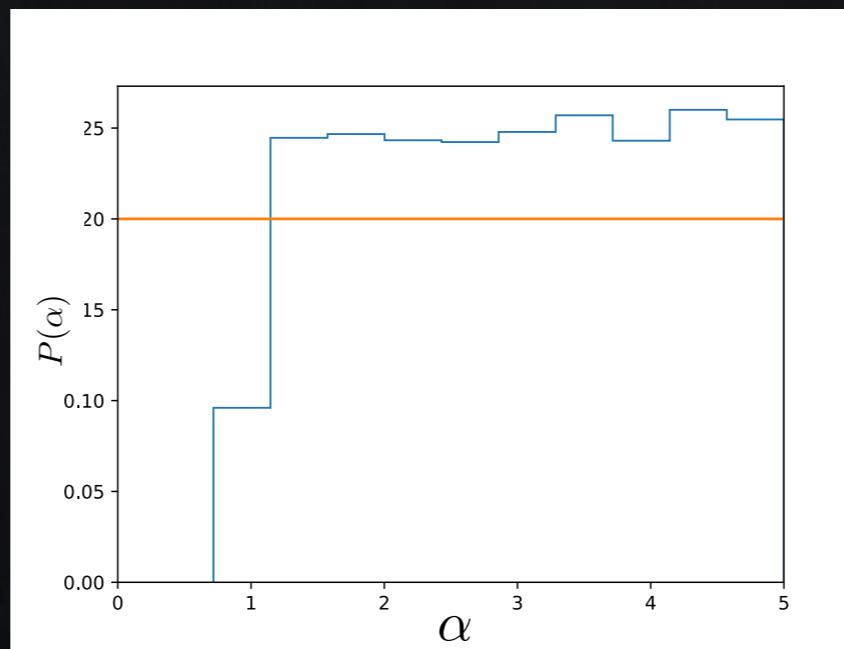
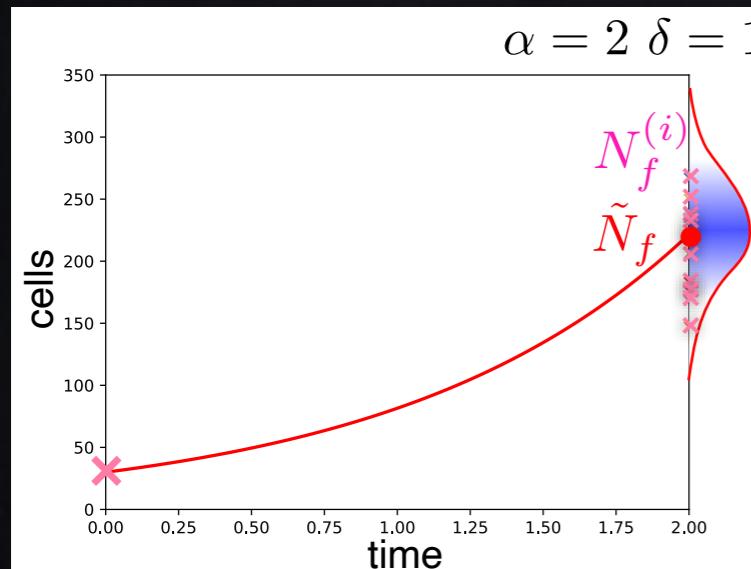


Example 1.2: Population growth



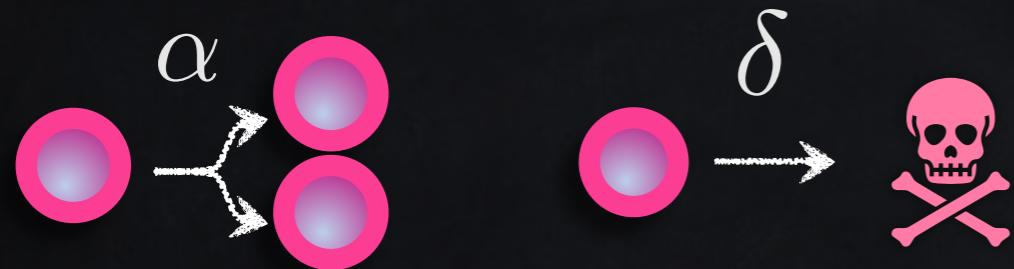
$$\frac{dN(t)}{dt} = \alpha N - \delta N = (\alpha - \delta)N$$

Posterior distributions



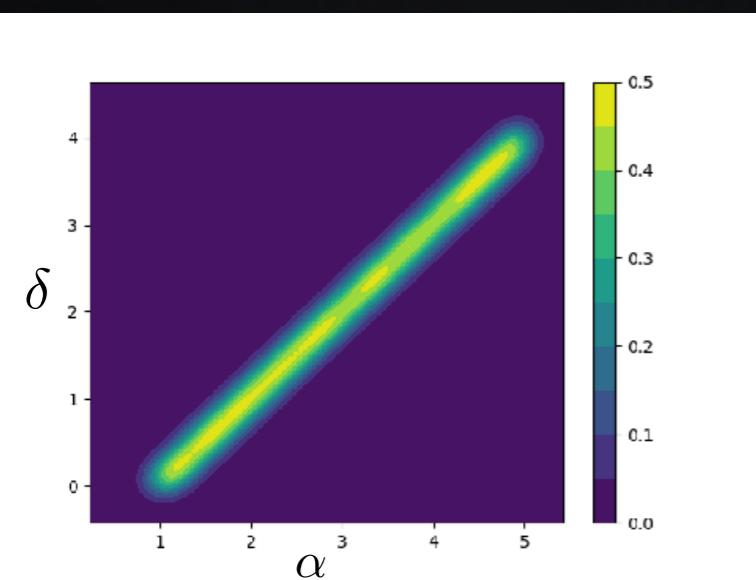
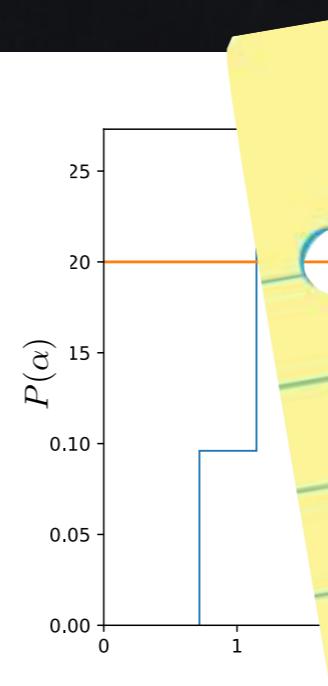
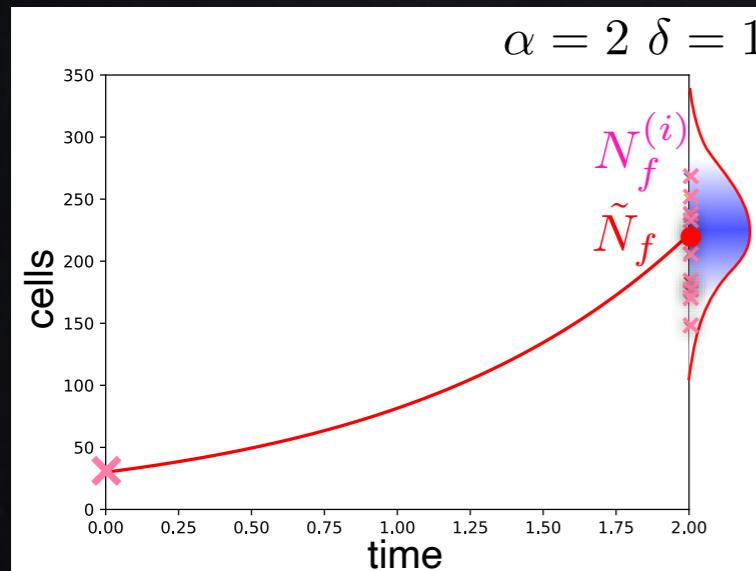
Bayesian inference yields how much information can we extract from the data with a model

Example 1.2: Population growth



$$\frac{dN(t)}{dt} = \alpha N - \delta N = (\alpha - \delta)N$$

Posterior distributions

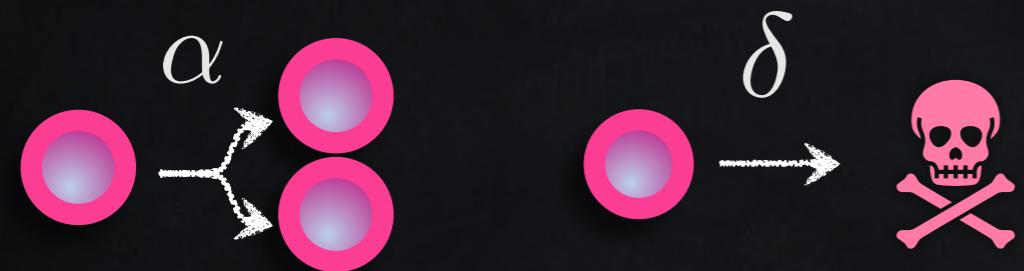


Bayesian inference
we extract from

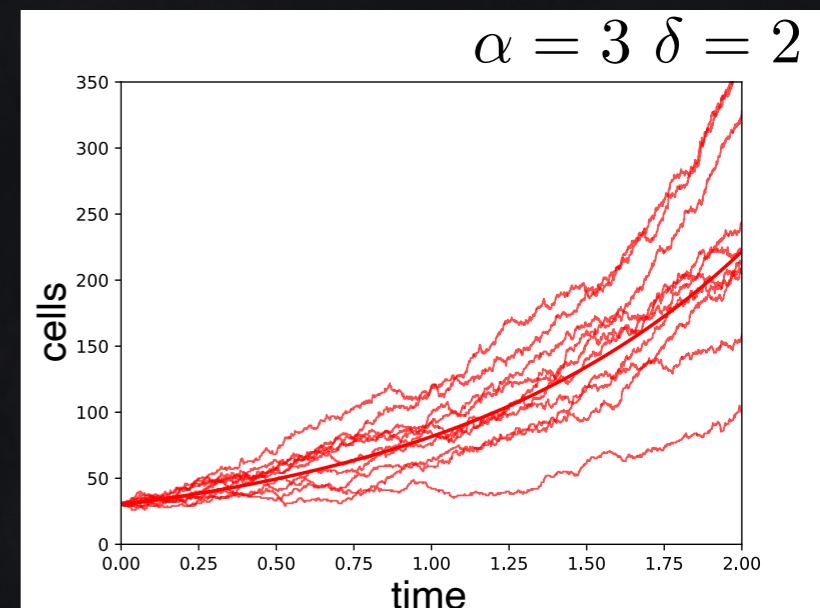
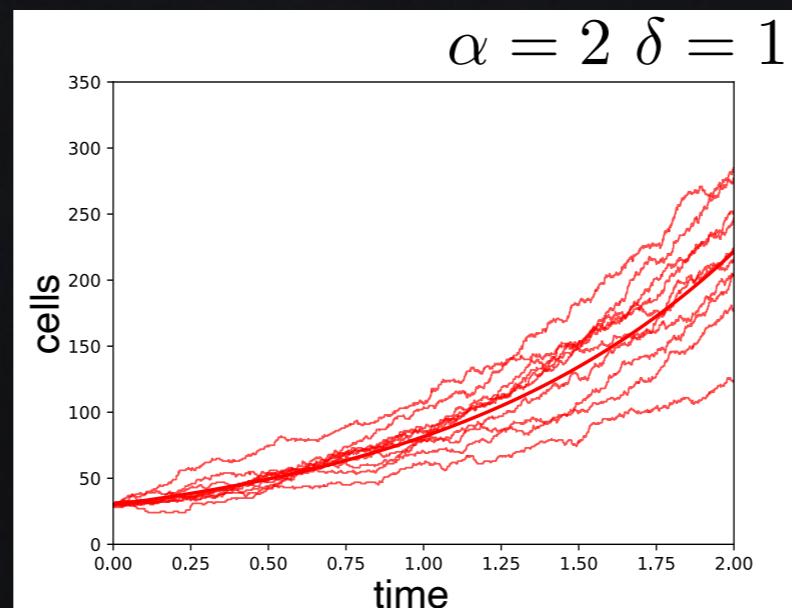
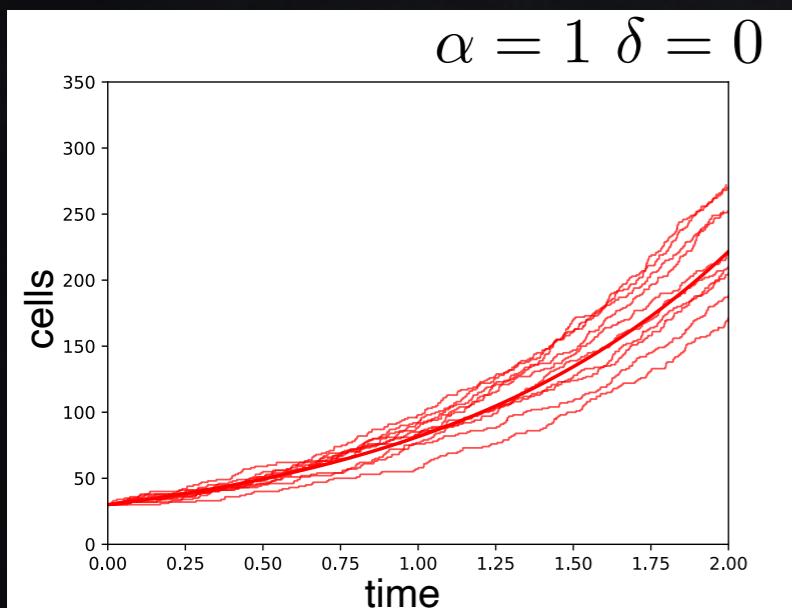
Question:

- would this result change if we had a Gaussian prior distribution?
- what experiment would you propose to obtain narrower posteriors of the rates?

Example 1.2: Population growth



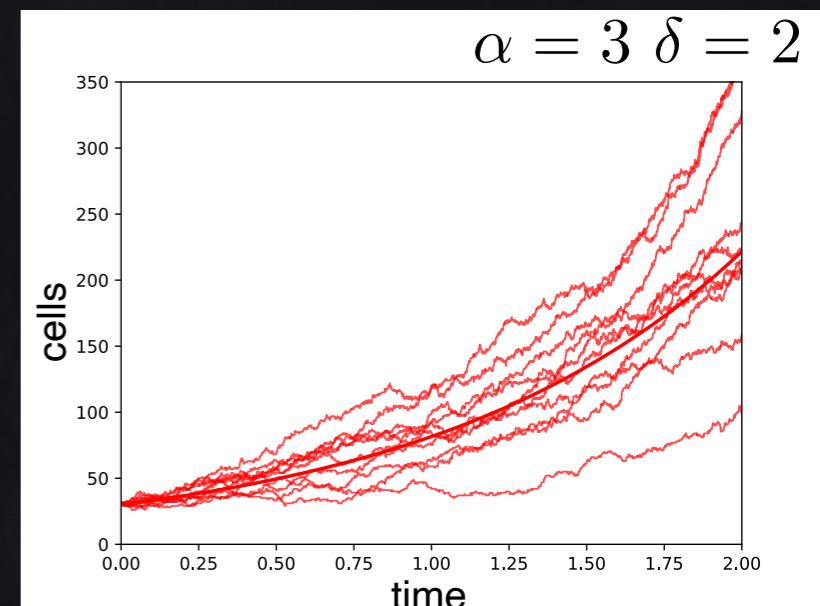
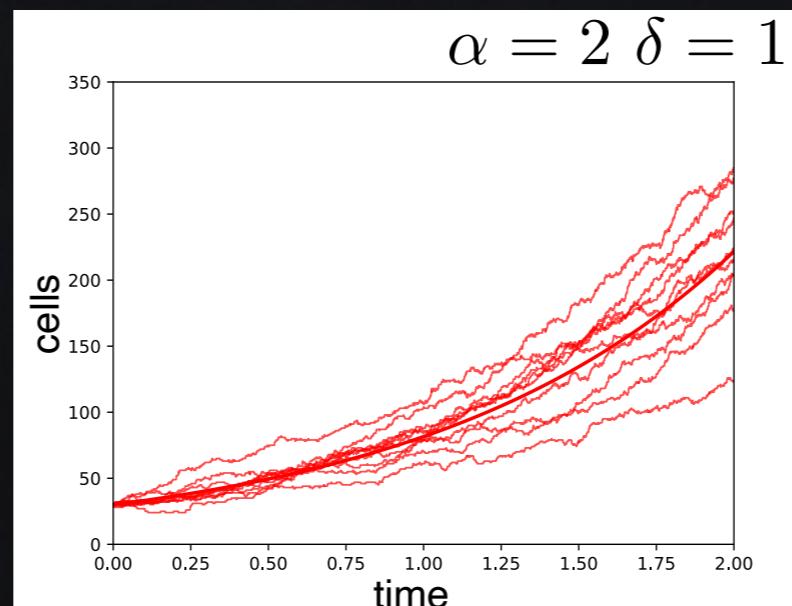
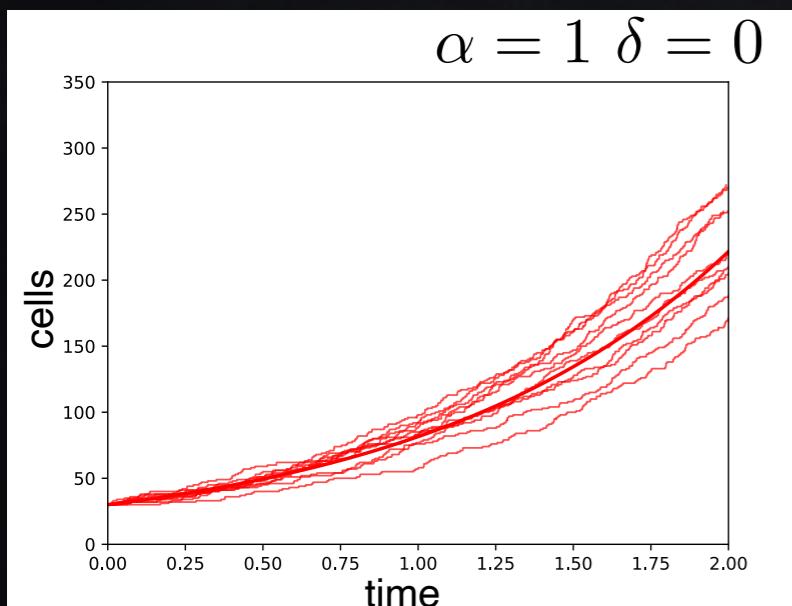
$$\frac{dN(t)}{dt} = \alpha N - \delta N = (\alpha - \delta)N$$



