Medical Microbiology

คณะแพทยศาสตร์ มหาวิทยาลัยขอนแก่น
FACULTY OF MEDICINE KHON KAEN UNIVERSITY

RCEID

Thesis Progression

Genomic characterization and pangenome analysis of *Mycobacterium kansasii*

Weawwadee Sribunrueang
1 st year M.Sc. student
09.30 AM-10.30 PM 18 February,2026

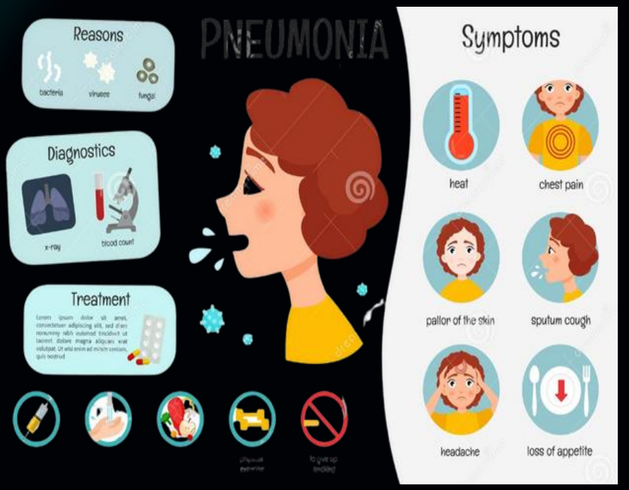
Advisor : Asst. Prof. Dr. Auttawit Sirichoat

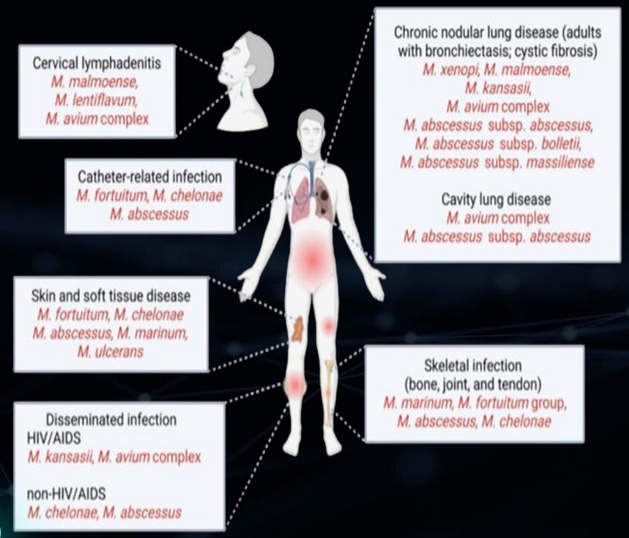
Introduction

NON-TUBERCULUS MYCOBACTERIA (NTM)

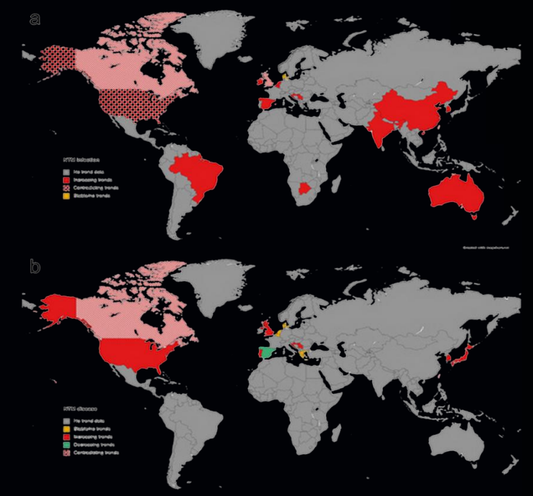
- Aerobic, Non-motile
- Acid-fast bacilli
- >180 species
- Rapid growing, Slow growing
- Found in environments
- Difficult to diagnosis
- Multi-drugs resistant
- Pneumonia
- Opportunistic infection (Parte et al., 2020)
- A global increase in reported NTM-associated infections and diseases

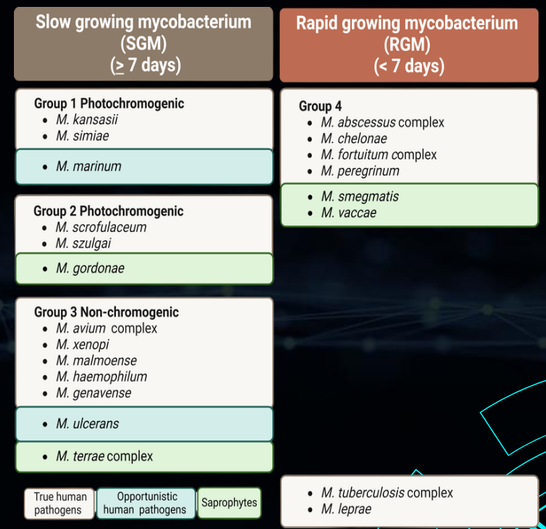
(Dahl et al. 2022)





(Baldwin, 2019)



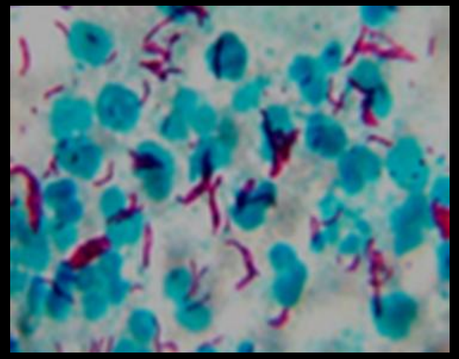


(RUNYON, 1959)

