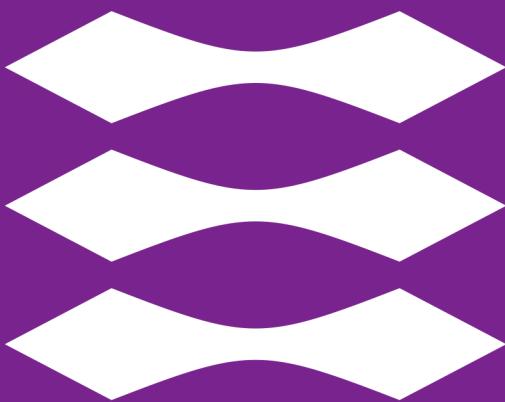


DTU



Infectious Disease Bioinformatics (23262)

Pimlapas Leekitcharoenphon (Shinny), pile@food.dtu.dk
Rolf Sommer Kaas, rkmo@food.dtu.dk
Philip Clausen, plan@food.dtu.dk
Frank Aarestrup, fmaa@food.dtu.dk

Research Group for Genomic Epidemiology



World Health Organization

EU & WHO Reference laboratory for
Antimicrobial Resistance and Genomics

- **Predicting and preventing infectious diseases in humans and supporting global detection, surveillance and control** (focus: antimicrobial resistance and zoonotic infections)

Infectious Disease Bioinformatics ?

Epidemiology

- The science that studies the patterns, causes, and effects of health and disease conditions in defined populations
- Questions;
 - What is it ?
 - Has it been seen before ?
 - How can we fight it ?
 - Is it an outbreak ?

Identification and Typing

- Any characterization below the (sub-) species level is termed “typing”
- Methods used for this characterization are per definition “typing methods”

Family

Genus

Species

(Subspecies)

