

00: Course Overview

CSCI4181/6802 Bioinformatics Algorithms
Finlay Maguire (finlay.maguire@dal.ca)

Why am I teaching this course?

Maguire Lab Overview

Genomic Epidemiology

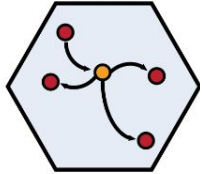
Sequencing
Data



Clinical
Data



Epidemiological
Data



Affiliations:



Computer Science
Community Health & Epidemiology
Institute of Comparative Genomics



Sunnybrook Research Institute
Shared Hospital Laboratory



Public Health Alliance for
Genomic Epidemiology

Maguire Lab Overview

Genomic Epidemiology

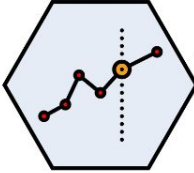
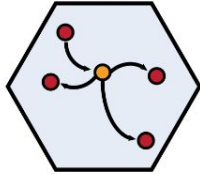
Sequencing
Data



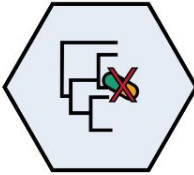
Clinical
Data



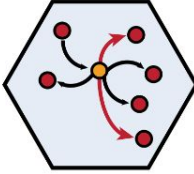
Epidemiological
Data



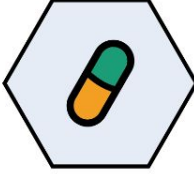
Epidemiological
Modelling



Pathogen
Surveillance



Outbreak
Prevention



Diagnostics
& Prescribing

Affiliations:



Computer Science
Community Health & Epidemiology
Institute of Comparative Genomics



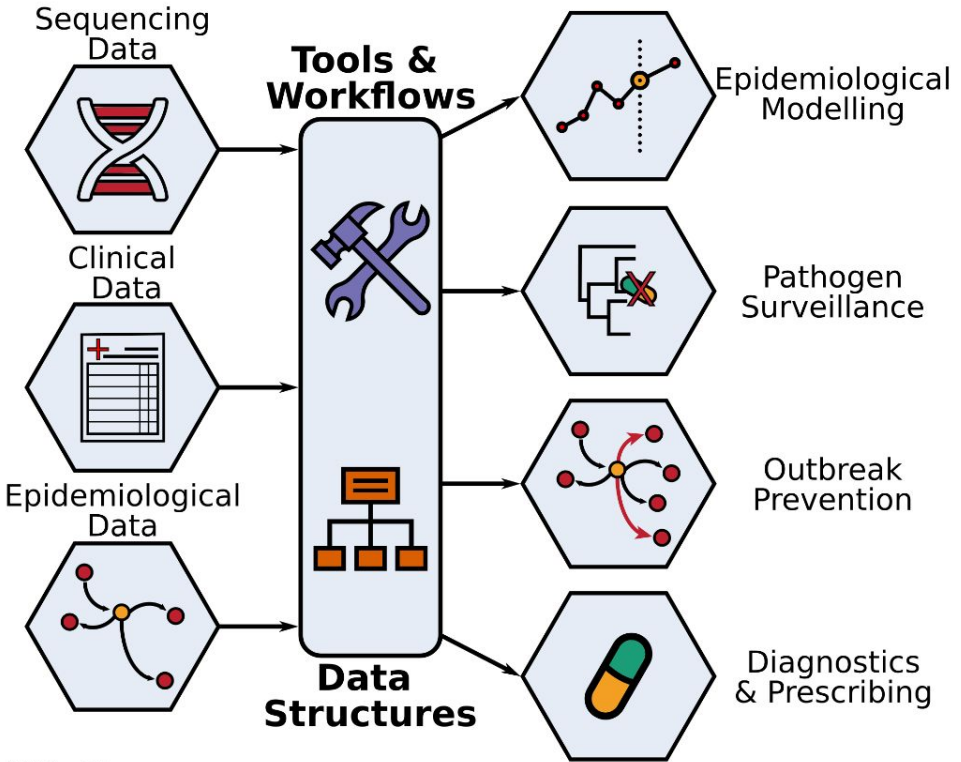
Sunnybrook Research Institute
Shared Hospital Laboratory



Public Health Alliance for
Genomic Epidemiology

Maguire Lab Overview

Genomic Epidemiology



Affiliations:



Computer Science
Community Health & Epidemiology
Institute of Comparative Genomics



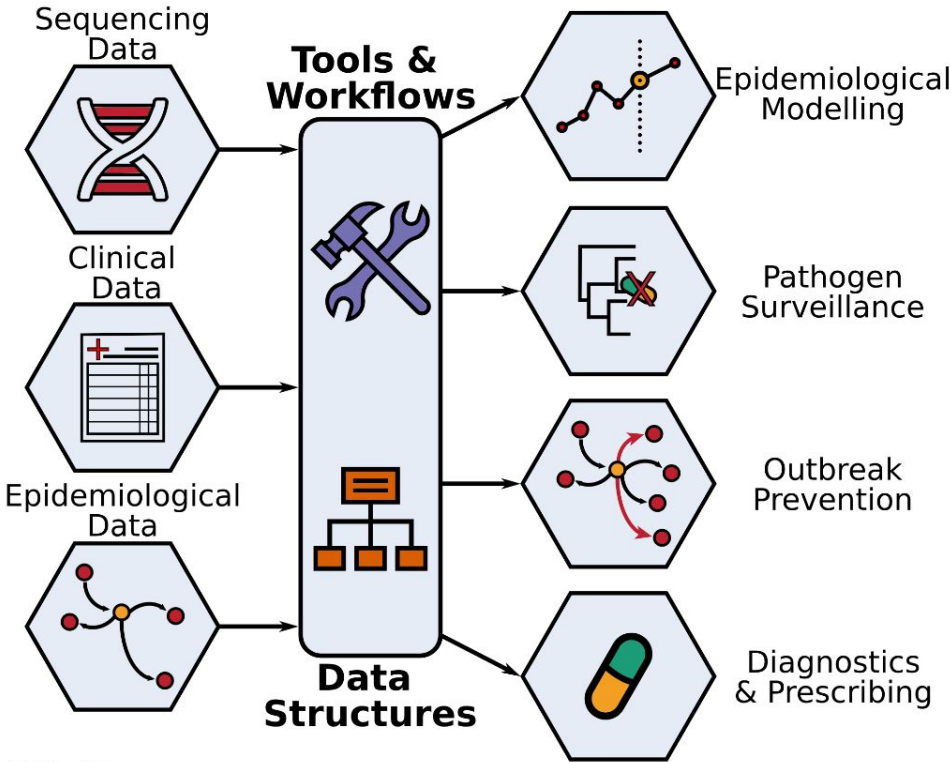
Sunnybrook Research Institute
Shared Hospital Laboratory



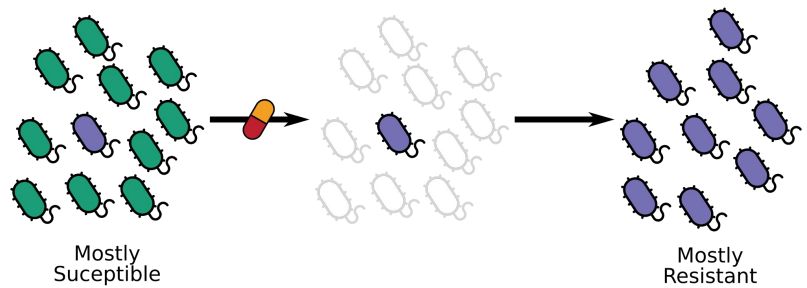
Public Health Alliance for
Genomic Epidemiology

Maguire Lab Overview

Genomic Epidemiology



Antimicrobial Resistance



Affiliations:



Computer Science
Community Health & Epidemiology
Institute of Comparative Genomics



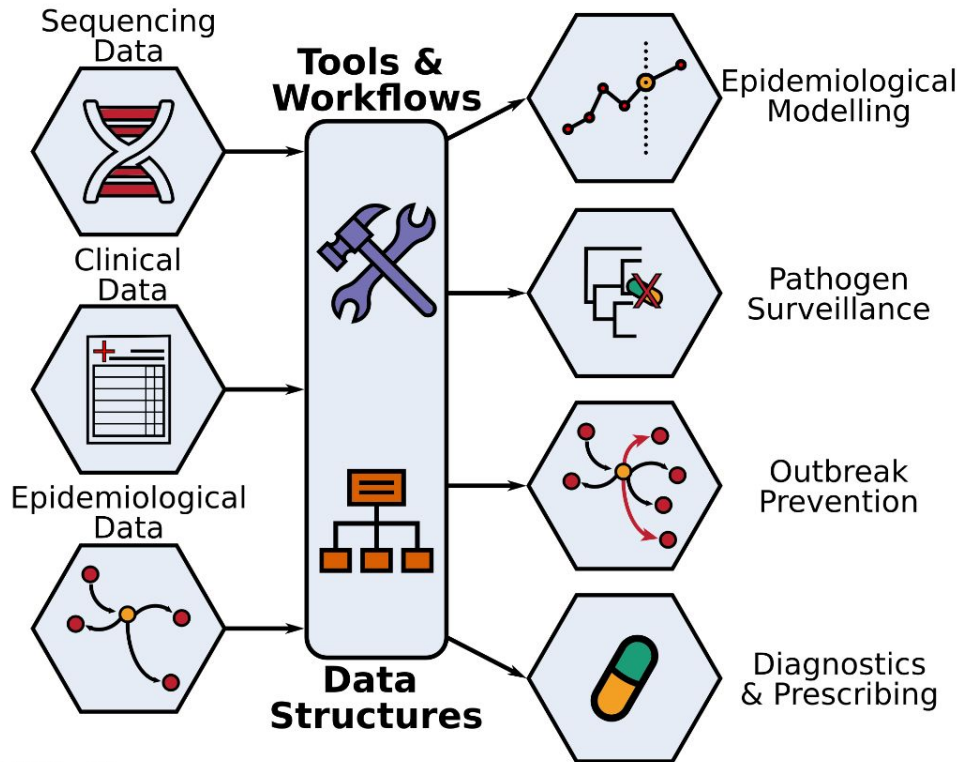
Sunnybrook Research Institute
Shared Hospital Laboratory



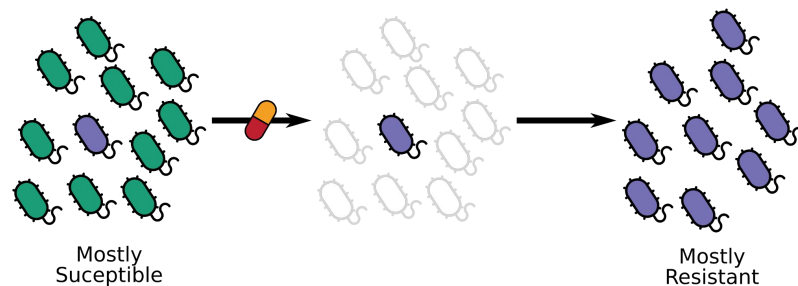
Public Health Alliance for
Genomic Epidemiology

Maguire Lab Overview

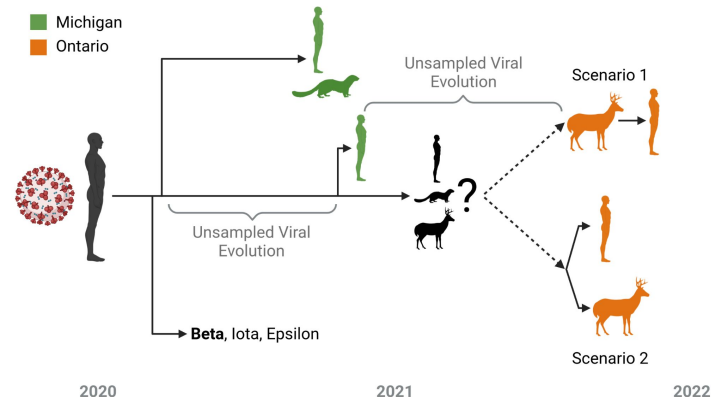
Genomic Epidemiology



Antimicrobial Resistance



SARS-CoV-2 Evolution



Affiliations:



Computer Science
Community Health & Epidemiology
Institute of Comparative Genomics



Sunnybrook Research Institute
Shared Hospital Laboratory

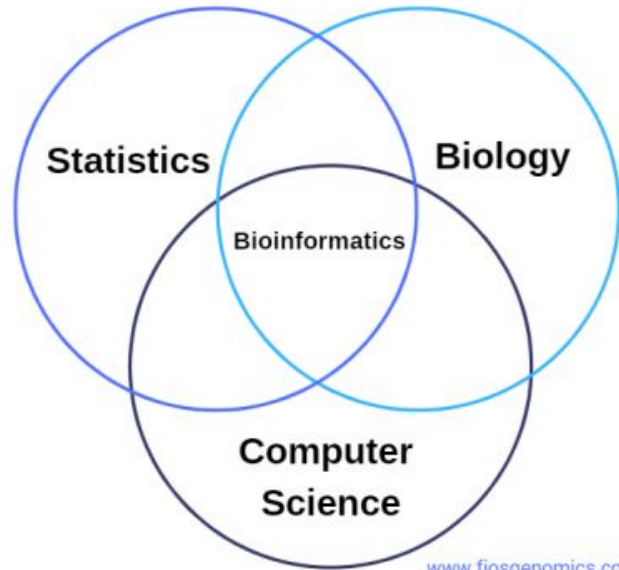


Public Health Alliance for
Genomic Epidemiology

What is bioinformatics?