Introduction

Mycobacterium kansasii

- Member of the *Mycobacterium kansasii* complex (MKC)
- **Slow-growing** non-tuberculous mycobacterium (NTM)
- Commonly found in **tap** and **municipal water systems** (Vaerewijck et al., 2005)
- **First** isolation location in **Kansas City**, USA (Hauduroy, 1955)
- Causing **pulmonary infections** similar to tuberculosis



Mycobacterium kansasii
Hauduroy 1955, ATCC 12478

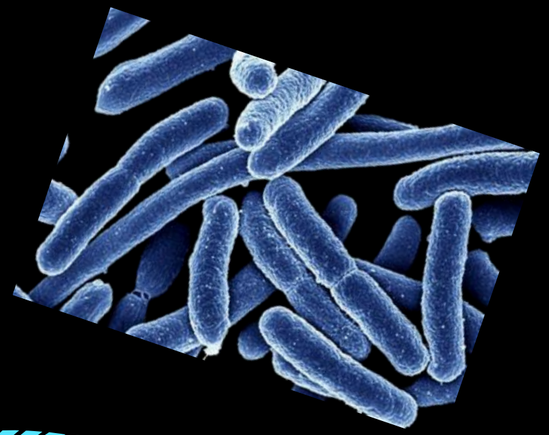
M.kansasii (a) nonpigmented, rough colonies in the dark on Middelbrook 7H11 agar

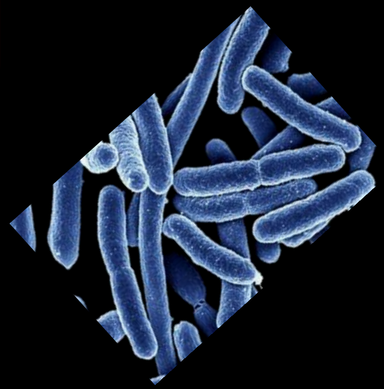
(Gutierrez et al., 2014)

Introduction

Member of the *Mycobacterium kansasii* complex (MKC)

- ***Mycobacterium kansasii*** **(subtype I) (major human pathogen)**





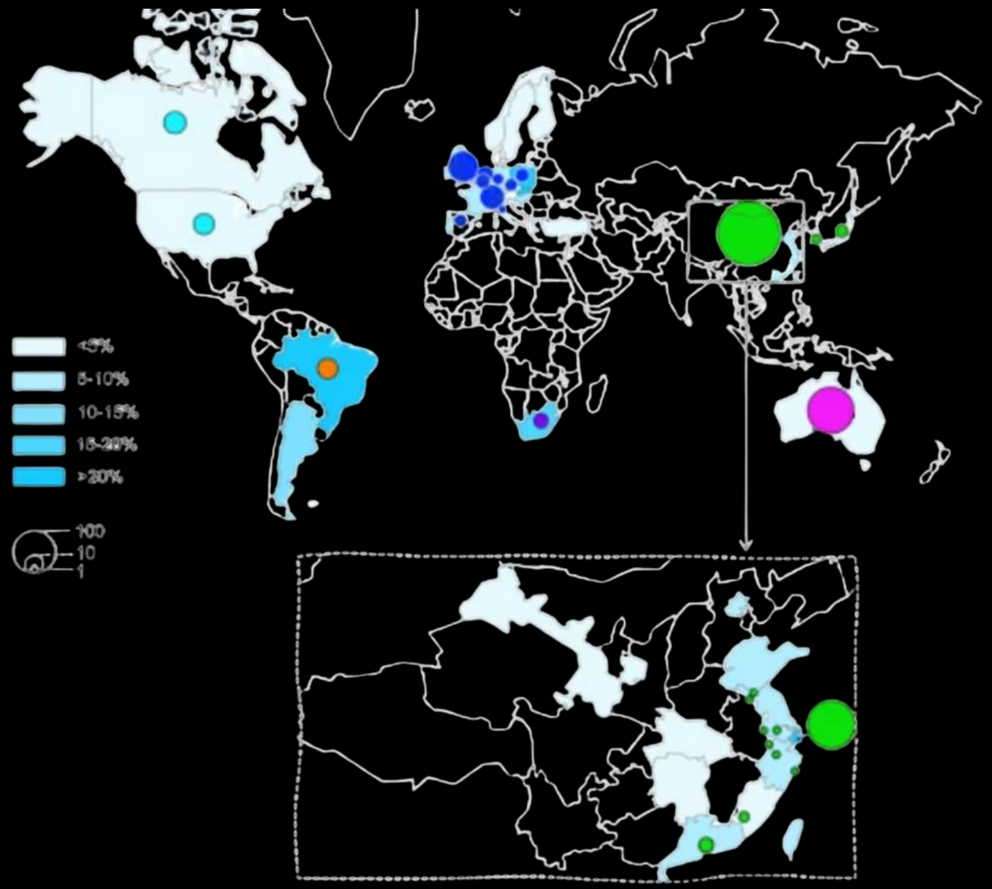
- ***Mycobacterium persicum*** **(subtype II)**
- ***Mycobacterium pseudokansasii*** **(subtype III)**
- ***Mycobacterium ostraviense*** **(subtype IV)**
- ***Mycobacterium innocens*** **(subtype V)**
- ***Mycobacterium attenuatum*** **(subtype VI)**
- ***Mycobacterium gastri*** (Luo et al., 2021)