Identification

Serovar

Phagetype

Ribotype

PFGE type

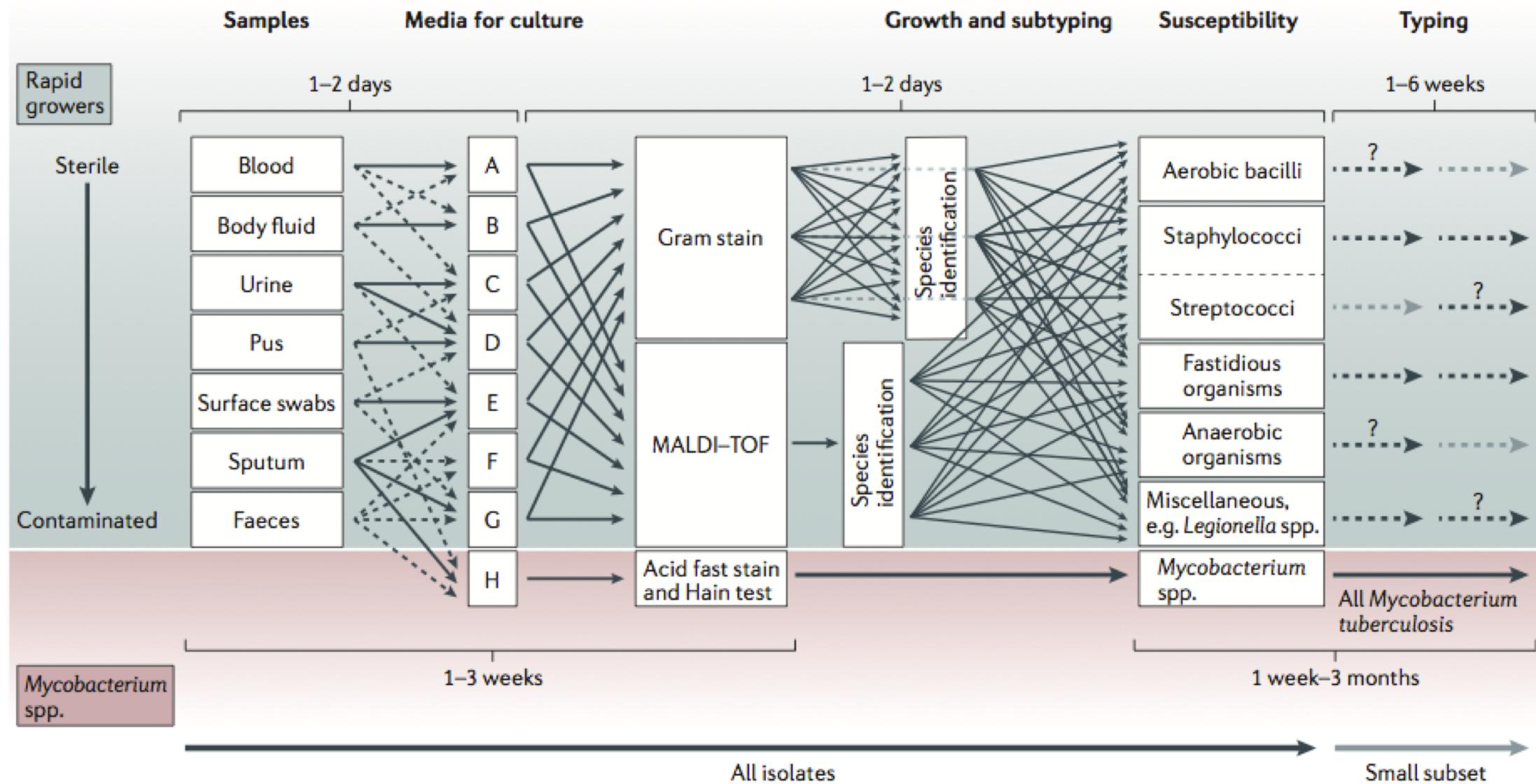
MLVA type

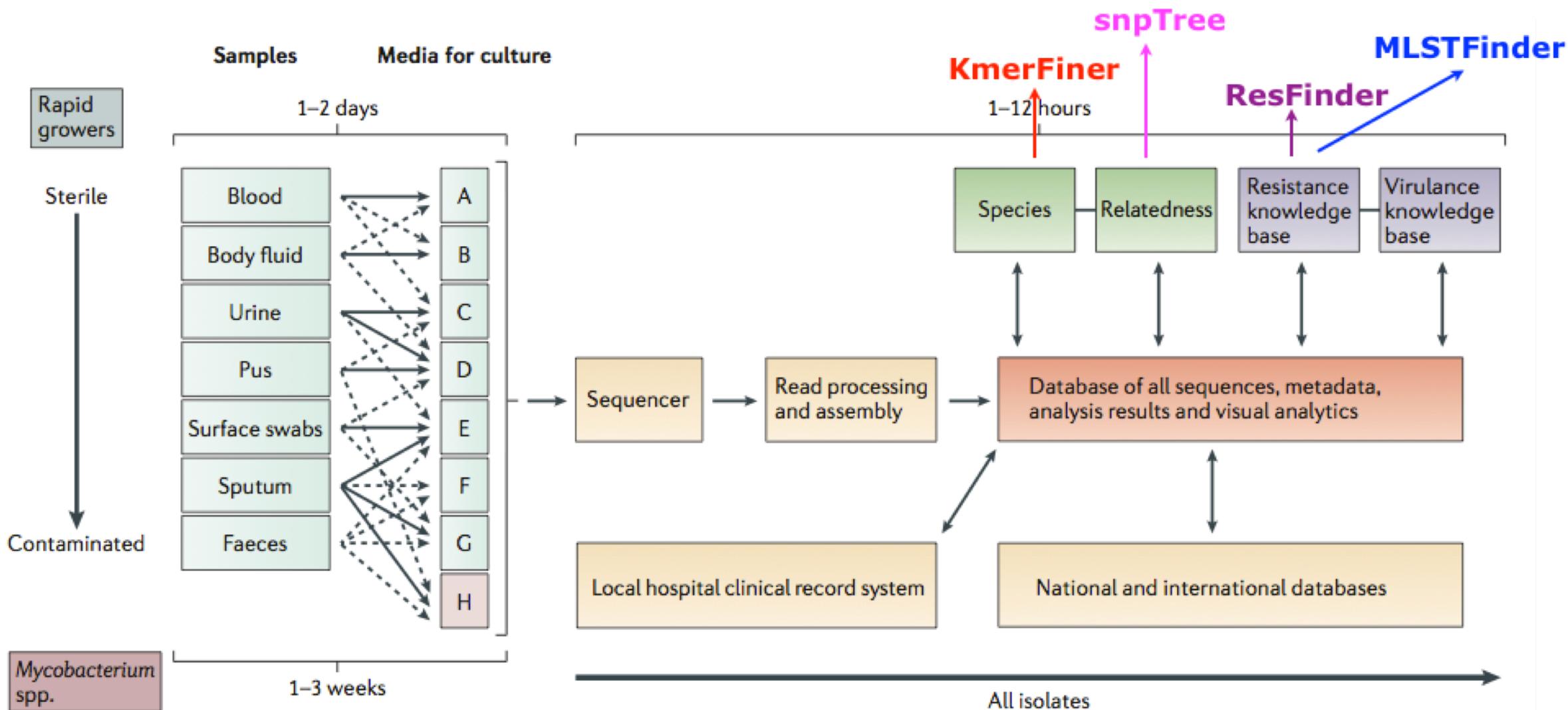
MLST type

DNA Microarray analysis

Whole genomic sequence

Typing





Pimplapas Leekitcharoenphon (Shinny)



- Researcher
- Bioinformatics background
- WGS analysis of foodborne pathogens, machine learning and metagenomics analysis for surveillance of AMR and infectious diseases.
- Onsite course on “23205 Fighting Infectious Diseases” (Fall)
- Online courses (COURSEREA) on AMR, WGS and Metagenomics

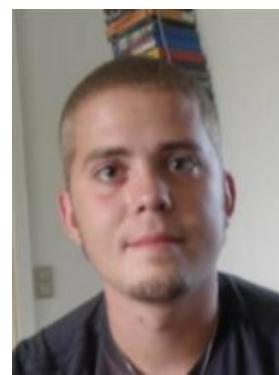


Rolf Sommer Kaas



- Senior Researcher
- Bioinformatics
- Analysis of bacterial isolates (AMR, virulence, phylogeny)
- Bioinformatics tool development
- Implementation of WGS at the Danish Veterinary and Food Administration

Philip T.L.C. Clausen

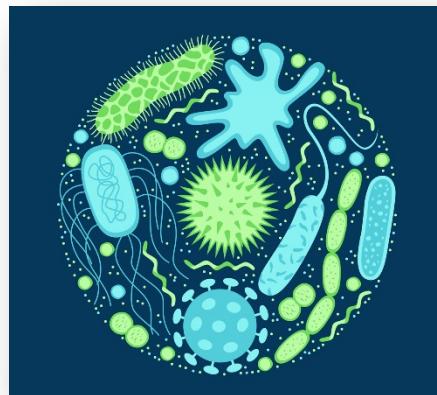


- B204/227
- Postdoc