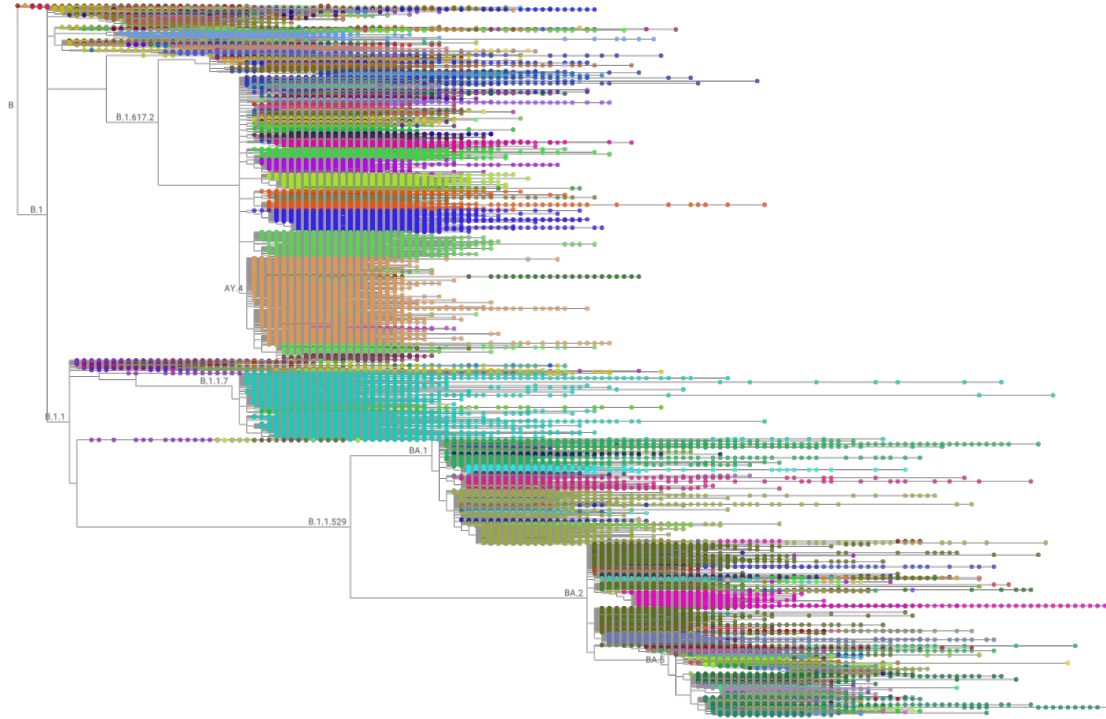
Using **computers** to understand **biology**

Using **computers** to understand **biology**



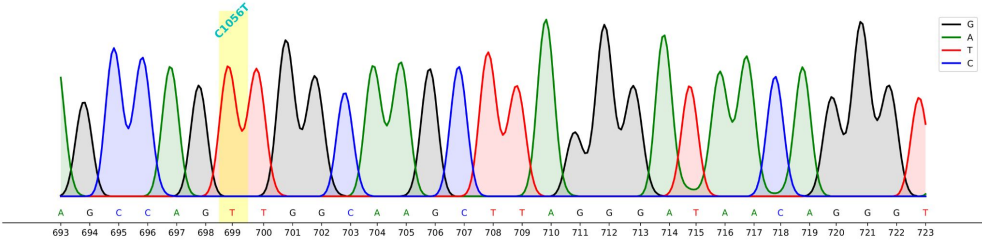
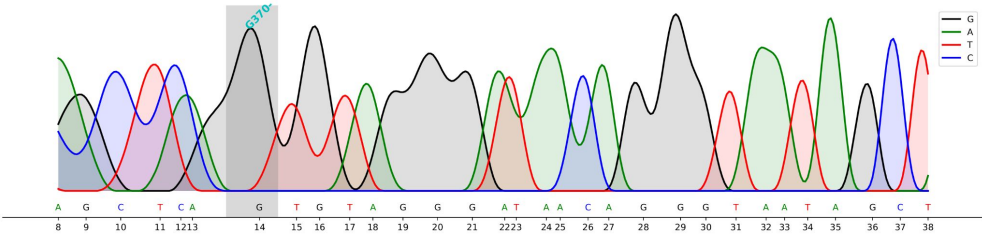
Why do we care about algorithms in
bioinformatics?

Lots of data



Cov2tree: taxonum 13.4 million SARS-CoV-2 Genomes

Lots of messy data



<https://github.com/y9c/cfutils>

AutoSave Off Book2 - Excel

File Home Insert Page Layout Formulas Data Review

Cut Copy Paste Format Painter Clipboard Font

Calibri 11 A^ A^ B I U Font Color

A3 3/1/2021

	A	B	C	D	E
1	MX1	HDAC5			
2	FZD1	MYC			
3	1-Mar	IL8			
4	PSEN2	1-Dec			
5	RBPJ	WNT5B			
6	PTPRN2	WNT6			
7	15-Sep	INF2			
8	CUL1	AGO2			
9					

<https://cosmosmagazine.com/science/excel-autocorrect-errors-still-plague-genetic-research/>

Lots of messy **biological** data

Biological data is special:

Lots of messy **biological** data

Biological data is special:

- Data is shaped by evolution i.e., “**what works at a given time**” not “what is pretty and easy to understand”

Lots of messy **biological** data

Biological data is special:

- Data is shaped by evolution i.e., “**what works at a given time**” not “what is pretty and easy to understand”
- Correct answer is often not obvious - we need **statistics**

Lots of messy **biological** data

Biological data is special:

- Data is shaped by evolution i.e., “**what works at a given time**” not “what is pretty and easy to understand”
- Correct answer is often not obvious - we need **statistics**
- Can't agree on many of the objects and questions!

Lots of messy **biological** data

Biological data is special:

- Data is shaped by evolution i.e., “**what works at a given time**” not “what is pretty and easy to understand”
- Correct answer is often not obvious - we need **statistics**
- Can't agree on many of the objects and questions!
Philosophy of Biology infuses what we do in many interesting ways

What are we actually going to learn?

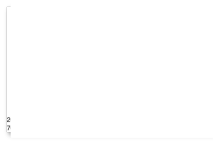
1. |

2. |

3. |

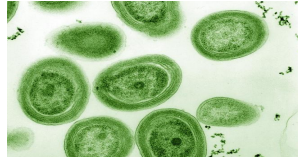
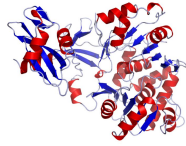
4. |

5. |



1. Introduction: Biological foundations

LEVEL	CHARACTERISTICS	EXAMPLES
1. CELL	Smallest unit of life; performs all life processes.	Animal cell, plant cell, bacterial cell.
2. TISSUE	Group of similar cells working together to perform a specific function.	Epithelial tissue, muscle tissue, nervous tissue.
3. ORGAN	Structure composed of different tissues working together to perform a specific function.	Heart, lung, kidney, stomach.
4. ORGAN SYSTEM	Group of organs working together to perform a specific function.	Circulatory system, digestive system, respiratory system.
5. ORGANISM	Complete living entity capable of performing all life processes.	Human, dog, cat, tree, bird.



2.

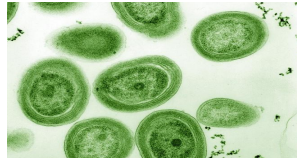
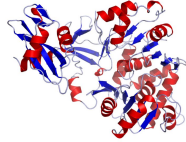
3.

4.

5.

1. Introduction: Biological foundations

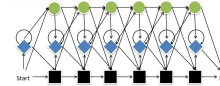
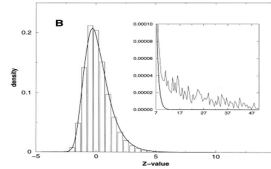
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----



2. Homology: Comparing biological sequences

K-ELQRAASLTIEV
KDEGQK--SLVIDV

	A	W	G	H	E
0	-5	-10	-15	-20	-25
A	3	2	-3	-8	-13
W	-10	-3	19	14	9
H	-15	-8	14	17	20
E	-20	-13	9	14	18
A	-25	-18	4	10	13



0	6	\$googo l
1	3	gol\$go o
2	0	googol \$
3	5	l\$goog o
4	2	ogol\$g o
5	4	o!\$goo g
6	1	oogol\$ g

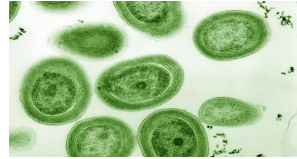
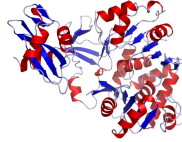
3.

4.

5.

1. Introduction: Biological foundations

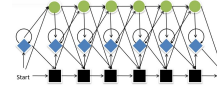
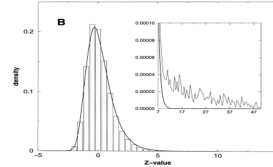
Protein	AA	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30								
...



2. Homology: Comparing biological sequences

K-ELQRAASLTIEV
KDEGQK--SLVIDV

Window	A	W	G	H	E	
A	0	-5	-10	-15	-20	-25
W	-5	2	-3	-8	-13	-18
G	-10	-3	19	14	9	4
H	-15	-8	14	17	20	15
E	-20	-13	9	14	18	24
A	-25	-18	4	10	13	19

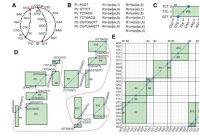
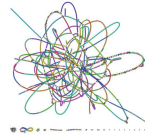


0	6	\$	g	o	o	g	o	l
1	3	g	o	l	\$	g	o	o
2	0	g	o	o	g	o	l	\$
3	5	l	\$	g	o	o	g	o
4	2	o	g	o	l	\$	g	o
5	4	o	l	\$	g	o	g	o
6	1	o	o	g	l	\$	g	o

3. Assembly: Recovering genomes from sequencing data

```
Reference: TCC TAGAGATCCCCCTCTTAGCGGATATAATACAGCCGAATGCTAGCGGAATGCCAGCAGAG
Reads:      CCTAGAGATCC
          GAGATCCGCTC
          ATCCCTCTTA
          CTTAGCGG
          TTAGCGGATATA
          TAGCGGATATA
          TATATACAGCC
          ATACAGCCAAATCT
          CCAATCTTAGC
          GAACTTAGCGG
          AACCTAGCGG
          ACCGAAATGCC
          GGAACTTAGCGG
          GAACTTAGCGG
          AATGCCAGCAGAG
          TTAGCGGATATA
          TCGCAGCAGAG

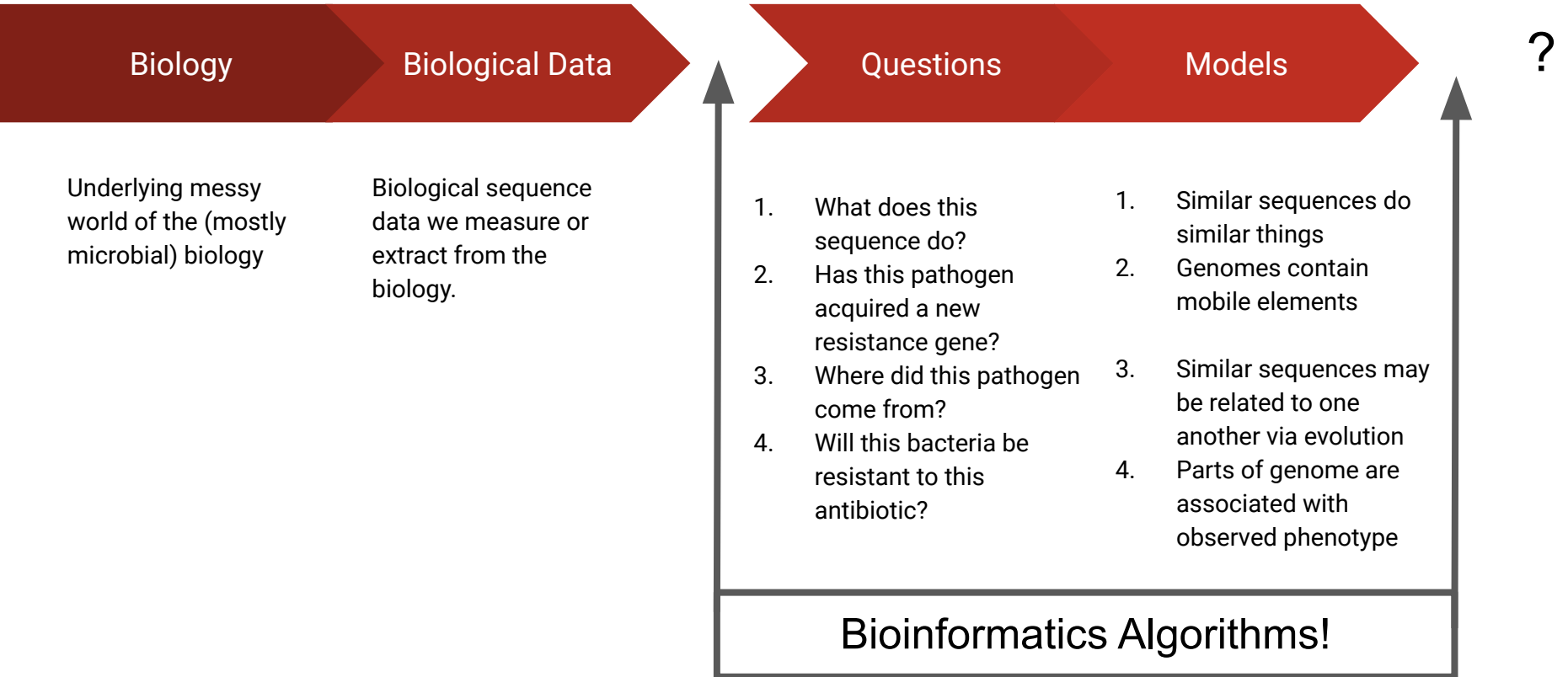
Contigs:   CCTAGAGATCCCCCTCTTAGCGG
          CTTAGCGGATATAATACAGCCGAATCTTAGCGG
          TCTTAGCGGAATGCCAGCAGAG
```



4.

5.

General Overarching Theme



How are we going to learn?

Lectures, practical assignments, and a paper review

- 22 lectures (every Tuesday and Thursday)
 - Guest lectures by Dr. Ryan Fink, Prof. Robert Beiko, and Alex Manuele

Lectures, practical assignments, and a paper review

- 22 lectures (every Tuesday and Thursday)
 - Guest lectures by Dr. Ryan Fink, Prof. Robert Beiko, and Alex Manuele
- 4 Practical Assignments (due via Brightspace 1 week after release)
 1. Alignment and Distant Homology (20%)
 2. Genome Assembly (20%)
 3. SARS-CoV-2 Genome Analysis (20%)
 4. Prediction of AMR Phenotypes (20%)

Lectures, practical assignments, and a paper review

- 22 lectures (every Tuesday and Thursday)
 - Guest lectures by Dr. Ryan Fink, Prof. Robert Beiko, and Alex Manuele
- 4 Practical Assignments (due via Brightspace 1 week after release)
 1. Alignment and Distant Homology (20%)
 2. Genome Assembly (20%)
 3. SARS-CoV-2 Genome Analysis (20%)
 4. Prediction of AMR Phenotypes (20%)
- Paper Review
 - Review of recent bioinformatics paper (selection due 2023/03/21)
 - Written Summary due 2023/04/04 (10%)
 - Oral Presentation 2023/04/04 to 2023/04/11 (10%)

Other Class Logistics

- Details on course website
- Contact for office hours
- **TA: Jee-in Kim (remote)**
- Assignment must be submitted via Brightspace as .docx or .pdf and named:

“BannerID_LastName_AssignmentX.{pdf,docx}”



[CSCI4181/6802 Bioinformatics Algorithms](#) / Winter 2022-2023

Course Description

Bioinformatics uses computational and statistical approaches to tackle questions of biological function and evolution. The goal of Algorithms in Bioinformatics is to introduce key applications of algorithms, data structures, and encodings to the analysis of large biological data sets. A recurring theme throughout the course will be the disconnect between algorithmic beauty and the horrifying realities of biological data. Every statistical model is violated and every classification comes with an asterisk, as we struggle with even the most basic concepts of 'gene' and 'species', and the challenges of understanding events that happened across ~3.5 billion years. In spite of these challenges, in this age of massive data sets we stand to learn a good deal if the computational tools we use are efficient, robust, properly validated, and correctly applied. The course covers major challenge areas in bioinformatics, each focused on an aspect of DNA or protein sequence analysis. The goal in each case is to define an overarching problem, and then explore different approaches that have been applied to solving that problem, with an emphasis on the match (or mismatch) between the algorithm and the underlying biological system.

