



**RESEAU NATIONAL DE SURVEILLANCE  
DES *YERSINIA* ENTEROPATHOGENES**

***UNE NOUVELLE METHODE  
D'IDENTIFICATION/CARACTERISATION  
DES SOUCHES DE YERSINIA***

Fascicule N° 22 – mai 2018



Institut Pasteur

## L'ancienne méthode phénotypique

- Utilisée jusqu'à présent, la caractérisation phénotypique des souches de *Yersinia* repose sur divers tests :
  - ✓ la mise en évidence de caractères biochimiques en utilisant des galeries API20E et 50CH.
  - ✓ la mise en évidence d'activités enzymatiques : tween-estérase et pyrazinamidase.
  - ✓ le sérotypage par séro-agglutination grâce à une batterie d'antisera.
- Cette caractérisation complète permet de confirmer le genre *Yersinia* des souches envoyées au CNR, d'identifier l'espèce et la sous-espèce (les différents biotypes de *Y. enterocolitica*) et leur caractère pathogène.

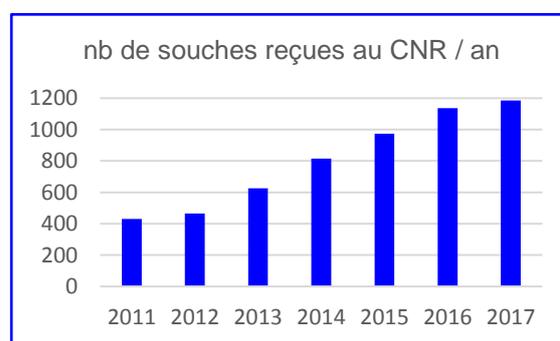
## Pourquoi une nouvelle méthode ?

- L'évolution des techniques de séquençage de l'ADN permet de séquencer rapidement et facilement le génome complet de souches bactériennes.
- A partir des données de séquençage il est possible de développer une méthode *in silico* permettant de caractériser les souches.
- Cette méthode d'analyse du génome permettra :
  - ✓ D'identifier plus précisément les espèces
  - ✓ De définir plus finement les sous-espèces

✓ D'effectuer une assignation taxonomique fiable y compris pour les souches présentant des atypies phénotypiques.

- L'utilisation d'une méthode basée sur le séquençage génomique des souches nécessitera moins de manipulations au laboratoire.

Cela permettra aux techniciens du CNR de faire face à l'explosion du nombre de souches reçues depuis l'année 2010.



- La caractérisation phénotypique complète d'une souche prend environ 1 semaine. L'utilisation d'une méthode basée sur le séquençage génomique des souches devrait permettre d'obtenir un résultat plus rapide.

## Principe de la méthode cgMLST

### • Extraction d'ADN :

La première étape consiste à mettre en culture la bactérie puis d'extraire l'ADN génomique avec un kit commercial d'extraction/purification d'ADN.

### • Séquençage de l'ADN génomique :

✓ L'ADN génomique est préparé pendant l'étape de construction de librairie pour pouvoir être séquencé.

✓ La librairie de chaque souche est déposée sur un séquenceur haut-débit et est ensuite séquencée.

✓ A l'Institut Pasteur, nous avons accès à la plateforme de microbiologie mutualisé (P2M) pour le séquençage du génome des souches reçues dans un cadre de santé publique.

✓ Le séquenceur utilisé sur la plateforme est un NextSeq500 (Illumina) permettant d'obtenir des coûts compétitifs en raison de la chimie utilisée et du multiplexage des ADN (le génome de plusieurs souches est séquencé en même temps).

✓ Les données de séquençage sont récupérées sous forme de données informatiques contenant la séquence génomique de chaque souche.

- Analyse bioinformatique :

✓ A partir des données de séquençage génomique, nous avons choisi de développer une méthode de core genome multi locus sequence typing (cgMLST) pour l'identification et la caractérisation des souches de *Yersinia*.

✓ La cgMLST est définie par une sélection de gènes présents dans toutes les souches de *Yersinia*.

✓ La comparaison de la séquence des gènes d'une souche à une base de données de référence permet de lui assigner un profil allélique.

✓ La comparaison du profil allélique à une base de données de référence permet d'effectuer l'assignation taxonomique de la souche (espèce et éventuellement sous-espèce).

## Développement de la cgMLST

- Constitution de la base de données de référence :

✓ Pour créer cette base, nous avons :

- Téléchargé les données de séquençage génomiques disponibles publiquement pour les *Yersinia*. Cela a représenté près de 547 génomes.

