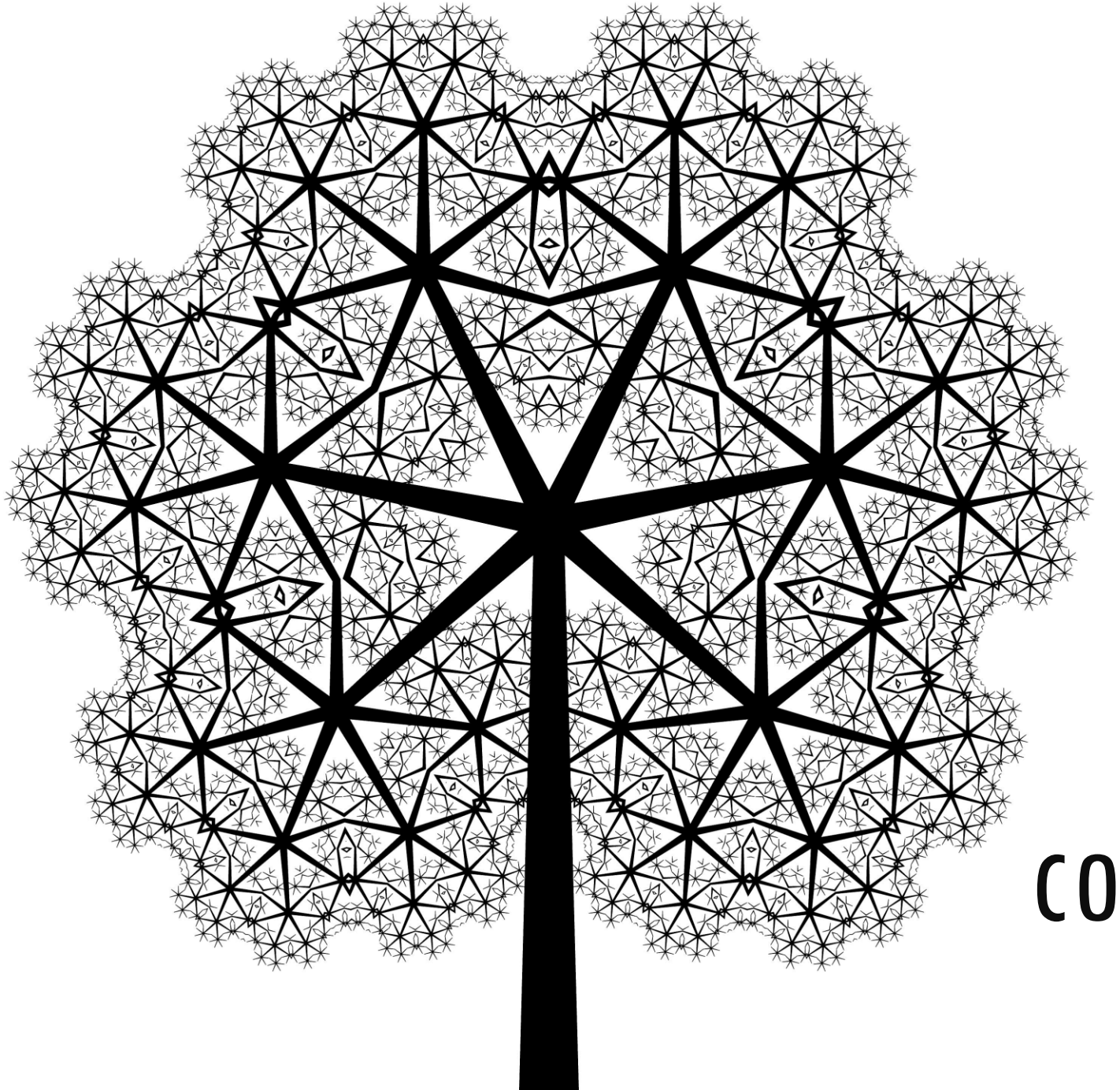
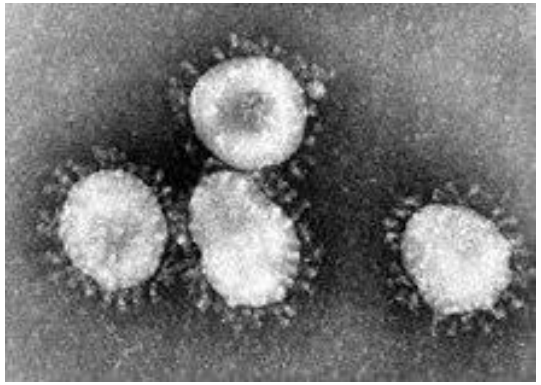
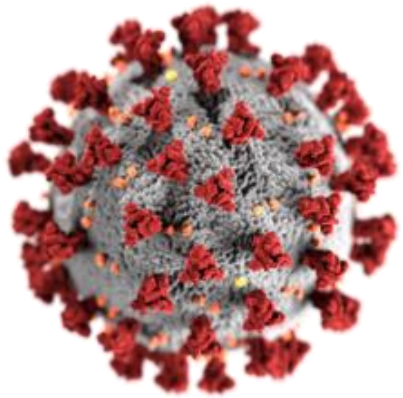


COVID19

Real time epidemic modelling

CORONAVIRUSES



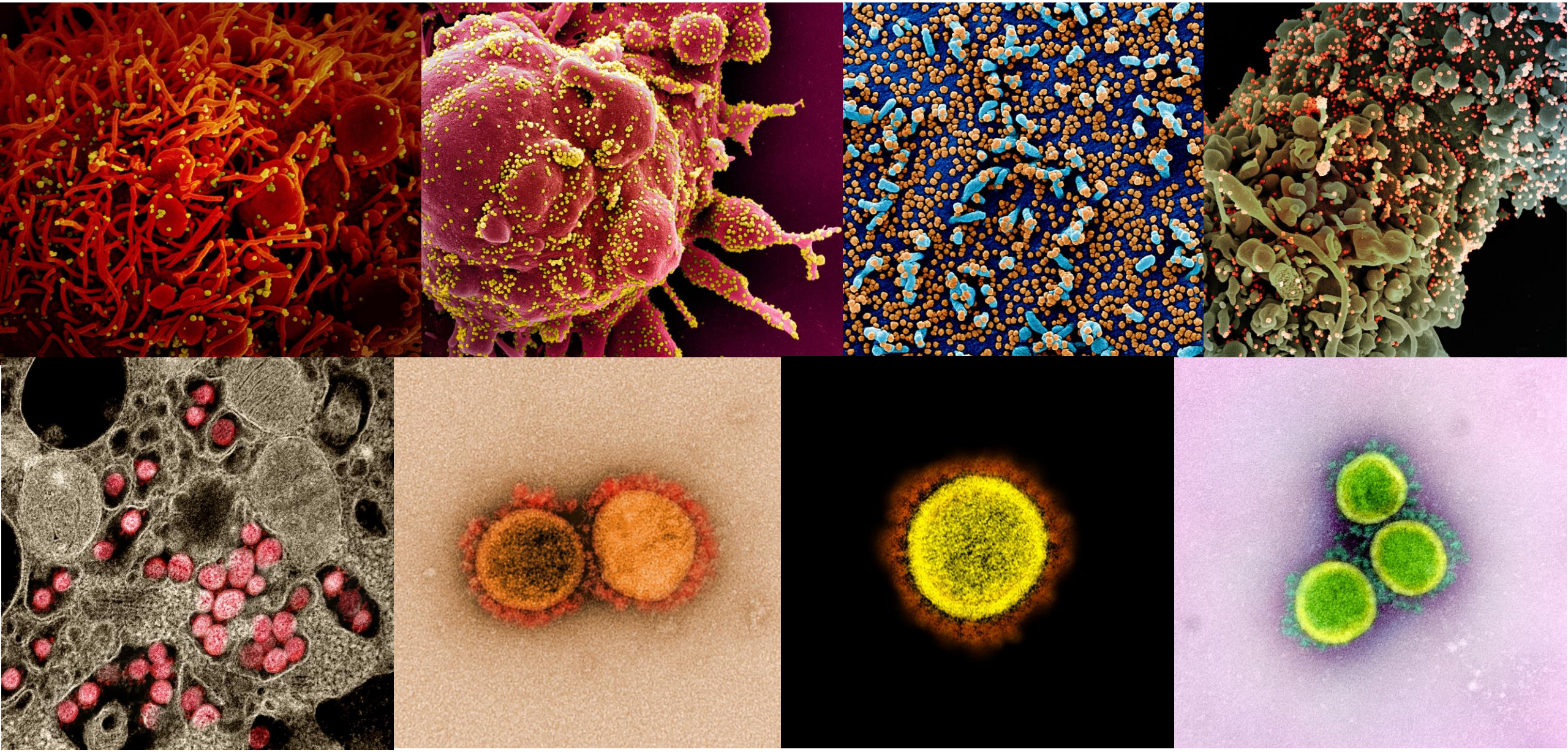


Coronaviruses are an extremely common class of virus.

- Responsible for a number of illnesses including the common cold, pneumonia, gastroenteritis
- Infect birds and mammals. Many harbor variants asymptomatically.
- Named for spike proteins resembling solar coronas : corona (Latin) “crown” or “wreath”
- We don’t know much about these viruses.

Coronavirus outbreaks can happen anywhere, any time.

- SARS (2002) and MERS (2012) were most recent epidemics
- Government & global health orgs routinely drill coronavirus emergencies
- COVID19 should not be surprising -- more or less inevitable and no one’s fault



SITUATION REPORT

The COVID19 pandemic originated near Hubei province, China sometime around November 2019 according to phylogenetic analyses.

The first identified case occurred in December 2019 in Wuhan.