- Bioinformatician with a Ph.D. in bioinformatical methods
- Algorithm and method development
- Microbial genomics, mostly focused on WGS of bacteria

Frank M. Aarestrup



- Veterinarian DVM with a PhD in veterinary microbiology, professor in microbiology
- Antimicrobial resistance
- Global surveillance
- Bioinformatics
- Microbial ecology



Baptiste Avot (Teaching assistant)

- Research Assistant
- Bioinformatics background



Other guest lecturers covering various topics

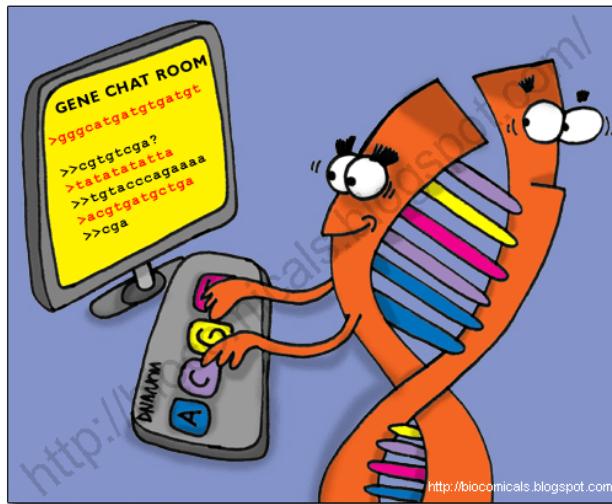
Today

- Introduction – setting the scene
- Expectations from students
- Course overview & time schedule
- Journal Club
- Set up Computerome account
- Sequencing technologies, sequencing quality, trimming and assembly WGS data

Small exercise – go to the survey under DTUlearn

Use 5-7 minutes to reflect

- Why did you select this course?
- What are your expectations to the course?
- Your programming skill?