Introduction

Trends of *M. kansasii* in global

- ***M. kansasii*** **infections have been documented worldwide**
- **High prevalence in England, Wales and South Africanminers**
  (Corbett et al., 1999)
- **Most common in sub-Saharan Africa**
  (Huang et al., 2017)
- ***M. kansasii*** **among NTM isolates is highest in Europe (~12.1%) lowest in North America (~2-3%)**
  (Narimisa et al., 2024)
- **In Thailand, the increasing incidence of NTM-related diseases including those caused by *M. kansasii***
  (สำนักวัณโรค, 2024)

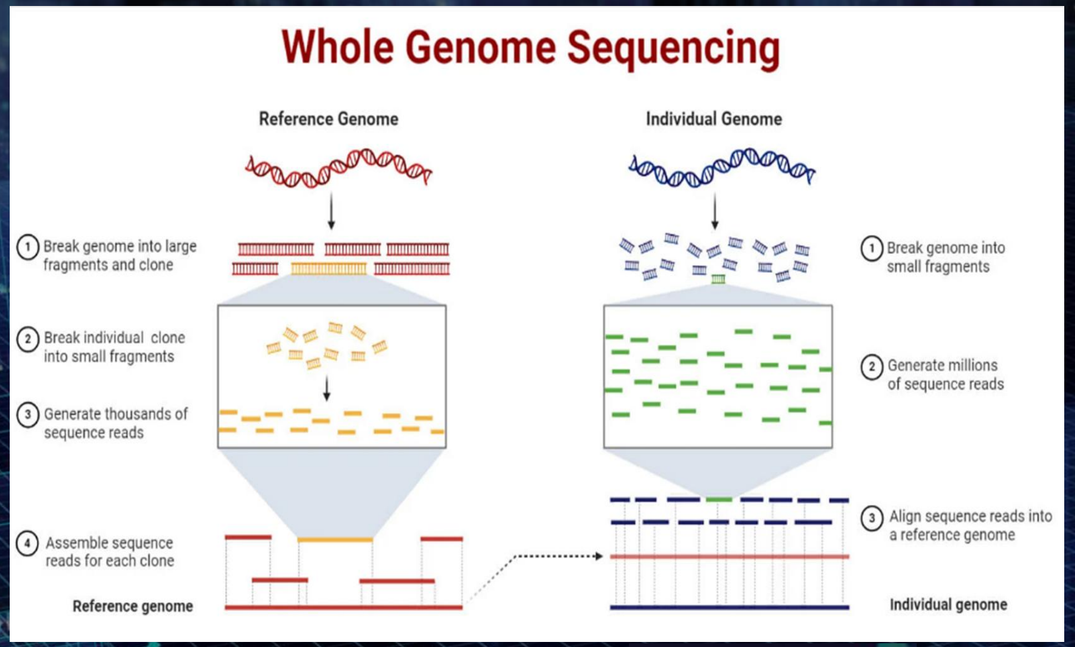
Global diversity of the *M. kansasii*



Geographical distribution of the 358 isolates in the study. The gradient blue colors indicate the prevalence of *M. kansasii* among NTM disease. (Luo et al., 2021)

Introduction

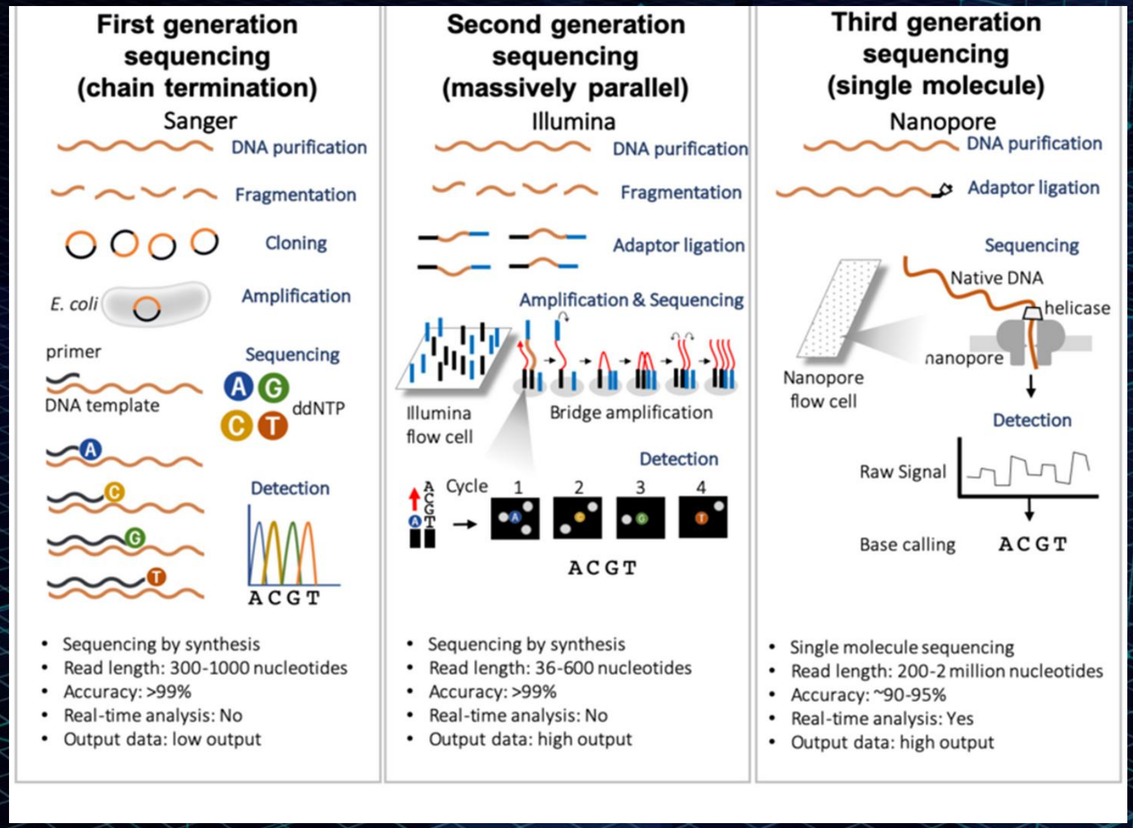
Whole genome sequencing (WGS) technology