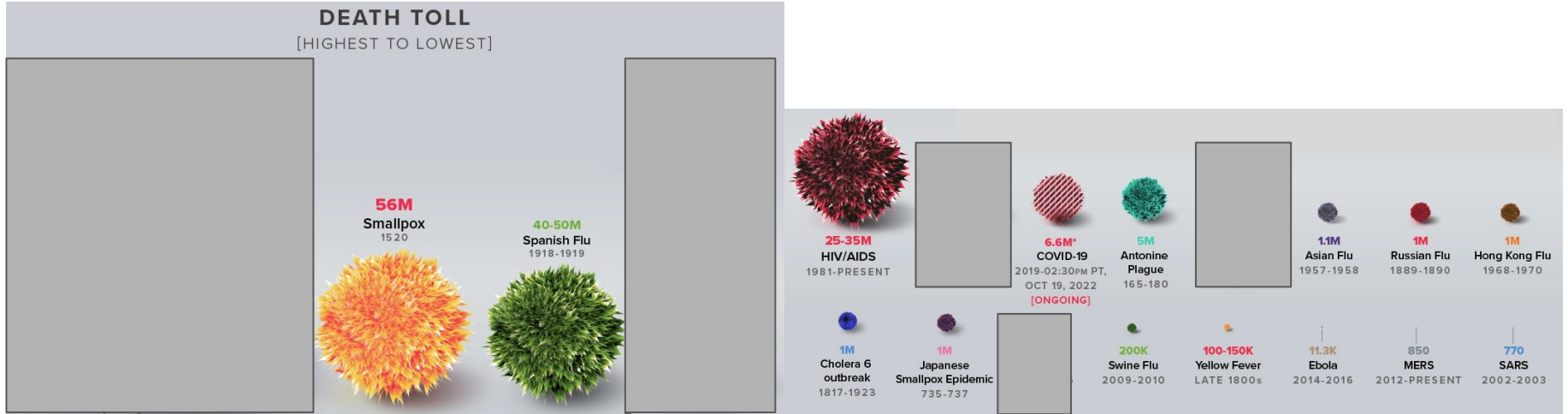
2023 Course Details

- Tuesday & Thursday: 8:35-9:55, 1201 Mona Campbell Building
- Office: 4239 Mona Campbell Building, Studley Campus
- Email: finlay.maguire@dal.ca
- Office Hours: No fixed hours, email us for an appointment
- BrightSpace for assignment/paper review submission.
- Syllabus

https://maguire-lab.github.io/bioinformatics_algorithms

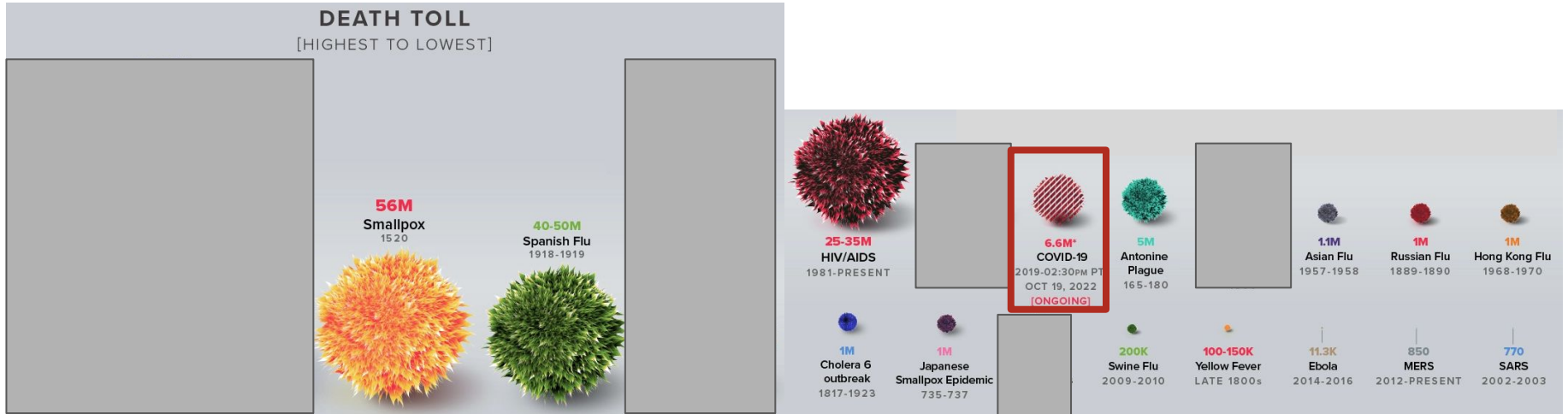
Why does any of this matter?

Pandemics matter to humans



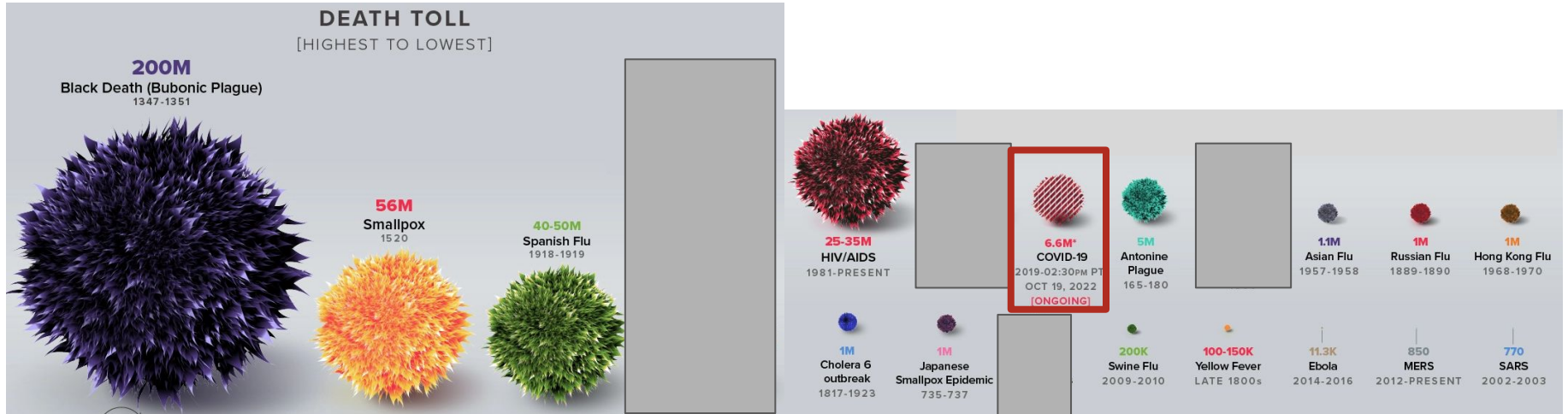
<https://www.visualcapitalist.com/history-of-pandemics-deadliest/>

Pandemics matter to humans



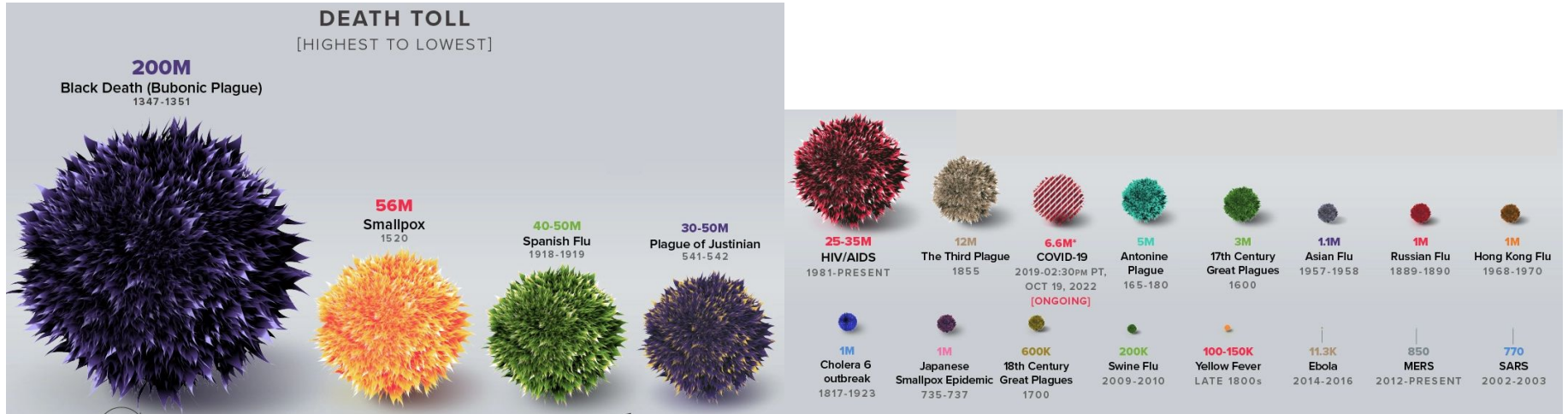
<https://www.visualcapitalist.com/history-of-pandemics-deadliest/>

Pandemics matter to humans



<https://www.visualcapitalist.com/history-of-pandemics-deadliest/>

Pandemics matter to humans



<https://www.visualcapitalist.com/history-of-pandemics-deadliest/>

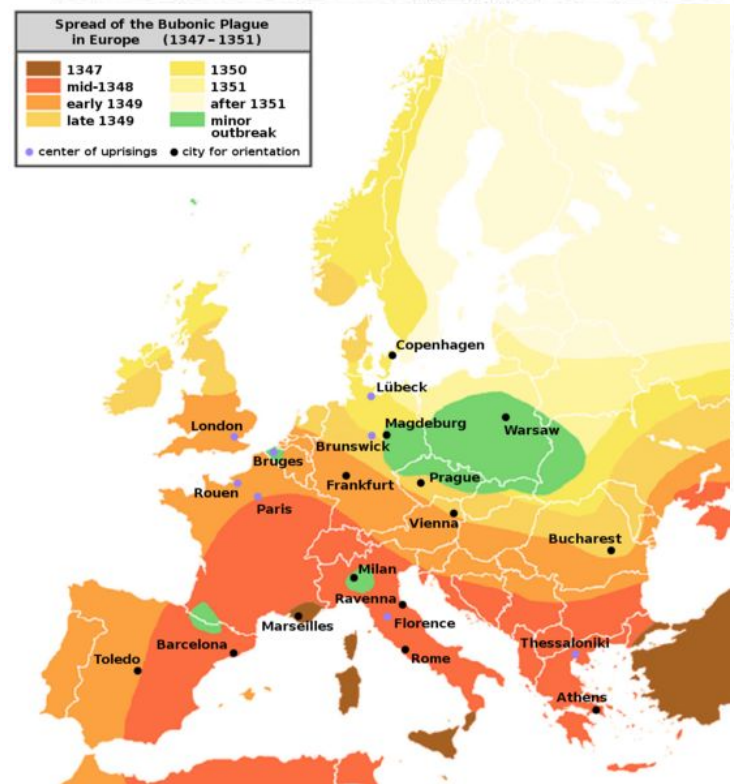
Plague Huge Impact on Human History

- Justinian Plague (541–549)



Plague Huge Impact on Human History

- Justinian Plague (541–549)
- Black Death (1347-1351 / 1330s to 1830s)



Plague Huge Impact on Human History

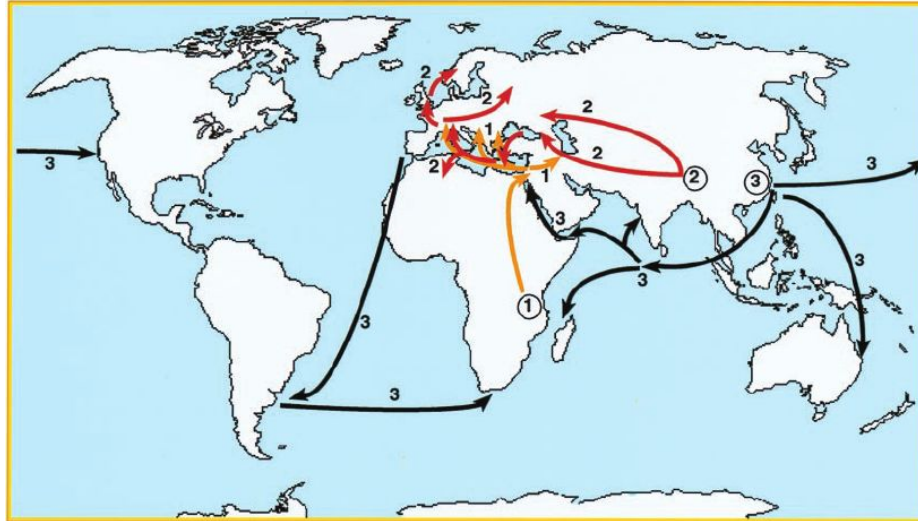
- Justinian Plague (541–549)
- Black Death (1347-1351 / 1330s to 1830s)



https://commons.wikimedia.org/wiki/File:Lord_haue_mercy_on_London.jpg

Plague Huge Impact on Human History

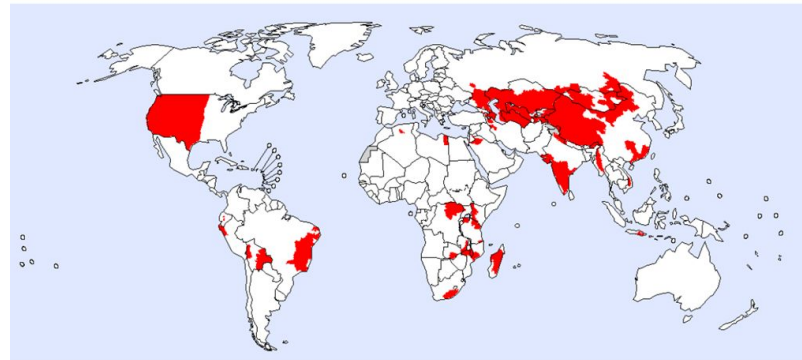
- Justinian Plague (541–549)
- Black Death (1347-1351 / 1330s to 1830s)
- 3rd Plague (1855-1960 / Today)



Plague Huge Impact on Human History

- Justinian Plague (541–549)
- Black Death (1347-1351 / 1330s to 1830s)
- 3rd Plague (1855-1960 / Today)

Global distribution of natural plague foci
as of March 2016



■ Areas* with potential plague natural foci based on historical data and current information

* First administrative level representation
Source: WHO/PED, as of 15 March 2016

The boundaries and names shown and the designations used on this map do not imply the expression of any opinion whatsoever on the part of the World Health Organization concerning the legal status of any country, territory, city or area or of its authorities, or concerning the determination of its frontiers or boundaries. Dotted lines on maps represent approximate border lines for which there may not yet be full agreement.
© WHO 2016. All rights reserved.