CLASSROOM
EXPECTATIONS

23262 – Infectious Disease Bioinformatics

Course Type	MSc
Points (ECTS)	5
Duration	13 weeks (Spring F3A)
Responsible	Pimplapas Leekitcharoenphon, Rolf Kaas, Philip Clausen, Frank Aarestrup
Assessment	Project Presentation (100%), Journal Club presentation (Pass/ Not pass)
Evaluation	7 step scale, internal examiner

Learning objectives

- Be able to use whole genome sequencing data from different sequencing technologies to check for quality, trimming, mapping, alignment and assembly using command-line based bioinformatics methods.
- Apply command line-based methods to analyse species, sub-typing, identify antimicrobial resistance and other important gene markers.
- Apply command line-based methods for constructing phylogeny for epidemiological investigations, long-term evolution and large scale phylogeny.
- Apply command line-based methods to determine major and minor mutations in pathogen genomes.
- Apply command line-based methods to identify and analyse core-genome, accessory genes and pan-genome.
- Be able to download and share sequencing data from/to public databases.
- Evaluate the single components of bioinformatic workflows and pipelines.
- Apply state-of-the-art methodologies and tools to analyse and interpret genomic data from humans and various reservoirs (animals, food and environment) and prepare such data for epidemiological analyses and modelling.

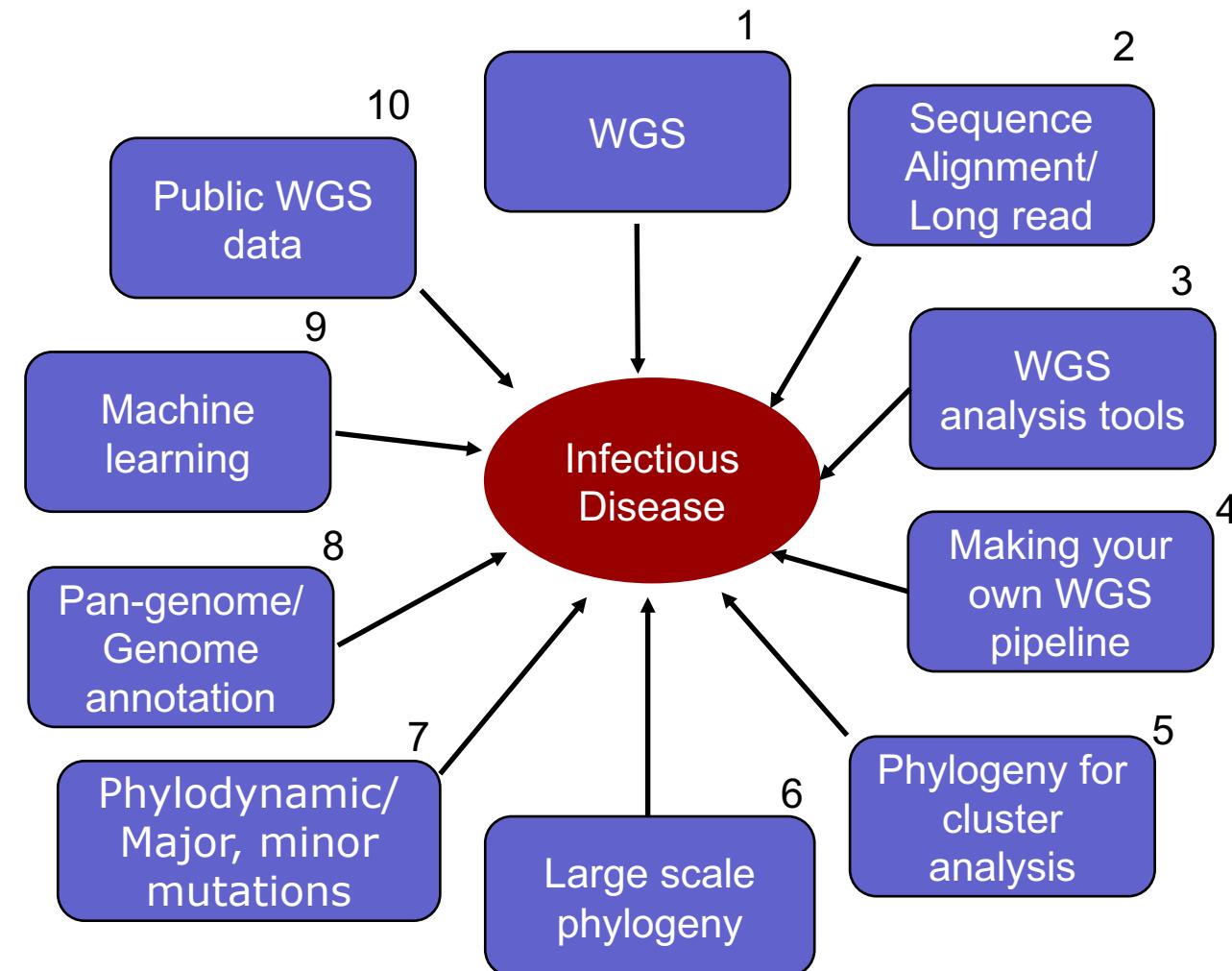
The program is available through
DTU Learn. Check it regularly, as
changes may occur