As of today there are:

2,543,327 cumulative confirmed cases

176,527 cumulative deaths

686,440 recoveries

SITUATION REPORT

The COVID19 pandemic originated near Hubei province, China sometime around November 2019 according to phylogenetic analyses.

The first identified case occurred in December 2019 in Wuhan.

As of today there are:

2,543,327 cumulative confirmed cases

176,527 cumulative deaths

686,440 recoveries

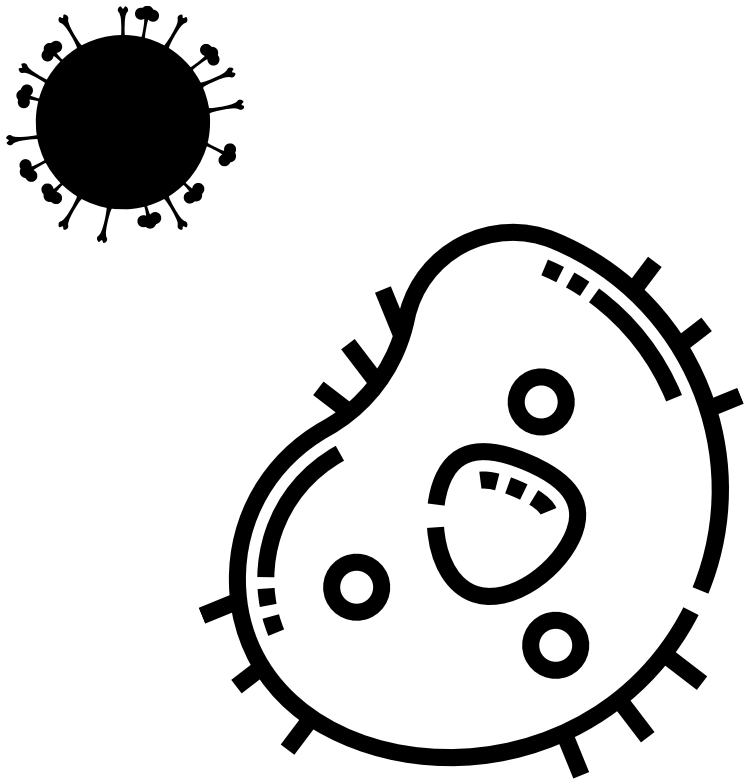
THE LIFE OF A VIRUS

A quick review of introductory concepts

ONE GRAIN OF RICE

A MATHEMATICAL FOLKTALE BY DEMI

ONE VIRUS

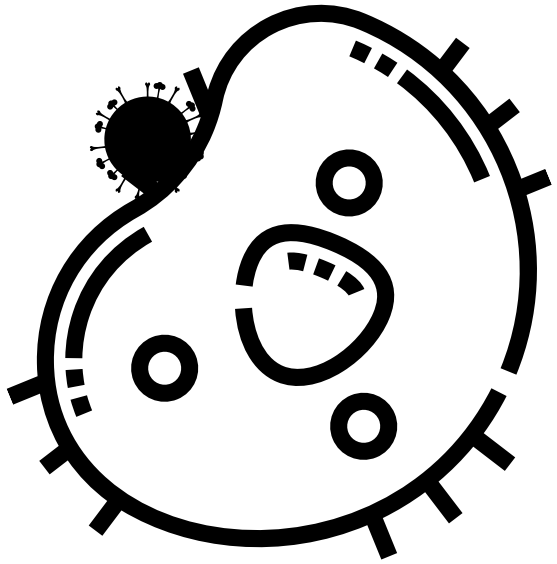


A single virus can initiate an infection if it finds the target cell.

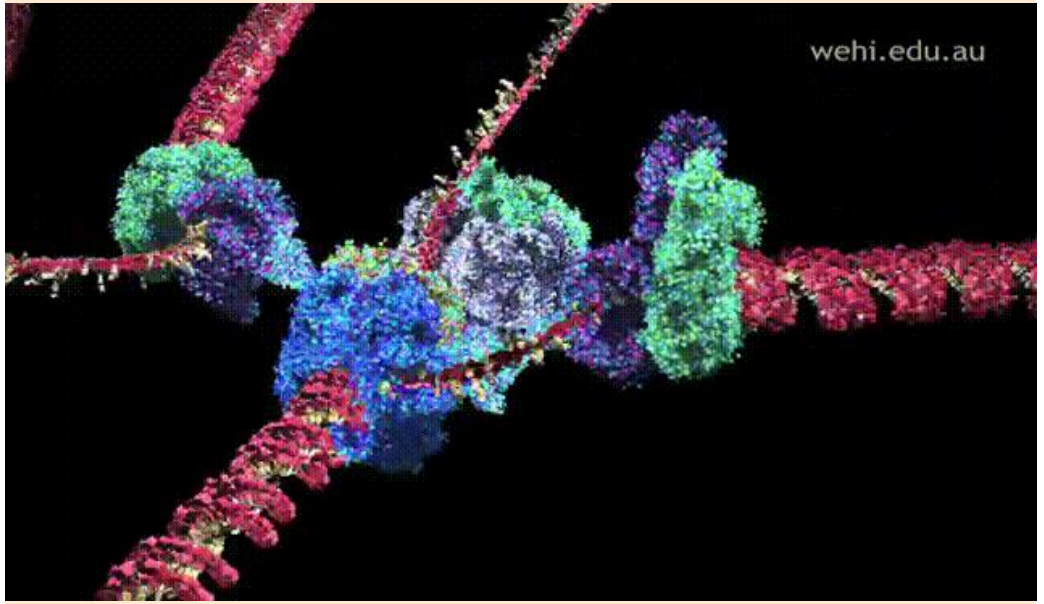
For SARSCoV2, target cells are those containing “ACE2” receptors that react with the external spike proteins.

These are generally lung or gastrointestinal membrane cells.

ONE VIRUS



After infiltrating the cell, it tricks it into replicating its RNA and assembling many copies of the virus.

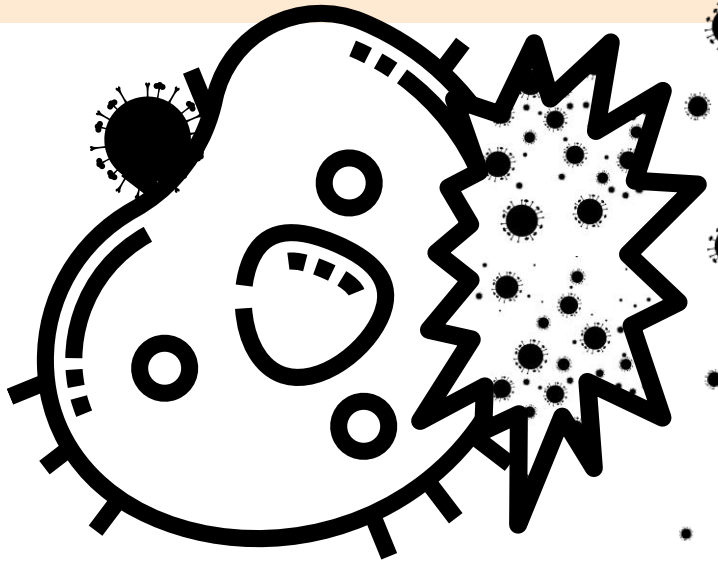


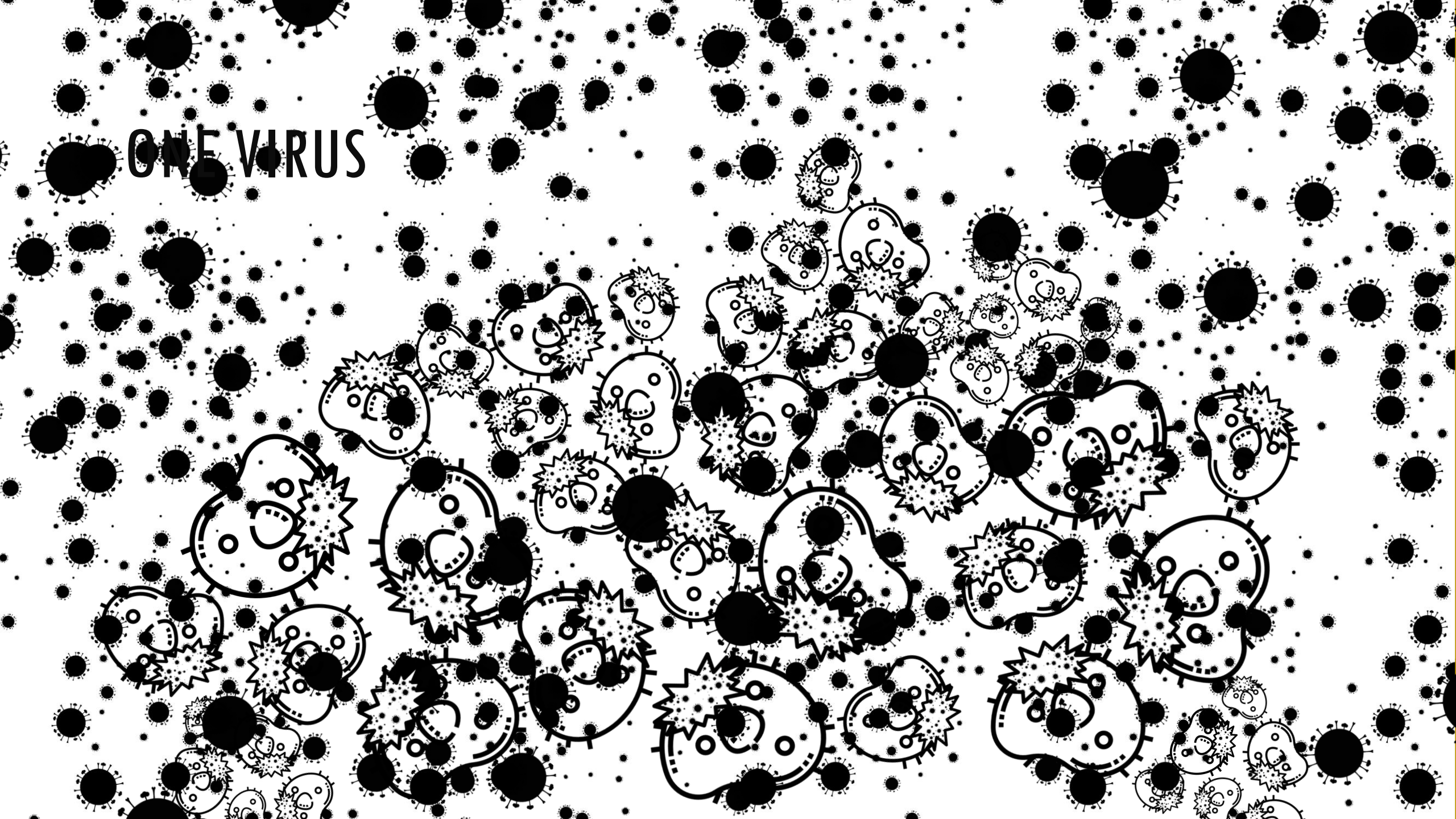
Mistakes made during this process can lead to mutations