Example 1.2: Population growth



$$\frac{dN(t)}{dt} = \alpha N - \delta N = (\alpha - \delta)N$$

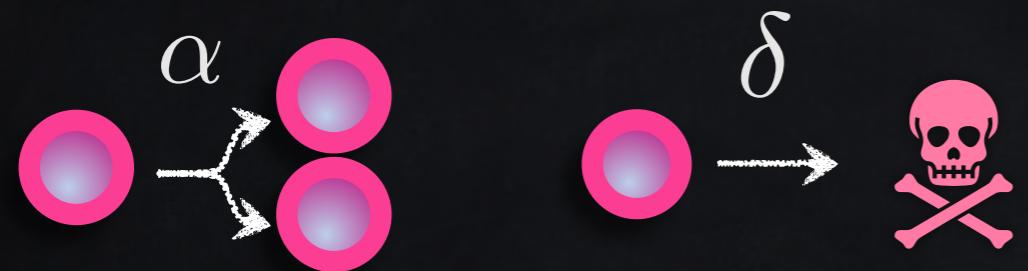


Stochastic theory says... $\sigma(t) = N_0 \frac{\alpha + \delta}{\alpha - \delta} e^{2(\alpha - \delta)t} \left(1 - e^{(\alpha - \delta)t} \right)$

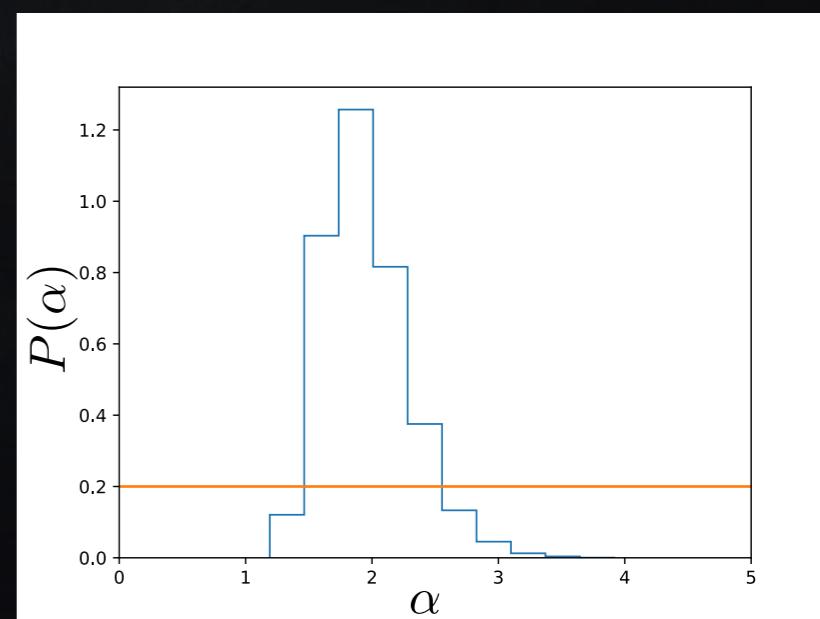
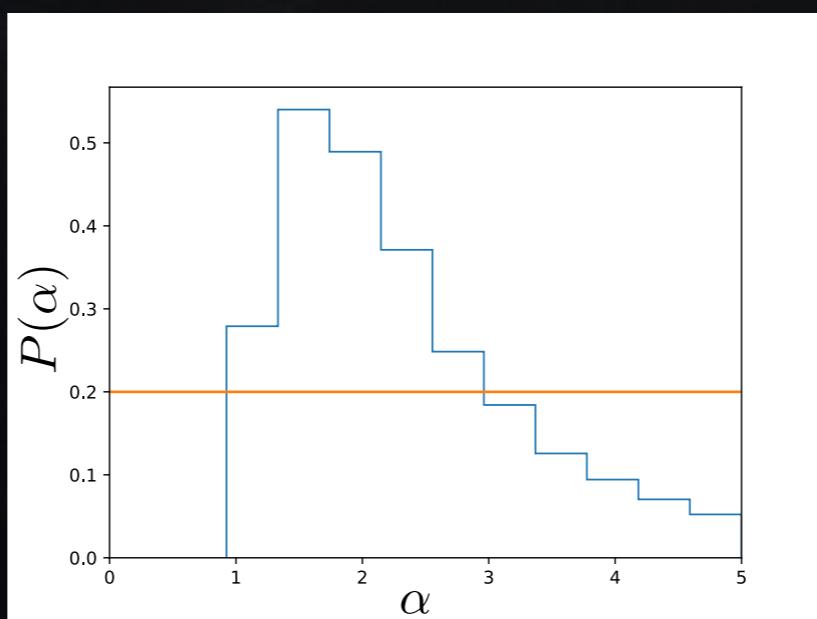
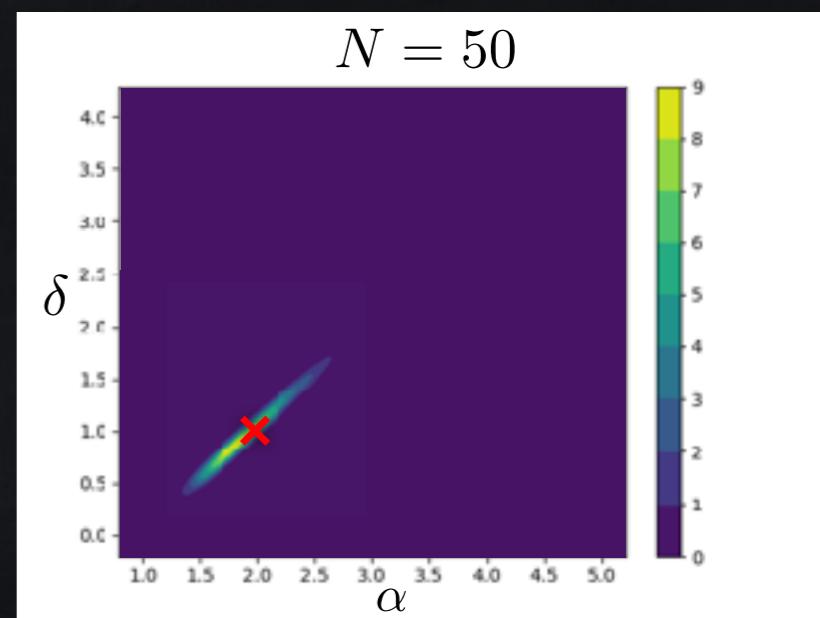
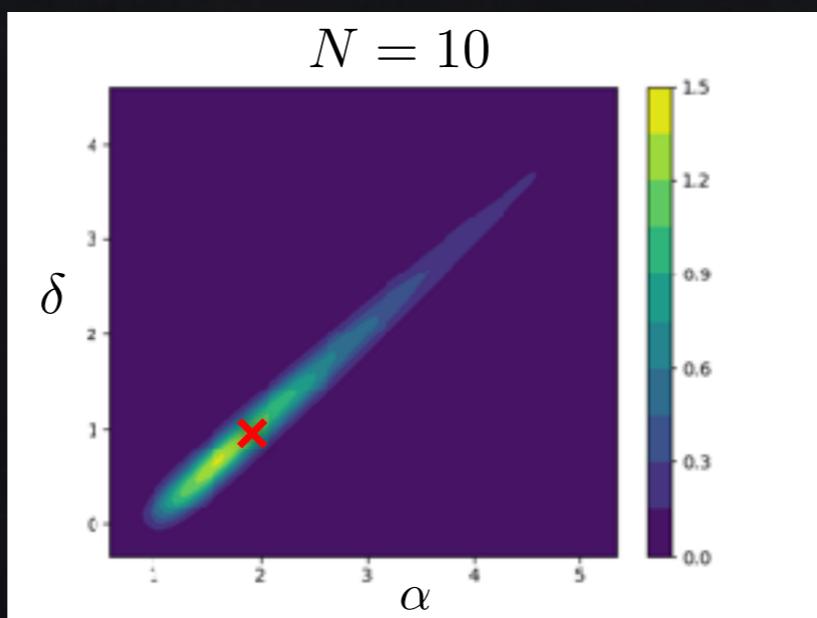
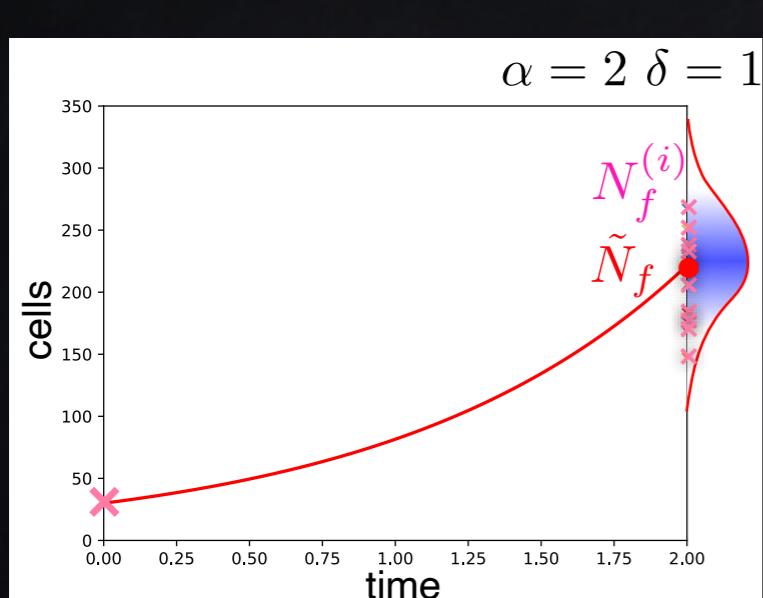
Likelihood

$$P(\text{data}|\alpha, \delta) = \prod_{i=1}^N P(\text{datapoint } i|\alpha, \delta) = \prod_{i=1}^N \mathcal{N}(\tilde{N}_f - \tilde{N}_f^{(i)}, \sigma)$$

Example 1.2: Population growth



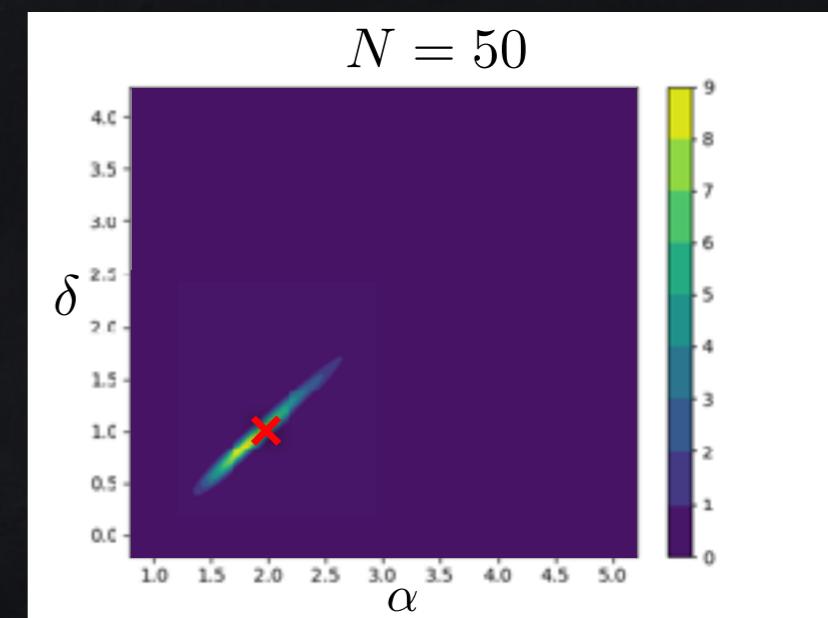
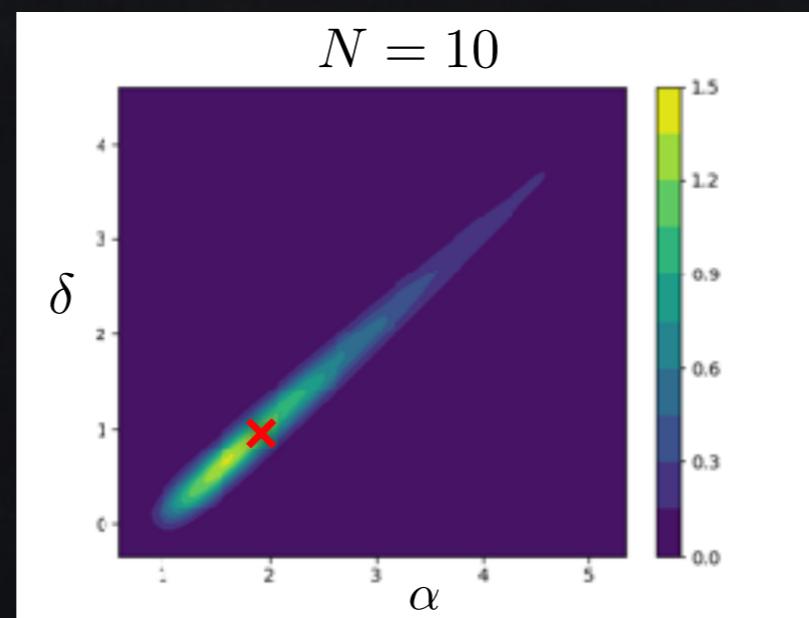
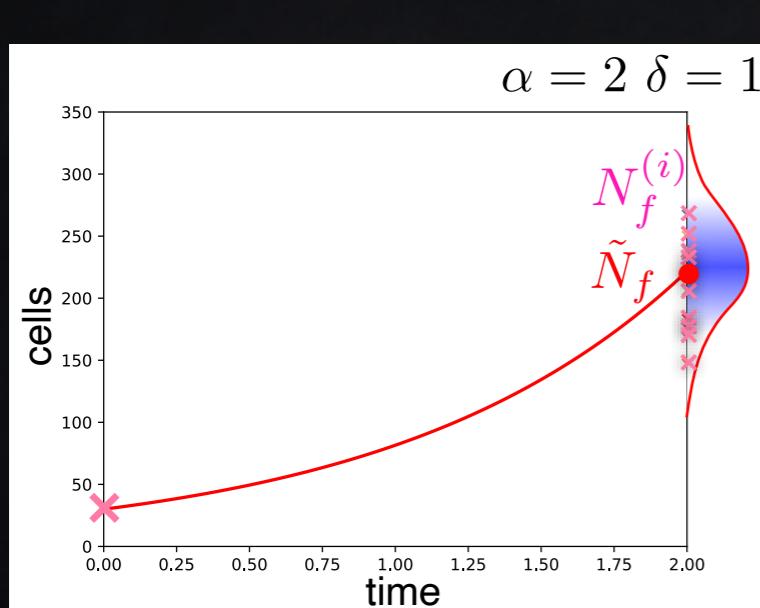
$$\frac{dN(t)}{dt} = \alpha N - \delta N = (\alpha - \delta)N$$



Example 1.2: Population growth



$$\frac{dN(t)}{dt} = \alpha N - \delta N = (\alpha - \delta)N$$



Bayesian inference returns how much information can we extract from the data with a model

Therefore it can help us to...

