

TRACKING VIRAL OUTBREAKS IN REAL TIME

John Huddleston
jlhudd@uw.edu

WHY ARE WE HERE?

WHY ARE WE HERE?

- **Connect** the fields of programming and biology

WHY ARE WE HERE?

- **Connect** the fields of programming and biology
- **Use** real tools for computational biology

WHY ARE WE HERE?

- **Connect** the fields of programming and biology
- **Use** real tools for computational biology
- **Explore** patterns in virus evolution

WHO AM I?

- Ph.D. student in a virus evolution lab at the Fred Hutch
- Professional programmer for 8 years
- M.S. in computer science and biology

WHO ARE YOU?

- Rising 10th and 11th grade students
- Programmers?
- Scientists?

THE PLAN FOR TODAY

1. Introductions
2. Reconstructing the Ebola outbreak with freely available tools
3. Exploring the Ebola outbreak with Nextstrain

OUTCOMES FOR THIS WORKSHOP

At the end of this workshop you will know how to:

- Identify viral sequences from online databases
- Construct a multiple sequence alignment
- Build a phylogenetic tree
- Interpret ancestral relationships between organisms in a tree

ORIGINS OF AN OUTBREAK



QUESTIONS WE NEED TO ANSWER

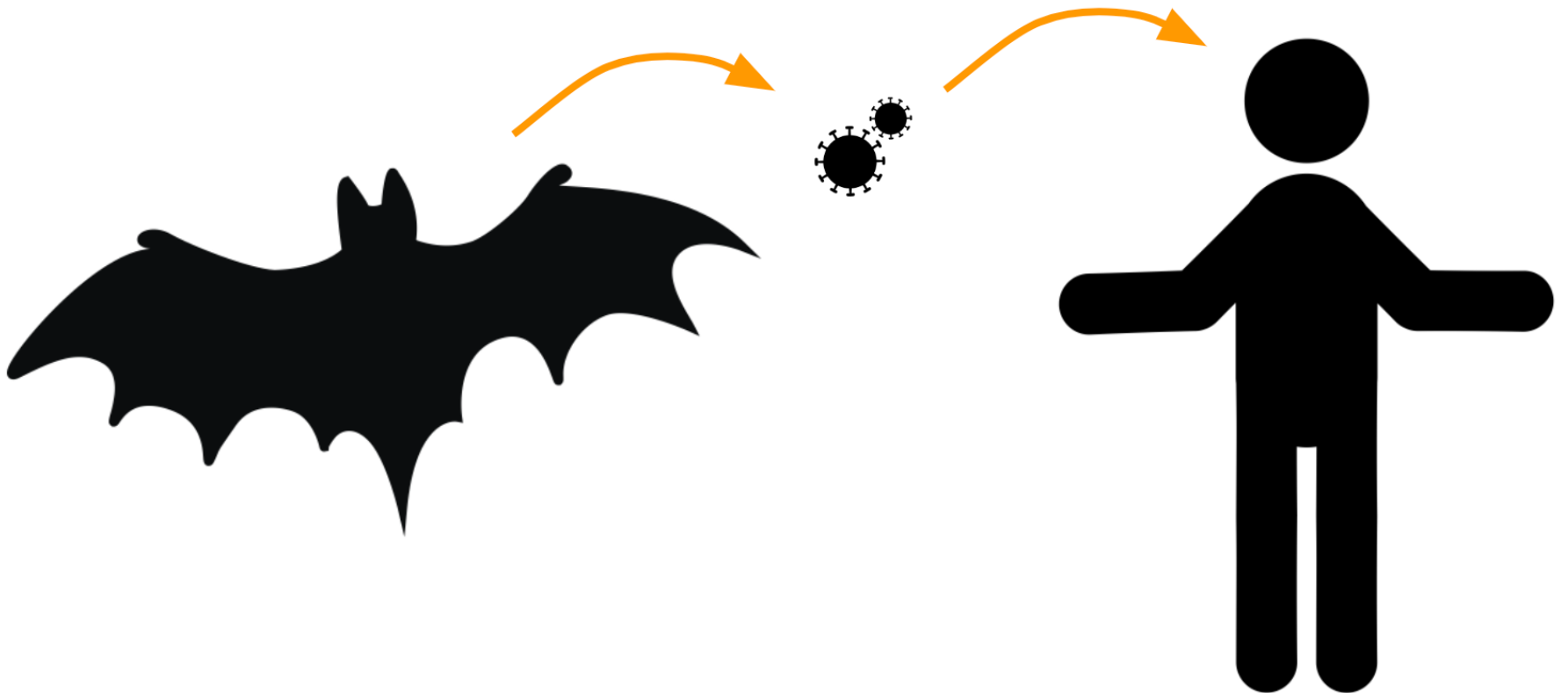
- What is the infectious unit?
- Where is the outbreak?
- How is it transmitted geographically?
- How is it transmitted between people?
- How is it evolving?

WHERE IS THE OUTBREAK?