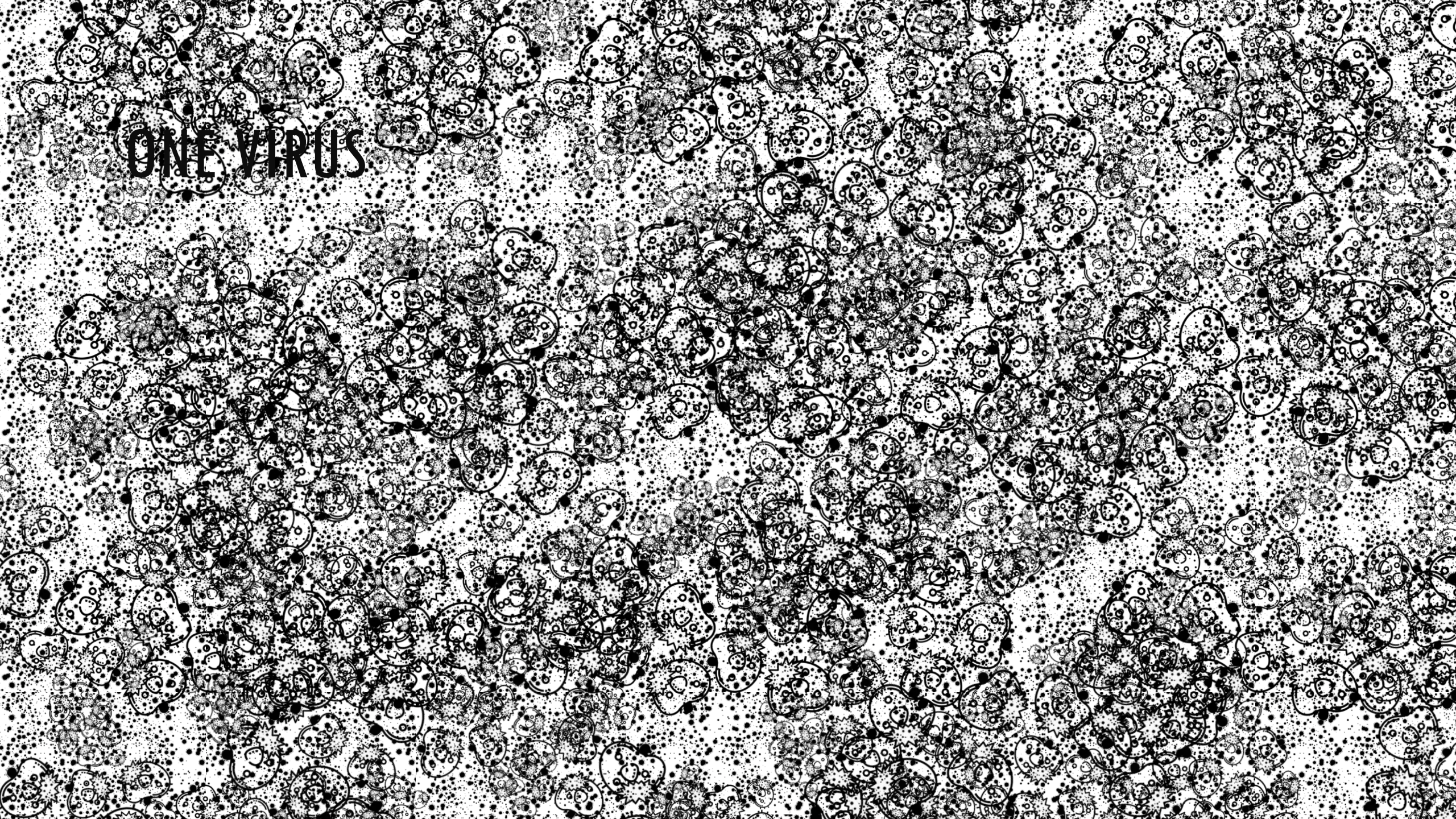
ONE VIRUS

The cell continues to produce viral replicas for the remainder of its life cycle, secreting them via exocytosis.

The cycle takes several hours, producing thousands of replicates each: $N_n \sim 10000^n$




ONE VIRUS


ONE VIRUS

ONE PERSON

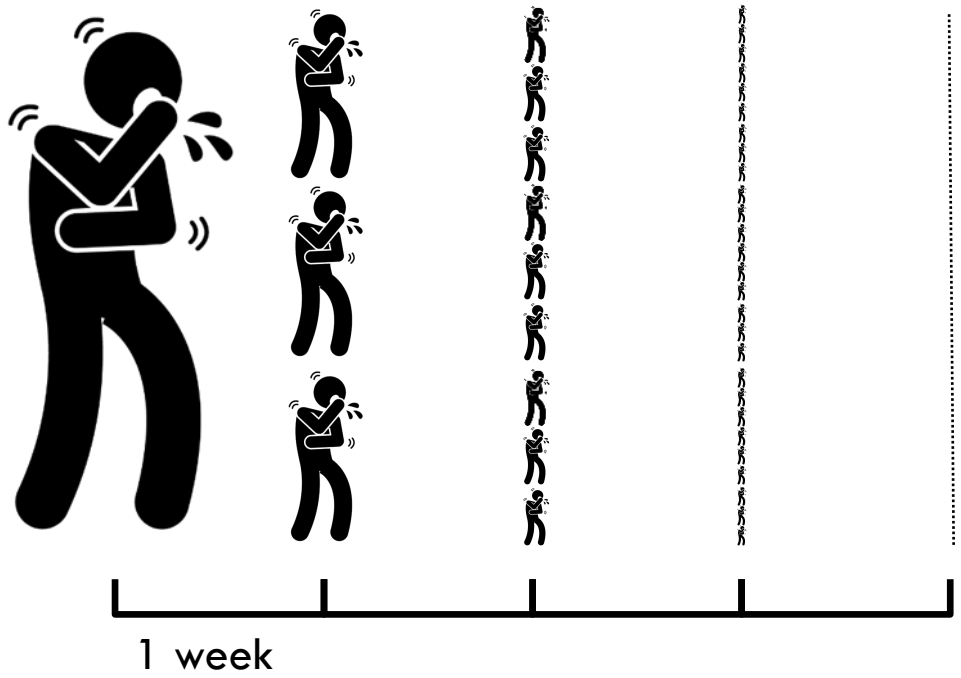
It takes between 3-14 days for symptoms to develop, dependent on a number of factors, possibly including but not limited to: initial viral load, individual immune response, blood type, infection pathway

ONE PERSON

Let's assume that over the course of the infection, one person transmits the virus to three new people on average.

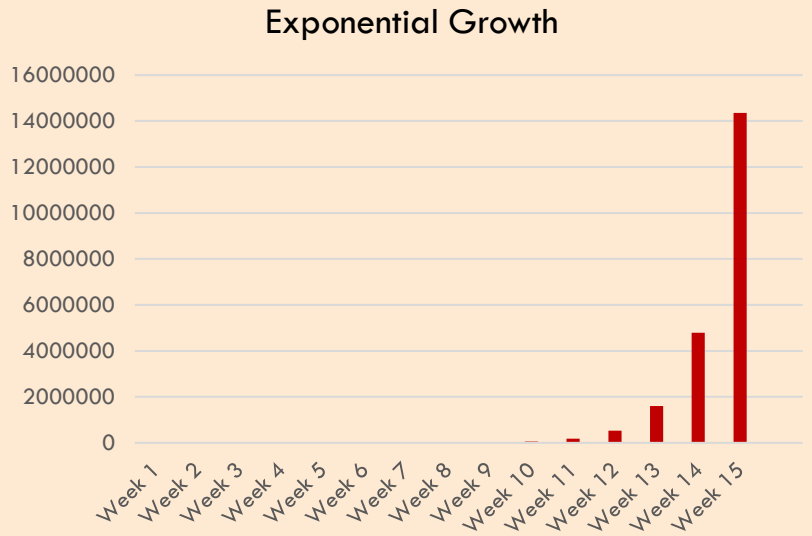
We'll take the course of infection to be one week for simplicity