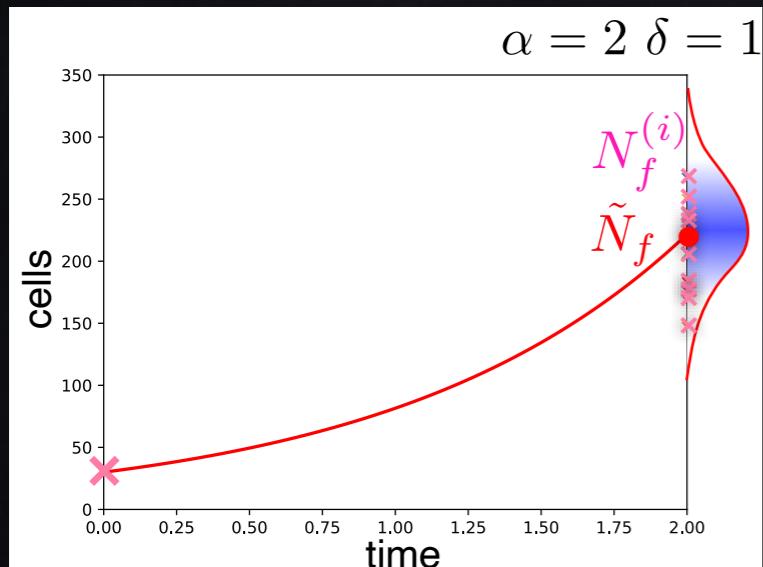


Design appropriate models e.g. include stochastic effects
Design better experiments e.g. number of replicates

Example 1.2: Population growth



$$\frac{dN(t)}{dt} = \alpha N - \delta N = (\alpha - \delta)N$$



Bayesian inference returns how much information can we extract from the data with a model

APPROXIMATE BAYESIAN COMPUTATION

Likelihood

$$\begin{aligned}
 P(\text{data}|\alpha, \delta) &= \prod_{i=1}^N P(\text{datapoint } i|\alpha, \delta) = \prod_{i=1}^N \mathcal{N}(\tilde{N}_f - \tilde{N}_f^{(i)}, \sigma) \\
 &= \prod_{i=1}^N \sum_{j=0}^{\min(N_f^{(i)}, N_0)} \binom{N_0}{j} \binom{N_0 + N_f^{(i)} - j - 1}{N_0 - 1} \alpha^{N_0 - j} \beta^{N_f^{(i)} - j} (1 - \alpha - \beta)^j
 \end{aligned}$$

✓ approximate
(Likelihood - free)

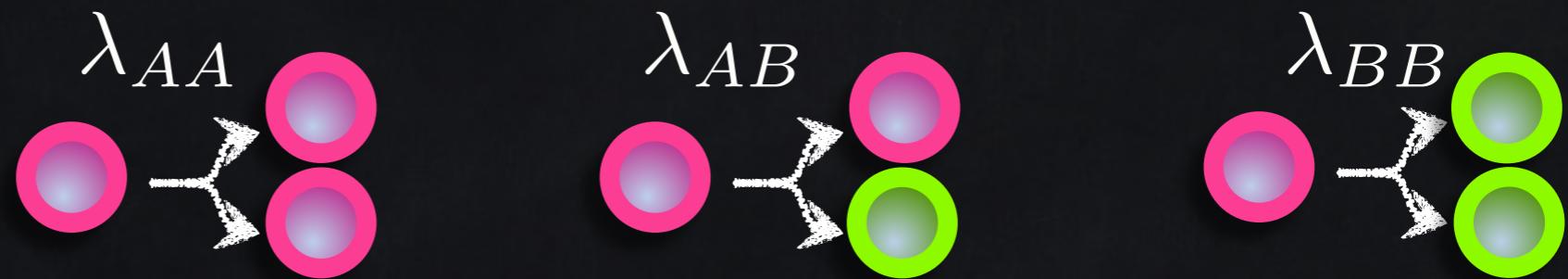
exact

Beyond

Example 1: Population growth

 Pluripotent

 Differentiated



Physical Biology

ACCEPTED MANUSCRIPT • OPEN ACCESS

Coupled differentiation and division of embryonic stem cells inferred from clonal snapshots

To cite this article before publication: Liam J Ruske et al 2020 *Phys. Biol.* in press <https://doi.org/10.1088/1478-3975/aba041>

Manuscript version: Accepted Manuscript

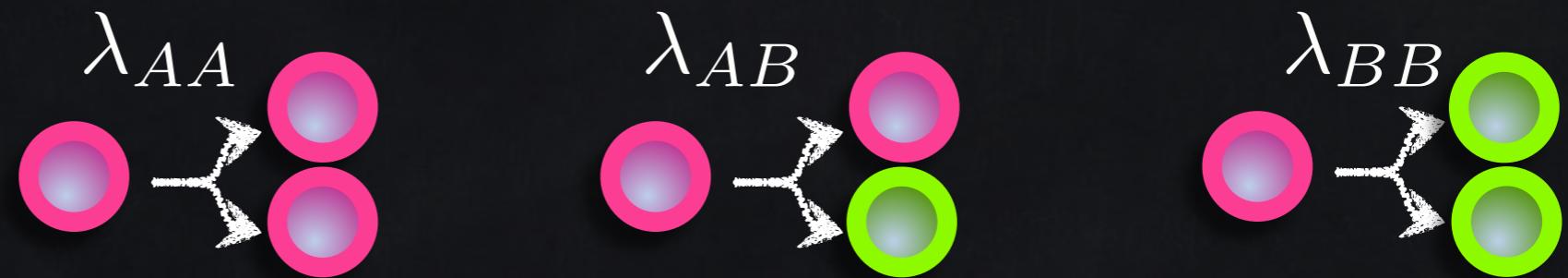
Accepted Manuscript is "the version of the article accepted for publication including all changes made as a result of the peer review process, and which may also include the addition to the article by IOP Publishing of a header, an article ID, a cover sheet and/or an 'Accepted Manuscript' watermark, but excluding any other editing, typesetting or other changes made by IOP Publishing and/or its licensors"

Beyond

Example 1: Population growth

Pluripotent

Differentiated



Can we distinguish differentiation mechanisms?

Physical Biology

ACCEPTED MANUSCRIPT • OPEN ACCESS

Coupled differentiation and division of embryonic stem cells inferred from clonal snapshots

To cite this article before publication: Liam J Ruske et al 2020 *Phys. Biol.* in press <https://doi.org/10.1088/1478-3975/aba041>

Manuscript version: Accepted Manuscript

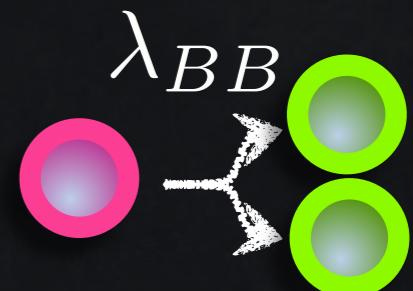
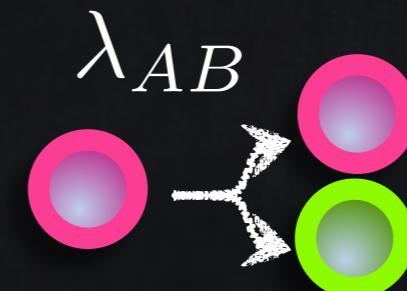
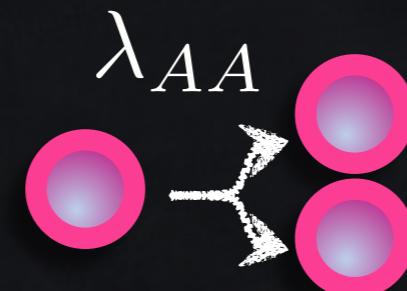
Accepted Manuscript is "the version of the article accepted for publication including all changes made as a result of the peer review process, and which may also include the addition to the article by IOP Publishing of a header, an article ID, a cover sheet and/or an 'Accepted Manuscript' watermark, but excluding any other editing, typesetting or other changes made by IOP Publishing and/or its licensors"

Beyond

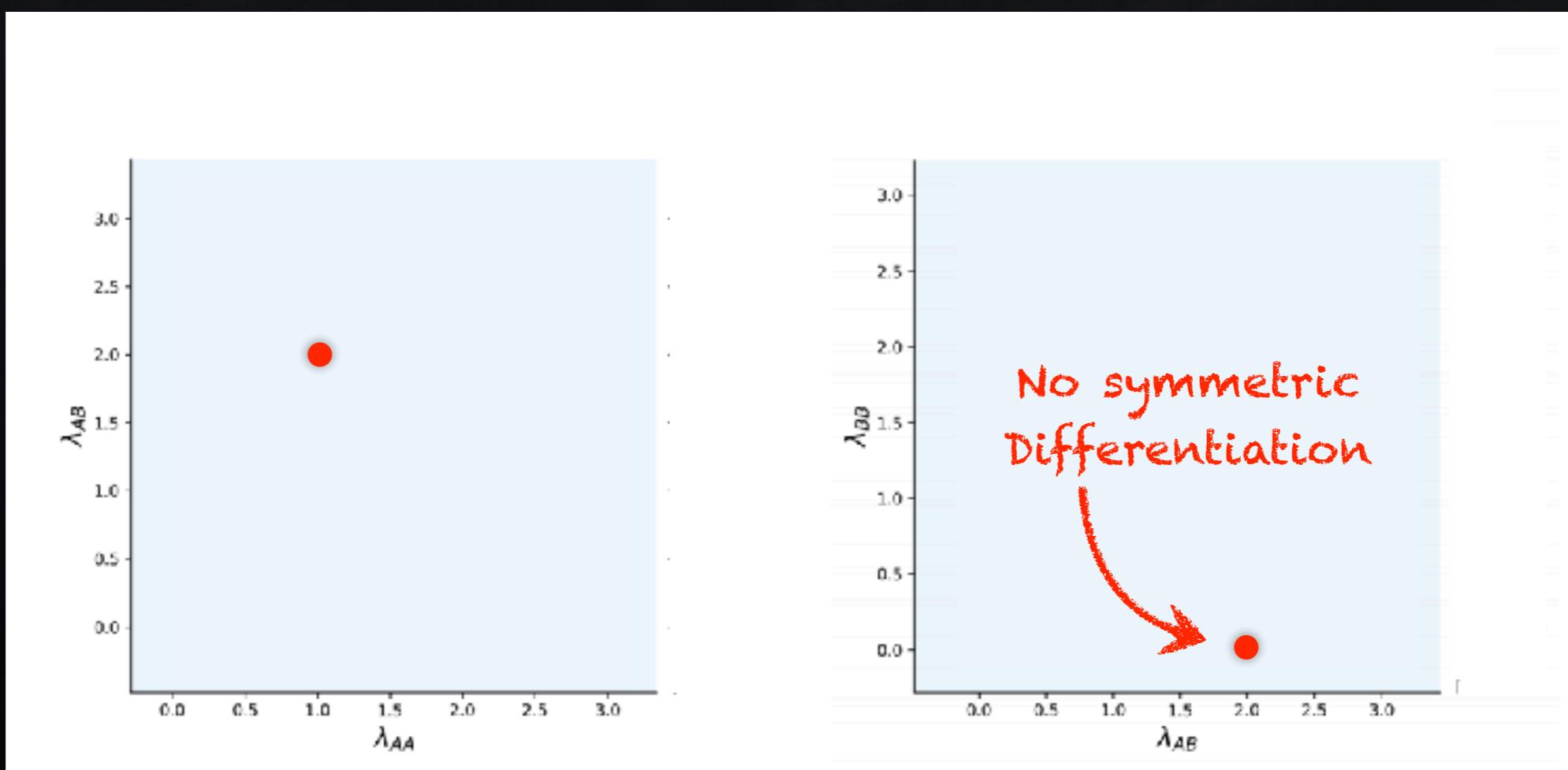
Example 1: Population growth

Pluripotent

Differentiated



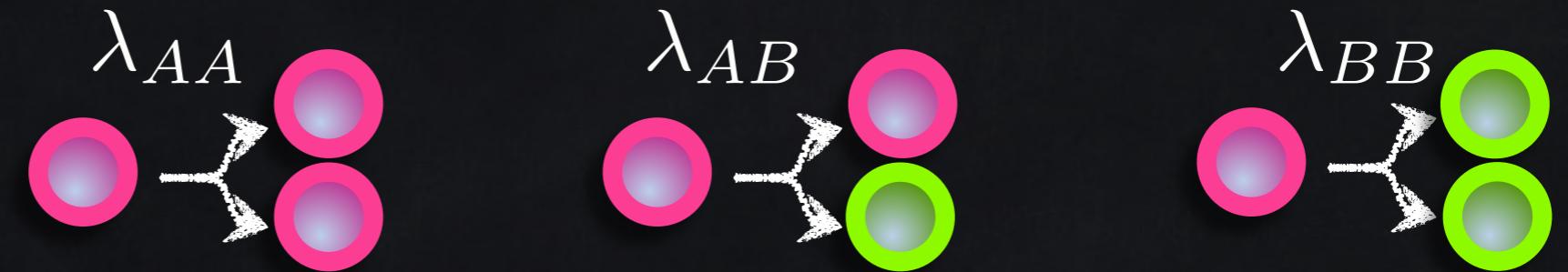
Can we distinguish differentiation mechanisms?



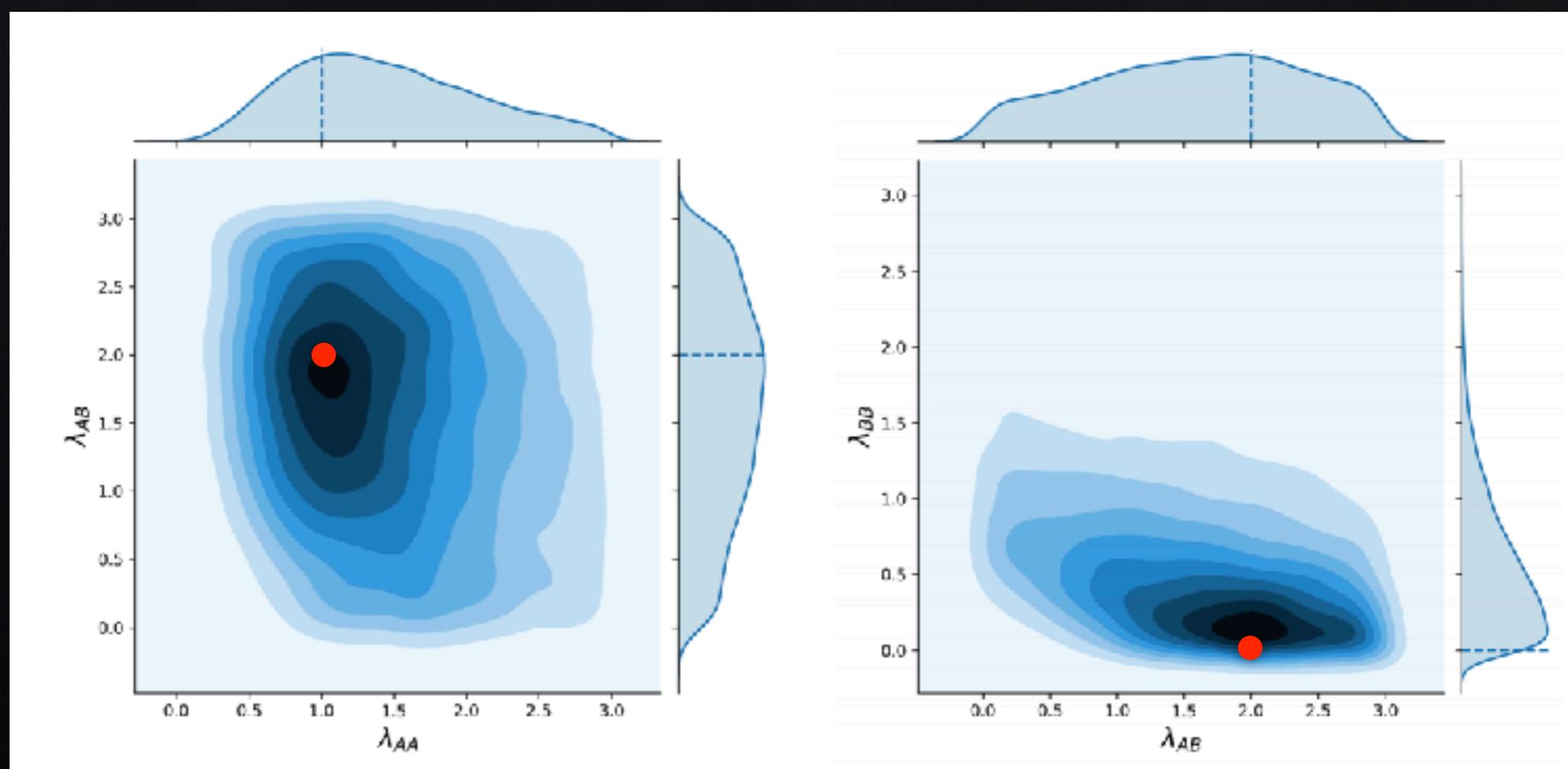
Beyond

Example 1: Population growth

- Pluripotent
- Differentiated



Can we distinguish differentiation mechanisms?