(Sanju Tamang:2024)

(Rodriguez et al. 2020)

NEXT GENERATION SEQUENCING (NGS) TECHNOLOGY



- Sequencing by synthesis
- Read length: 300-1000 nucleotides
- Accuracy: >99%
- Real-time analysis: No
- Output data: low output

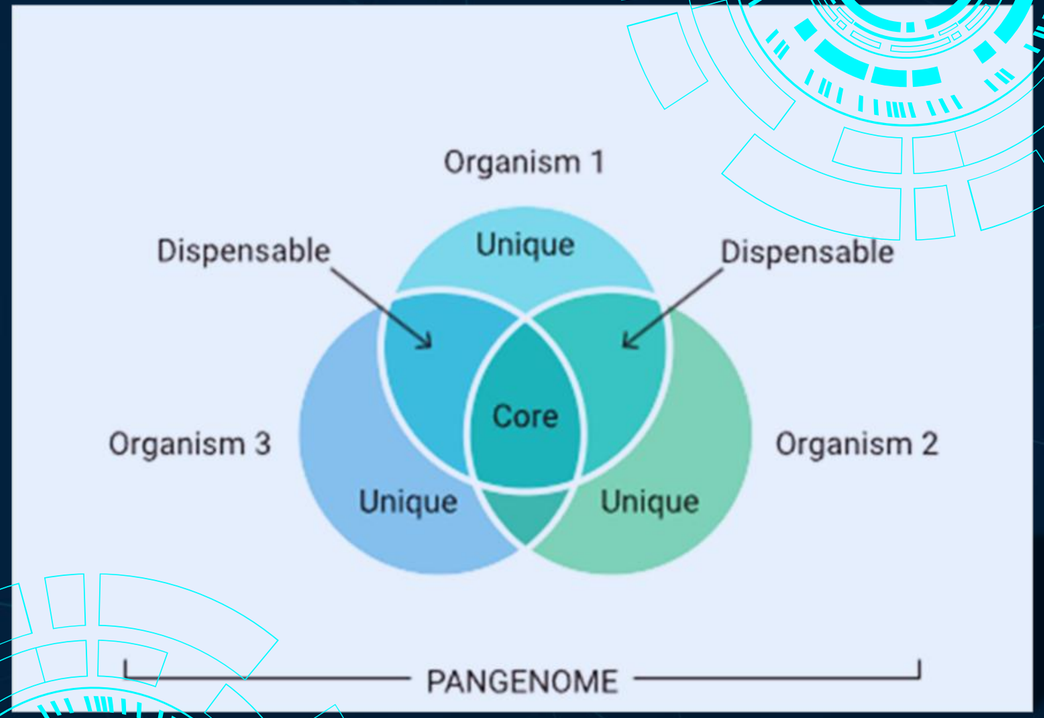
- Sequencing by synthesis
- Read length: 36-600 nucleotides
- Accuracy: >99%
- Real-time analysis: No
- Output data: high output

- Single molecule sequencing
- Read length: 200-2 million nucleotides
- Accuracy: ~90-95%
- Real-time analysis: Yes
- Output data: high output

Introduction

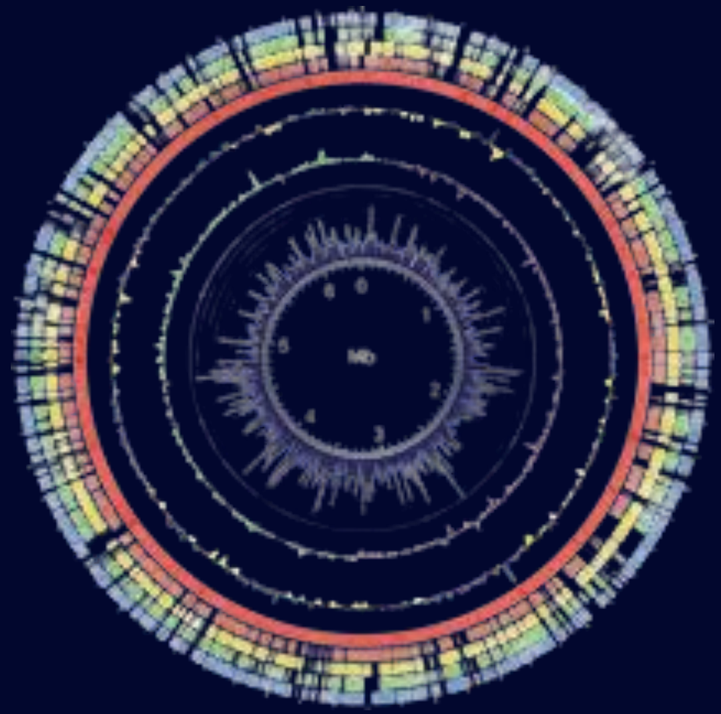
PANGENOME ANALYSIS

- A **comparative genomics** approach that examines all genes across multiple isolates of the same species
- **Identifies** both **shared** (conserved) and **variable** genes among related genomes
- Categorized into **core, accessory** and **unique** genomes (Jagielski et al., 2019)