ONE PERSON

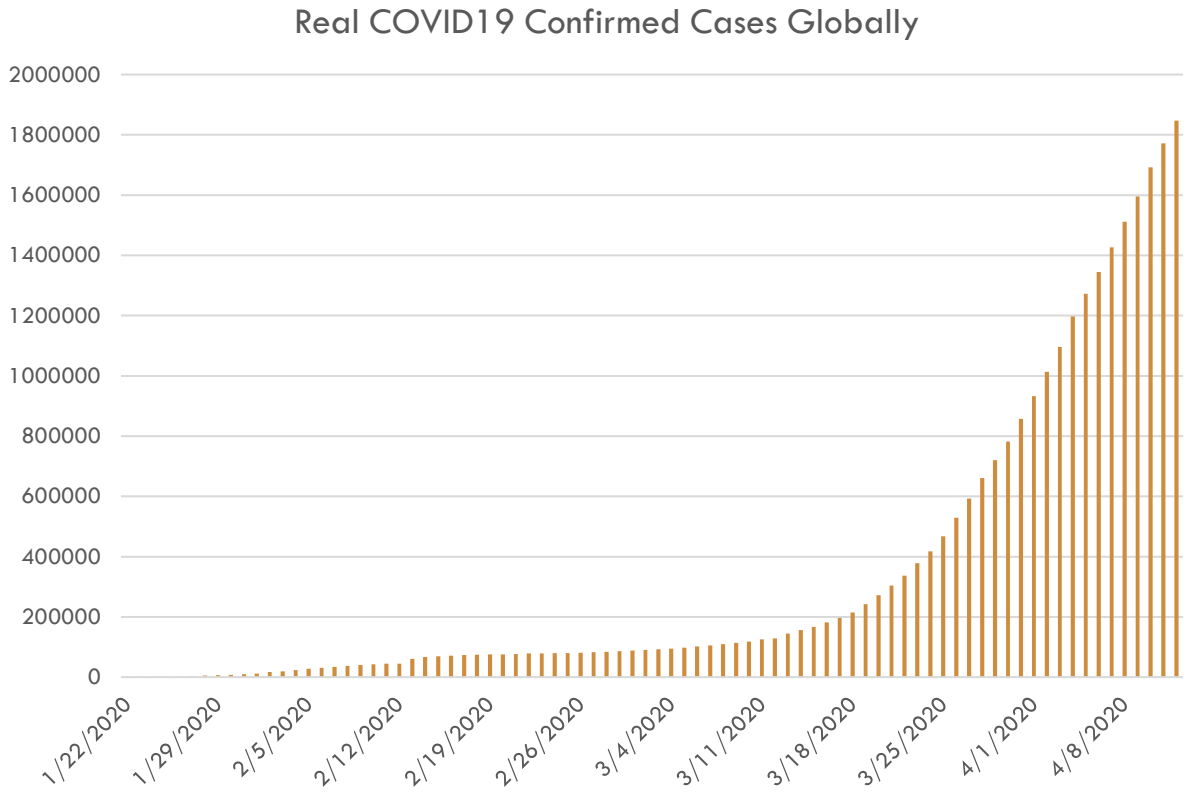


Early in the epidemic, infected grow exponentially in time!

$$N \sim 3^n$$

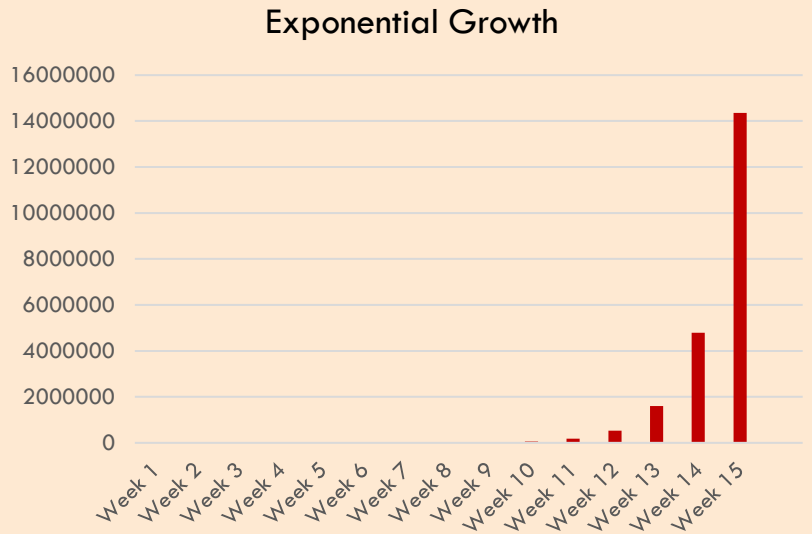


ONE PERSON

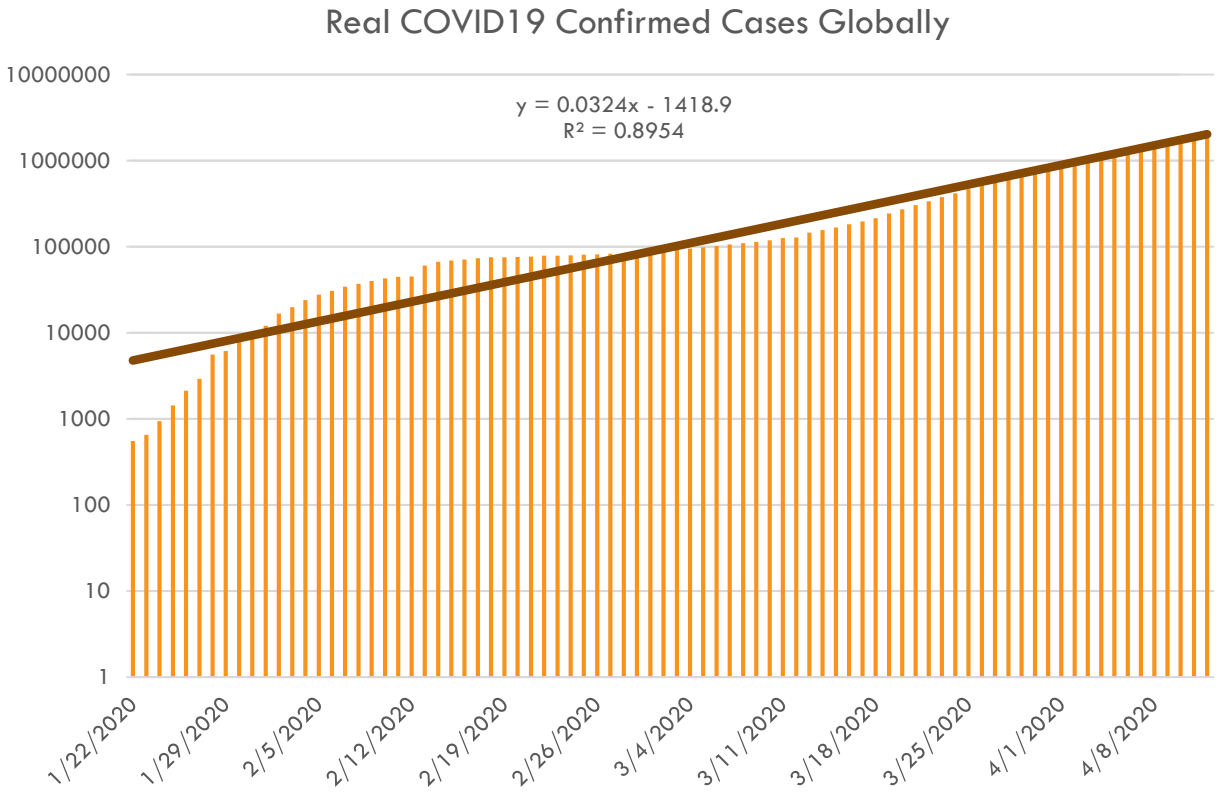


Compare shape to real data from Johns Hopkins

(https://github.com/CSSEGISandData/COVID-19)



ONE PERSON



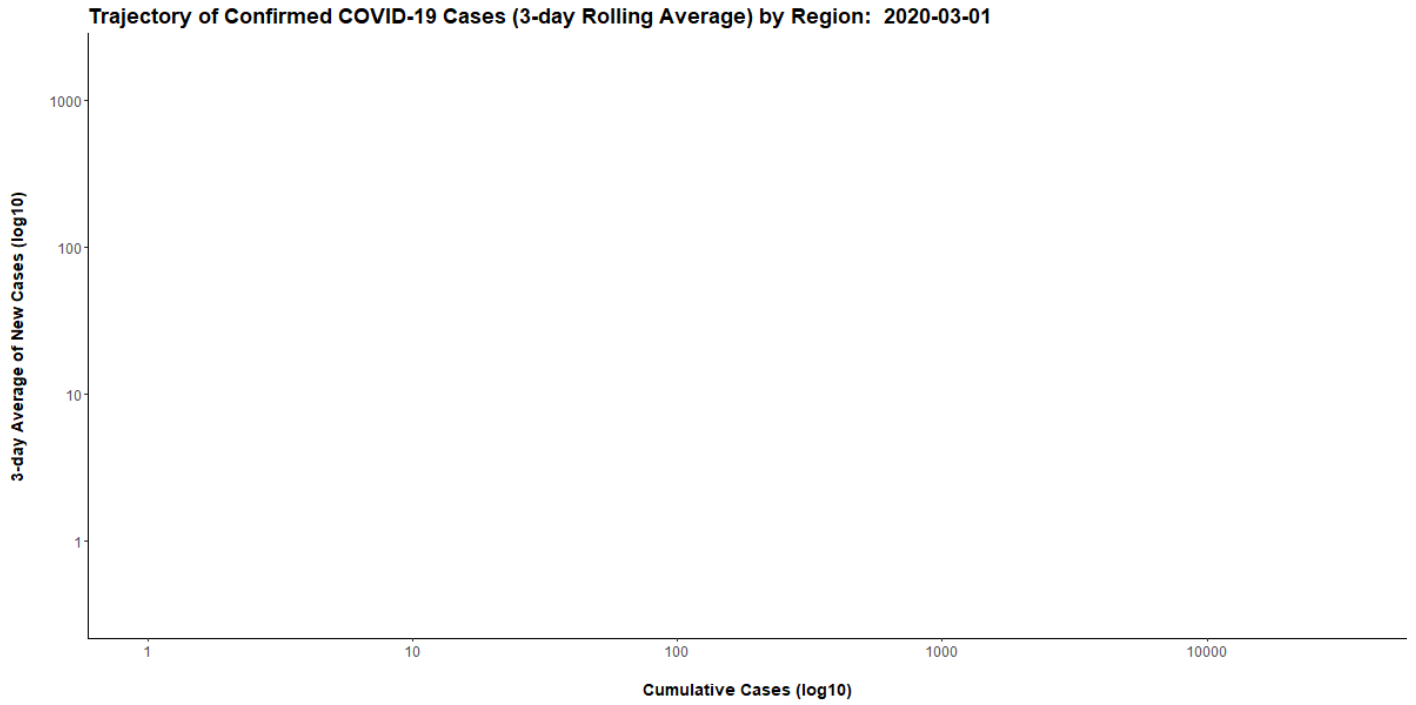
Compare shape to real data from Johns Hopkins

(Better to plot Logarithmically)

QUESTIONS

Finite number of humans on Earth!