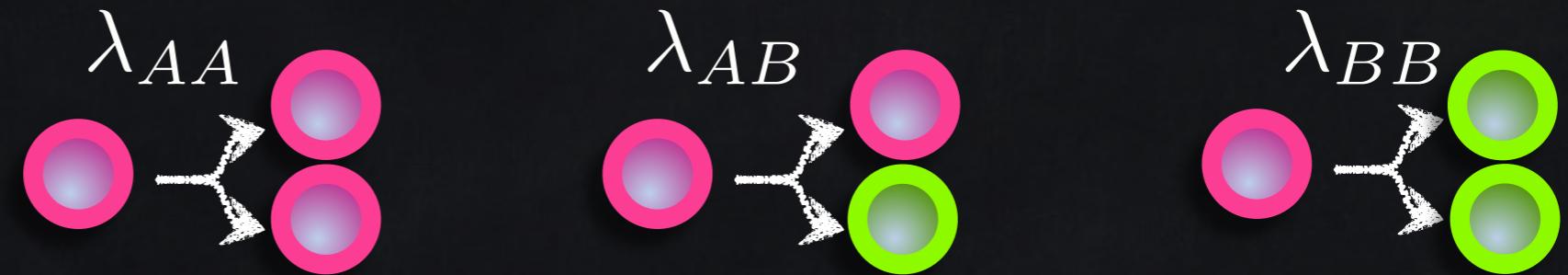


Beyond

Example 1: Population growth

Pluripotent

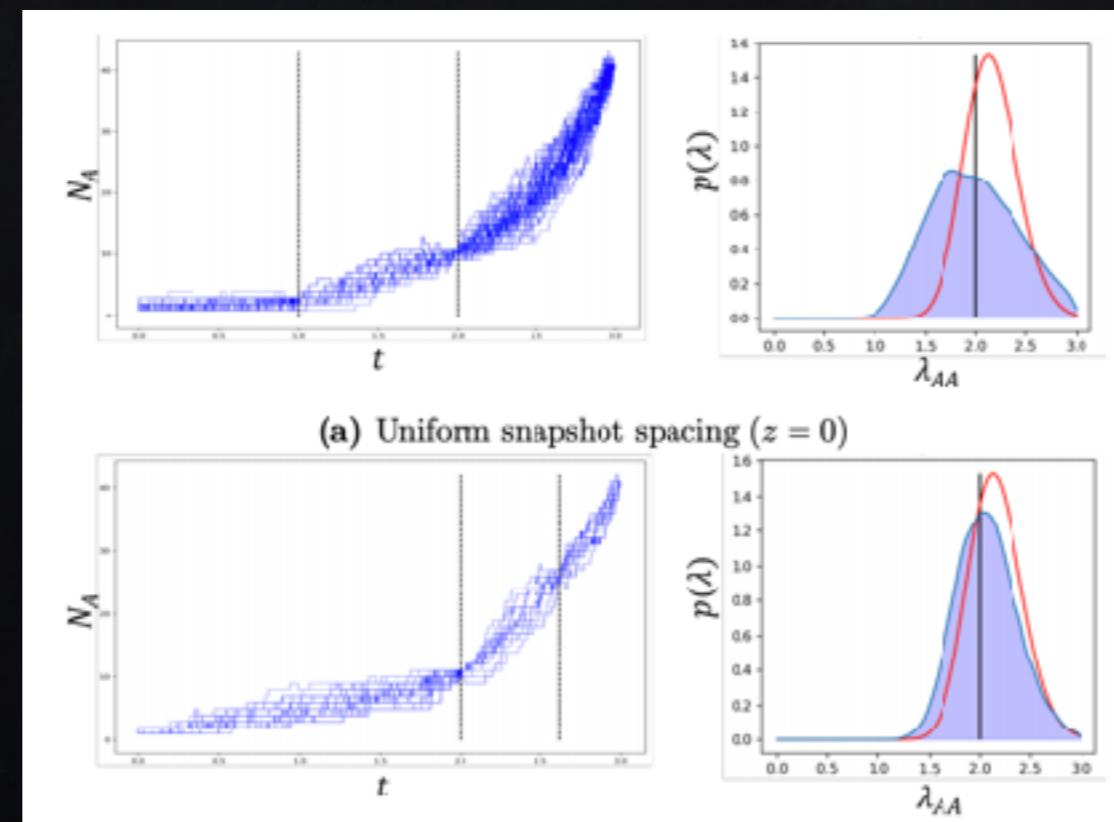
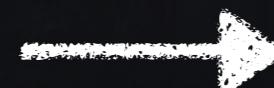
Differentiated



Can we distinguish differentiation mechanisms?

Is there an optimal set of sampling times?

Uniform
sampling



Optimal
sampling



Bayesian inference methods for ecology

Theory

What is Bayesian inference?

Probability 101

Bayes' theorem in biology

Priors and likelihood

Interpretation of posteriors

Markov Chain Monte Carlo

Applications

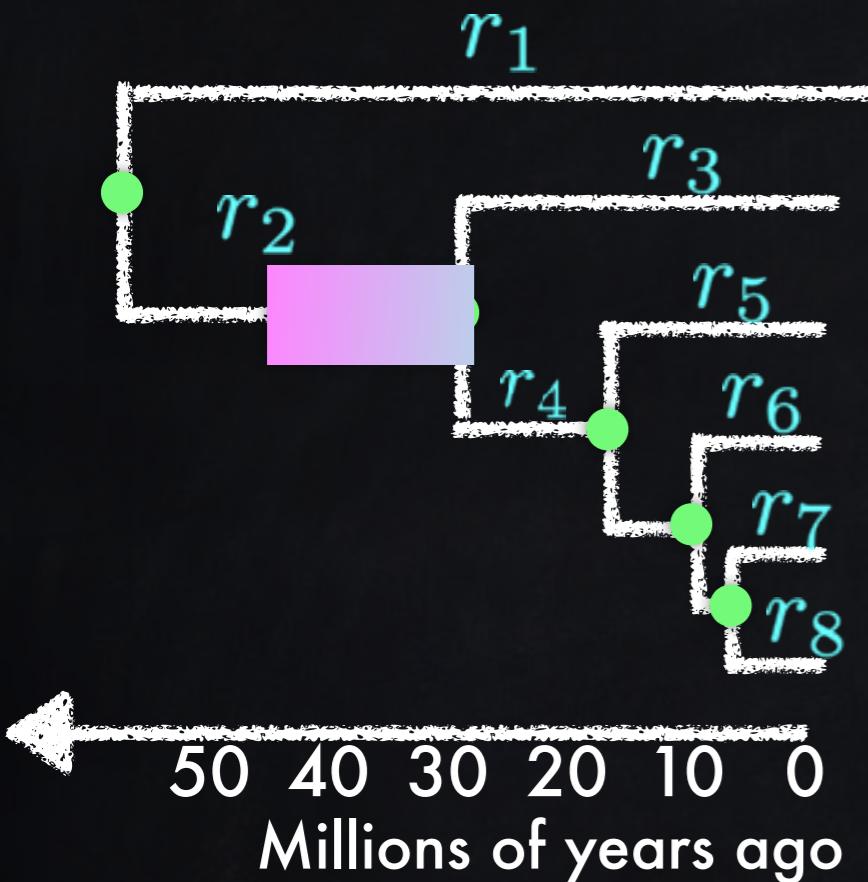
(to cell/developmental biology)

Proliferation and
differentiation of a
stem cell population

Phylogeny and model
selection

Morphogen patterning
of embryonic tissues

Phylogeny



$$P(t_i, r_i, \theta | \text{data})$$

Mutation rates model parameters

Data

Marmoset
Macaque
Orangutan
Gorilla
Chimpanzee
Human

DNA alignment
Fossils/Events

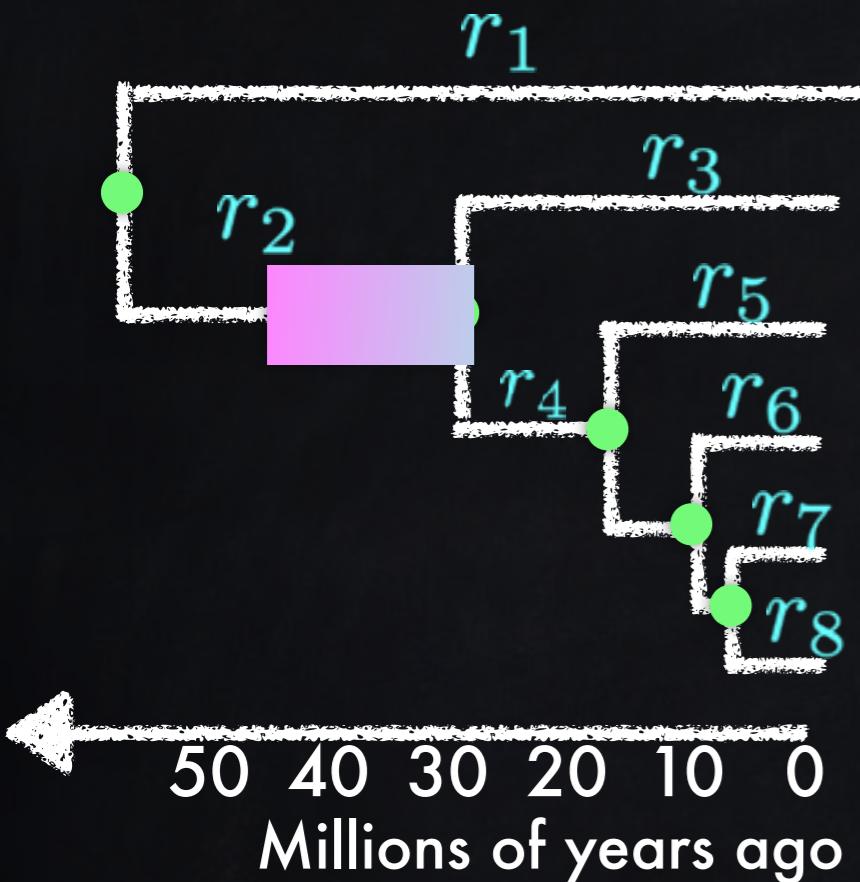
Uncertainty

Mutation variability
Inaccurate bounds

$$\text{Tree} = \{ t_i \}$$
$$P(t_i | \text{data})$$

Speciation times

Phylogeny



Data

Marmoset
Macaque
Orangutan
Gorilla
Chimpanzee
Human

DNA alignment
Fossils/Events

Uncertainty

Mutation variability
Inaccurate bounds

Tree = $\{t_i\}$

Speciation times

$P(t_i | data)$

Mutation rates model
parameters

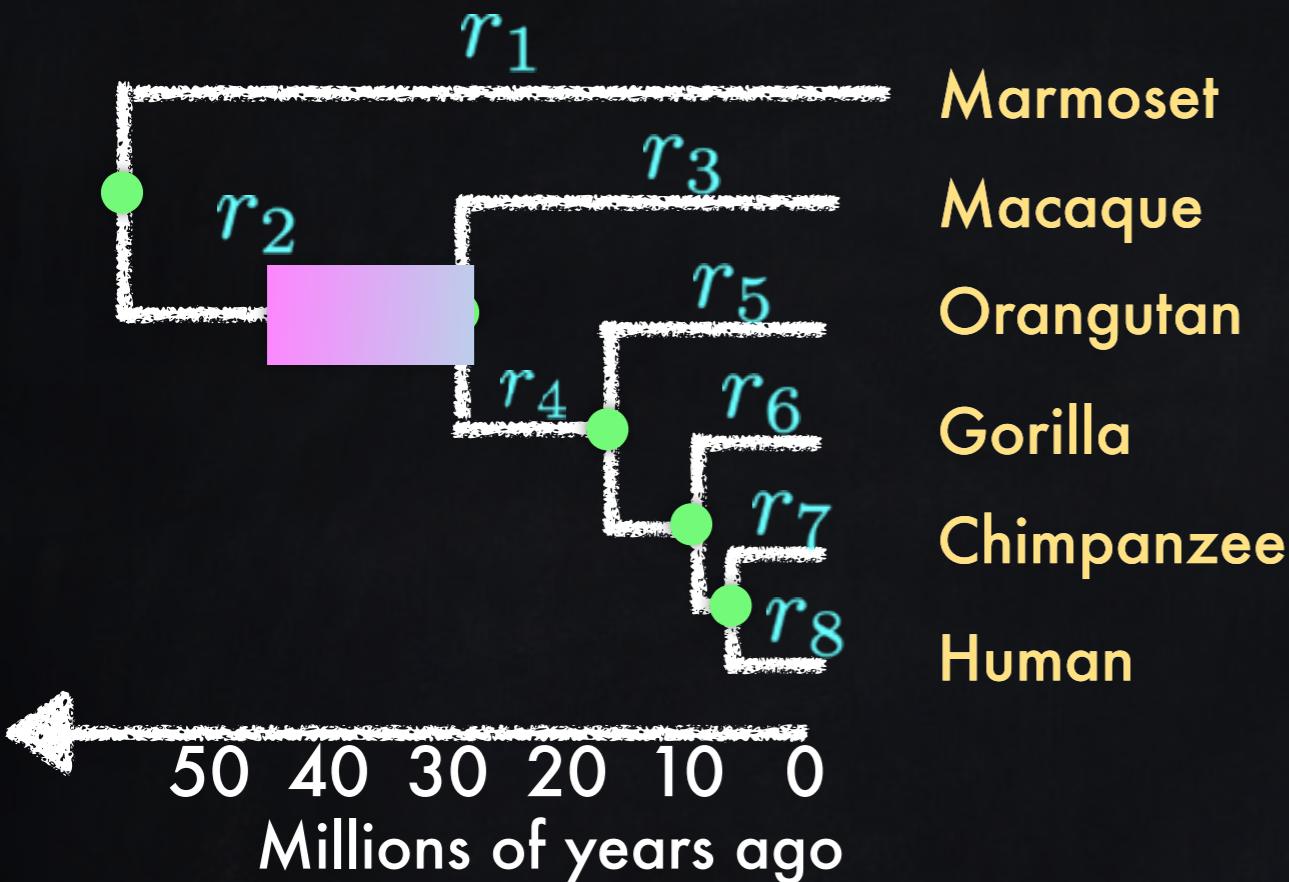
$$P(t_i, r_i, \theta | data) \propto P(data | r_i, t_i, \theta) P(r_i, t_i, \theta) \\ \propto P(data | r_i, t_i) P(r_i | t_i, \theta) P(t_i | \theta) P(\theta)$$

Nucleotide
substitution
model

Molecular clock
model

Divergence time
model

Phylogeny



Bayesian inference allows for nested models that capture the complexity of underlying structure/correlations:
Hierarchical models

$$\text{Tree} = \{ t_i \} \xrightarrow{\text{Speciation times}} P(t_i | \text{data})$$

model
parameters

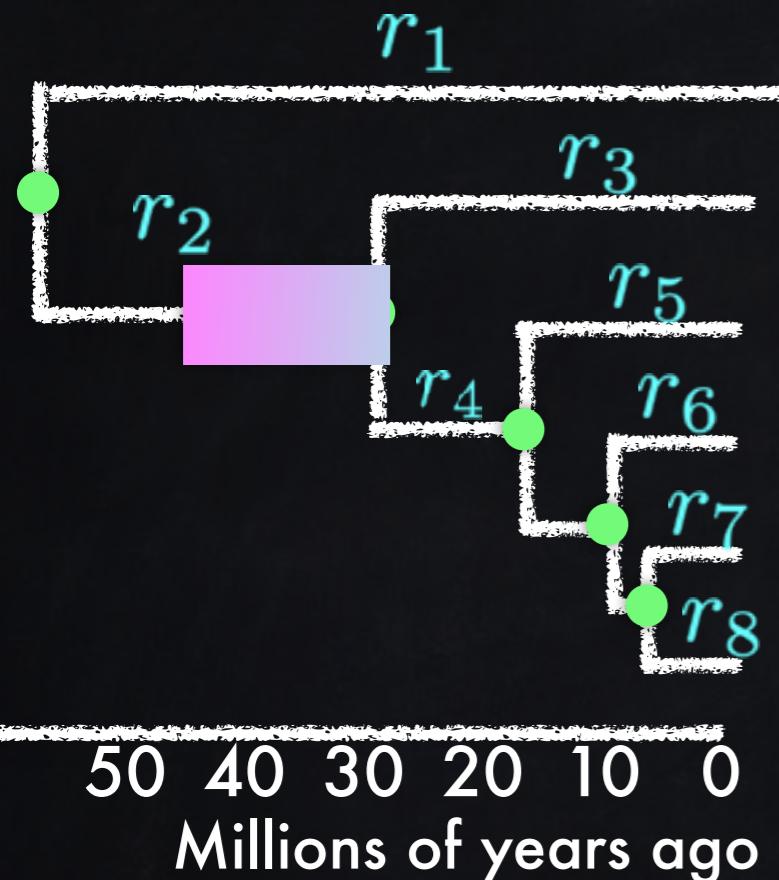
$$P(t_i, r_i, \theta | \text{data}) \propto P(\text{data} | r_i, t_i, \theta) P(r_i, t_i, \theta)$$
$$\propto P(\text{data} | r_i, t_i) P(r_i | t_i, \theta) P(t_i | \theta) P(\theta)$$

Mutation rates

Nucleotide substitution model

Molecular clock Divergence time model

Phylogeny



Marmoset
Macaque
Orangutan
Gorilla
Chimpanzee
Human



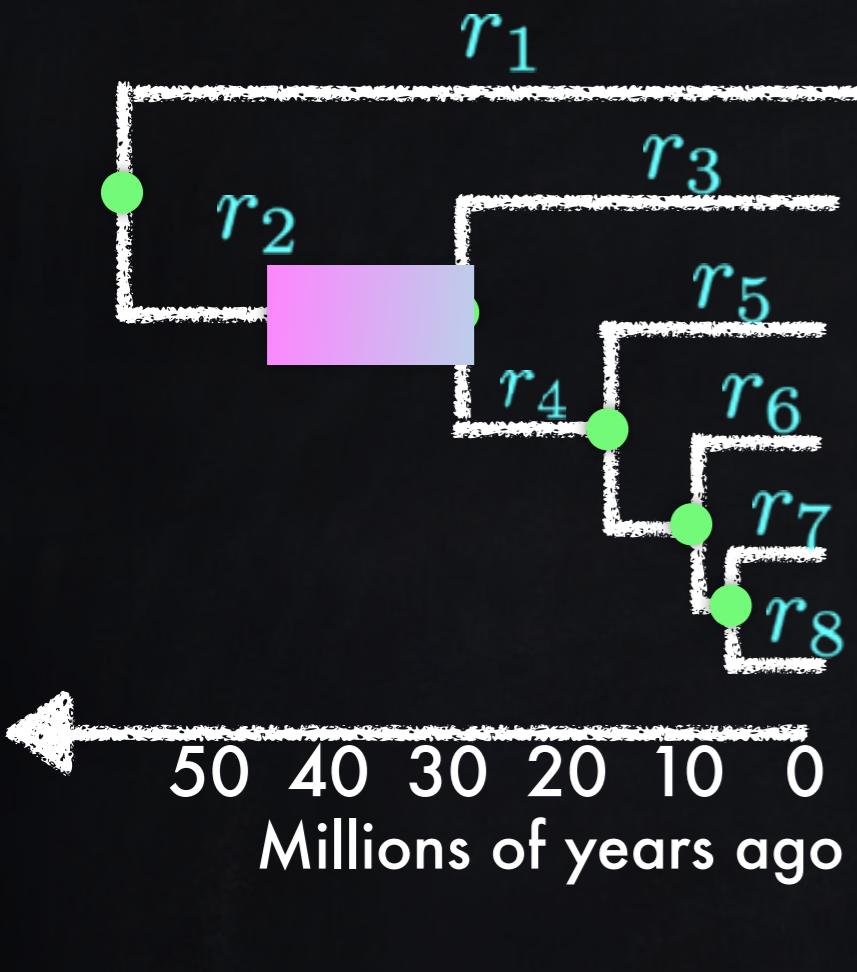
Bayesian inference allows for nested models that capture the complexity of underlying structure/correlations:
Hierarchical models