Introduction

Literature Reviews



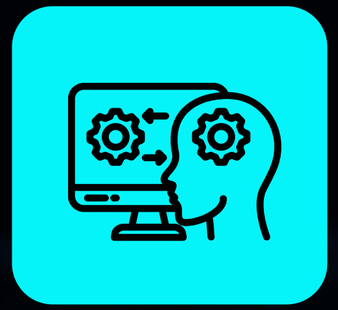
There are many studies on the genome of M. kansasii

- Luo et al. study in 2021 about Population genomics
- Tagini et al. study in 2021 about pathogenicity determination
- Keith et al. in 2023 about essentiality genome
- Sur S. in 2023 study about Pangenome structure of *M. kansasii*
- Machado et al. in 2024 study about Large-scale genomic

HYPOTHESIS

- *M. kansasii* exhibit significant genetic diversity
- Pangenome analysis will reveal distinct genomic features and structural variations that contribute to this diversity
- Enhancing our understanding of the evolutionary and adaptive characteristics of *M. kansasii*

OBJECTIVES



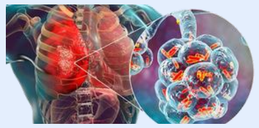
1. To investigate the genetic diversity among isolates of the *M. kansasii* isolates

2. To characterize the genome structure of *M. kansasii* using pangenome analysis

CONCEPTUAL FRAMEWORK

Problem

> 180 species

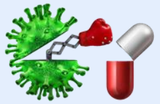


Pneumonia & Opportunistic infection

Difficult to diagnosis

Studies on *M. kansasii* in Thailand is limited

Increasing global trends of pulmonary infections with NTM



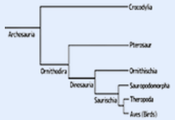
Multi-drugs resistant

Process

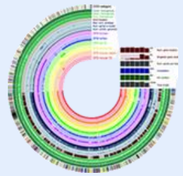
Data collection from Srinagarind Hospital by RCEID

Data processing and quantity check

Species identification and characterization

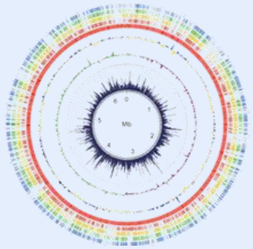


Phylogenetic analysis



Pangenome analysis

Output



Genome structure and genetic diversity

Genome database of *M. kansasii*

Bioinformatics and big data analysis skills

Research publication

Outcome & Impact

Academic development and highly skilled

Management public health control

Precision & Accuracy

Drug development

Diagnosis development

EXPERIMENTAL DESIGN

Isolates in this study kindly provided by the staff at the Research and Diagnostic Center for Emerging Infectious Disease (RCEID), Khon Kaen University, Khon Kaen, Thailand

RCEID

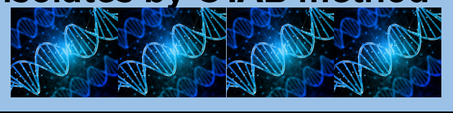
Data collection from Srinagrind Hospital

NTM isolates from Srinagarind Hospital

→ *M. kansasii* Identification by LPA

GenoType Mycobacterium CM VER 2.0

→ Genomic extraction from *M. kansasii* isolates by CTAB method



→ Whole-genome sequencing by Illumina platform

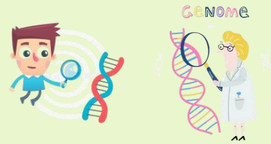
illumina

Bioinformatic analysis

1. Sequencing quality assessment

→ 2. Genome assembly and identification of *M. kansasii* species

→ 3. Genome characterization

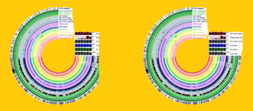


4. Public genome retrieval and species confirmation

NIH National Library of Medicine
National Center for Biotechnology Information

→ 5. Phylogenetic analysis

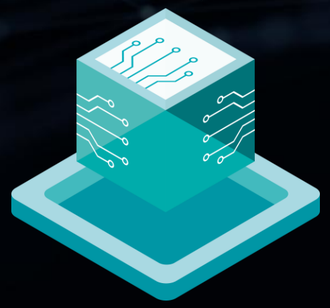
→ 6. Pangenome analysis



NTM isolates from Srinagarind Hospital

Data collection from Srinagrind Hospital

Isolates in this study kindly provided by the staff at the Research and Diagnostic Center for Emerging Infectious Disease (RCEID), Khon Kaen University, Khon Kaen, Thailand

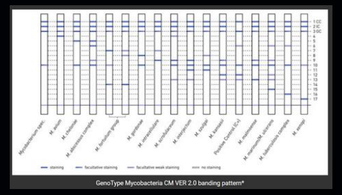


RCEID

- Six *M. kansasii* were obtained from the Clinical Micro Lab, MD, KKU 2012 - 2016
- Sputum, pleural fluids, lymph node tissues