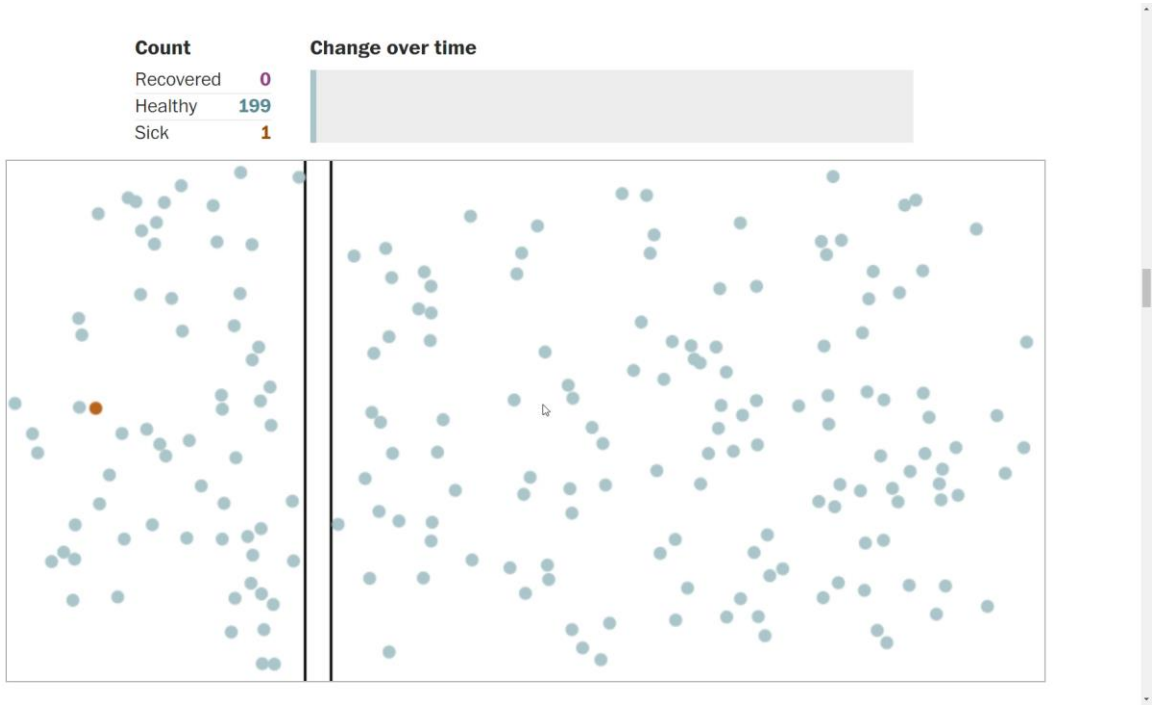
- Can't exponentially grow indefinitely – what is late phase behavior?
- What most significantly affects this trend?



QUESTIONS

How can we understand interventions?

- How do different interventions change behavior?
- When and under what conditions are different interventions effective?



QUESTIONS

Finite number of humans on Earth!

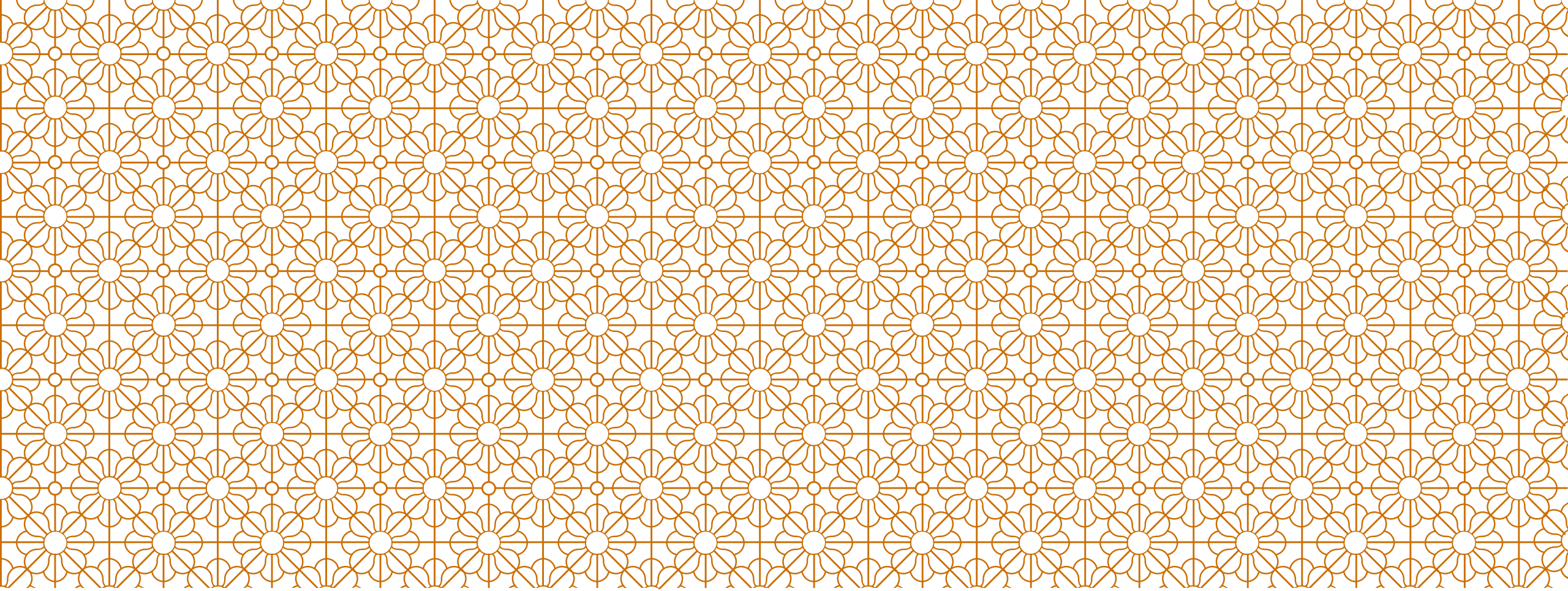
- Can't exponentially grow indefinitely – what is late phase behavior?
- What most significantly affects this trend?

How can we understand interventions?

- How do different interventions change behavior?
- When and under what conditions are different interventions effective?

What lies in the far future?

- Why are some diseases endemic – i.e. return periodically – like the flu? Will COVID19 become endemic?



MODELING EPIDEMICS

Growth, Spread, Diversity

SIMPLE EPIDEMIOLOGY

Starting Point: Simple exponential growth

First equation you solve in an ODE course: Change in population is proportional population size

EQUATION:

$$\frac{d}{dt} I(t) = k\, I(t)$$

SOLUTION:

$$\int \frac{dI}{I} = \int dt\, k$$

$$\ln\left(\frac{I(t)}{I_0}\right) = k\,(t - t_0)$$

$$\boxed{I(t) = I_0 e^{k(t-t_0)}}$$

PLOT:

SIMPLE EPIDEMIOLOGY

Correction 1: Unrealistic Numbers of New Infections!

The number of people "susceptible" to the virus monotonically decreases with infection!

STRATEGY:

- Keep track of them with S(t)
- Increase in I(t) $\propto$ S% and I(t)
- I' = - S'

NEW EQUATIONS:

$$\frac{d}{dt}S(t) = -k\frac{S(t)}{N}I(t)$$

$$\frac{d}{dt}I(t) = \quad k\frac{S(t)}{N}I(t)$$

SOLUTION:

$$S(t) = S_0 - I(t)$$

$$I(t) = \frac{I_0 e^{\frac{I_0 k}{N}}}{e^{\frac{I_0 k}{N}} - e^{I_0 S_0}}$$

SIMPLE EPIDEMIOLOGY

Correction 2: Unrealistic Numbers of Old Infections!

The number of infected decreases via recovery or death!