Mutation rates model
parameters

$$\begin{aligned} P(t_i, r_i, \theta | \text{data}) &\propto P(\text{data} | r_i, t_i, \theta) P(r_i, t_i, \theta) \\ &\propto P(\text{data} | r_i, t_i) P(r_i | t_i, \theta) P(t_i | \theta) P(\theta) \\ &\propto P(\text{data} | r_i, t_i) P(r_i | t_i, \theta, \theta_M) P(t_i | \theta) P(\theta, \theta_M) \end{aligned}$$

Model selection parameter

Phylogeny



Mutation rates \uparrow \rightarrow model parameters

$$\begin{aligned}
 P(t_i, r_i, \theta | data) &\propto P(data | r_i, t_i, \theta) P(r_i, t_i, \theta) \\
 &\propto P(data | r_i, t_i) P(r_i | t_i, \theta) P(t_i | \theta) P(\theta) \\
 &\propto P(data | r_i, t_i) P(r_i | t_i, \theta, \theta_M) P(t_i | \theta) P(\theta, \theta_M)
 \end{aligned}$$

Model selection parameter

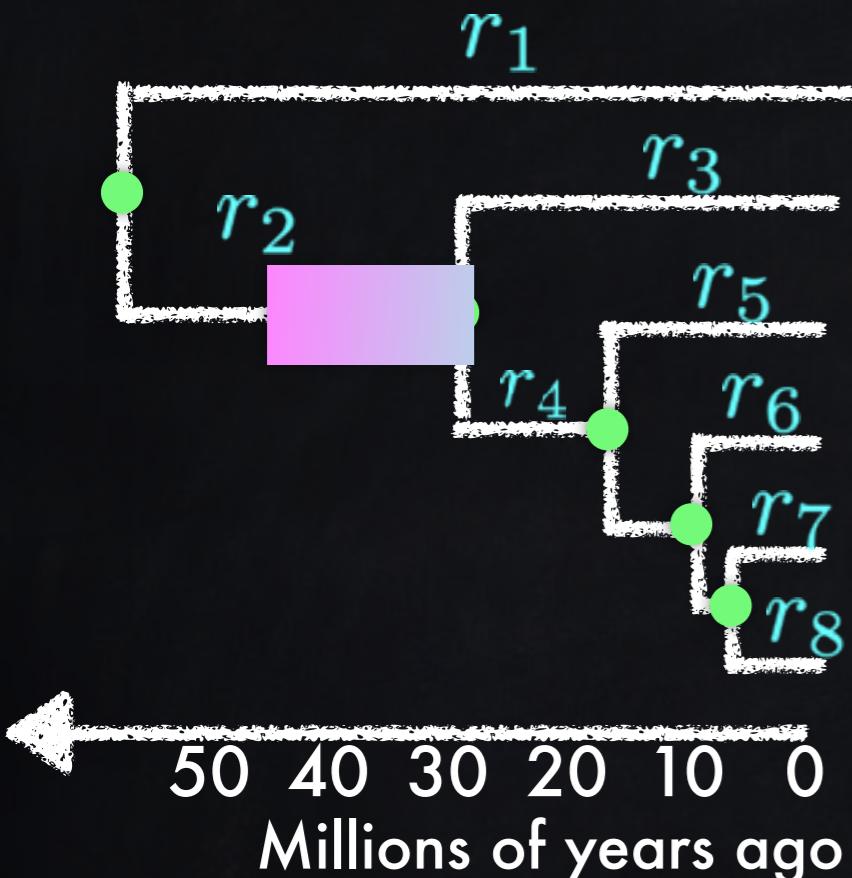


Bayesian inference allows for nested models that capture the complexity of underlying structure/correlations:
Hierarchical models

$$\frac{P(\theta_1 | data)}{P(\theta_2 | data)} = \text{Bayes' factor}$$

Credibility among models

Phylogeny



Marmoset
Macaque
Orangutan
Gorilla
Chimpanzee
Human



Bayesian inference allows for nested models that capture the complexity of underlying structure/correlations:
Hierarchical models



Bayesian inference allows model selection



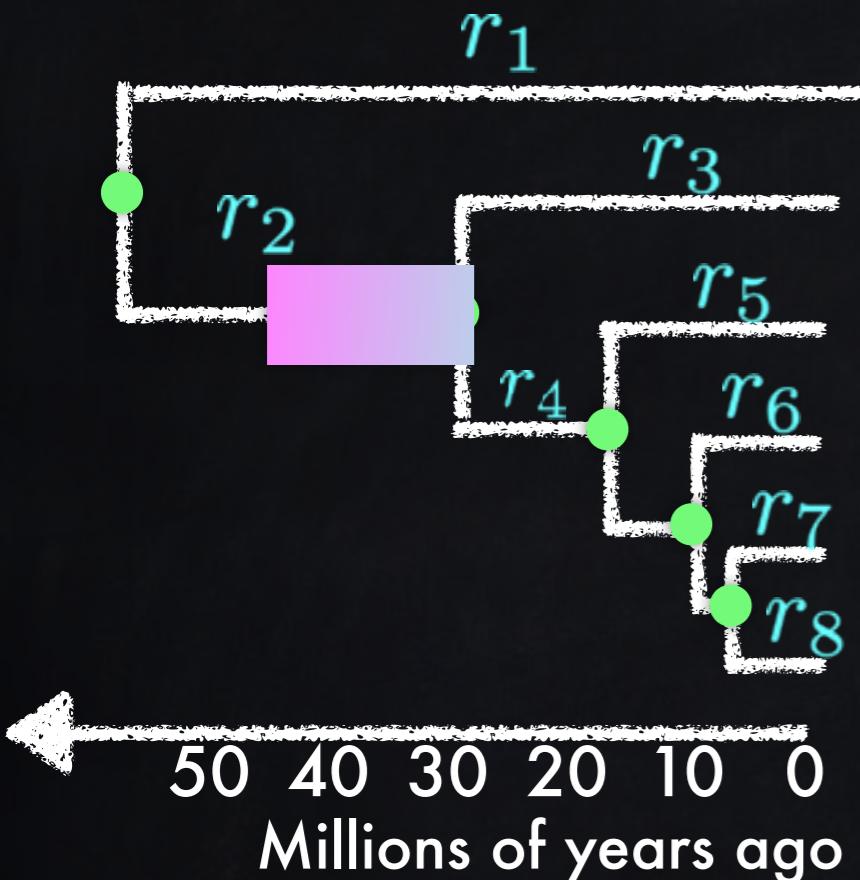
$\frac{P(\theta_1|data)}{P(\theta_2|data)} = \text{Bayes' factor}$
Credibility among models

Mutation rates model parameters

$$\begin{aligned}
 P(t_i, r_i, \theta | data) &\propto P(data | r_i, t_i, \theta) P(r_i, t_i, \theta) \\
 &\propto P(data | r_i, t_i) P(r_i | t_i, \theta) P(t_i | \theta) P(\theta) \\
 &\propto P(data | r_i, t_i) P(r_i | t_i, \theta, \theta_M) P(t_i | \theta) P(\theta, \theta_M)
 \end{aligned}$$

Model selection parameter

Phylogeny



$$\begin{aligned} P(t_i, r_i, \theta | data) &\propto P(data | r_i, t_i, \theta) P(r_i, t_i | \theta) \\ &\propto P(data | r_i, t_i) P(r_i | t_i, \theta) \\ &\propto P(data | r_i, t_i) P(r_i | t_i, \theta, \text{model parameters}) \end{aligned}$$

Model sele

Bayesian inference allows for nested that capture the complexity of underlying structure/correlations:
Hierarchical models

Bayes mode



The Molecular Evolutionary Clock

Theory and Practice

Example 3: Developmental GRNs / Neural tube patterning



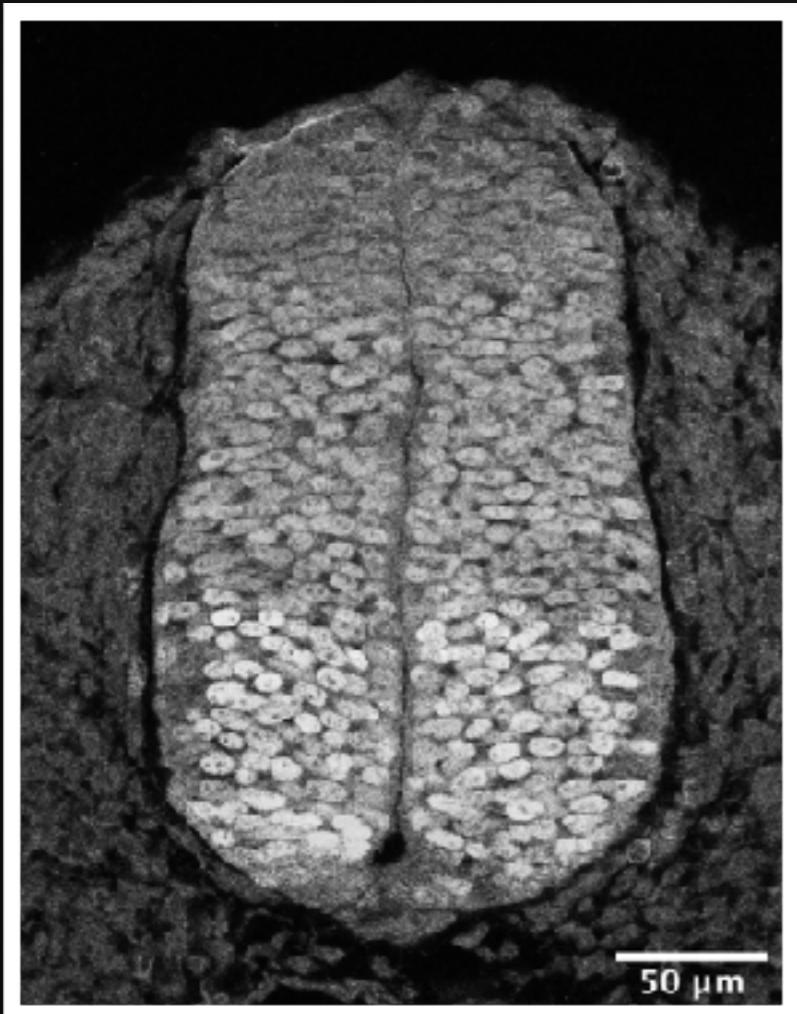
Which?

How?

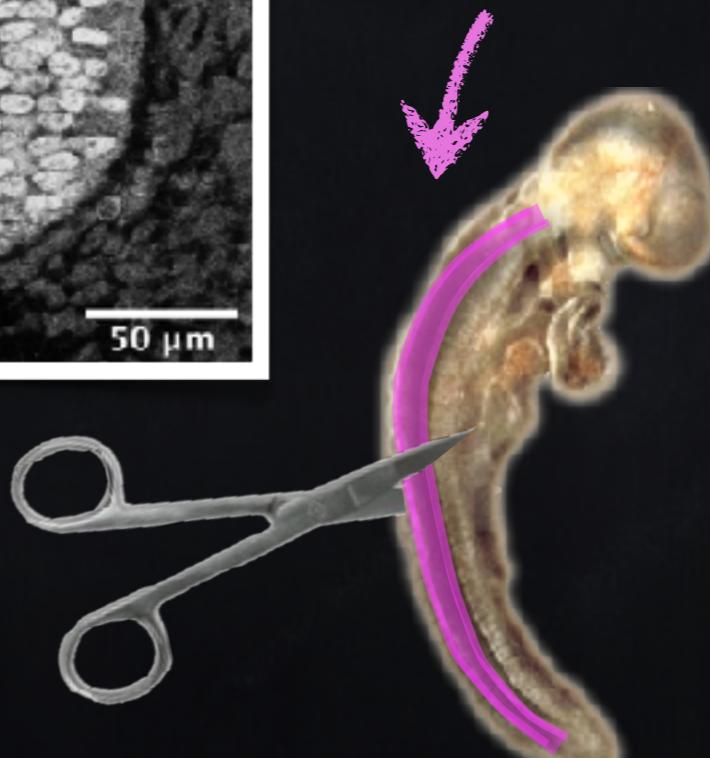
Where?

When?

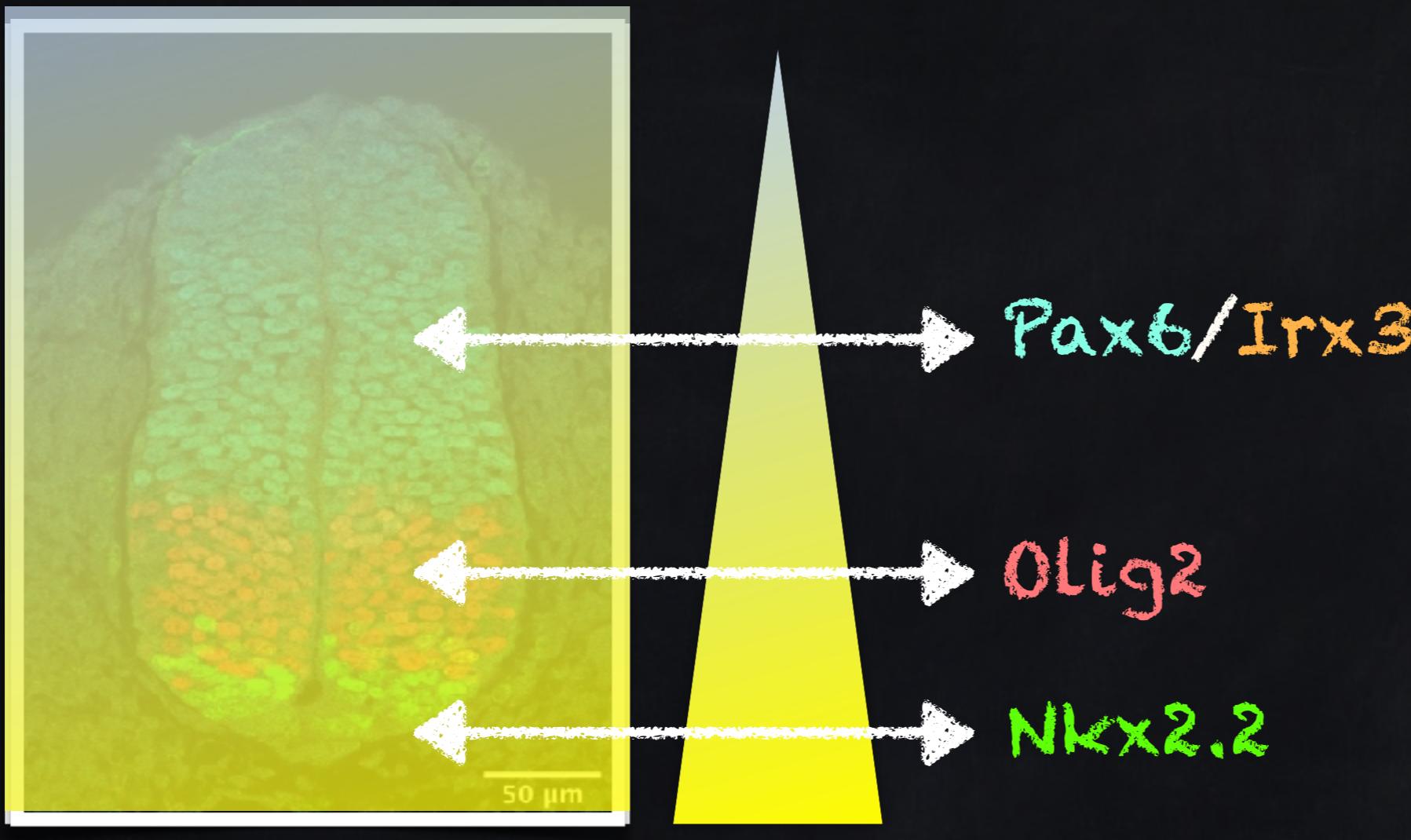
Example 3: Developmental GRNs / Neural tube patterning



