
DMP du projet "Spectral database of the subspecies of the *Mycobacterium abscessus* complex (MALDI-TOF MS)"

Plan de gestion de données créé à l'aide de DMP OPIDoR, basé sur le modèle "Science Europe - DMP template (english)" fourni par Science Europe.

Plan Details

Plan title	DMP du projet "Spectral database of the subspecies of the <i>Mycobacterium abscessus</i> complex (MALDI-TOF MS)"
Language	fra
Creation date	2021-03-27
Last modification date	2021-06-16

Project Details

Project title Spectral database of the subspecies of the *Mycobacterium abscessus* complex (MALDI-TOF MS)

Abstract 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

Start date 2020-01-01

End date 2021-12-31

Research outputs :

1. Mass spectra of *Mycobacterium abscessus* complex (Mass spectra of *Mycobacterium abscessus* complex)

Contributors

Name	Affiliation	Roles
Alexandre Godmer		<ul style="list-style-type: none">• Coordinateur du projet• Personne contact pour les données• Responsable du plan de gestion de données

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.

DMP du projet "Spectral database of the subspecies of the *Mycobacterium abscessus* complex (MALDI-TOF MS)"

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

1a. How will new data be collected or produced and/or how will existing data be re-used?

- 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 and will be analyzed by MALDI-TOF mass spectrometry according MycoEx protocol (Bruker) and molecular method (DNA/DNA hybridization, using GenoType NTM-DR (Hain Lifescience, Nehren, Germany) according to the manufacturer's instructions) for identification. The spectra were obtained according to the following steps :
 - Each of the 41 strains was cultured in aerobic atmosphere at 37°C for 7±2 days on blood agar (COH, bioMérieux®). Then, one colony was extracted according to the MycoEx protocol (Bruker®). For each of the extracts, 8 technical replicates were realized and analyzed by MALDI-TOF MS (Bruker®). Dried spots were overlaid with 1µL of MALDI matrix (α-HCCA).
 - Data acquisition was performed using a Microflex LT (Bruker® Daltonics) mass spectrometer equipped with a N2 laser (λ =377 nm). Instrument parameters used were as follows: a mass range between 200-20000 Da, ion source 1: 20 kV, ion source 2: 18.5 kV, lens: 8.45 kV, pulsed ion extraction: 330 ns, laser frequency: 20.0 Hz. Spectra were obtained after 500 shots. Each spot was analyzed three times. In total 24 spectra were obtained for each extraction.
 - Spectra acquired for each isolate were visualized and analyzed using Flex Analysis software (Bruker® Daltonics), and spectra with low quality peaks were removed. A minimum of 20 spectra per extraction was necessary to validate the extraction.

1b. What data (for example the kind, formats, and volumes), will be collected or produced?

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

2a. What metadata and documentation (for example the methodology of data collection and way of organising data) will accompany the data?

- **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.
-

2b. What data quality control measures will be used?

- 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

3a. How will data and metadata be stored and backed up during the research?

- The data will be saved on a dedicated research hospital computer at Centre d'Immunologie et des maladies infectieuses (CIMI Paris) ; INSERM U1135.
 - A back-up is performed on external hard drive every month.
-

3b. How will data security and protection of sensitive data be taken care during the research?

- Sensitive patient data will not be collected.
 - Each strain will be assigned a random number.
-

4. Legal and ethical requirements, codes of conduct

4a. If personal data are processed, how will compliance with legislation on personal data and on security be ensured?

No personal data are processed.

4b. How will other legal issues, such as intellectual property rights and ownership, be managed? What legislation is applicable?

- 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.**
-

4c. What ethical issues and codes of conduct are there, and how will they be taken into account?

- 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

5a. How and when will data be shared? Are there possible restrictions to data sharing or embargo reasons?

- 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
-

5b. How will data for preservation be selected, and where data will be preserved long-term (for example a data repository or archive)?

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

5c. What methods or software tools are needed to access and use data?

- 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.
-

5d. How will the application of a unique and persistent identifier (such as a Digital Object Identifier (DOI)) to each data set be ensured?

- A deposit number is being obtained.

6. Data management responsibilities and resources

6a. Who (for example role, position, and institution) will be responsible for data management (i.e. the data steward)?

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

6b. What resources (for example financial and time) will be dedicated to data management and ensuring that data will be FAIR (Findable, Accessible, Interoperable, Re-usable)?

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