M. kansasii identification

- LPA, GenoType Mycobacterium CM VER 2.0



Genomic extraction

- CTAB method

Whole-genome sequencing

illumina

- DNA was subsequently sent to NovogeneAIT (Singapore) for whole-genome sequencing using the Illumina HiSeq 2500 platform
- Paired-end sequencing (2 × 150 bp)

Bioinformatic analysis

Sequencing quality assessment

Raw sequencing reads (FastQ)

Quality check using FastQC v0.12.1

(Andrews,2010)

Trimming and filtering with Trimmomatic v0.39

(LEADING:3, TRAILING:3, SLIDINGWINDOW:4:15, MINLEN:75)

Trimmomatic (Bolger et al., 2014)

Bioinformatic analysis

Genome assembly and species identification of *M. kansasii*

***De novo* assembly**

Unicycler v0.5.1

(Wick et al., 2017)

Scaffold improvement

RagTag v2.1.0

(Alonge et al., 2022)

Polishing

Pilon v1.24

(Walker et al., 2014)

Assembly quality

QUAST v5.3.0

(Gurevich et al., 2013).

Genome completeness & contamination

CheckM v1.2.3

(Parks et al., 2015

Inclusion criteria

≥95% completeness

≤5% contamination

Results : Characteristics and studied isolated

Method

- Bacterial isolated
- Source
- Collection date

QUAST

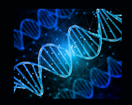
- Genome size(bp)
- GC content
- NO of contige
- Contig N50
- Contig L50

BV-BRC

- No. of CDs
- No.of subsystems
- tRNAs
- RRNA

General genome features of the 6 *Mycobacterium kansasii* isolates

Feature/protein coding genes	Isolate					
	MKA80499	MKA80686	MKA80709	MKA82243	MKA82252	MKA82253
Source	Sputum	Sputum	Sputum	Pleural fluid	Lymph node (TISSUE)	Sputum
Collection date (d/m/y)	20/2/2013	11/3/2013	12/2/2013	17/9/2016	19/9/2016	19/9/2016
Genome size (bp)	6,505,392	6,425,212	6,416,927	6,538,639	6,533,506	6,536,265
GC content(%)	66.09	66.2	66.2	66.14	66.14	66.13
NO. of contigs	24	28	28	26	25	27
Contig N50 (bp)	6,476,335	6,384,716	6,374,337	6,513,204	6,508,363	6,509,982
Contig L50	1	1	1	1	1	1
No. of CDSs	6,159	5,997	6,003	6,142	6,167	6,159
No. of subsystems	310	309	309	309	310	309
tRNAs	50	50	51	49	49	49
rRNA	3	3	3	3	3	3



Genome identification of *M. kansasii*

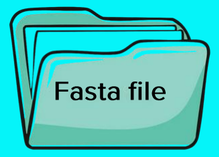
Species identification using multiple tools

- NTM Profiler
- Type Strains Genome Sever (TYGS)
- Pathogenwatch
- GTDB-Tk
- FastANI
- ANIclustermap
- PubMLST

- A total of 6 isolates were confirmed as *M. kansasii*
- ST38 predominant lineage among the sampled isolates