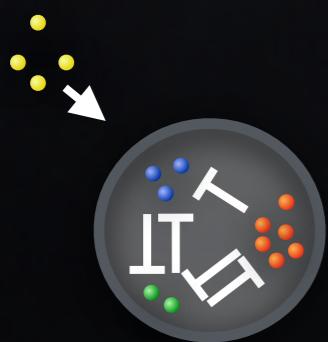
Neural tube



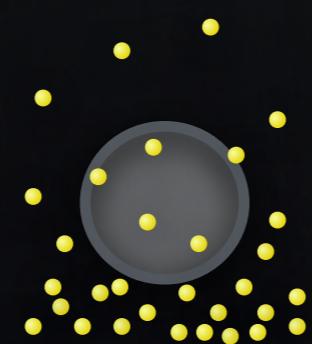
Example 3: Developmental GRNs / Neural tube patterning



Sonic
Hedgehog



Single cell
regulation

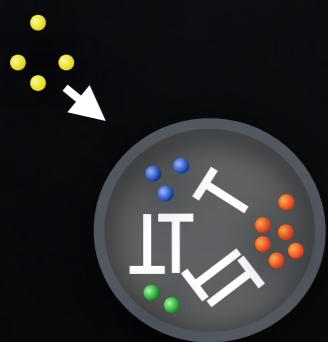
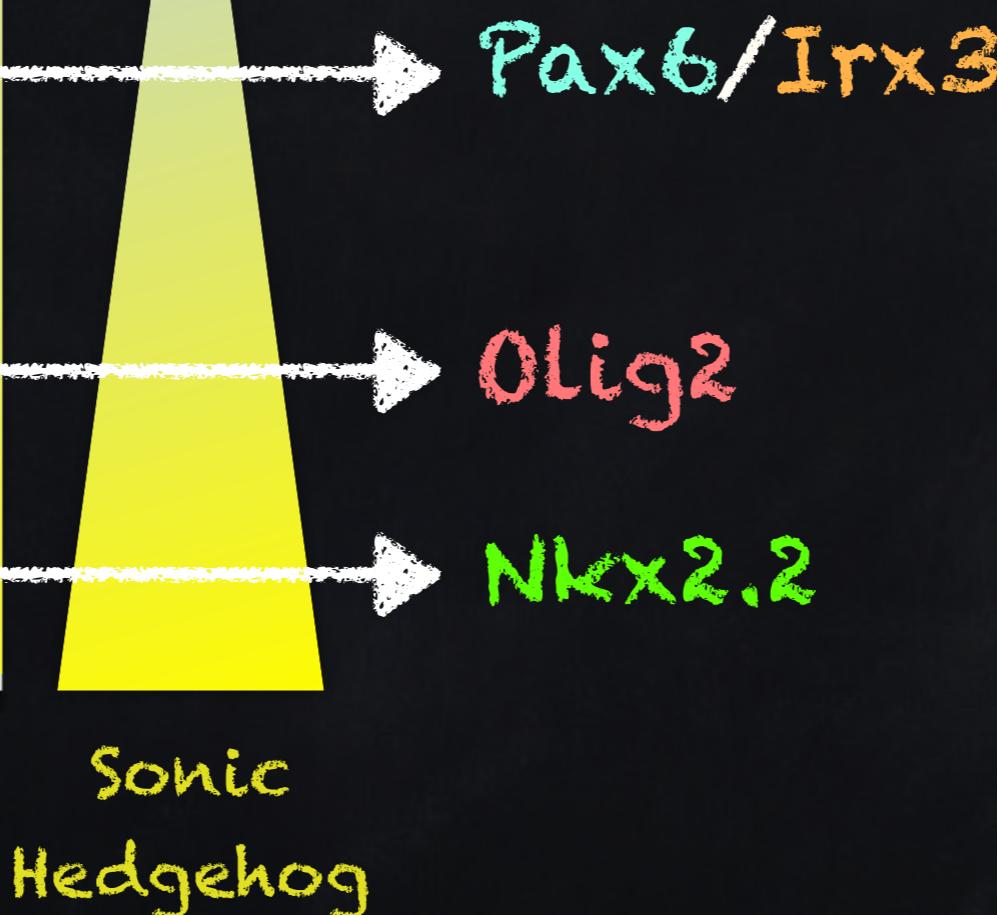
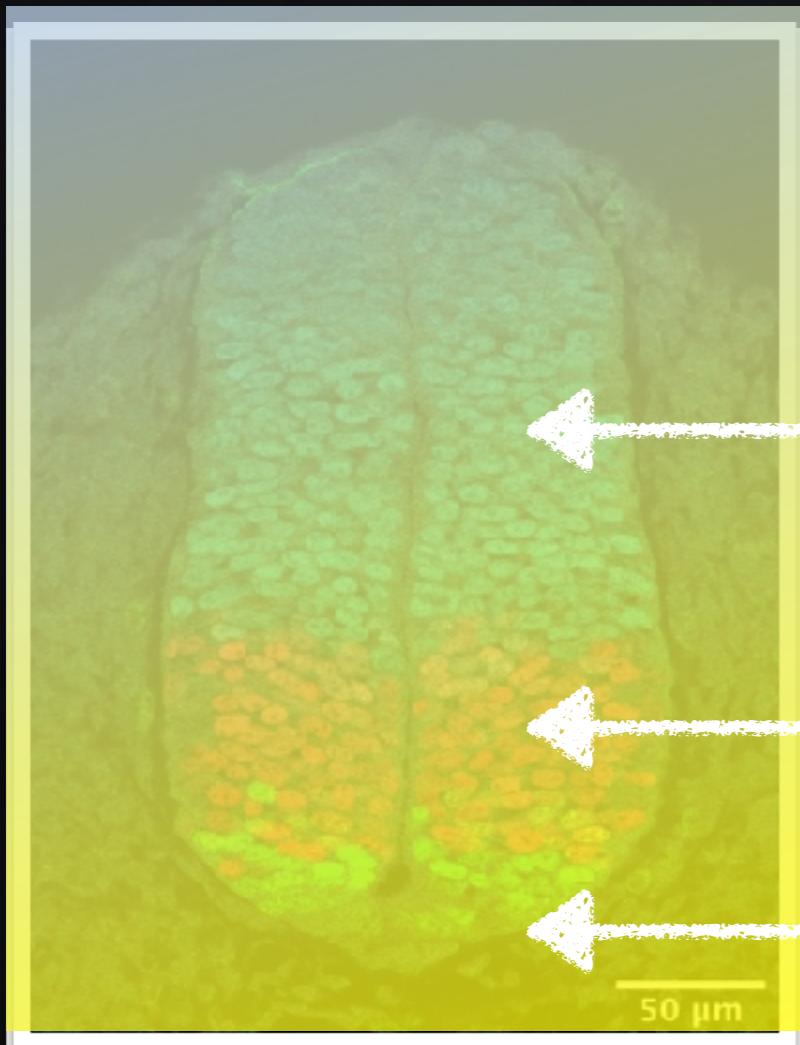


Signal
diffusion



Tissue
dynamics

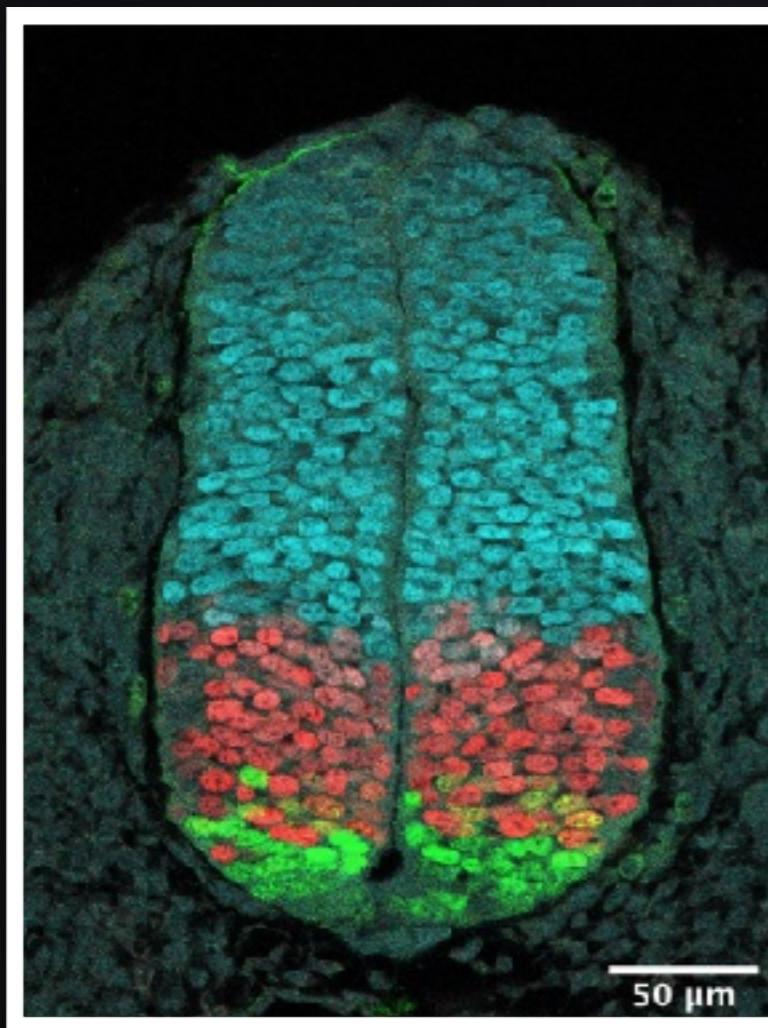
Example 3: Developmental GRNs / Neural tube patterning



Single cell
regulation

How does gene regulation
control the patterning mechanism?

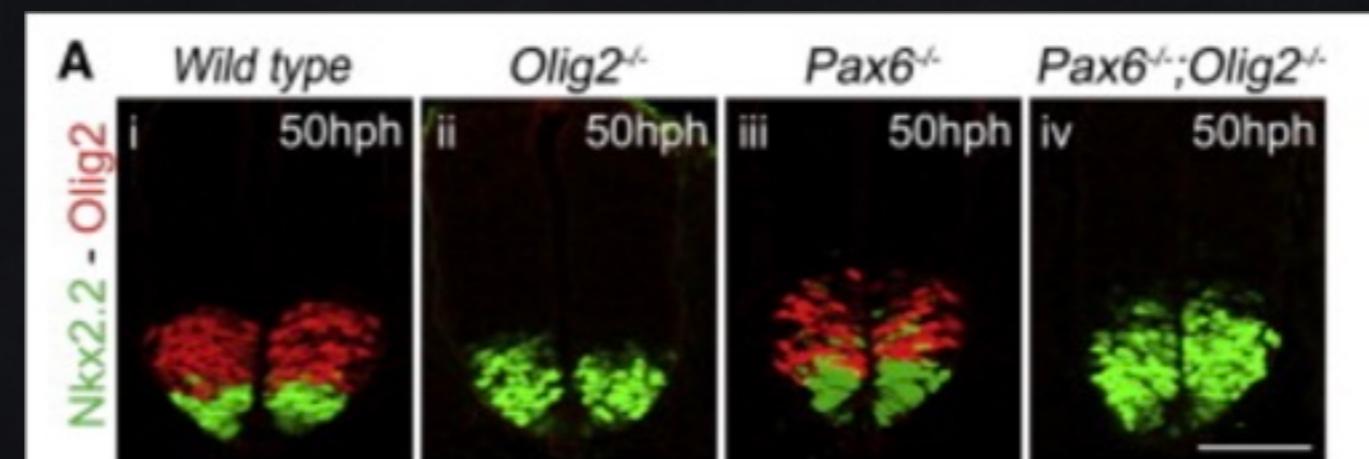
Example 3: Developmental GRNs / Neural tube patterning



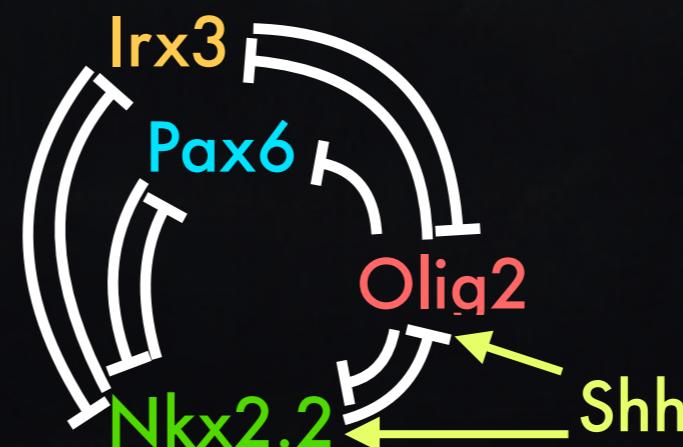
How does gene regulation control the patterning mechanism?

What do we know?

Mutant patterns



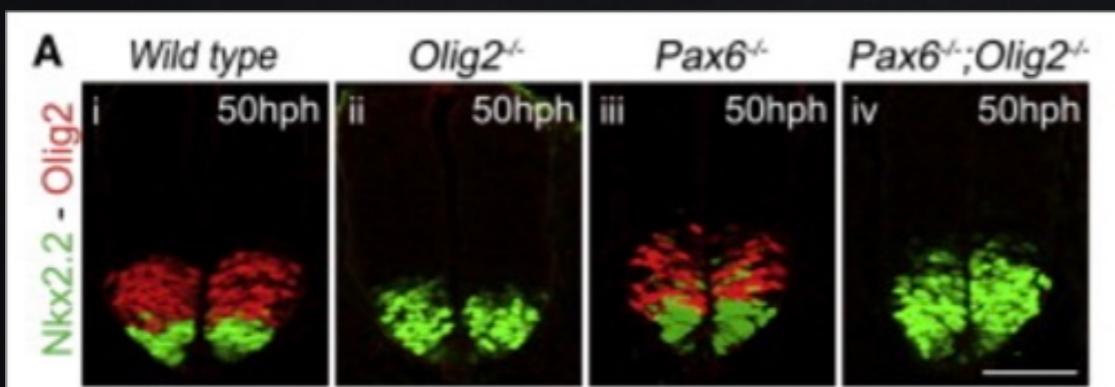
GRN topology (causality)



Example 3: Developmental GRNs / Neural tube patterning

What do we know?

Mutant patterns



GRN topology (causality)



How do we encode it in a quantitative model?

Set of ODEs

Production Degradation

$$\dot{x}_I = f_I(x_N, x_O) - \delta_I x_I$$

$$\dot{x}_P = f_P(x_N, x_O) - \delta_P x_P$$

$$\dot{x}_O = f_O(x_N, x_I, S) - \delta_O x_O$$

$$\dot{x}_N = f_N(x_O, x_P, x_I, S) - \delta_N x_N$$

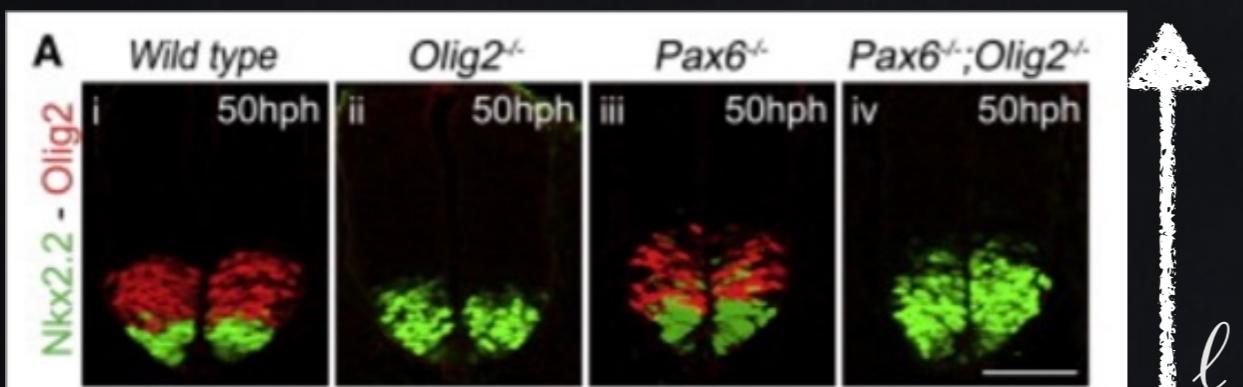
Question:

- Try to write your own regulatory functions $f()$

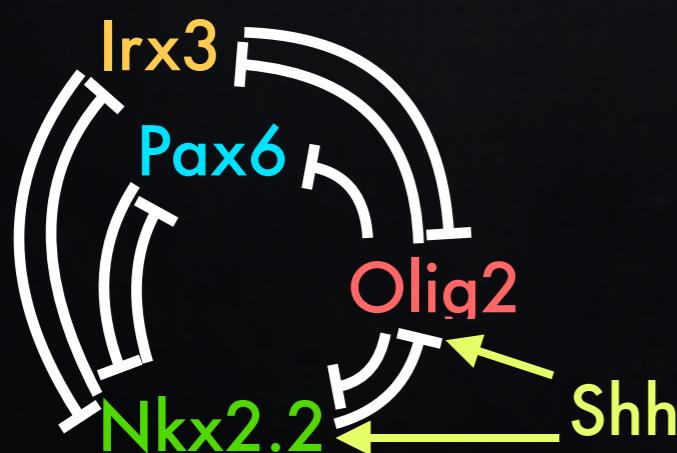
Example 3: Developmental GRNs / Neural tube patterning

What do we know?

Mutant patterns



GRN topology (causality)



How do we encode it
in a quantitative model?

