
Characterization of the subspecies of the *Mycobacterium abscessus* complex

Plan de gestion de données créé à l'aide de DMP OPIDoR

Créateur du PGD : Alexandre Godmer

Affiliation du créateur principal : Sorbonne Université

Modèle du PGD : Science Europe - DMP template (english)

Dernière modification du PGD : 16/06/2021

Résumé du projet :

Subspecies of the *Mycobacterium abscessus* complex (MABSC) are opportunistic pathogenic bacteria which are difficult to cure because of natural inducible resistance to clarithromycin (CLA) in certain subspecies. Identification of the 3 subspecies (*M. abscessus*, *M. bolletii* and *M. massiliense*) of MABSC is unsatisfactory by mass spectrometry (MALDI-TOF MS).

The objective of this work is to improve the performance of MALDI-TOF MS for the identification of subspecies of *M. abscessus complex*

Chercheur Principal : Alexandre Godmer

Contact pour les Données : Alexandre Godmer

Droits d'auteur

Le(s) créateur(s) de ce plan accepte(nt) que tout ou partie de texte de ce plan soit réutilisé et personnalisé si nécessaire pour un autre plan. Vous n'avez pas besoin de citer le(s) créateur(s) en tant que source. L'utilisation de toute partie de texte de ce plan n'implique pas que le(s) créateur(s) soutien(nen)t ou aient une quelconque relation avec votre projet ou votre soumission.

Characterization of the subspecies of the *Mycobacterium abscessus* complex

1. Data description and collection or re-use of existing data

- Clinical strains of *Mycobacterium abscessus* complex will be routinely collected prospectively in the bacteriology laboratory of APHP.6 (Sorbonne Université) and associated CNR for Mycobacteria.

- We used a collection of 41 strains of MABSC corresponding to 1001 mass spectra :

- 15 strains of *Mycobacterium abscessus* subsp. *abscessus* (633 mass spectra)
- 9 strains of *Mycobacterium abscessus* subsp. *massiliense* (204 mass spectra)
- 7 strains of *Mycobacterium abscessus* subsp. *bolletii* (164 mass spectra)

- Each strain will be analyzed by MALDI-TOF mass spectrometry according MycoEx protocol (Bruker) and molecular method (DNA/DNA hybridization) for identification.

- All codes for analysis are available at : <https://agodmer.github.io/StageDubii2021/>

Data are :

- **Metadata** : tables containing random name of the strain, molecular identification of the subspecies from *Mycobacterium abscessus* and if it's possible molecular detection of clarithromycin susceptibility.
- **Raw data** : spectral characteristics (mass-to-charge ratio (m/z) and intensities) in a .csv or .flex format for each strain. A minimum of 20 mass spectra could be collected for one strain.
 - Our database contains 41 strains of the MABSC complex provided by the CNR of non tuberculosis mycobacteria (15 strains of *M. abscessus* subsp. *abscessus*; 9 *M. abscessus* subsp. *massiliense* and 7 strains of *M. abscessus* subsp. *bolletii* .
 - A total of 1001 spectra were collected with 633 spectra of the *abscessus* subspecies, 164 spectra of *bolletii* and 204 spectra of *massiliense*.

The total raw data and the metadata files will have a total size less than 2GB.

2. Documentation and data quality

Patient information will not be collected.

Only the molecular identification of the strains will be collected and the precise identification of the strain as well as its phenotypes of resistance to antibiotics.

The molecular identification of the strains (DNA / DNA hybridization) corresponding to subspecies identification will be listed (gold standard).

3. Storage and backup during the research process

The data will be saved on a dedicated research hospital computer at Centre d'Immunologie et des maladies infectieuses (CIMI) ; INSERM U1135.

A back-up is performed on external hard drive every month.

Sensitive patient data will not be collected.

Each strain will be assigned a random number.

4. Legal and ethical requirements, codes of conduct

No personal data are processed.

- The database has been the subject of an invention declaration.
- The Sorbonne University is the owner of the database.
- These questions are managed by Satt lutech (<https://www.sattlutech.com/>). This organisation linked to Sorbonne University
- Please contact : *medecine- drv@sorbonne-universite.fr* for data access.

- These questions are managed by Satt lutech (<https://www.sattlutech.com/>). This organisation linked to Sorbonne University
- Please contact : *medecine- drv@sorbonne-universite.fr* for access.

5. Data sharing and long-term preservation

- The data will be available after agreement with the Satt lutech.
- Please contact : *medecine- drv@sorbonne-universite.fr* for data access (reference : X21 – AUBRY/GODMER (MABSC)). This organisation linked to Sorbonne University

The data will be saved on a dedicated research hospital computer at Centre d'Immunologie et des maladies infectieuses (CIMI) ; INSERM U1135.

- Access to data is managed by Satt lutech (<https://www.sattlutech.com/>). All data could be available after prior agreement with this structure. This organisation linked to Sorbonne University.
- All codes for analysis are available at : <https://agodmer.github.io/StageDubii2021/>

- A deposit number is being obtained.

6. Data management responsibilities and resources

- These questions are managed by Satt lutech (<https://www.sattlutech.com/>). This organisation linked to Sorbonne University.

- Access to data is managed by Satt lutech (<https://www.sattlutech.com/>). All data could be available after prior agreement with this structure. This organisation linked to Sorbonne University
- All codes for analysis are available at : <https://agodmer.github.io/StageDubii2021/>