STRATEGY:

- Keep track of them with R(t)
- Decrease in I(t) $\propto$ I(t)
- Increase in R(t) $\propto$ I(t)

NEW EQUATIONS:

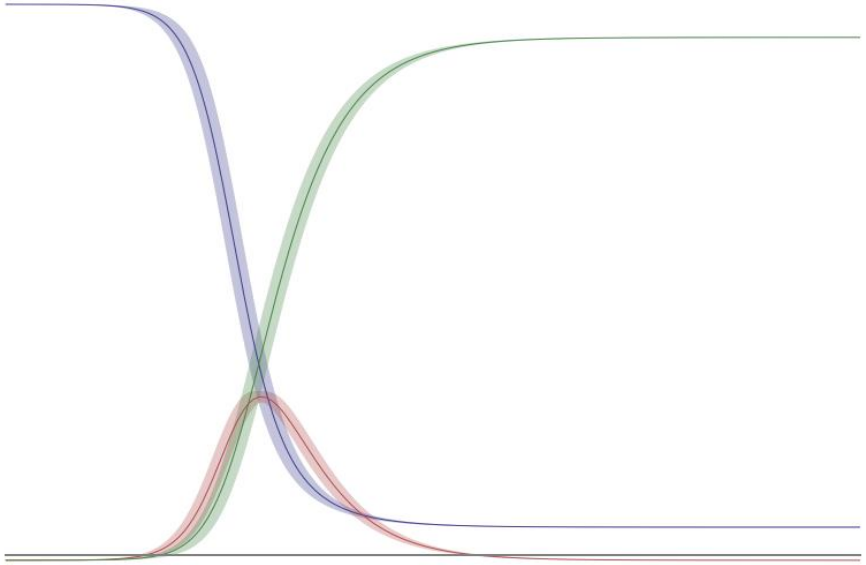
$$\frac{d}{dt}S(t) = -k\frac{S(t)}{N}I(t)$$

$$\frac{d}{dt}I(t) = k\frac{S(t)}{N}I(t) - \gamma\, I(t)$$

$$\frac{d}{dt}R(t) = \gamma\, I(t)$$

SOLUTION:

Nonlinear: No general solution!



SIMPLE EPIDEMIOLOGY

Correction 2: Unrealistic Numbers of Old Infections!

The number of infected decreases via recovery or death!

"SIR MODEL":

- Starting Point for epi models
- Special constant:

$$R_0 = \frac{k}{\gamma}$$

"Basic Reproduction Number"

- R_0 estimates # new infections per current infection
- What happens when $R_0<1$?

NEW EQUATIONS:

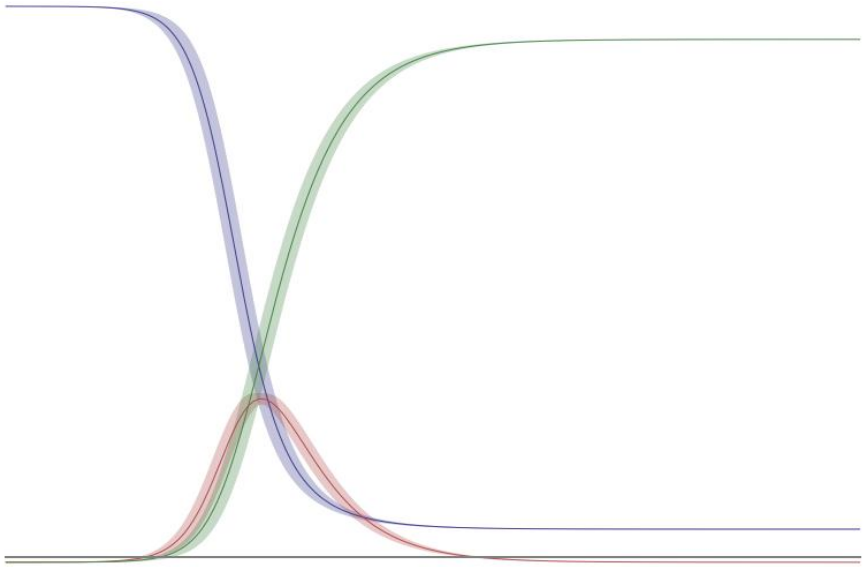
$$\frac{d}{dt}S(t) = -k\frac{S(t)}{N}I(t)$$

$$\frac{d}{dt}I(t) = k\frac{S(t)}{N}I(t) - \gamma\, I(t)$$

$$\frac{d}{dt}R(t) = \gamma\, I(t)$$

SOLUTION:

Nonlinear: No general solution!



SIMPLE EPIDEMIOLOGY

Correction N: Lots of variations!

Lots of variations to study— invent your own or visit Wikipedia "Compartmental Models in Epidemiology"

REALISM:

- Exposed + Age classes
- Spatial diffusion
- Random R_0
- Seasonal Forcing
- Mutant spread (*)

MITIGATION:

- Quarantine Measures
- Contact Tracing
- Vaccination
- Herd immunity

FORECASTING:

- Resurgent Waves from Imports
- Containment v.s. Endemic (*)

SIMPLE EPIDEMIOLOGY

Important points:

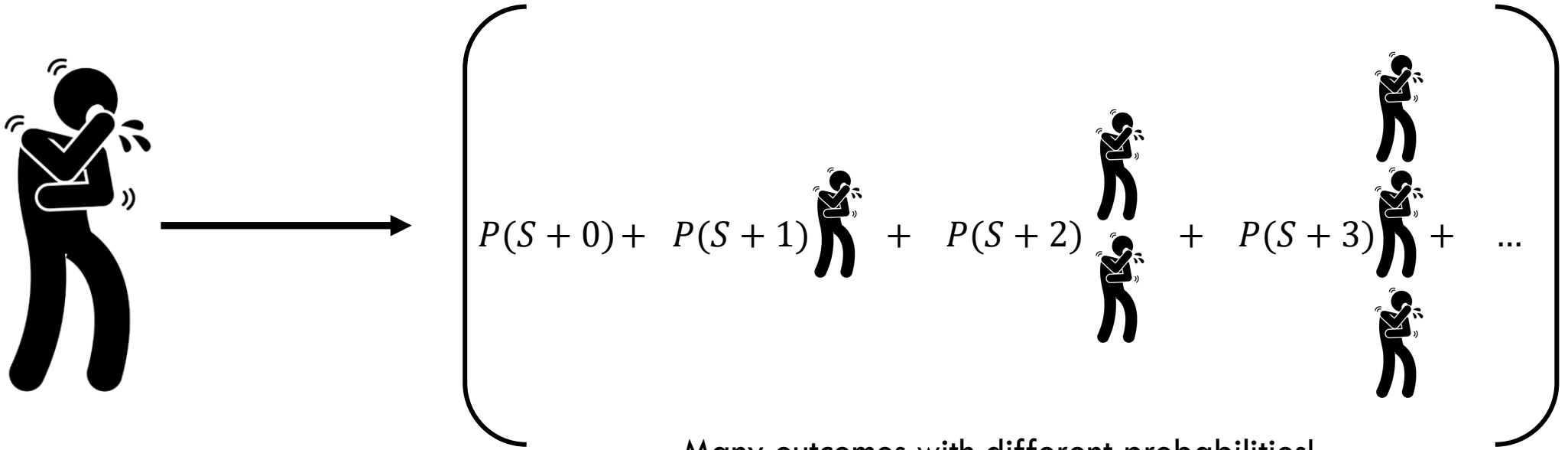
- The SIR model is a "Mean Field" equation: best applicable to large, dense cities
- Does not capture random stochastic behaviour, or directly tell you about variance
- Long term predictions are likely qualitatively wrong, especially when ignoring spatial diffusion*

*(the model has an upper critical dimension of 2 – a subtle technical point beyond the scope of undergraduate level courses!)

EPIDEMIOLOGY CONT'D

Defining a Microscopic Model: Real life is **stochastic**

A more realistic simulation needs to account for randomness.



Many outcomes with different probabilities!

EPIDEMIOLOGY CONT'D

Revisit simple exponential growth

Let's learn to define a model stochastically through a *master equation. Recall:*

$$dI(t) = k\, I(t)dt$$

COMPONENTS OF A STOCHASTIC MODEL:

- A set of possible states
- Probability per unit time (rate) of transitions between two given states
- Probability of being in a given state

IN SIMPLE EXP GROWTH:

- The states are labeled by the # of infected, I
- Transition probability from a state I to a state I+1:
$$p_{I+1|I}(\Delta t) = k\, I\Delta t$$
- The probability of being in state I at time t
$$P_I(t)$$

EPIDEMIOLOGY CONT'D

Construction:

Starting with $I_0 = 1$ at $t = 0$, we want to know the probability of being in state I at time t

$P_I(t) =?$

1. Markov Assumption: This can only depend on the state of the system immediately before time t

EPIDEMIOLOGY CONT'D

Construction:

Starting with $I_0 = 1$ at $t = 0$, we want to know the probability of being in state I at time t

$P_I(t) =?$

1. Markov Assumption: This can only depend on the state of the system immediately before time t
2. "Immediately" means a time step, δt, so small that only one transition can happen

EPIDEMIOLOGY CONT'D

Construction:

Starting with $I_0 = 1$ at $t = 0$, we want to know the probability of being in state I at time t

$P_I(t) =?$

1. Markov Assumption: This can only depend on the state of the system immediately before time t
2. "Immediately" means a time step, δt, so small that only one transition can happen

Conclusion:

At $t - \delta t$ we were either in a state with I or I-1

EPIDEMIOLOGY CONT'D

Construction:

Starting with $I_0 = 1$ at $t = 0$, we want to know the probability of being in state I at time t

$$P_I(t) = p_{I|I-1}(\delta t)P_{I-1}(t - \delta t) + p_{I|I}(\delta t)P_I(t - \delta t)$$

1. Markov Assumption: This can only depend on the state of the system immediately before time t
2. "Immediately" means a time step, δt, so small that only one transition can happen

Conclusion:

At $t - \delta t$ we were either in a state with I or I-1

EPIDEMIOLOGY CONT'D

Construction:

Starting with $I_0 = 1$ at $t = 0$, we want to know the probability of being in state I at time t

The probability that we gain a new infected is proportional to number of infected:

$$\mathrm{k}\delta t(I-1)$$

$$P_I(t) = p_{I|I-1}(\delta t)P_{I-1}(t-\delta t) + p_{I|I}(\delta t)P_I(t-\delta t)$$