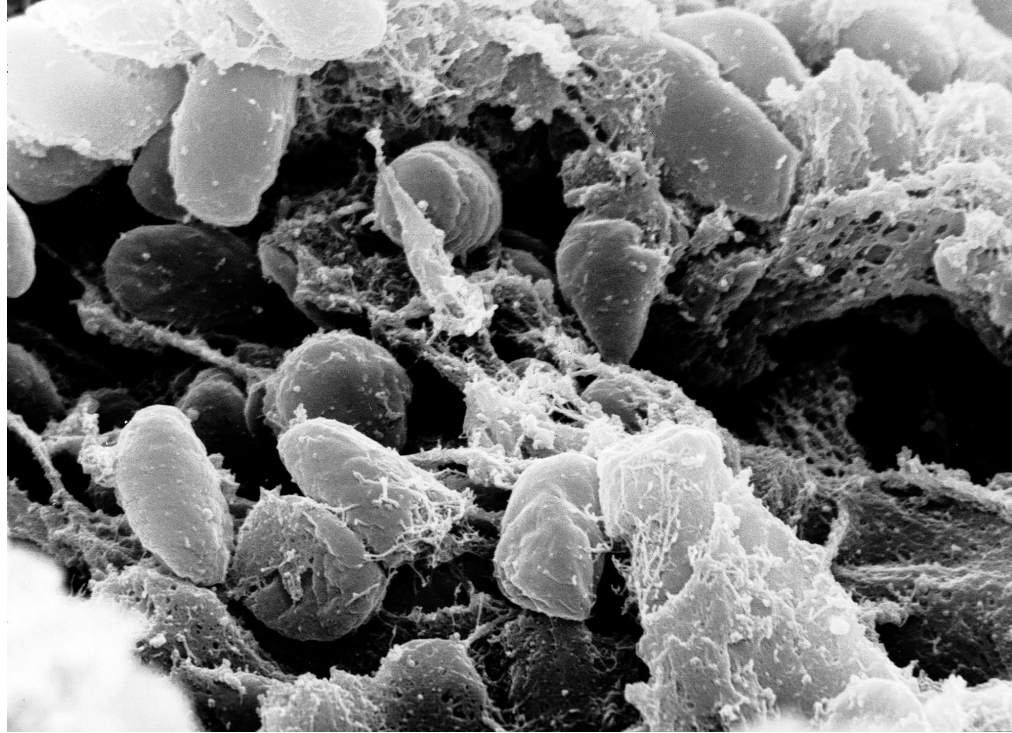


Working out the cause of the plague

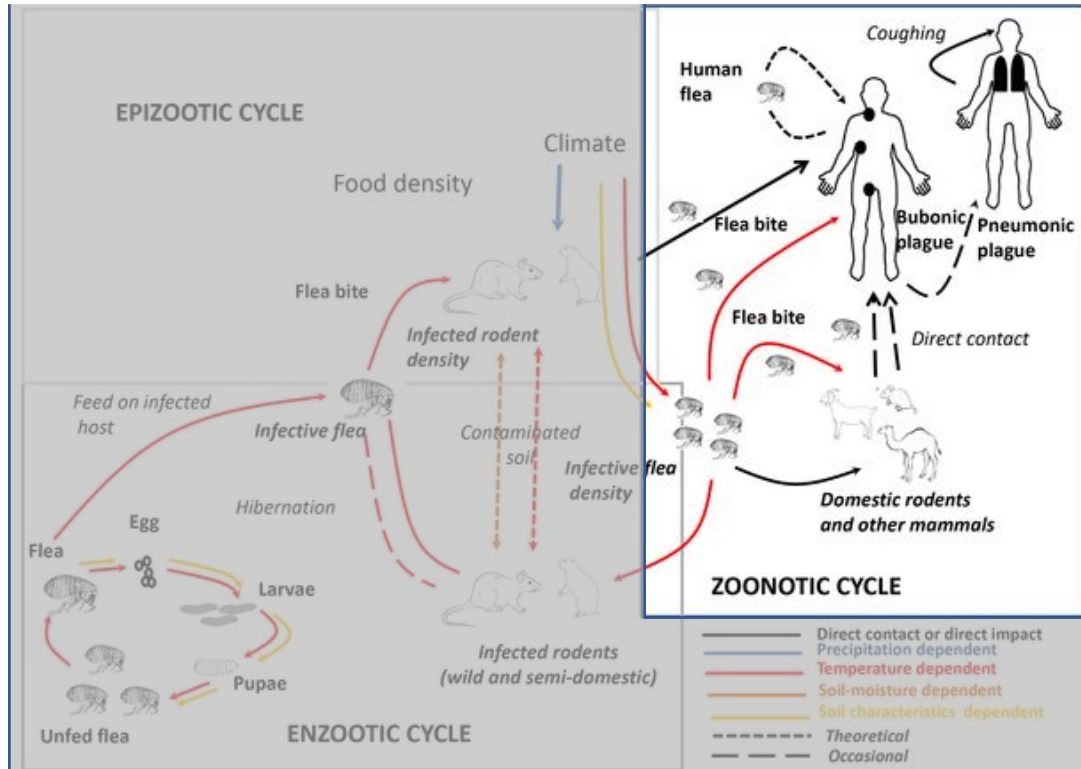
- **1546 Fracastro:** disease from invisible “seeds”
- **1683 Van Leeuwenhoek:** microscopic creatures
- **1857 Pasteur:** microbes are living creatures that don't spontaneously appear
- **1876 Koch:** microbes cause disease (Koch's Postulates)
- **1894 Shiibasaburō & Yersin:** Isolation of microbe in 3rd Plague
- **1896 Simon:** Isolation of microbe in fleas



Bubonic plague is caused by *Yersinia pestis*

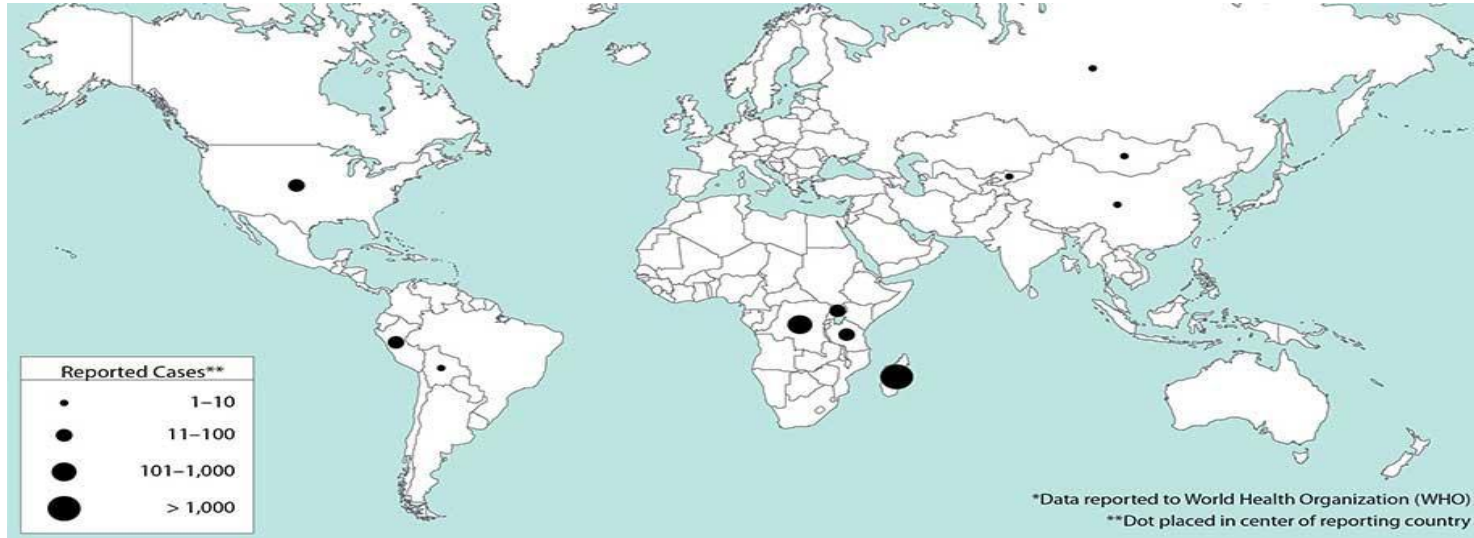


Complex life cycle including fleas and rodents



Vallès X, Stenseth NC, Demeure C, Horby P, Mead PS, et al. (2020) Human plague: An old scourge that needs new answers. PLOS Neglected Tropical Diseases 14(8): e0008251.

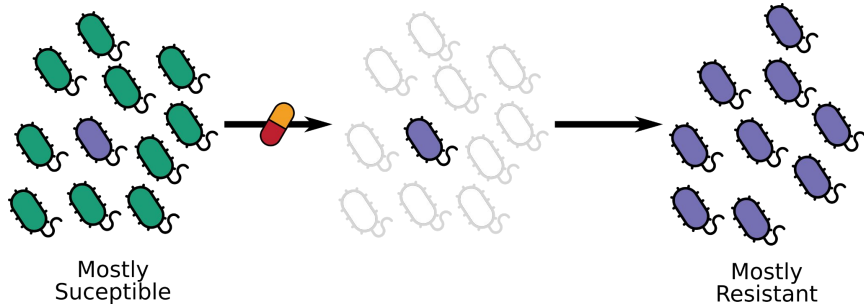
Still causes disease but is treatable



Fatality rate (untreated) = **40%-70%**

Fatality rate (treated) = **5%-15%**

Becoming less treatable



Becoming less treatable

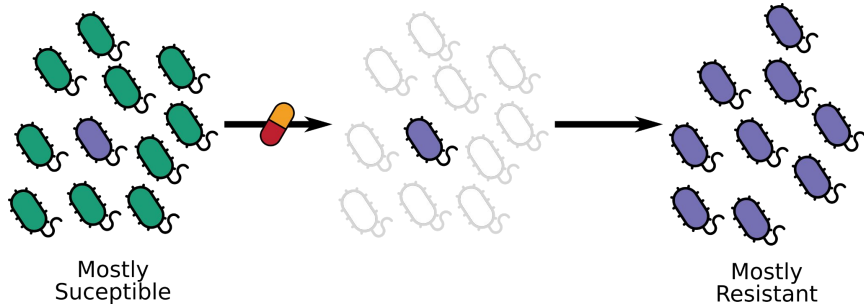


TABLE 2 Antimicrobial MIC distributions for *Y. pestis* isolates in this study ($n = 392$)

Antimicrobial	No. of isolates with MIC ($\mu\text{g/ml}$) ^a							
	0.06	0.12	0.25	0.5	1	2		
Gentamicin		3	260	110	18	1		
<u>Streptomycin</u>				2	6	119	263	2
<u>Tetracycline</u>			11	154	224	3		
Doxycycline			4	66	245	77		
Ciprofloxacin	371	20	1					
Levofloxacin	385	7						
<u>Chloramphenicol</u>					34	86	201	71
Trimethoprim-sulfamethoxazole	9	333	49	1				

Urich et al. (2012) Antimicrob Agents Chemotherapy

Becoming less treatable

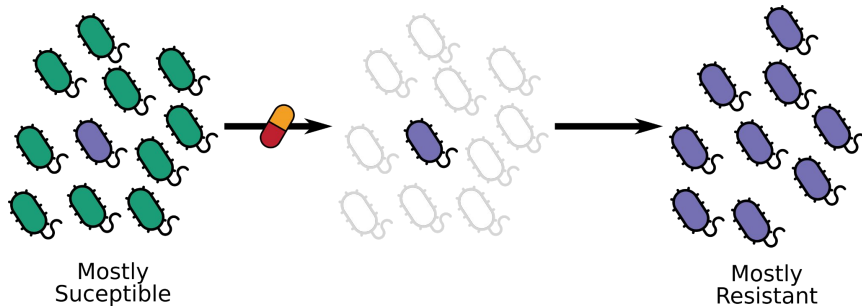


TABLE 2 Antimicrobial MIC distributions for *Y. pestis* isolates in this study ($n = 392$)

Antimicrobial	No. of isolates with MIC ($\mu\text{g/ml}$) ^a							
	0.06	0.12	0.25	0.5	1			
Gentamicin		3	260	110	18	1		
<u>Streptomycin</u>				2	6	119	263	2
<u>Tetracycline</u>				11	154	224	3	
Doxycycline			4	66	245	77		
Ciprofloxacin	371	20	1					
Levofloxacin	385	7						
<u>Chloramphenicol</u>					34	86	201	71
Trimethoprim-sulfamethoxazole	9	333	49	1				

Urich et al. (2012) Antimicrob Agents Chemotherapy

MULTIDRUG RESISTANCE IN *YERSINIA PESTIS* MEDIATED BY A TRANSFERABLE PLASMID

MARC GALIMAND, PH.D., ANNIE GUIYOULE, GUY GERBAUD, BRUNO RASOAMANANA, M.D., SUZANNE CHANTEAU, PH.D., ELISABETH CARNIEL, M.D., PH.D., AND PATRICE COURVALIN, M.D.

Plasmid-mediated doxycycline resistance in a *Yersinia pestis* strain isolated from a rat

Nicolas Cabanel ^a, Christiane Bouchier ^b, Minoarisoa Rajerison ^c, Elisabeth Carniel ^{a,*}

Transferable Plasmid-Mediated Resistance to Streptomycin in a Clinical Isolate of *Yersinia pestis*

Annie Guiyoule,^a Guy Gerbaud,^a Carmen Buchrieser,^a Marc Galimand,^a Lila Rahalison,^a Suzanne Chanteau,^a Patrice Courvalin,^a and Elisabeth Carniel^a
^aInstitut Pasteur, Paris, France; and ^bInstitut Pasteur, Antananarivo, Madagascar

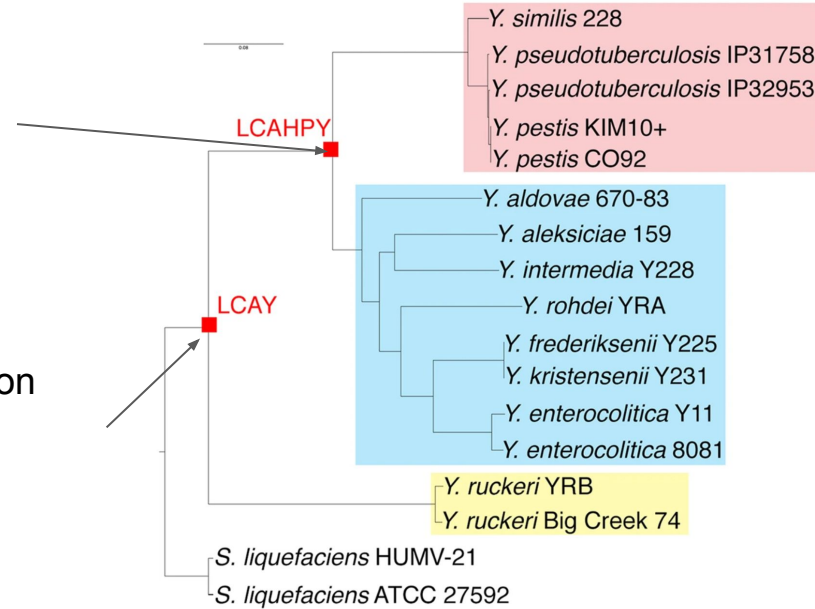
How can we control *Yersinia pestis*?