Results : Functional subsystem annotation

Method

6 x 

⬇

BV-BRC

⬇

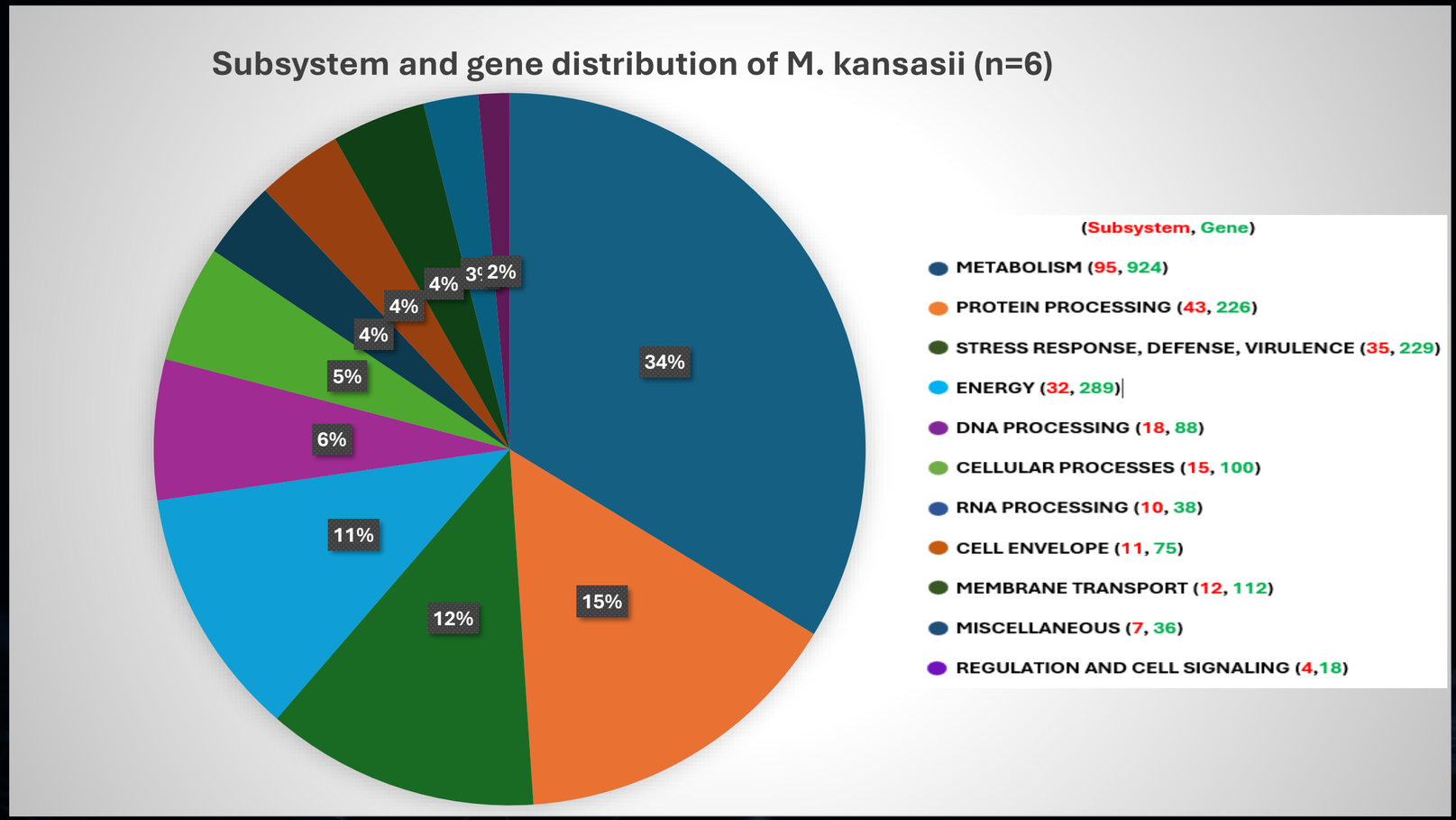
RAST toolkit

⬇

Gene prediction + subsystem annotation

⬇

Function categories subsystem chats



Majority of genes were associated with metabolic processes(34%), followed by protein processing(15%), stress response(12%), energy production(11%), DNA-related functions(6%), cellular processes(5%) and RNA processing(4%).