Set of ODEs

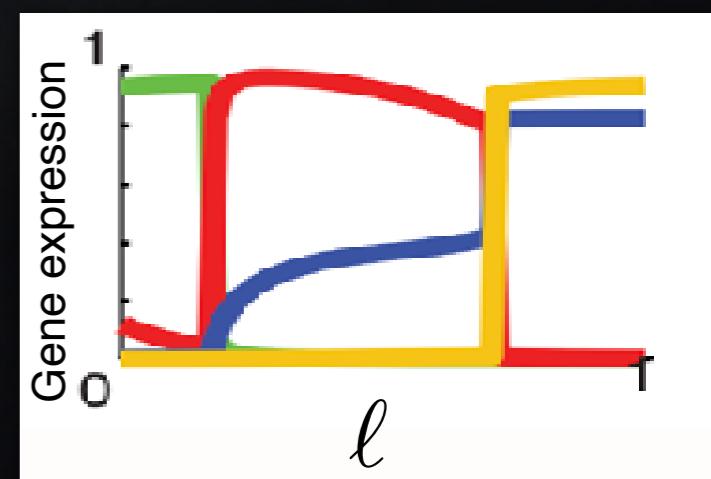
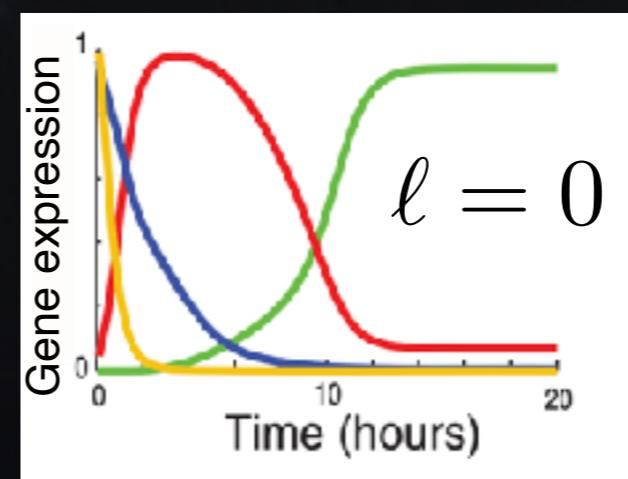
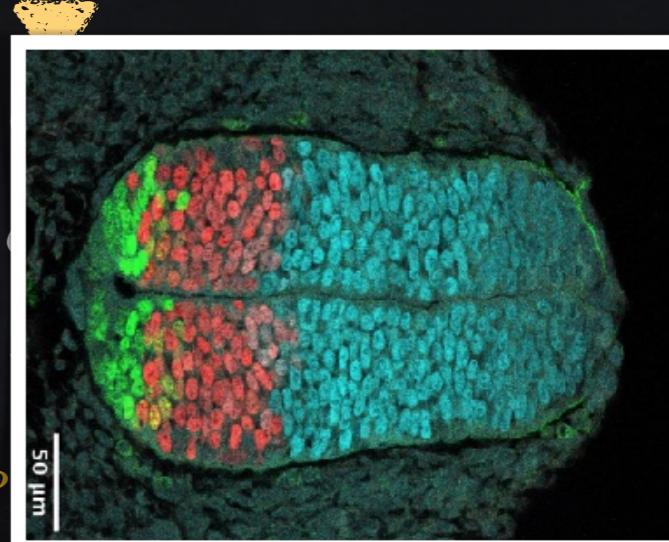
Production Degradation

$$\dot{x}_I = f_I(x_N, x_O) -$$

$$\dot{x}_P = f_P(x_N, x_O) -$$

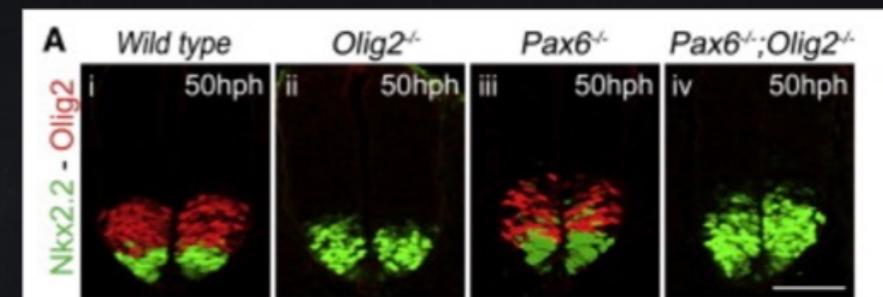
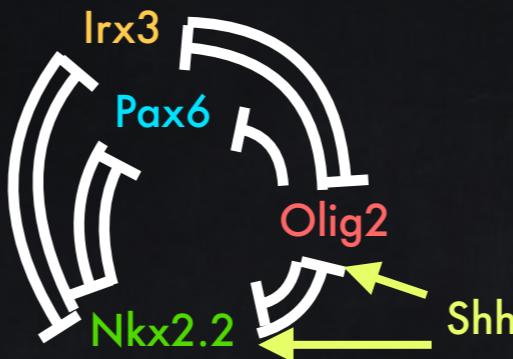
$$\dot{x}_O = f_O(x_N, x_I, S)$$

$$\dot{x}_N = f_N(x_O, x_P, x_I)$$



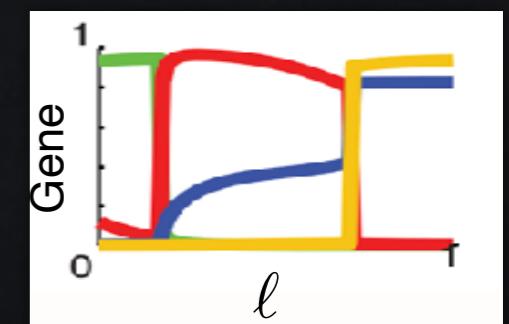
Example 3: Developmental GRNs / Neural tube patterning

What do we know?

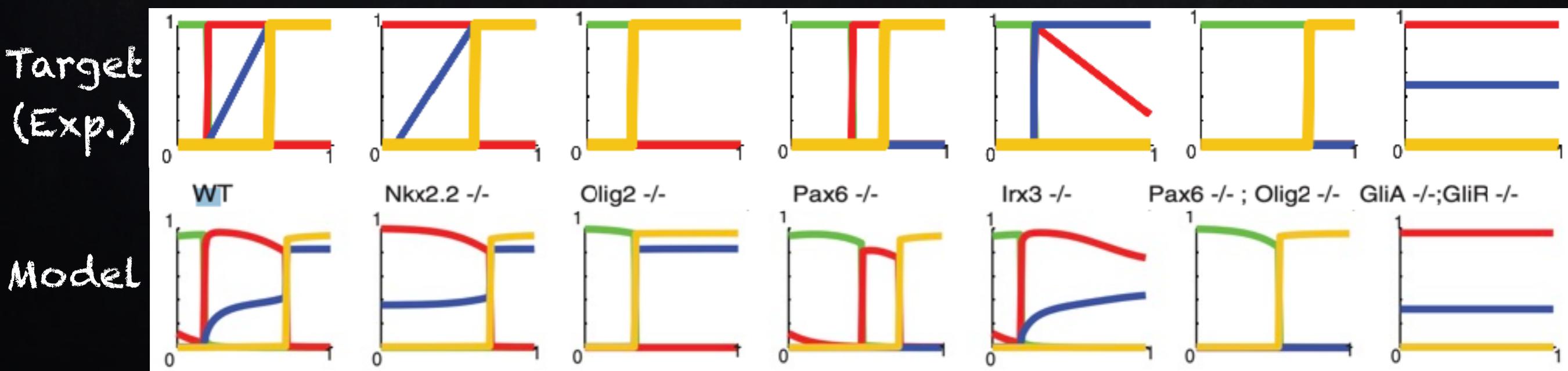


How do we encode it in a quantitative model?

$$\begin{aligned}\dot{x}_I &= f_I(x_N, x_O) - \delta_I x_I \\ \dot{x}_P &= f_P(x_N, x_O) - \delta_P x_P \\ \dot{x}_O &= f_O(x_N, x_I, S) - \delta_O x_O \\ \dot{x}_N &= f_N(x_O, x_P, x_I, S) - \delta_N x_N\end{aligned}$$

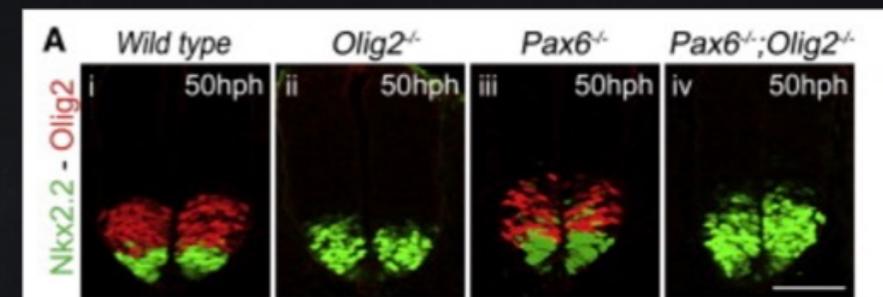
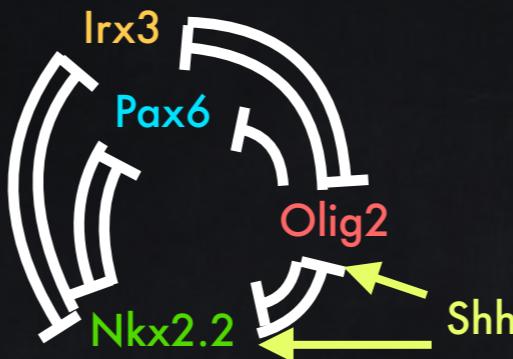


How do I compare the data with the model? (Likelihood)?



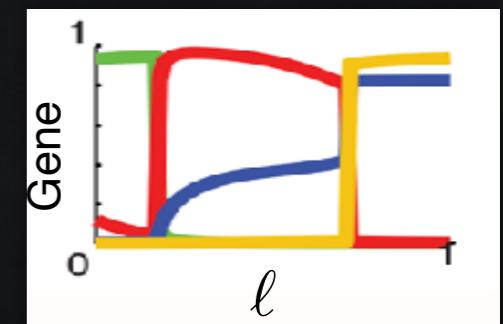
Example 3: Developmental GRNs / Neural tube patterning

What do we know?

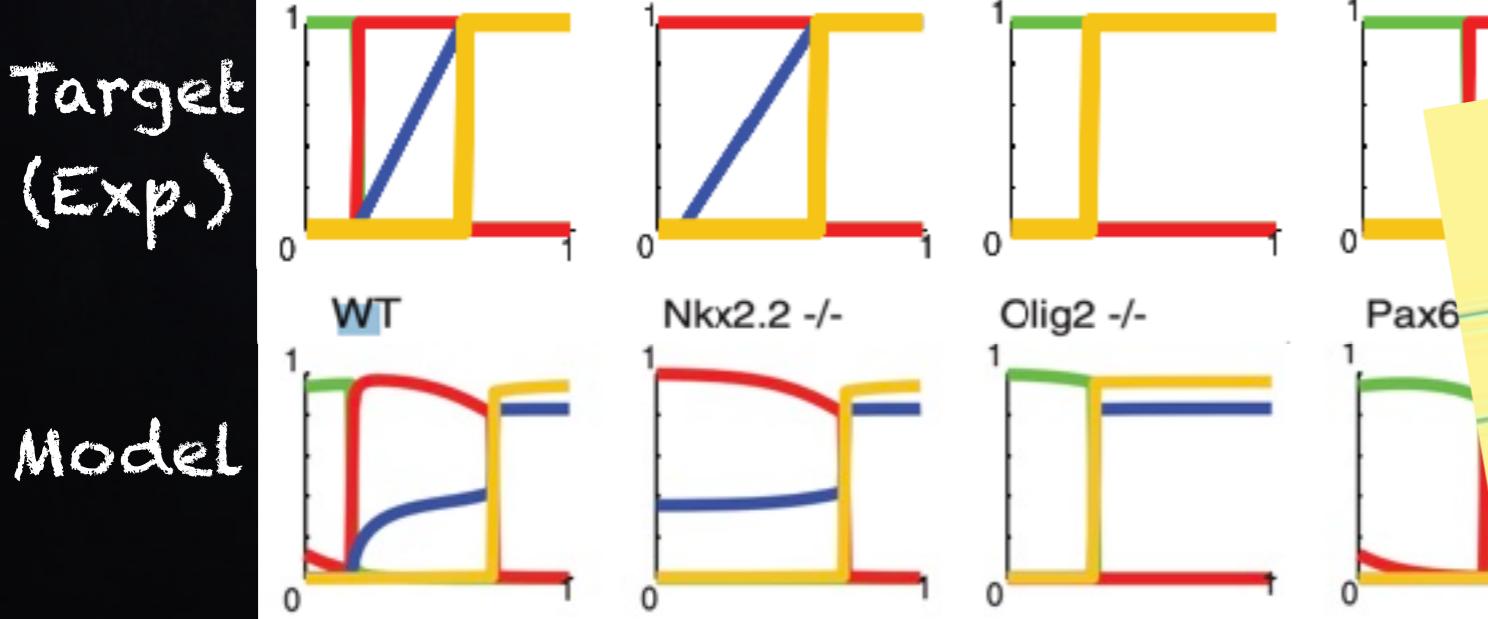


How do we encode it in a quantitative model?

$$\begin{aligned}\dot{x}_I &= f_I(x_N, x_O) - \delta_I x_I \\ \dot{x}_P &= f_P(x_N, x_O) - \delta_P x_P \\ \dot{x}_O &= f_O(x_N, x_I, S) - \delta_O x_O \\ \dot{x}_N &= f_N(x_O, x_P, x_I, S) - \delta_N x_N\end{aligned}$$



How do I compare the data with the model? (Likelihood)?

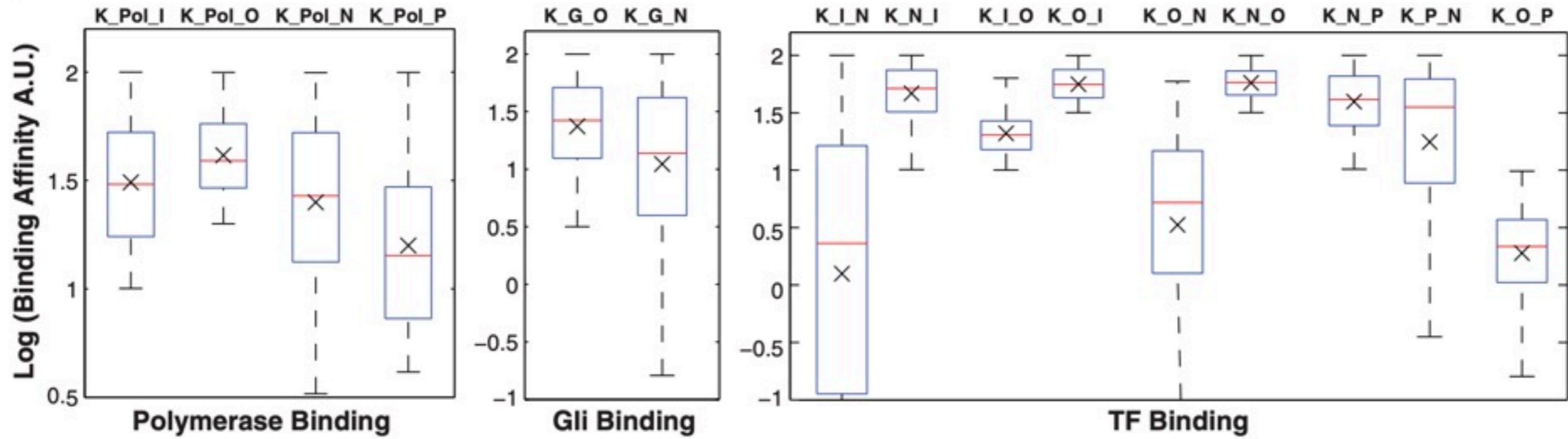


Question:

- Try to write your own distance function

Example 3: Developmental GRNs / Neural tube patterning

F



© 2014. Published by The Company of Biologists Ltd | *Development* (2014) 141, 3868-3878 doi:10.1242/dev.112573