Need to learn about pathogen to control it

How did *Y. pestis* evolve?

Last common ancestor of human pathogenic *Yersinia*

Last common ancestor of *Yersinia*

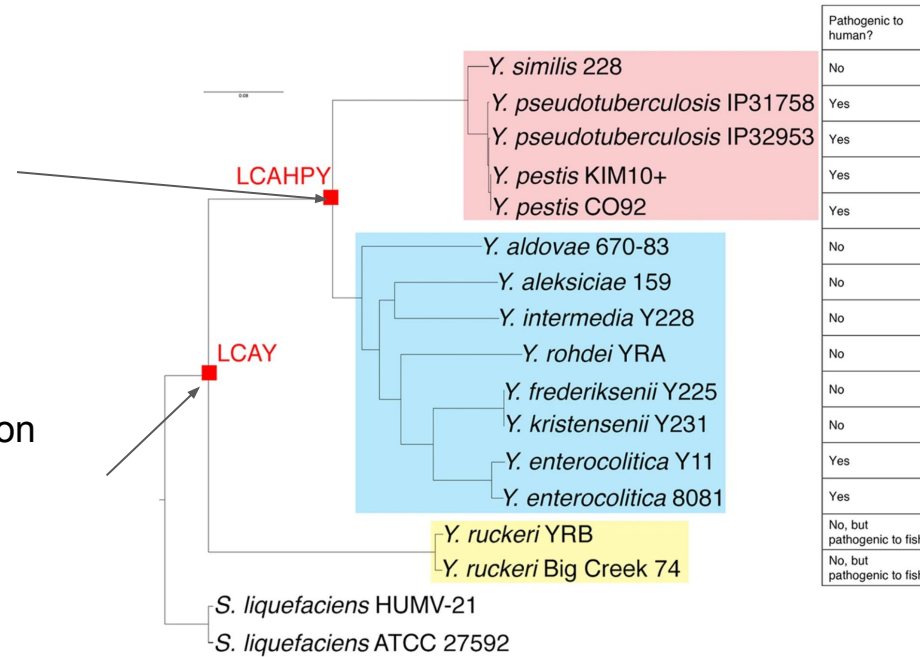


Tan et al. (2016) Sci Rep <https://www.nature.com/articles/srep36116>

How did *Y. pestis* evolve?

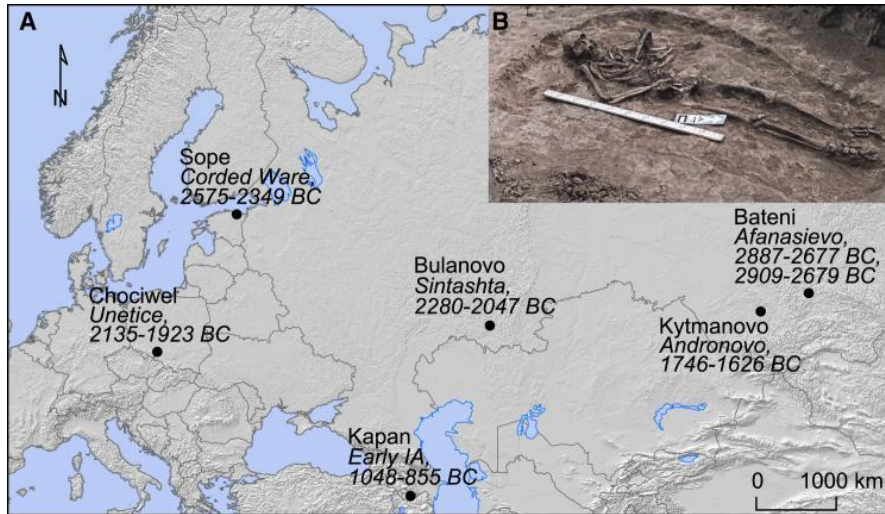
Last common ancestor of human pathogenic *Yersinia*

Last common ancestor of *Yersinia*

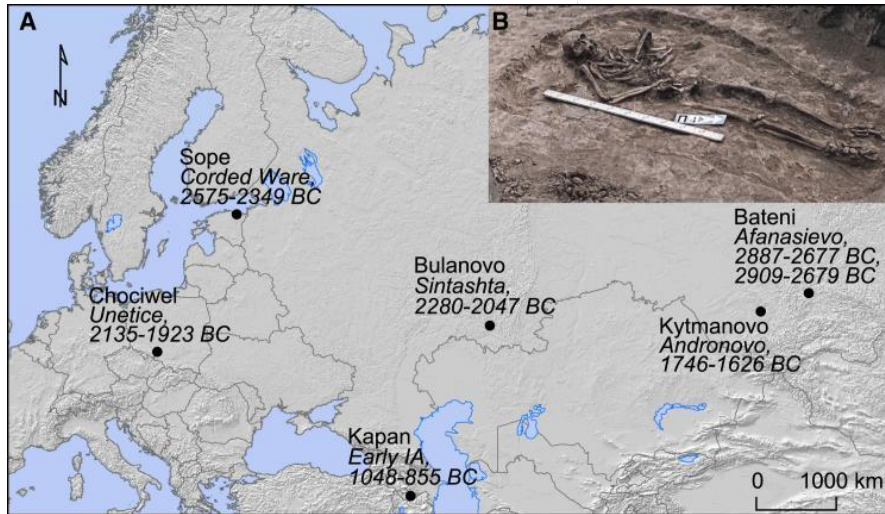
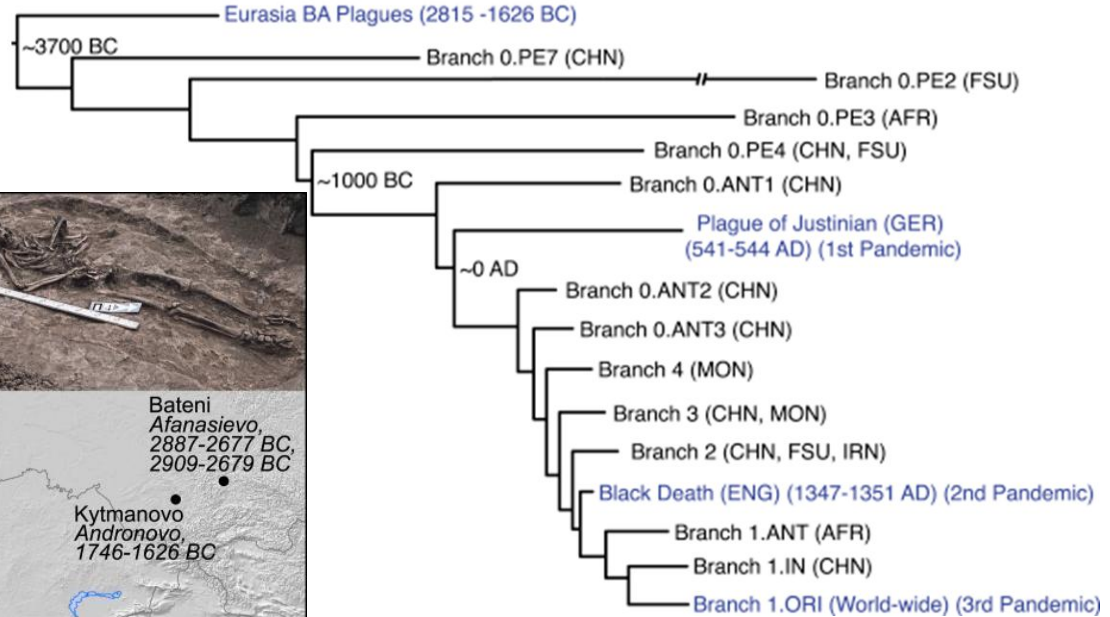


Tan et al. (2016) Sci Rep <https://www.nature.com/articles/srep36116>

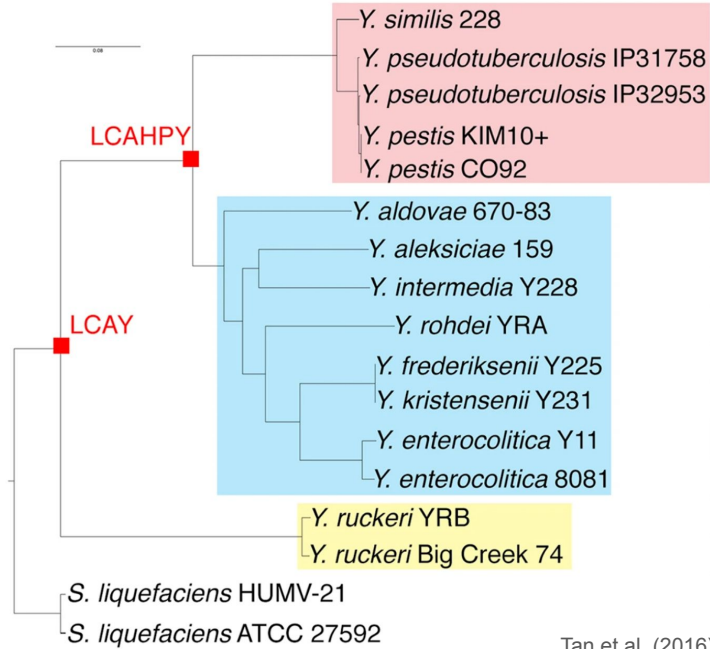
How old is *Y. pestis*?



How old is *Y. pestis*?



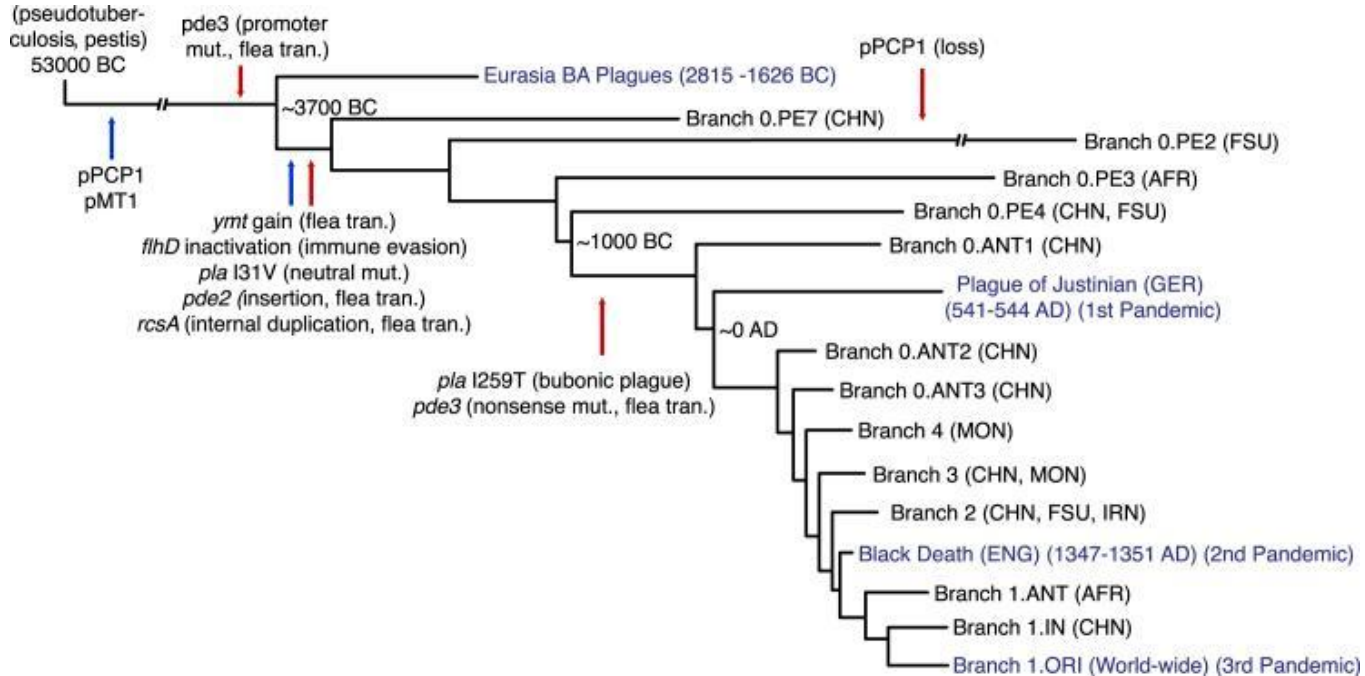
What makes it so virulent?



Pathogenic to human?	Presence of pYV plasmid?	Presence of CRISPR spacer to become immune to pYV or pYE854 plasmid?	Presence of <i>inv</i> homolog and N-terminal?	Number of copy of <i>ail</i> homolog?
No	No	Yes	Yes	4
Yes	Yes	No	Yes	4
Yes	Yes	Immune to pYV of <i>Y. enterocolitica</i>	Yes	4
Yes	Yes	Immune to pYV of <i>Y. enterocolitica</i>	No	4
Yes	Yes	Immune to pYV of <i>Y. enterocolitica</i>	No	4
No	No	No	No	1
No	No	No	No	1
No	No	No	No	1
No	No	Yes	No	1
No	No	Yes	No	1
Yes	Yes	No	Yes	2
Yes	Yes	No	Yes	2
No, but pathogenic to fish	No	No	No	1
No, but pathogenic to fish	No	No	No	1

Tan et al. (2016) Sci Rep <https://www.nature.com/articles/srep36116>

How did it become so deadly?



How do we know this?

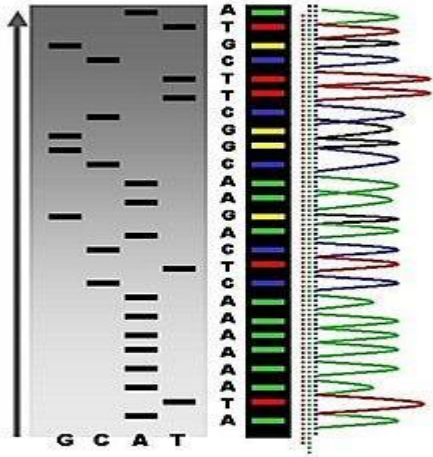


Molecular biology

How do we know this?