- Effectué le séquençage génomique de 743 souches de toutes les espèces et sous-espèces de *Yersinia* décrites.

✓ Le choix des souches a été effectué afin d'avoir le maximum de diversité possible et un nombre minimum de 20 souches par espèce quand cela a été possible.

→ Au final, nous avons constitué une base de données de référence de 1280 génomes.

Espèce	Données publiques	Séquençage par le CNR
<i>Y. aldovae</i>	6	15
<i>Y. aleksiciae</i>	2	11
<i>Y. bercovieri</i>	3	16
<i>Y. enterocolitica</i>	124	154
<i>Y. entomophaga</i>	1	1
<i>Y. frederiksenii</i>	23	4
<i>Y. intermedia</i>	17	4
<i>Y. kristensenii</i>	13	7
<i>Y. massiliensis</i>	2	5
<i>Y. mollaretii</i>	9	9
<i>Y. nurmii</i>	1	0
<i>Y. pekkanenii</i>	2	0
<i>Y. pestis</i>	280	20
<i>Y. pseudotuberculosis</i>	44	445

<i>Y. rohdei</i>	6	9
<i>Y. ruckeri</i>	9	11
<i>Y. similis</i>	3	17
<i>Y. wautersii</i>	2	5
Total	547	733

- **Définition du schéma cgMLST :**

- ✓ La comparaison des 1280 génomes de la base de données de référence a permis d'identifier 500 gènes communs à toutes les souches de toutes les espèces de *Yersinia*.

- ✓ Un schéma cgMLST avec ces 500 gènes a été créé dans une base de données sur le serveur de l'Institut Pasteur.

- ✓ Cette base de données utilise le programme BIGSdb permettant de stocker et d'analyser les données de séquences pour des souches bactériennes.

- ✓ A chaque souche enregistrée dans la base de données, son génome lui est associé puis le programme BIGSdb effectue un scan à la recherche des 500 gènes du schéma cgMLST :

- Si une variante du gène existe déjà dans la base de données, un numéro d'allèle identique est attribué à ce gène.

- Si la variante du gène trouvé n'existe pas, le programme ajoute cette variante dans la base de données et lui attribue un nouveau numéro d'allèle.

➔ Ainsi, un profil allélique est attribué à chaque souche.

souche	Gène 1	Gène 2	Gène 3	Gène X
X	1	1	4	3
Y	1	4	12	3
Z	2	2	4	6

- **Programme d'assignation taxonomique :**

- ✓ Un bioinformaticien de l'Institut Pasteur a développé un programme permettant d'effectuer une assignation taxonomique (attribution d'espèce et/ou sous-espèce).

- ✓ Ce programme compare le profil allélique d'une souche au plus proche profil de la base de données de référence et lui attribue la même assignation taxonomique si la distance entre ces profils est inférieure à un seuil défini.

### Structure de la population des *Yersinia*

- Une étude phylogénétique menée sur l'ensemble des souches de la base de données de référence a permis :

- ✓ De confirmer l'existence des 18 espèces distinctes actuellement décrites.

- ✓ D'identifier au moins 5 nouvelles espèces de *Yersinia*. Elles semblent toutes a priori non pathogènes, en attendant leur description.

- ✓ De montrer que les espèces *Y. frederiksenii* et *Y. kristensenii* sont composées de divers génotypes.

- ✓ De définir 13 génotypes au sein de l'espèce *Y. enterocolitica* correspondant à des biotypes ou biosérotypes de l'ancienne caractérisation phénotypique (cf corrélation des méthodes en page 6).

- ✓ D'identifier 31 génotypes différents pour l'espèce *Y. pseudotuberculosis* n'ayant aucun lien avec l'ancienne caractérisation basée sur le sérotype.

## Validation de la cgMLST

- Nous avons ensuite validé la cgMLST *Yersinia* avec l'ensemble des souches reçues au CNR de janvier 2016 à juillet 2017.
- Pendant cette période, le séquençage du génome des souches a été effectué en parallèle de la caractérisation phénotypique.
- Nous avons comparé les résultats des 2 méthodes sur 1843 souches :
  - ✓ Pour 1834 souches, il y a eu une parfaite corrélation entre les caractérisations phénotypique et génotypique.
  - ✓ Il y a eu discordance pour seulement 9 souches :

### - souches non pathogènes :

4 souches phénotypiquement identifiées comme des *Y. enterocolitica* 1A/O:5 qui sont en fait des *Y. frederiksenii*.