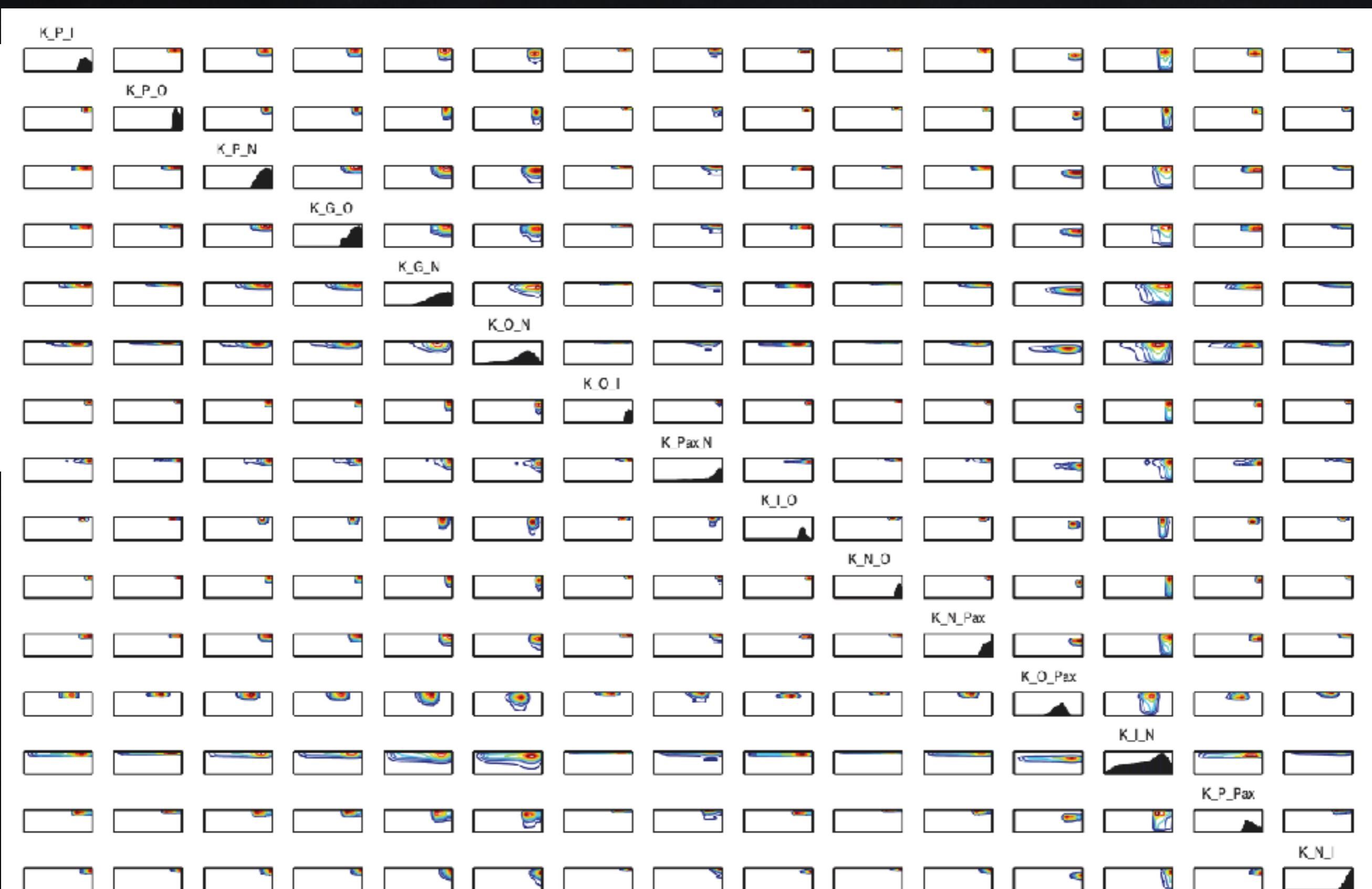
RESEARCH ARTICLE

A theoretical framework for the regulation of Shh morphogen-controlled gene expression

Michael Cohen¹, Karen M. Page², Ruben Perez-Carrasco², Chris P. Barnes³ and James Briscoe^{1,*}

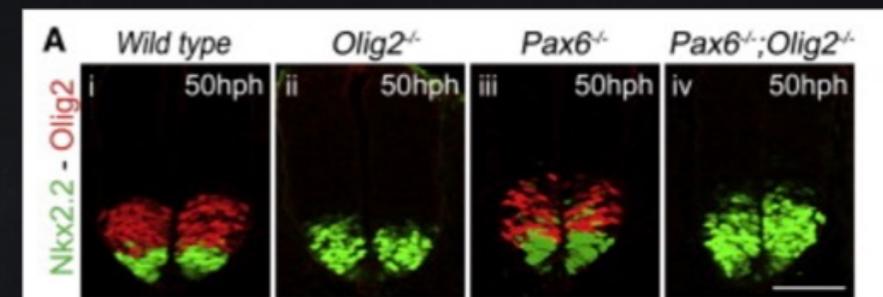
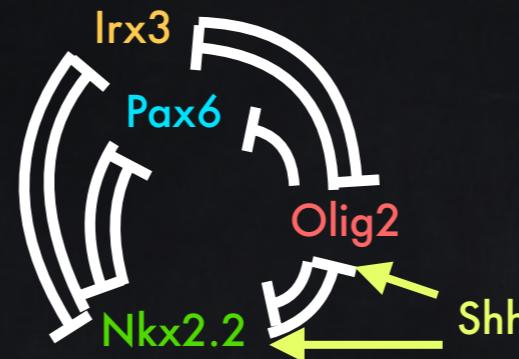
¹ The affinity threshold model came from experiments in the early 1990s. ² The model was developed in 2008. ³ Present address: University of Cambridge, Cambridge, UK.

Example 3: Developmental GRNs / Neural tube patterning



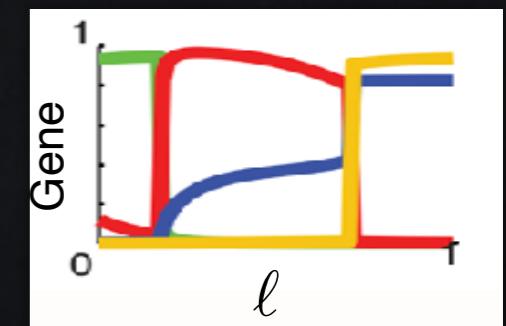
Example 3: Developmental GRNs / Neural tube patterning

What do we know?

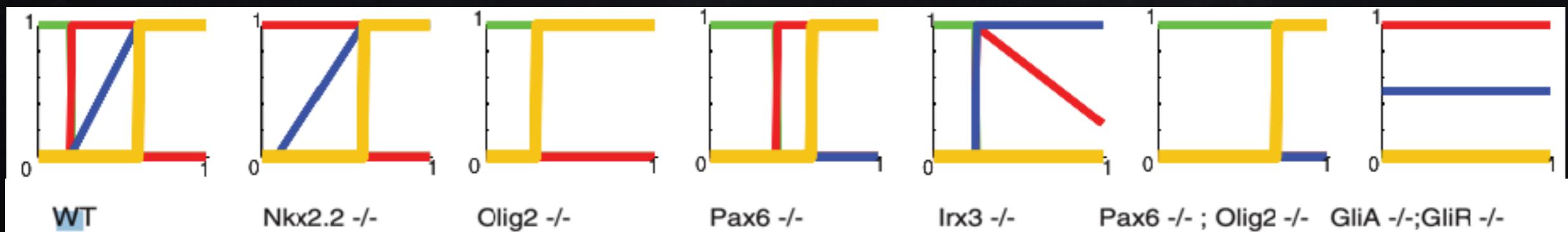


How do we encode it in a quantitative model?

$$\begin{aligned}\dot{x}_I &= f_I(x_N, x_O) - \delta_I x_I \\ \dot{x}_P &= f_P(x_N, x_O) - \delta_P x_P \\ \dot{x}_O &= f_O(x_N, x_I, S) - \delta_O x_O \\ \dot{x}_N &= f_N(x_O, x_P, x_I, S) - \delta_N x_N\end{aligned}$$



How do I compare the data with the model (Likelihood)?



Results in a credible set of parameters that encode the network and can be used to test hypotheses in the future!

Beyond Example 3: Developmental GRNs / Neural tube patterning

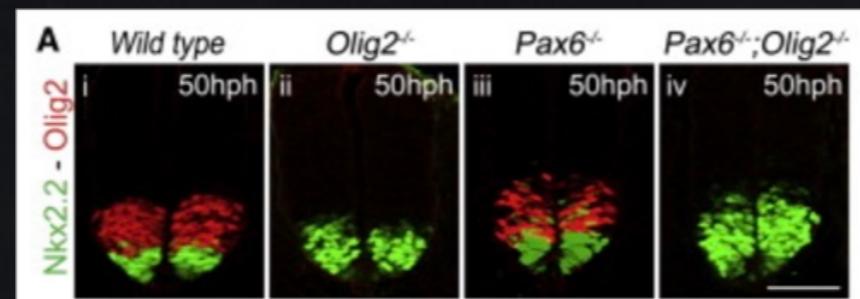
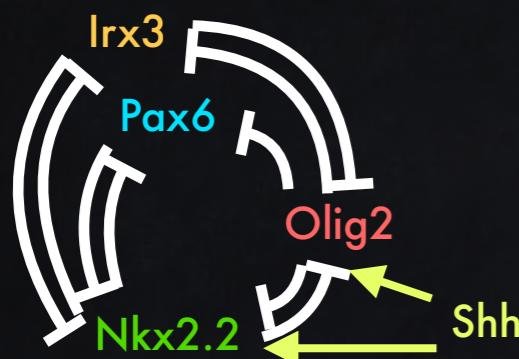


Human
30 days



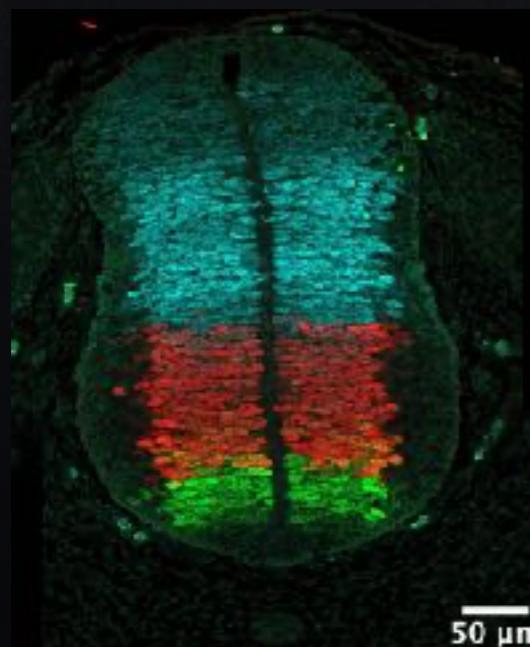
Mouse
9 days

Beyond Example 3: Developmental GRNs / Neural tube patterning

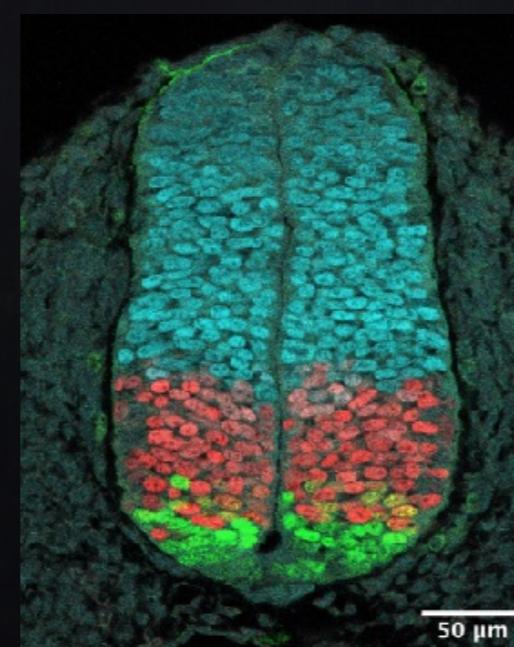


$$\begin{aligned}\dot{x}_I &= f_I(x_N, x_O) - \delta_I x_I \\ \dot{x}_P &= f_P(x_N, x_O) - \delta_P x_P \\ \dot{x}_O &= f_O(x_N, x_P, S) - \delta_O x_O \\ \dot{x}_N &= f_N(x_O, x_P, x_P, S) - \delta_1 x_N\end{aligned}$$

Results in a credible set of parameters that encode the network and can be used to test hypotheses in the future!

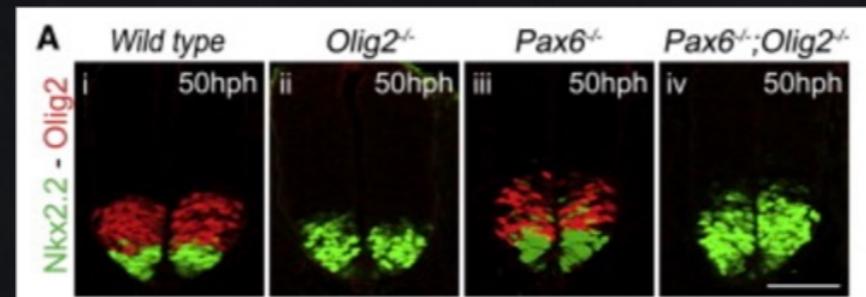
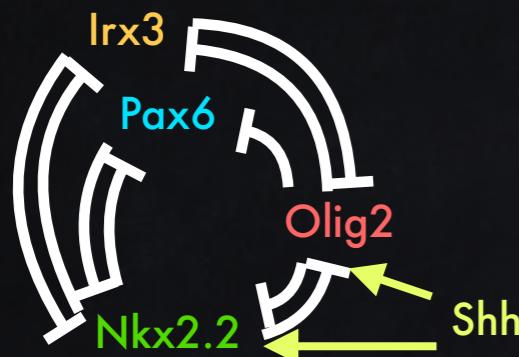


human



mouse

Beyond Example 3: Developmental GRNs / Neural tube patterning



$$\dot{x}_I = f_I(x_N, x_O) - \delta_I x_I$$
$$\dot{x}_P = f_P(x_N, x_O) - \delta_P x_P$$
$$\dot{x}_O = f_O(x_N, x_P, S) - \delta_O x_O$$
$$\dot{x}_N = f_N(x_O, x_P, x_P, S) - \delta_1 x_N$$

Results in a credible set of parameters that encode the network and can be used to test hypotheses in the future!

RESEARCH ARTICLE

Species-specific pace of development is associated with differences in protein stability

Teresa Rayon^{1,*}, Despina Stamatakis^{1,†}, Ruben Perez-Carrasco^{1,2,3,†}, Lorena Garcia-Perez¹, Christopher Barrington...

See all authors and affiliations

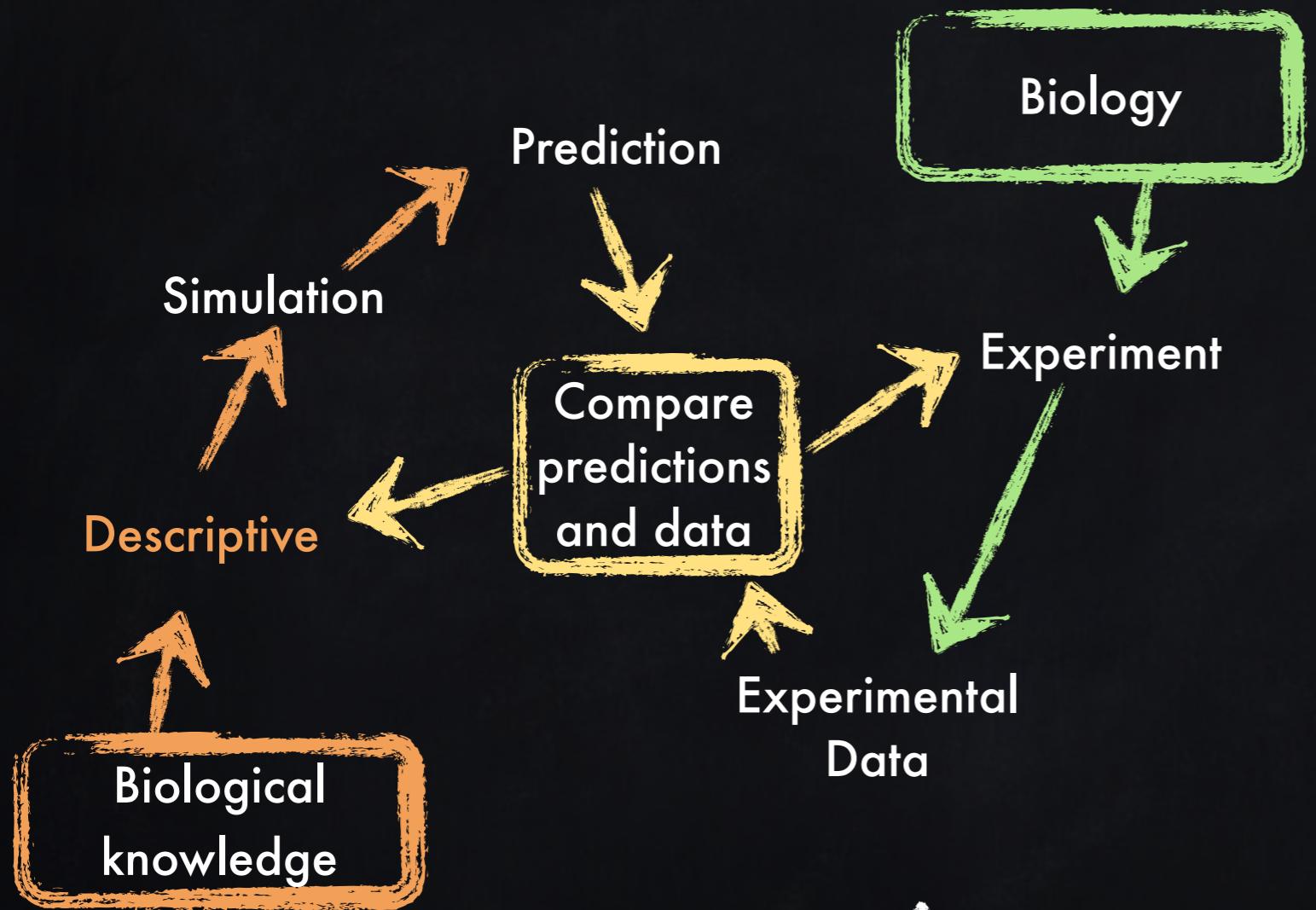
Science 18 Sep 2020:
Vol. 369, Issue 6510, eaba7667
DOI: 10.1126/science.eaba7667

Which? Where?

How?

When?

Summary



$$\frac{P(\delta)P(data|\delta)}{P(data)} = P(\delta|data) \quad \delta_2$$