This is just the probability that nothing happens in a time δt:

$$1 - \mathrm{k}\delta t\, I$$

1. Markov Assumption: This can only depend on the state of the system immediately before time t
2. "Immediately" means a time step, δt, so small that only one transition can happen

Conclusion:

At $t - \delta t$ we were either in a state with I or I-1

EPIDEMIOLOGY CONT'D

Construction:

Starting with $I_0 = 1$ at $t = 0$, we want to know the probability of being in state I at time t

The probability that we gain a new infected is proportional to number of infected:

$$\mathrm{k}\delta t(I-1)$$

$$P_I(t) = k\delta t(I-1)P_{I-1}(t-\delta t) + (1-k\delta t\, I)P_I(t-\delta t)$$

This is just the probability that nothing happens in a time δt:

$$1 - \mathrm{k}\delta t\, I$$

1. Markov Assumption: This can only depend on the state of the system immediately before time t
2. "Immediately" means a time step, δt, so small that only one transition can happen

Conclusion:

At $t - \delta t$ we were either in a state with I or I-1

EPIDEMIOLOGY CONT'D

Construction:

Starting with $I_0 = 1$ at $t = 0$, we want to know the probability of being in state I at time t

$$P_I(t) = k\delta t(I-1)P_{I-1}(t-\delta t) + (1 - k\delta t\, I)P_I(t-\delta t)$$

Rearrange the equation

$$\frac{P_I(t) - P_I(t-\delta t)}{\delta t} = k(I-1)P_{I-1}(t-\delta t) - kIP_I(t-\delta t)$$

Left hand side is the definition of a derivative, take limit

$$\frac{\mathrm{d}}{\mathrm{d}t}P_I(t) = k(I-1)P_{I-1}(t) - kIP_I(t)$$

Result is a "Master Equation"

EPIDEMIOLOGY CONT'D

The average solution is the ODE:

We take a quick detour to demonstrate that the mean solution is equivalent to the simple epi ODE

$$\frac{\mathrm{d}}{\mathrm{d}t}P_I(t) = k(I-1)P_{I-1}(t) - kIP_I(t)$$

$$\partial_t \bar{I}(t) = \sum_I I \frac{\mathrm{d}}{\mathrm{d}t}P_I(t) = \sum_I I\, k(I-1)P_{I-1}(t) - \sum_I kI^2 P_I(t)$$

$$= \sum_I (I+1)kIP_I(t) - \sum_I kI^2 P_I(t)$$

$$= \sum_I kIP_I(t)$$

$$\boxed{\partial_t \bar{I}(t) = k\,\bar{I}(t)}$$

Expectation value (avg) is defined as

$$\bar{x} = \sum_i x_i p_i$$

Shift indices

Simplify

Woah!

EPIDEMIOLOGY CONT'D

The Master Equation:

Aptly named, the Master Equation is all that's needed to calculate any statistical quantity

$$\frac{\mathrm{d}}{\mathrm{d}t}P_I(t) = k(I-1)P_{I-1}(t) - kIP_I(t)$$

Probabilities for particular states can be calculated recursively.

Example 1: Probability that I=1 at time t

$$\frac{\mathrm{d}}{\mathrm{d}t}P_1(t) = k(0)P_0(t) - kP_1(t)$$

$$= -kP_1(t)$$

Solving ODE with known initial condition $P_1(0) = 1$

$$P_1(t) = e^{-k\,t}$$

EPIDEMIOLOGY CONT'D

The Master Equation:

Aptly named, the Master Equation is all that's needed to calculate any statistical quantity

$$\frac{\mathrm{d}}{\mathrm{d}t}P_I(t) = k(I-1)P_{I-1}(t) - kIP_I(t)$$

Probabilities for particular states can be calculated recursively.

Example 2: Probability that I=2 at time t

$$\frac{\mathrm{d}}{\mathrm{d}t}P_2(t) = kP_1(t) - 2kP_2(t)$$

$$= ke^{-k\,t} - 2kP_2(t)$$

Solving ODE with known initial condition $P_1(0) = 1$

$$P_2(t) = \frac{kt}{2}\, e^{-k\,t}$$

(Hint: For any n, the result will be Poisson Statistics!)

EPIDEMIOLOGY CONT'D

The Master Equation:

Most importantly, the Master Equation is an *Algorithm*

$$\frac{\mathrm{d}}{\mathrm{d}t}P_I(t) = k(I-1)P_{I-1}(t) - kIP_I(t)$$

```
for t in range(1,t_max):
    random_number = random.random()
    if random_number>(1-k*I[t-1]):
        I[t]= I[t-1]+1
    else:
        I[t]=I[t-1]
```

Once you have the Master Equation, you merely have to translate it to code.

These are extremely useful tools for any simulation work.

EPIDEMIOLOGY CONT'D

SIR Master Equation:

Be sure to derive this in the group notebook!

$$\frac{\mathrm{d}}{\mathrm{d}t}P_{S,I,R}(t) = \frac{k}{N}(I-1)(S+1)P_{S+1,I-1,R}(t) + \gamma(I+1)P_{S,I+1,R-1}(t) - (\frac{k}{N}IS + \gamma I)P_{SIR}(t)$$

It looks complicated but deriving this following the simple case is an instructive and clarifying exercise.

This will be the starting point for your algorithm!

Once you understand how to derive this, developing variations will be simple

EPIDEMIOLOGY CONT'D

SIR Master Equation:

Be sure to derive this in the group notebook!

$$\frac{\mathrm{d}}{\mathrm{d}t}P_{S,I,R}(t) = \frac{k}{N}(I-1)(S+1)P_{S+1,I-1,R}(t) + \gamma(I+1)P_{S,I+1,R-1}(t) - (\frac{k}{N}IS + \gamma I)P_{SIR}(t)$$

Suggested Variations

- Same as ODE version, sans the spatial dependence (this is a shockingly difficult extension)
- Statistical analysis is easy, can show how variance changes with different initial conditions, extensions
- Stochastic mutant growth and spread (*)
- More advanced studies (potentially **publication worthy**) for highly motivated students with programming experience are possible

THIS WEEK'S AGENDA

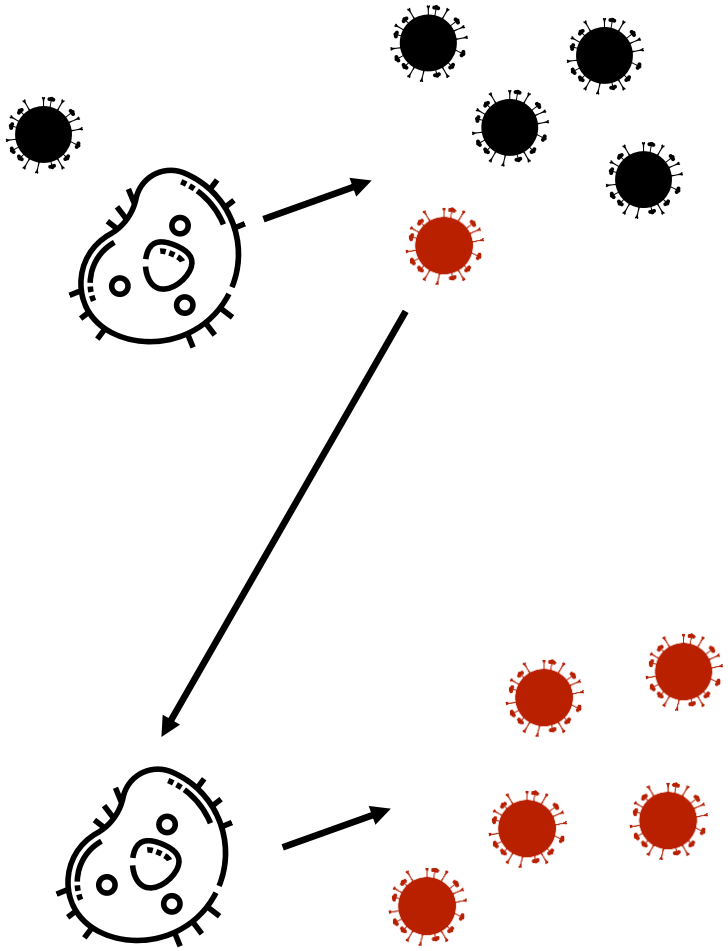
1. !! Decide which method you want to pursue !!
2. In the Group Notebook:
   - Derive SI and SIR models (ODE people)
   - Derive their Master Equations (Markov people)
   - Research possible implementations (ODE) or devise a for-loop in pseudocode (Markov) and summarize how these work, possible issues etc.
   - Find some sources of data that you will later compare your model to
   - Do some writing to summarize the project so e.g. another person in this course would be able to follow the notebook
3. On your own:
   - Start working on your basic algorithm and test proposed approaches
   - Try to write good code that is well-commented can be augmented easily
   - Try to get to a point where you can produce graphs for S, I, and R.