Week	Date	Time	Content	Instructor
1	01-02-22	8.00-9.00 9.00-9.30	----- Introduction to the course	Pimlapas Leekitcharoenphon
		9.30-10.00	Set up Computerome account	Baptiste Avot
		10.00-12.00	Sequencing technologies, sequencing quality, trimming and assembly of WGS data	Rolf Kaas/ Philip Clausen
2	08-02-22	8.00-10.00 10.00-12.00	Recap/ Sequence Alignment and mapping Handling long read sequencing data	Philip Clausen Philip Clausen
3	15-02-22	8.00-9.00 9.00-12.00	Journal club Recap/ Bioinformatics tools for WGS analysis of infectious diseases	Pimlapas Leekitcharoenphon/ Alfred Ferrer Florensa Pimlapas Leekitcharoenphon/ Baptiste Avot
4	22-02-22	8.00-12.00	Recap/ Making Finder tools for WGS analysis of infectious diseases	Nikiforos Pyrounakis Alfred Ferrer Florensa Rolf Kaas
5	01-03-22	8.00-9.00 9.00-11.00 11.00-12.00	Journal club Recap/ Phylogeny for outbreak investigation using SNP tree method phylogeny visualization	Rolf Kaas/ Pimlapas Leekitcharoenphon Rolf Kaas Pimlapas Leekitcharoenphon
6	08-03-22	8.00-10.00 10.00-12.00	Recap/ Phylogeny for outbreak investigation using cgMLST/wgMLST tree Large scale phylogeny for real time outbreak investigation	Pimlapas Leekitcharoenphon Judit Szarvas
7	15-03-22	8.00-10.00 10.00-12.00	Recap/ Phylogenetic inference from sequence data for epidemiology of infectious disease Determination of major and minor mutations in sequencing data	Pimlapas Leekitcharoenphon/ Baptiste Avot Philip Clausen/Malte Hallgren
8	22-03-22	8.00-9.00 9.00-11.00 11.00-12.00	Journal club Recap/ Pan-genome analysis for infectious disease Genome annotation	Pimlapas Leekitcharoenphon/ Philip Clausen Pimlapas Leekitcharoenphon/ Baptiste Avot Joseph Martin

The program is available through
DTU Learn. Check it regularly, as
changes may occur