Molecular biology

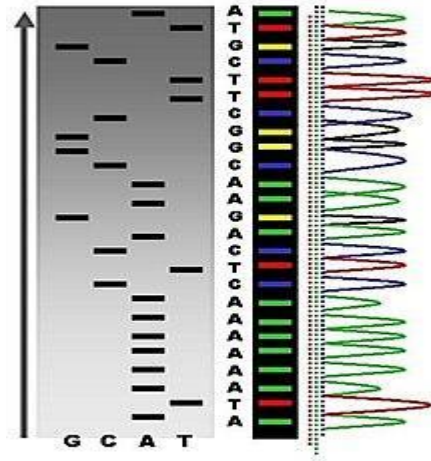


DNA sequencing

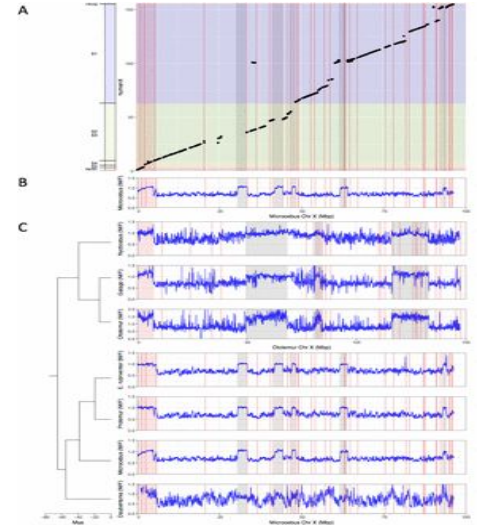
How do we know this?



Molecular biology



DNA sequencing



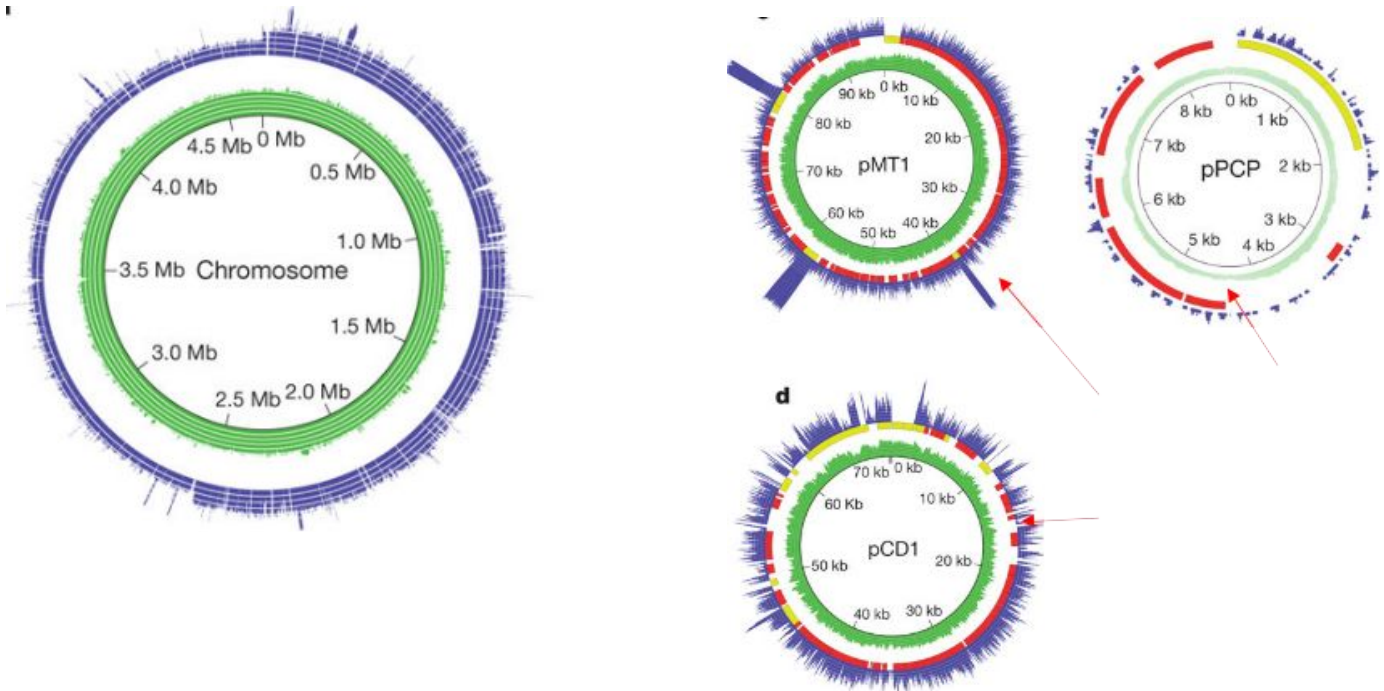
Bioinformatics!

We can sequence the Black Death Genome

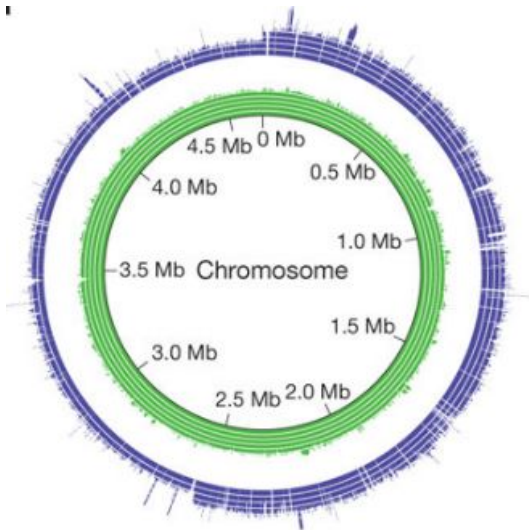


LONDON, 1593. By JOHN NORDEN.

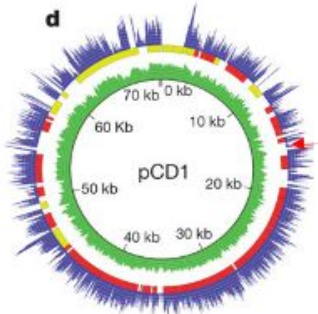
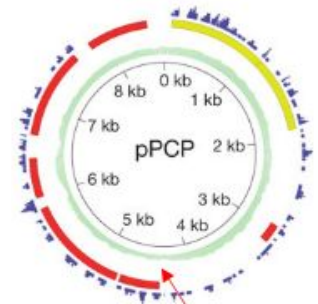
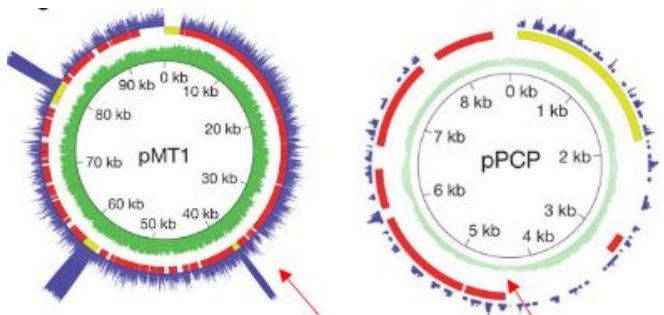
Need assembly algorithms to get genome



Need homology algorithms to decipher genome



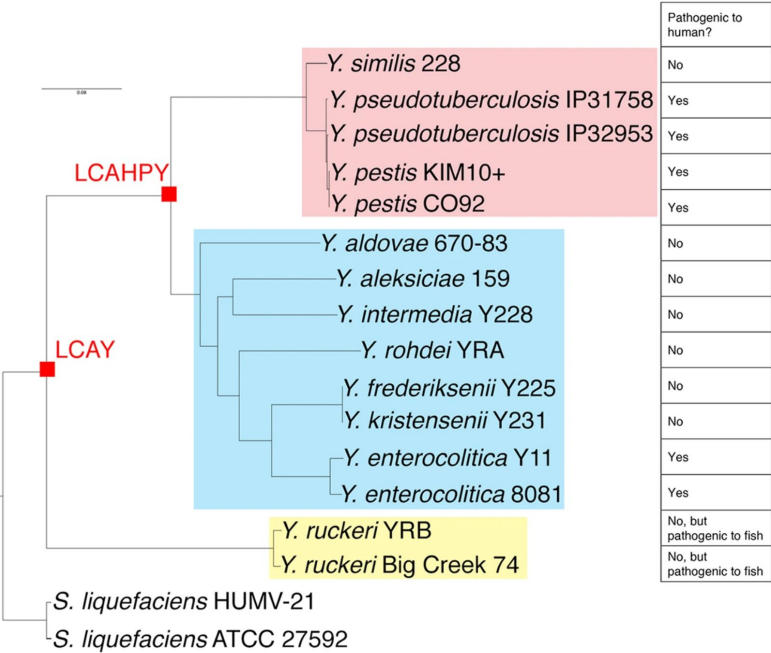
Main Chromosome



F1 (antigen; vaccine target) Pla (host invasion)
Antihost proteins and cellular delivery mechanism

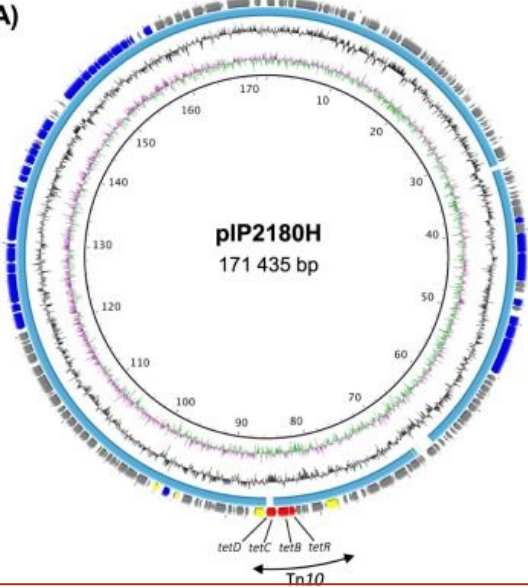
Plasmids

Need phylogenetic algorithms to trace



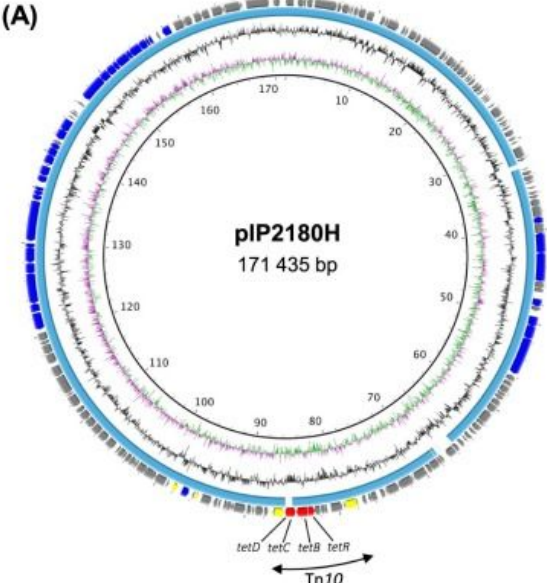
Need ML algorithms to predict resistance

(A)

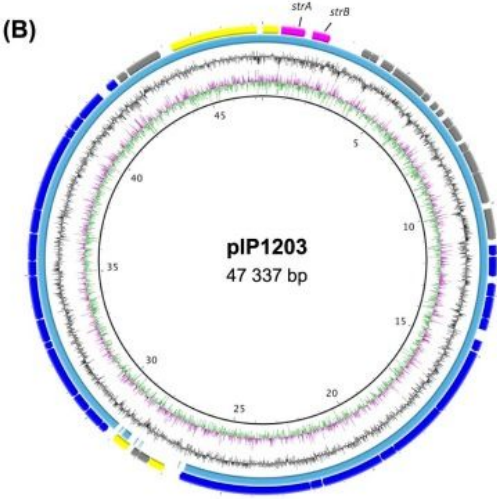


Doxycycline resistance genes (from *Salmonella*)

Need ML algorithms to predict resistance



Doxycycline resistance genes (from *Salmonella*)



Streptomycin resistance genes (from *Acidovorax*)

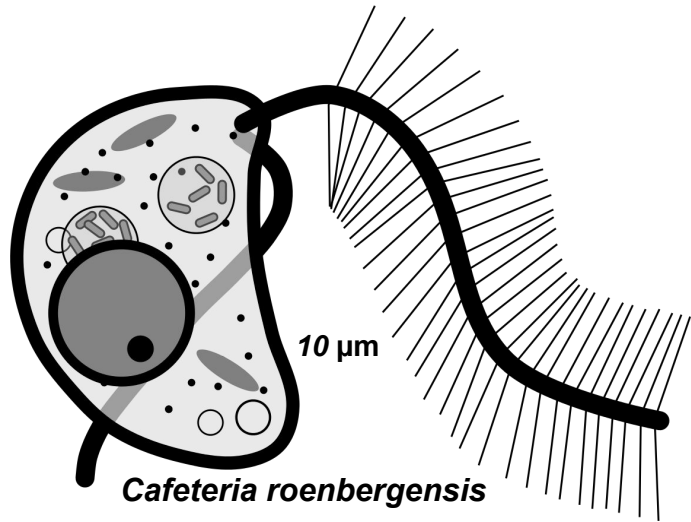
Summary!

- Bioinformatics is using **computers** to understand **biology**
- **Algorithms** are vital to doing this effectively
- We need **bioinformatic algorithms** to solve important problems including in human health

01: Life at Resolution: Organisms, Genomes, Sequences, and so on

CSCI4181/6802 Bioinformatics Algorithms
Finlay Maguire (finlay.maguire@dal.ca)

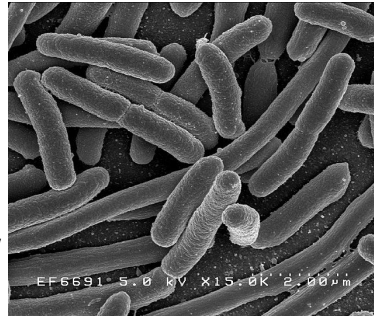
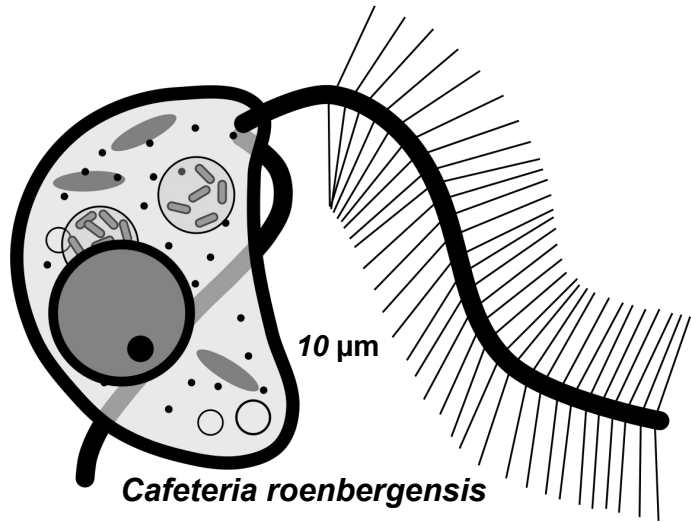
Even single microbial life covers a large range



<https://www.cam.ac.uk/sites/www.cam.ac.uk/files/styles/content-885x432/public/news/research/news/mitochondria.jpg?itok=CIFE3yjc>

https://upload.wikimedia.org/wikipedia/commons/f/f7/Giant_virus_CroV_with_its_virophage_Mavirus.png