TODAY

I will stay on Zoom until 5:00 pm to work with everyone and answer questions. We can:

- Talk about technical/mathematical details
- Look through algorithms together
- Search for data sources
- Start filling in the group notebook

ANCESTRY

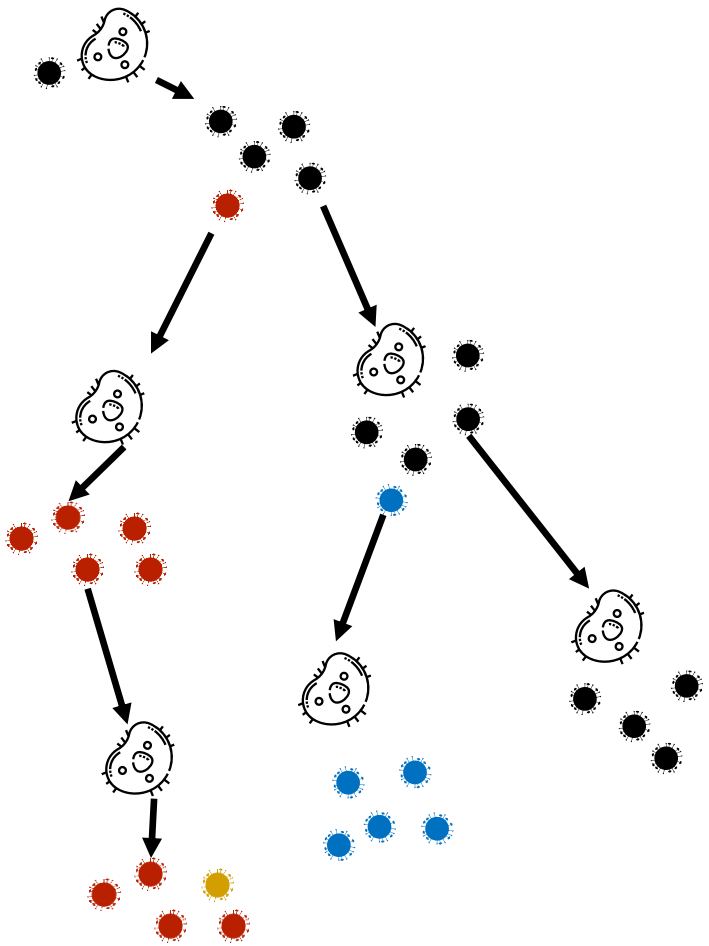


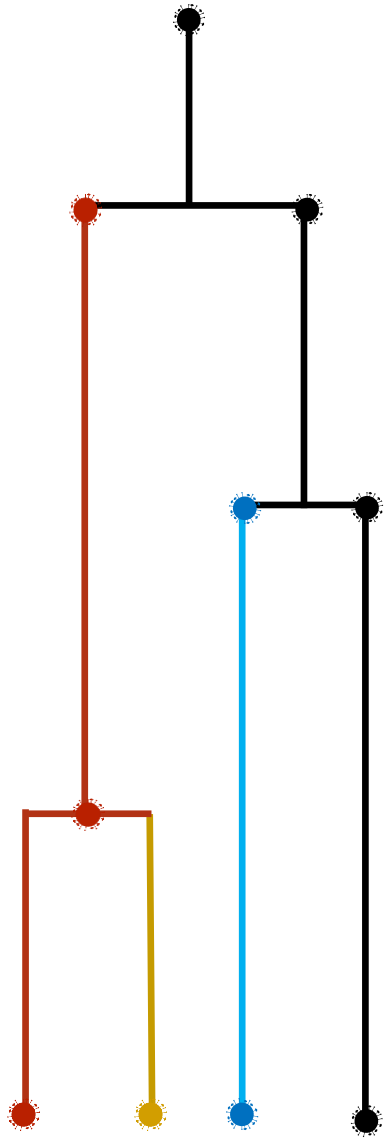
When viruses replicate, there is some chance that an error places the incorrect nucleotide.

This creates a “mutant”

A mutant that is capable of functioning and reproducing can fix in the population

ANCESTRY



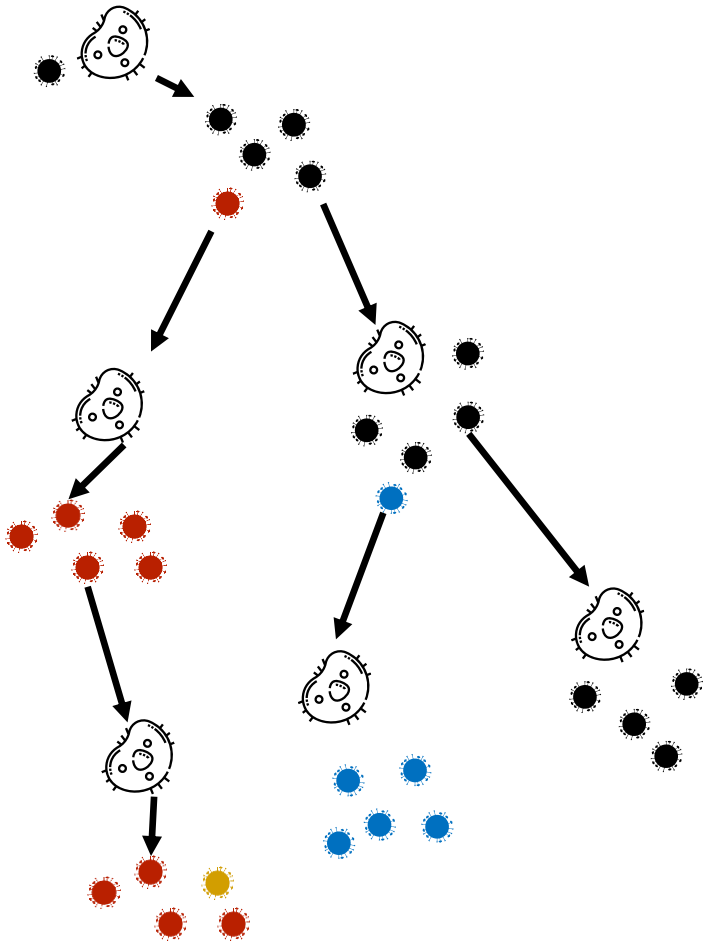


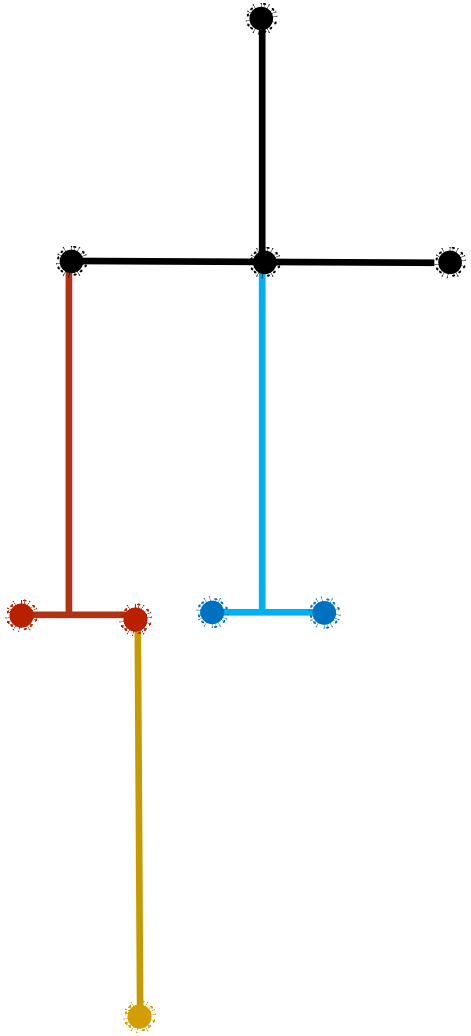
The history of these mutations at the genetic level is called a phylogeny.

It is usually represented on a tree

Here we show a time resolved tree

ANCESTRY

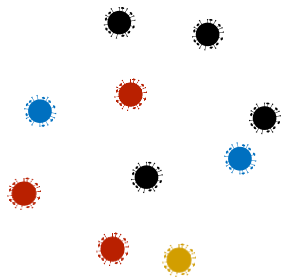




When the time of events isn't clear, trees can be plotted against "divergence", or mutations per nucleotide.

This is the most common display

ANCESTRY



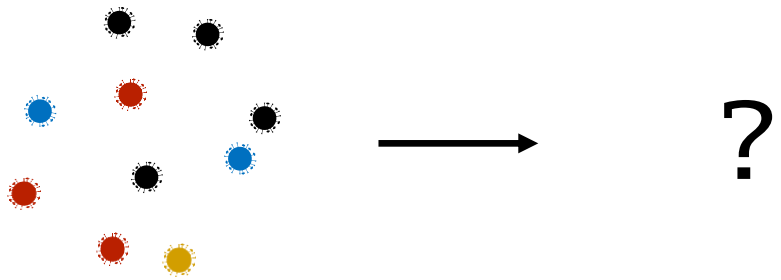
?

Given a set of sequences, we'd like to reconstruct this tree.

Problem:
Sequences are just random strings of letters
(A, T , C, G)

How do we infer relationships?

ANCESTRY



?

Problem:
Sequences (genomes) are just random strings of letters
(A, T , C, G)

How do we infer relationships?

Idea 1:
Define a notion of “distance” between two genomes based on the number of differences between then

ANCESTRY

The Jukes-Cantor Model:

Assume that the nucleotide mistakes are a markov process

$$P_n(t) = \sum_{n'} p_{n|n'} P_{n'}(t - \delta t) + \left(1 - \sum_{n'} p_{n',n}(\delta t)\right) P_n(t - \delta t)$$

$$\partial_t P_n(t) = \mu - 4\mu P_n(t)$$

$$\vec{P}\,(t) = \begin{pmatrix} P_A(t) \\ P_T(t) \\ P_C(t) \\ P_G(t) \end{pmatrix} = \begin{pmatrix} 1 - \frac{4\mu}{4} & \frac{\mu}{4} & \frac{\mu}{4} & \frac{\mu}{4} \\ 1 - \frac{4\mu}{4} & \frac{\mu}{4} & \frac{\mu}{4} & \frac{\mu}{4} \\ 1 - \frac{4\mu}{4} & \frac{\mu}{4} & \frac{\mu}{4} & \frac{\mu}{4} \\ 1 - \frac{4\mu}{4} & \frac{\mu}{4} & \frac{\mu}{4} & \frac{\mu}{4} \end{pmatrix}$$