Results : Virulence factor profiling

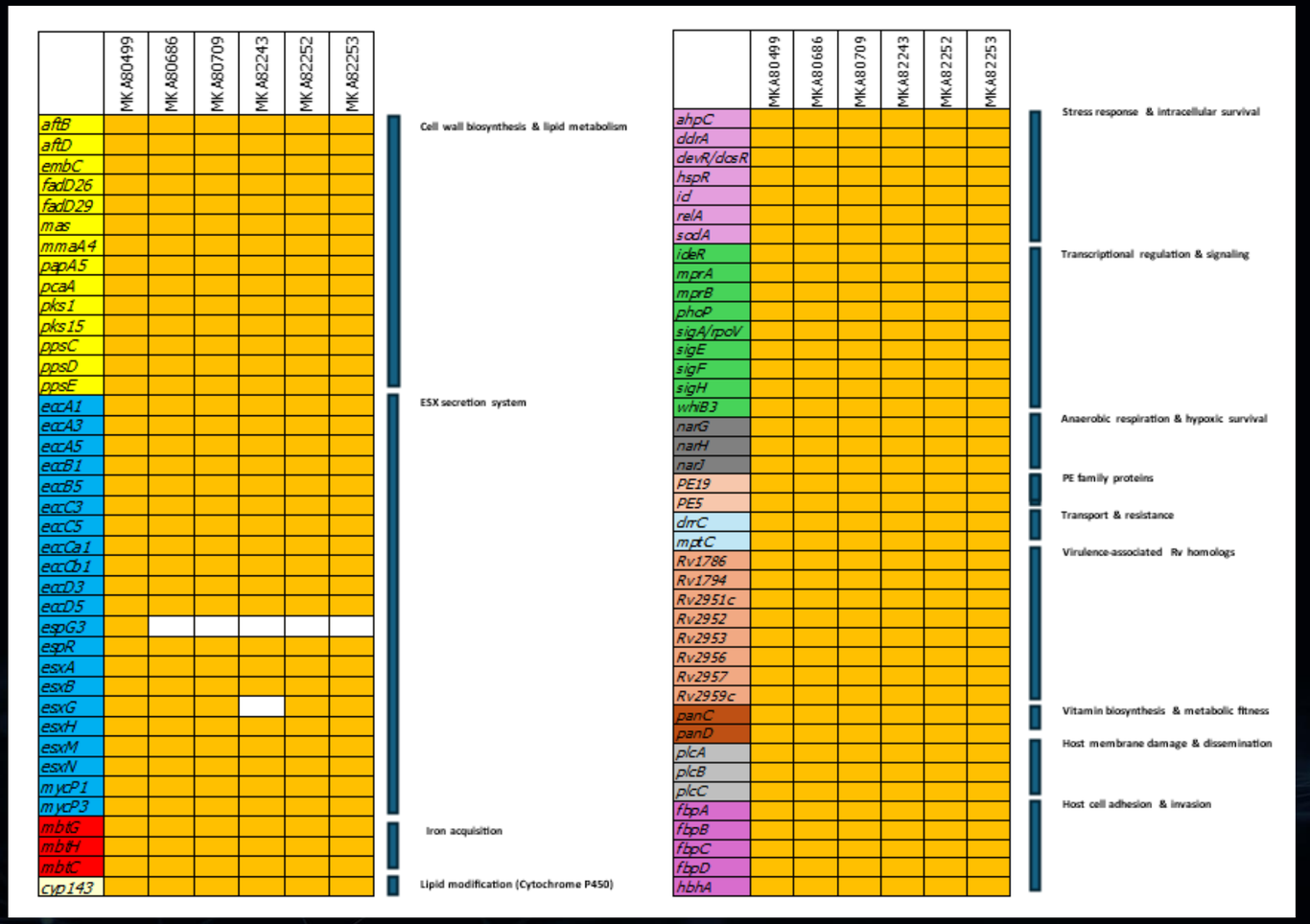
Method

6 x 

BV-BRC → **Genome annotation**

ABRicate → **VFDB** → **VF gene**

Know VF gene



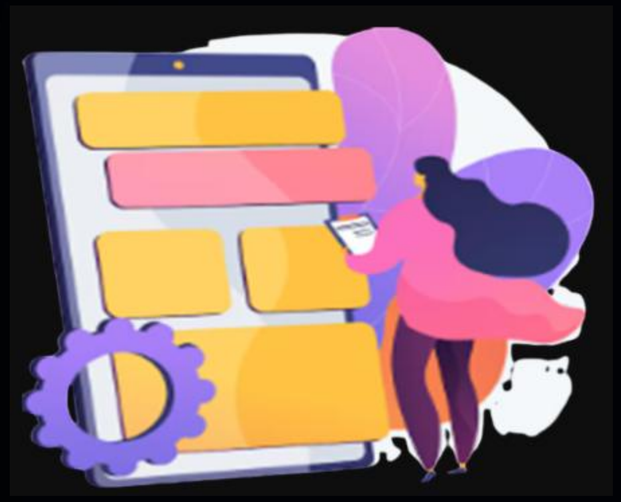
Detect around 80 gene per genome, conserved repertoire of genes associated with cell wall biosynthesis and lipid metabolism), ESX secretion systems), stress response and persistence, and intracellular survival

Conclusion

- All 6 isolates were confirmed as *M. kansasii* using multiple genome-based identification tool.
- Multilocus sequence typing revealed that ST38 was the predominant lineage.
- Functional annotation showed a high representation of metabolism and cellular process-related genes.
- Virulence gene screening demonstrated major pathways involved in cell wall biosynthesis, secretion system and stress response . Only minor variations within the ESX-associated gene cluster.

📦 These findings provide novel genomic insights into the population structure and biological characteristics of *M. kansasii* isolates circulating in Thailand.

Future plan



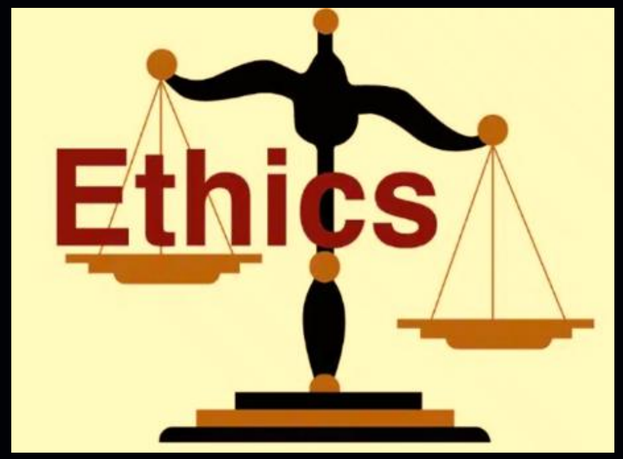
- **Public genome retrieval and species confirmation**

⬇

- **Phylogenetic analysis**

⬇

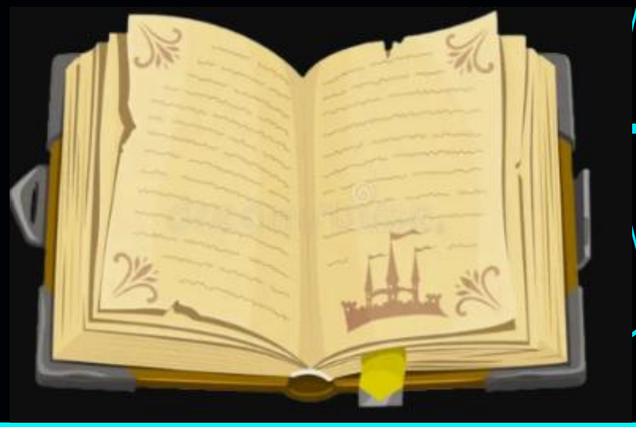
- **Pangenome analysis**



Ethics approval

⬇

EC No.xxxxxxx



- **Manuscript and preparation**

⬇

- **Thesis preparation and examination**

Thesis plan

Activities	2025		2026				2027	
	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2
1.Course work								
2.Literature review and planning								
3.Proposal examination								
4.Data collection of *M.kansasii*								
5.Sequencing quality assessment								
6.Genome assembly and identification of *M. kansasii* species								
7.Genome characterization								
8.Public genome retrieval and species confirmation								
9.Phylogenetic analysis								
10.Pangenome analysis								
11.Manuscript and preparation and submission								
12.Thesis preparation and examination								

= finished work, = ongoing work , = future work

ACKNOWLEDGEMENT

Advisor

Asst. Prof. Dr. Auttawit Sirichoat


Thank You
FOR YOUR KIND ATTENTION