Even single microbial life covers a large range

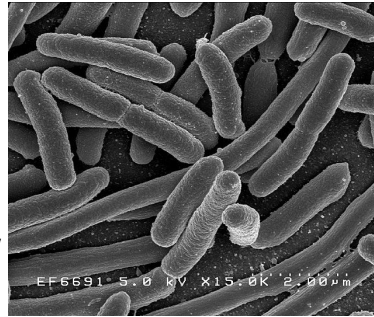
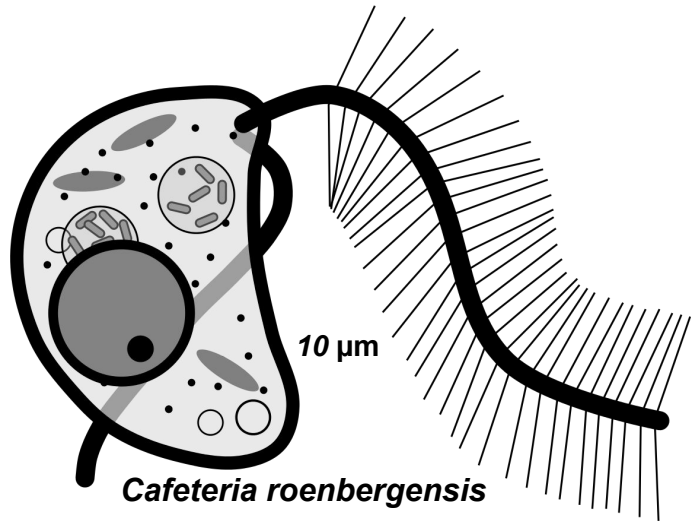


Escherichia coli

<https://www.cam.ac.uk/sites/www.cam.ac.uk/files/styles/content-885x432/public/news/research/news/mitochondria.jpg?tok=CIFE3yjc>

https://upload.wikimedia.org/wikipedia/commons/f/f7/Giant_virus_CroV_with_its_virophage_Mavirus.png

Even single microbial life covers a large range

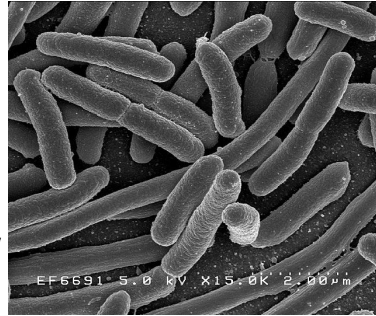
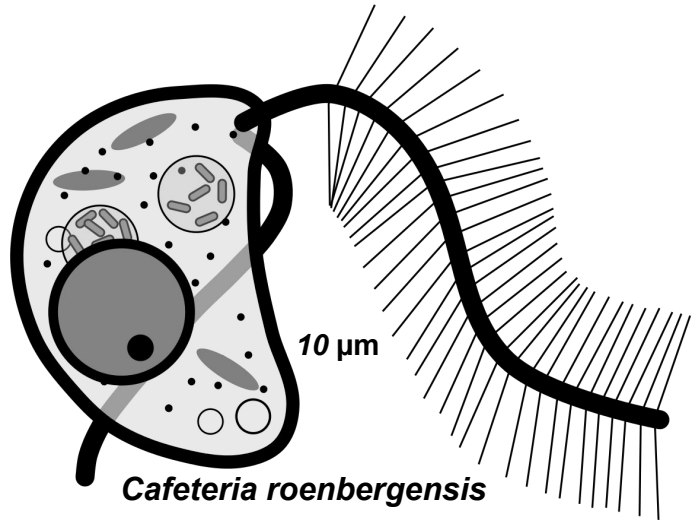


2 μm
Escherichia coli



0.5 μm
Mitochondria

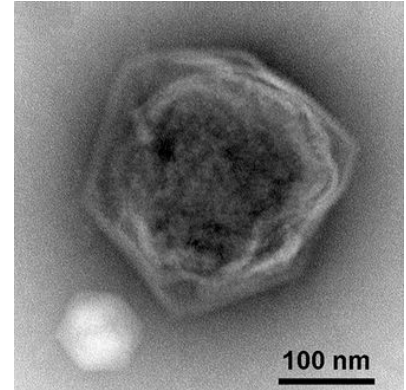
Even single microbial life covers a large range



Escherichia coli



Mitochondria



Cafeteria roenbergensis virus
Mavirus virophage

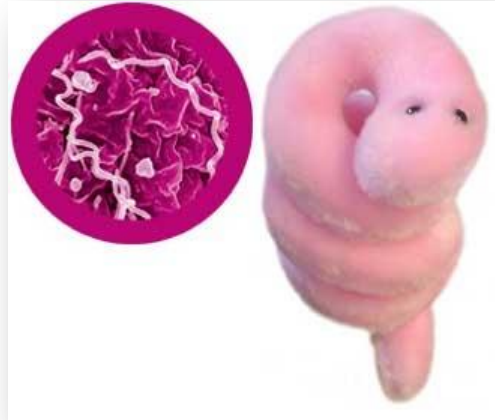
Overview

1. All living organisms have several key essential **properties**
2. Life can be viewed as a **hierarchical structure** with many levels of organization from **genome** (including genomic elements) to the **biosphere**
3. The levels we cannot observe with the naked eye are as (or more) **diverse** as the levels we can observe

Essential properties of an organism

Cellularity

Unicellular



Treponema pallidum
(www.teachersource.com)

Multicellular

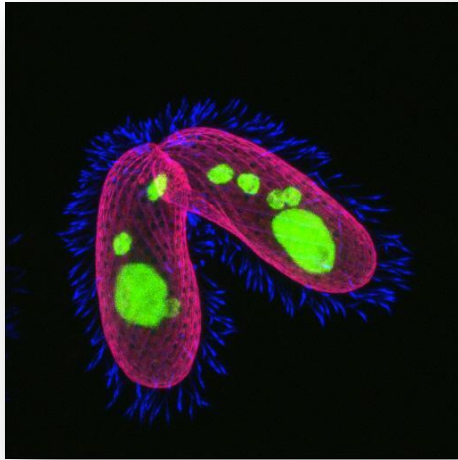


Caenorhabditis elegans (959 cells)
(www.ucl.ac.uk)

Essential properties of an organism

Reproduction

Sexual



Tetrahymena thermophila
(www.isleepinadrawer.com)

Asexual



Amoeba proteus
(www.teachnet.ie)

Essential properties of an organism

Biochemical Processes and Pathways, such as...

Fermentation



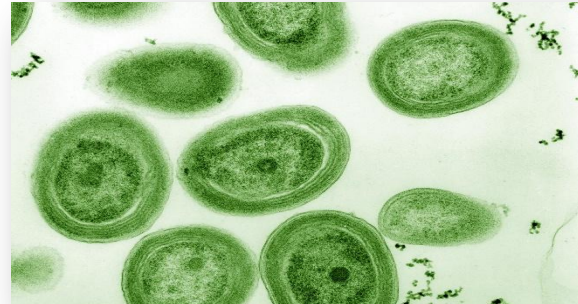
Toxin degradation



Antibiotic synthesis



Photosynthesis

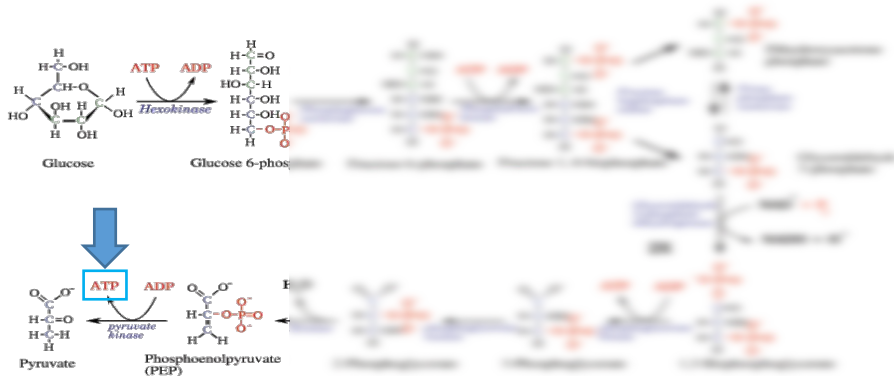


<http://en.wikipedia.org/wiki/File:Gimchi.jpg>
http://en.wikipedia.org/wiki/File:NOVAMOXIN_antibiotic.jpg
https://www.thestar.com/news/2007/01/28/sydney_tar_ponds_to_be_buried.html
https://commons.wikimedia.org/wiki/File:Prochlorococcus_marinus.jpg

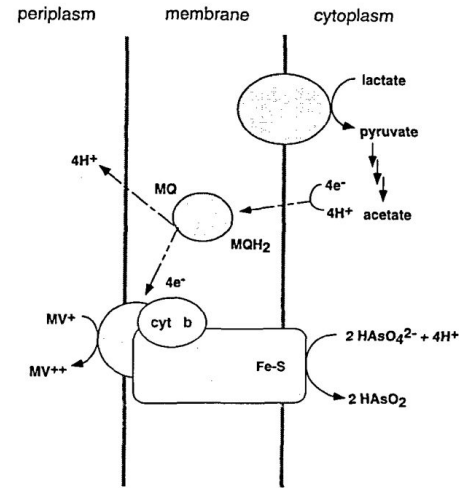
Essential properties of an organism

Respiration

Glucose - boring!



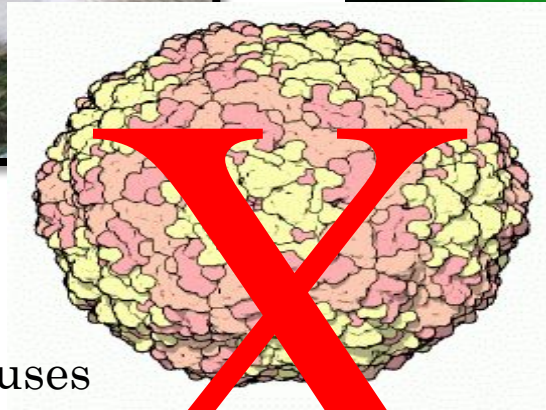
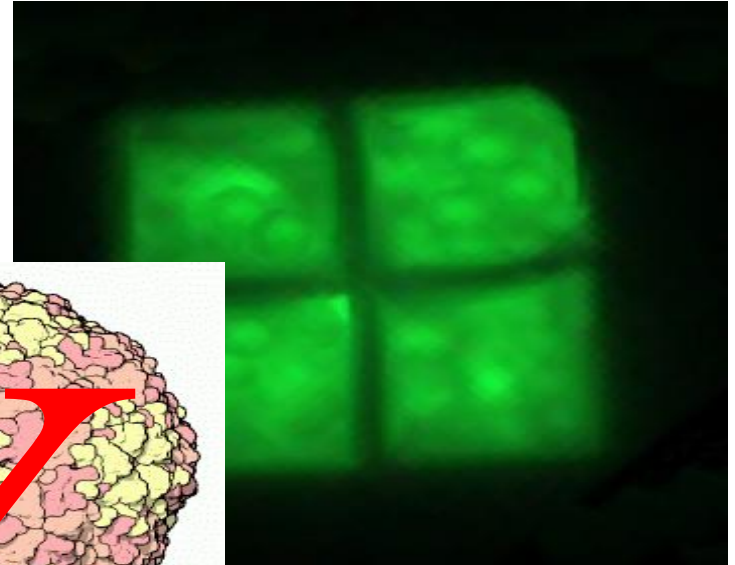
Arsenic - interesting!





Eukaryotes

Prokaryotes (Bacteria and Archaea)



Viruses
(nonliving)
(some have RNA genomes!)

Biosphere

Communities and Ecosystems

Populations

Organisms

Cells

Pathways and Systems

Proteins

Genes

Genomes



Groupings



Building blocks

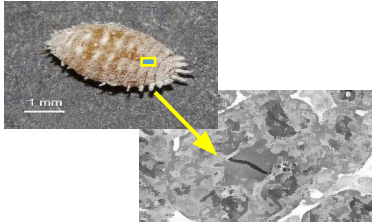
Genome:

The complete set of heritable genetic material

(DNA for all known cellular organisms)

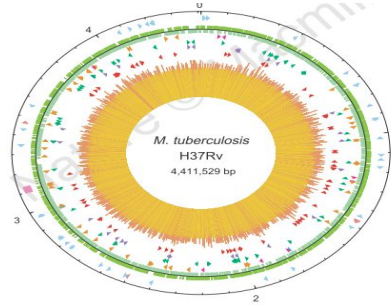
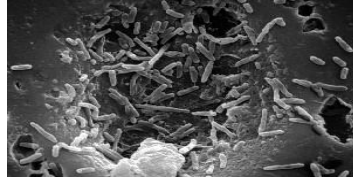
Your Genome and You

Tremblaya princeps



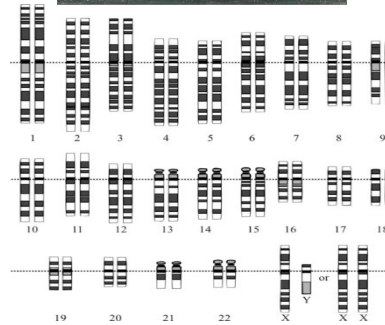
○

Mycobacterium tuberculosis H37Rv



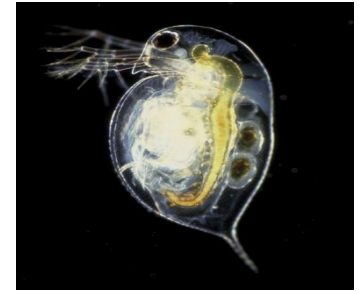
1 chromosome
110 genes
138,931 nucleotides

1 chromosome
4,000 genes
4.4 million nucleotides



23 chromosomes
+ mitochondrion
20,000 genes
3.1 billion nucleotides
(times two!)