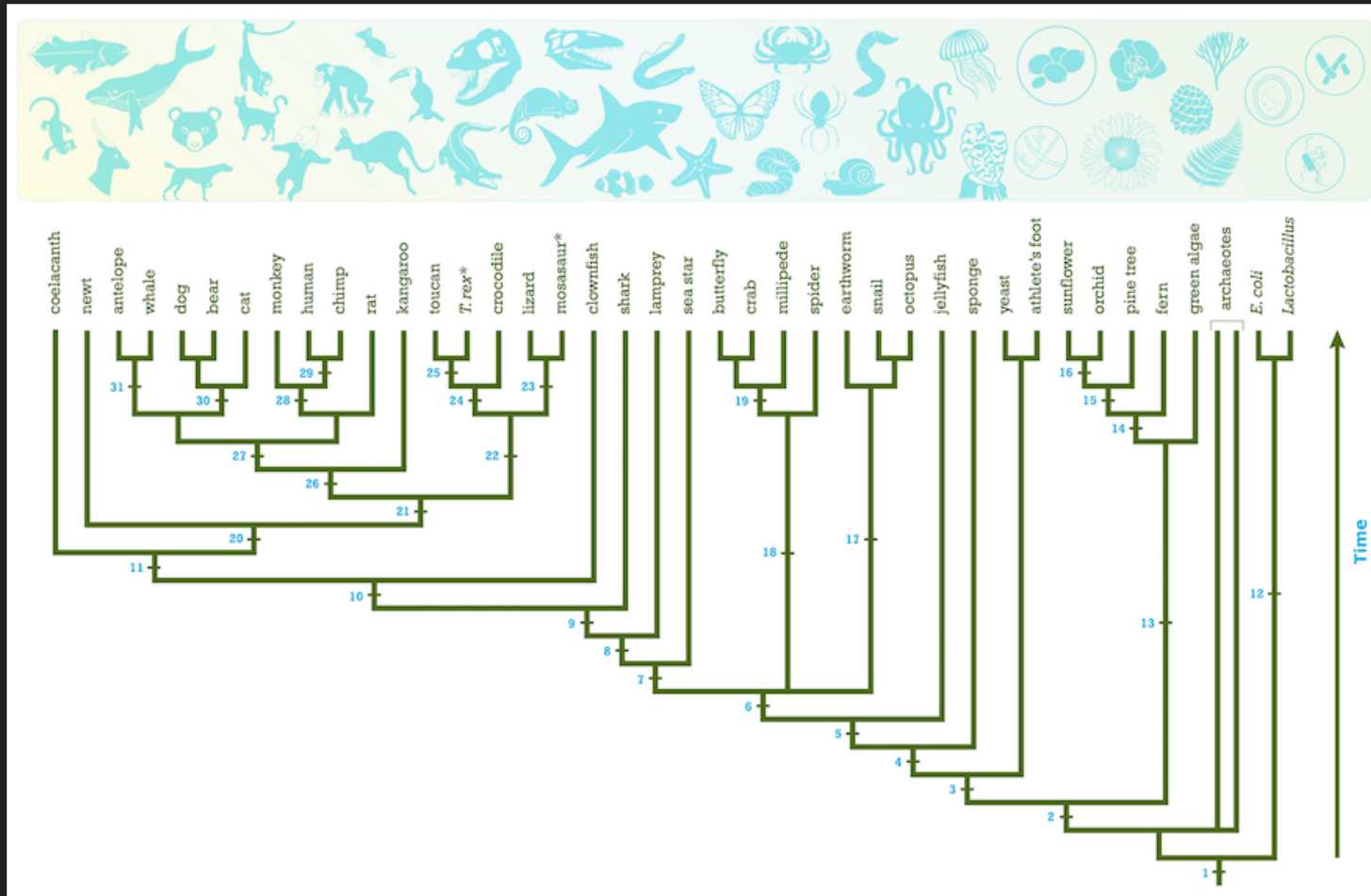
HOW IS THE VIRUS TRANSMITTED?

HOW IS THE VIRUS TRANSMITTED?



MUTATIONS CAN ALTER PROTEINS

MUTATIONS INFORM ANCESTRAL RELATIONSHIPS



EXPLORE ONLINE TOOLS

<https://www.ncbi.nlm.nih.gov/genomes/VirusVariation/Database>

Or, Google:

ncbi virus variation database

FILTER EBOLA SEQUENCES

Select sequence type

☒ Protein ☐ Nucleotide ☒ Full-length sequences only

Define search set

Species

Bundibugyo ebolavirus
Tai Forest ebolavirus
Sudan ebolavirus
Reston ebolavirus
Zaire ebolavirus

Host

Unknown
Bat
Chimpanzee
Gorilla
Human

Region/Country

any
regions
Africa
Asia
Europe

Genome region

Polymerase complex protein
Matrix protein
Second secreted glycoprotein
Small secreted glycoprotein
Spike glycoprotein

Isolation source

any
abdominal cavity
blood
cerebrospinal fluid
heart

Collection date: to

Release date: to
Year Month Day Year Month Day

Additional filters  

Add query

Show results

☒ Collapse identical sequences

ALIGN SEQUENCES

1. Build sequence alignment
2. Browse alignment
3. Download alignment ("Fasta plus gaps")
4. View alignment at <http://msa.biojs.net/app/>

BUILD A PHYLOGENETIC TREE

1. Build phylogenetic tree
2. Collapse view to "viewport"
3. Explore patterns

EXPLORE EBOLA OUTBREAK WITH NEXTSTRAIN

<https://nextstrain.org/ebola>

EXPLORE EBOLA OUTBREAK WITH NEXTSTRAIN

<https://nextstrain.org/ebola>

- Where did Ebola emerge? Where did it expand most rapidly?
- When did the outbreak start?
- What mutations were associated with the outbreak?
- How does the mutation rate of Ebola compare to other viruses?

BEHIND THE SCENES OF NEXTSTRAIN

<https://github.com/nextstrain>

CLOSING THOUGHTS

CLOSING THOUGHTS

Computational biology can provide insight into evolutionary processes happening in our lifetime

CLOSING THOUGHTS

Computational biology can provide insight into evolutionary processes happening in our lifetime

Many tools are available to everyone to use

CLOSING THOUGHTS

Computational biology can provide insight into evolutionary processes happening in our lifetime

Many tools are available to everyone to use

You can contribute to the tools and/or the science

THANK YOU!

More at <https://github.com/huddlej/workshop-2018-hutch-explorers>