9	29-03-22	8.00-9.00	Journal club	Alexander/ Patrick
		9.00-10.00	Recap (Pimplapas)/ Machine learning modelling for infectious disease	Patrick Njage/ Alexander Gmeiner
		10.00-11.00	Processing, analysing and preparing genomic data for epidemiological analyses and modelling (including computer exercise)	Alexander Gmeiner
		10.30-12.00	Machine learning exercise	Patrick Njage/ Alexander Gmeiner
10	05-04-22	8.00-9.00	Recap (Pimplapas)/ Importance of sharing sequencing data for preventing and controlling infectious disease	Pimplapas Frank Aarestrup
		9.00-10.00	How to download sequencing data of infectious disease from public database	Judit Szarvas
		10.00-11.00	How to share sequencing data to public database	Pimplapas Leekitcharoenphon
		11.00-12.00	Introduction to dataset for project work	Pimplapas Leekitcharoenphon
11	19-04-22	8.00-12.00	Project work	Pimplapas /Rolf/Philip
12	26-04-22	8.00-12.00	Project work	Pimplapas /Rolf/Philip
13	03-05-22	8.00-12.00	Final exam (project presentation)	Pimplapas /Rolf/Philip + external censor

Infectious Disease Bioinformatics



Computerome account

- Computer exercises require Computerome account
- If you have or do not have Computerome account, please send the following information to Baptiste (teaching assistant, bavja@food.dtu.dk) and Pimlapas (pile@food.dtu.dk)
 - **Name:**
 - **Student ID:**
 - **Email:**
 - **Phone number:**

Course material

Literature, notes, slides, videos, etc.

DTU Learn

- <https://learn.inside.dtu.dk/d2l/le/content/102833/Home>

Relevant online MOOC courses (OPTIONAL)

Antimicrobial Resistance & Methods

- <https://www.coursera.org/learn/antimicrobial-resistance>

Whole Genome Sequencing (WGS)

- <https://www.coursera.org/learn/wgs-bacteria>

Journal Club

Reading and interpreting infectious disease bioinformatics related articles

How to read a scientific paper

1. Read the introduction (not the abstract)
2. Identify the main research question
3. Summarise the background in a few sentences
4. Identify the specific research questions
5. Identify the approach
6. Draw the diagram for experiments.
Showing methods
7. Summarise results for each experiment
8. Do the results answer the specific question(s)
9. Read the conclusion/discussion/
interpretation section
10. Now read the abstract
11. What do other researchers say about this paper?

Journal Club

Reading and interpreting infectious disease bioinformatics related articles

- The presenting team will **submit 1 item on the day of the presentation:**
 - **Presentation slides** (as team – at DTU Learn)
- The opposing team will **submit 1 item before the day of the presentation :**
 - **A minimum of 5 questions** prepared and submitted beforehand (as team – at DTU Learn)



Team work – Journal Club - articles

8 papers – 1 paper per group



The Scientific Paper

Papers are uploaded to DTU Learn: Content < Journal Club

Journal Club Groups

Let us know which Journal Club Team you will join asap and no later than next Tuesday, February 8th. *The journal club team is the same team for the project presentation.*

- Select your group under 'My Course < Groups' at DTU Learn

You find the list of papers for Journal Club under '[Content < Journal Club](#)', where paper #1 is allocated to Team 1, paper #2 to Team 2 and so forth

Journal club guidelines and five question form are under '[Content < Journal Club](#)'

Journal club - 23262 Infectious Disease Bioinformatics

8:00 – 9:00

Date	B210-H168	B202-R1005	B202-R1013
15th February	Topic: Team: "" Presenter: Team Opponent: Team Moderator:	Topic: Team X: "" Presenter: Team Opponent: Team Moderator:	-
1st March	Topic: Team : "" Presenter: Team Opponent: Team Moderator:	Topic: Team : "" Presenter: Team Opponent: Team Moderator:	-
22nd March	Topic: Team : "" Presenter: Team Opponent: Team Moderator:	-	Topic: Team : "" Presenter: Team Opponent: Team Moderator:
29th March	Topic: Team : "" Presenter: Team Opponent: Team Moderator:	Topic: Team : "" Presenter: Team Opponent: Team Moderator:	-

Project presentation

The journal club team is the same team for the project presentation.

WGS dataset will be available to choose (on week 10)

The team can also come up with their own WGS dataset

The team will

- Form research question(s)
- Conduct bioinformatics analysis and/or develop/improve pipeline/database
- Interpret and conclude results
- Present the project (on Week 13)
- Answer questions individually

DTU