Daphnia pulex



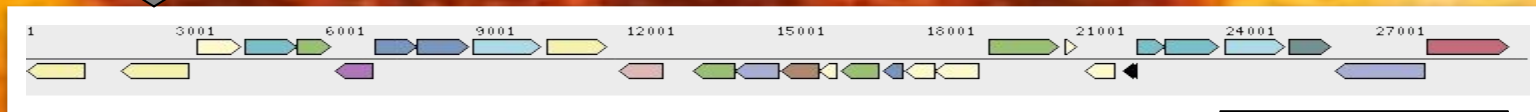
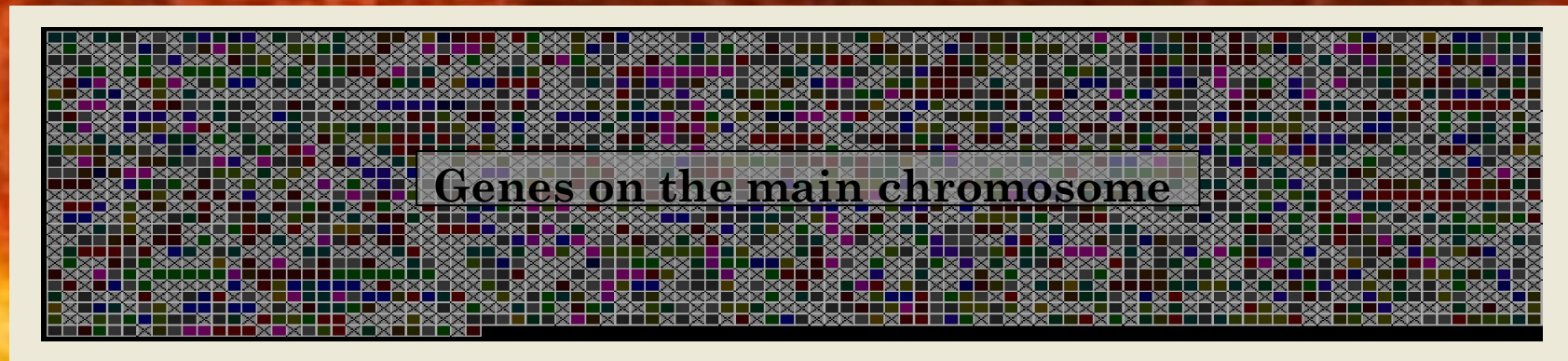
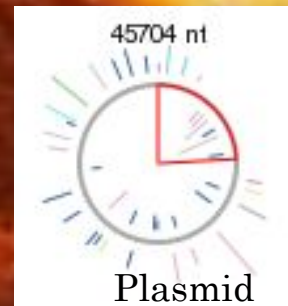
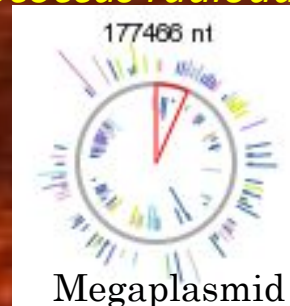
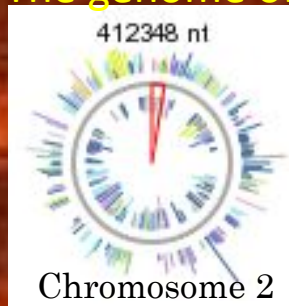
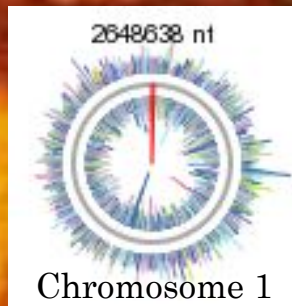
12 chromosomes
+ mitochondrion
31,000 genes
200 million nucleotides

Paris japonica

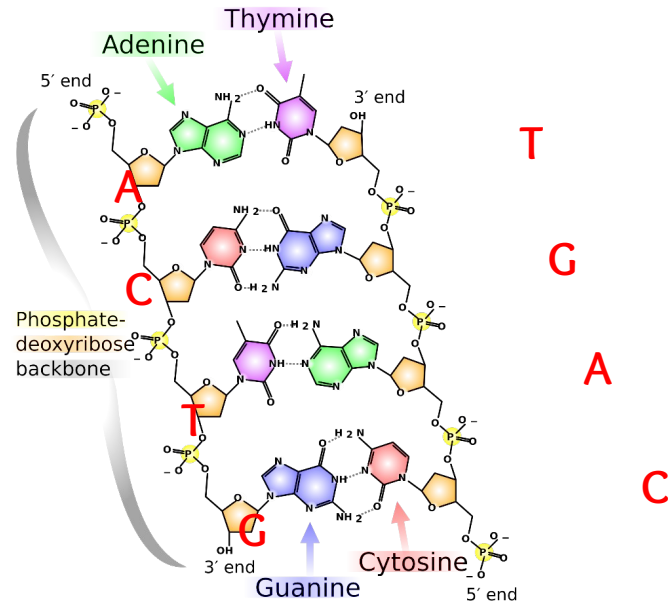


?? chromosomes
??? genes
150 billion nucleotides

The genome of *Deinococcus radiodurans*



The DNA sequence of a gene



5' - ATGCGTTACTTCGAAATGGCAACCCACTCGGGGACTTCCTCCAACGGTTGA- 3'
3' - TACGCAATGAAGCTTTACCGTTGGGTGAGCCCCTGAAGGAGGTTGCCAACT- 5'

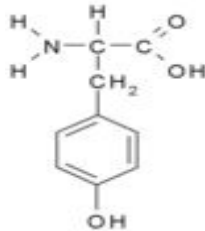
DNA to protein

DNA is read in triplets

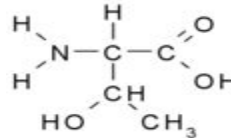
ATG CGT TAC TTC GAA ATG GCA ACC CAC TCG GGG ACT TCC TCC AAC GGT TGA



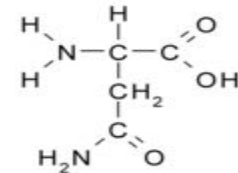
M A Y F E M A T H S G T S S N G *



Tyrosine



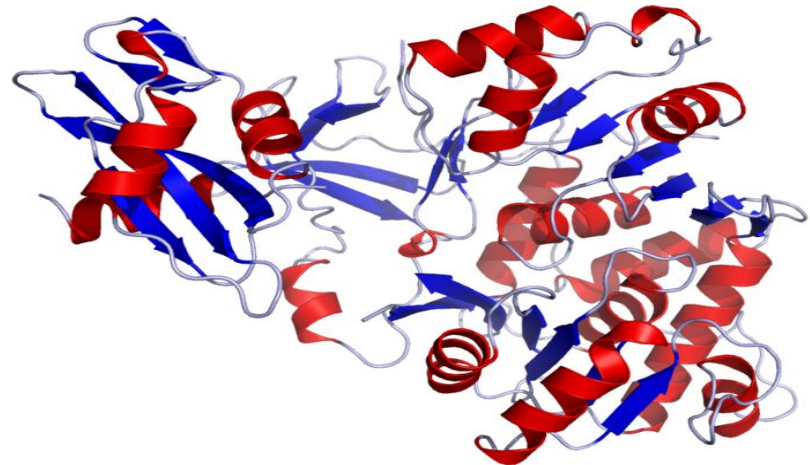
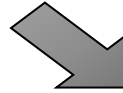
Threonine



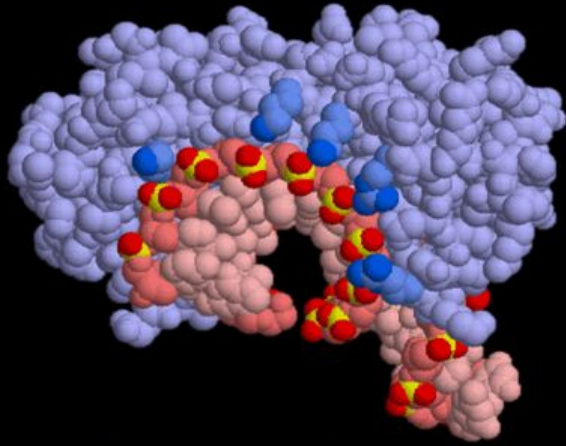
Asparagine

Protein sequence and structure

M A Y F E M A T H S G T S S N G *

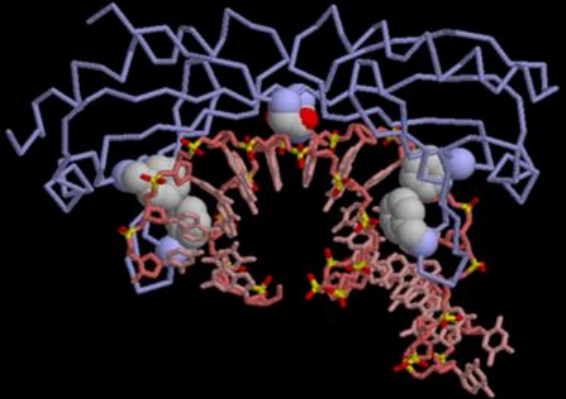


A DNA-protein complex

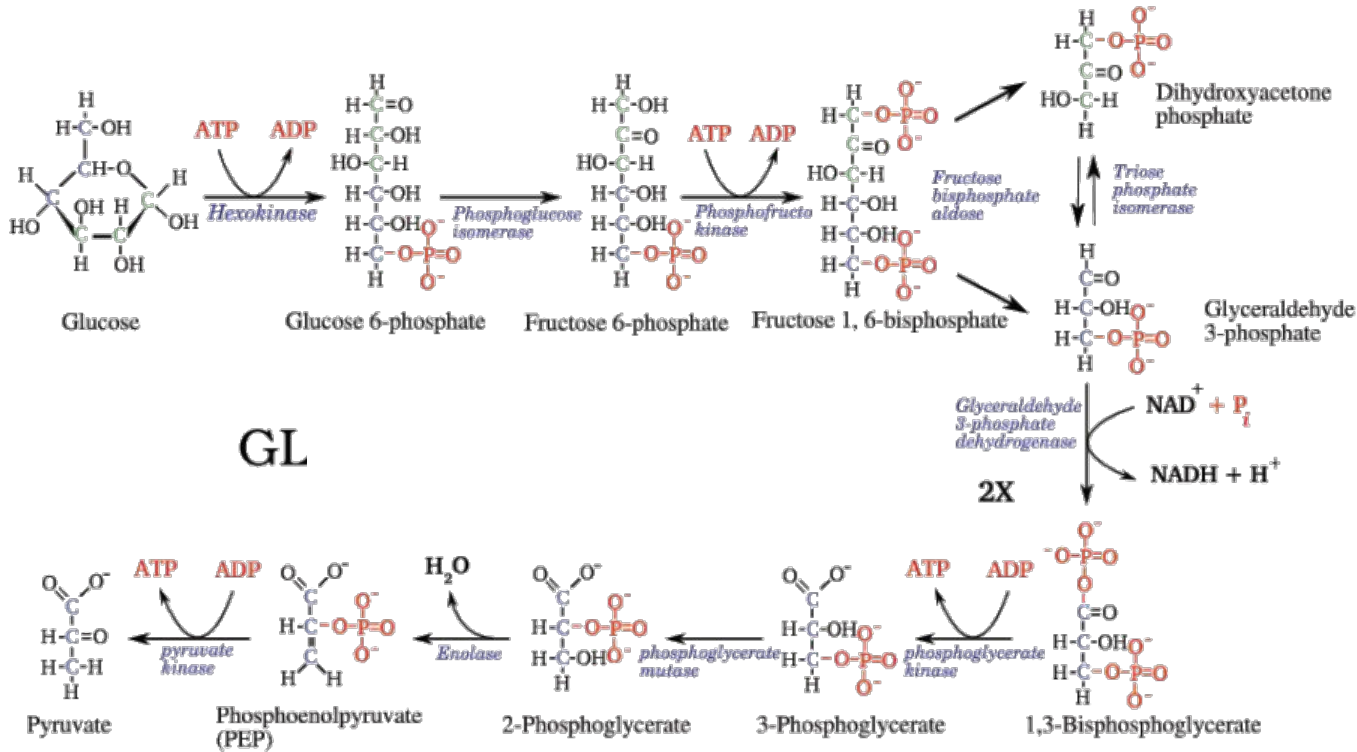


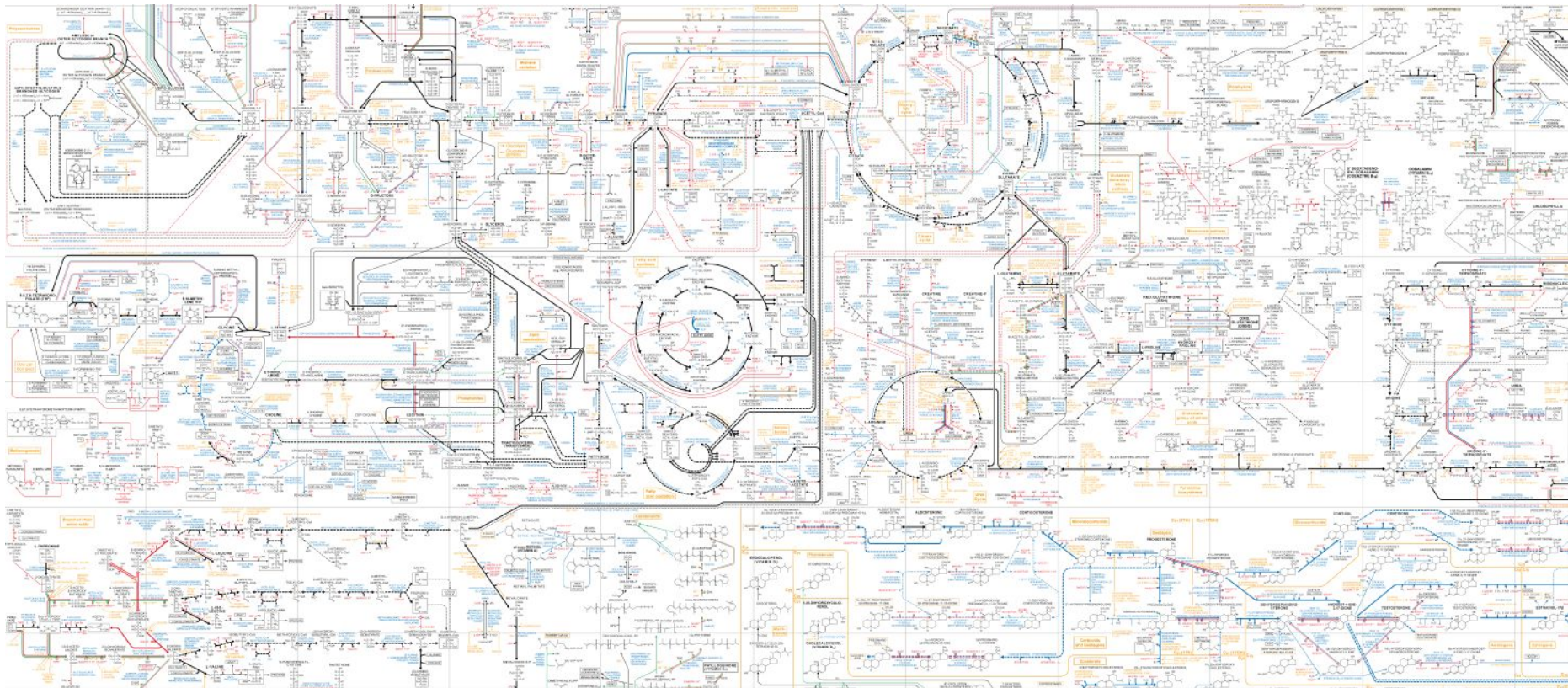
← DNA-binding protein
("TATA-box binding protein")

← DNA (note the recurring pattern;
yellow = phosphate)



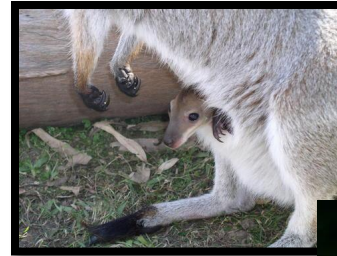
Metabolism – Proteins working together





Pathways (metabolism
+ self-replication
+ signalling)

=



Populations



Communities and Ecosystems



Overview

1. All living organisms have several key essential **properties**
2. Life can be viewed as a **hierarchical structure** with many levels of organization from **genome** (including genomic elements) to the **biosphere**
3. The levels we cannot observe with the naked eye are as (or more) **diverse** as the levels we can observe