3 souches phénotypiquement identifiées comme des *Y. frederiksenii* qui sont en fait des *Y. massiliensis*.

### - souches pathogènes :

1 souche phénotypiquement identifiée comme *Y. enterocolitica* 3/O:3 qui est en fait du génotype 4 (ou biosérotype 4/O:3).

1 souche phénotypiquement identifiée comme *Y. enterocolitica* 4/O:3 qui est en fait du génotype 3-3b (ou biosérotype 3/O:3).

→ Il n'y a pas eu de discordance sur le caractère pathogène des souches.

- Les performances de la cgMLST *Yersinia* sont :

- ✓ Répétabilité : 100%
- ✓ Reproductibilité : 100%
- ✓ Sensibilité : 100%
- ✓ Spécificité (testée sur d'autres entérobactéries) : 100%

→ La validation nous a permis d'utiliser la cgMLST en routine dès le mois de décembre 2017.

## Conclusion

- La technique cgMLST nécessite moins de manipulations au laboratoire ce qui permet de faire face à l'explosion du nombre souches reçues au CNR.
- L'identification des souches (espèce et/ou sous-espèce) est plus fiable qu'avec la caractérisation phénotypique.
- L'utilisation de la cgMLST élimine les problèmes d'atypie rencontrés avec la caractérisation phénotypique.
- La cgMLST permet d'identifier l'émergence de nouvelles espèces.
- Actuellement le résultat est obtenu en 7 jours pour la cgMLST comme pour l'ancienne caractérisation phénotypique. La prévision d'un plus grand nombre de run de séquençage à l'avenir sur la plateforme P2M permettra de raccourcir ce délai.
- Le séquençage génomique permet aussi de développer d'autres méthodes d'analyses *in silico* permettant d'étudier en la proximité génétique des souches (investigation de cas groupés).

## Corrélation entre les caractérisations génotypique et phénotypique

- L'assignation taxonomique obtenue avec la cgMLST permet de confirmer le genre *Yersinia*, d'attribuer une espèce et éventuellement un génotype (sous-espèce).
  - Il existe une correspondance entre les 2 méthodes de caractérisation :
- ✓ Souches pathogènes

Caractérisation génotypique (cgMLST)		Caractérisation phénotypique	
Espèce	Génotype	Espèce	Biosérotype
<i>Y. enterocolitica</i>	1B	<i>Y. enterocolitica</i>	1B
<i>Y. enterocolitica</i>	2/3-9a	<i>Y. enterocolitica</i>	2/O:9
<i>Y. enterocolitica</i>	2/3-9b	<i>Y. enterocolitica</i>	
<i>Y. enterocolitica</i>	2/3-5a	<i>Y. enterocolitica</i>	2/O:5,27 ou 3/O:5,27
<i>Y. enterocolitica</i>	2/3-5b	<i>Y. enterocolitica</i>	
<i>Y. enterocolitica</i>	3-3a	<i>Y. enterocolitica</i>	3/O:3
<i>Y. enterocolitica</i>	3-3b	<i>Y. enterocolitica</i>	
<i>Y. enterocolitica</i>	3-3c	<i>Y. enterocolitica</i>	
<i>Y. enterocolitica</i>	3-3d	<i>Y. enterocolitica</i>	
<i>Y. enterocolitica</i>	4	<i>Y. enterocolitica</i>	4/O:3
<i>Y. enterocolitica</i>	5	<i>Y. enterocolitica</i>	5
<i>Y. pseudotuberculosis</i>	1 à 32	<i>Y. pseudotuberculosis</i>	I à V

- ✓ Souches non pathogènes

Caractérisation génotypique (cgMLST)		Caractérisation phénotypique	
Espèce	Génotype	Espèce	Biosérotype
<i>Y. enterocolitica</i>	1Aa	<i>Y. enterocolitica</i>	1A
<i>Y. enterocolitica</i>	1Ab		
<i>Y. frederiksenii</i>	1 à 3	<i>Y. frederiksenii</i>	Divers sérotypes
<i>Y. kristensenii</i>	1 à 2	<i>Y. kristensenii</i>	Divers sérotypes

- Les 14 autres espèces du genre *Yersinia* sont répondues sans génotype.



## EVOLUTION DU CNR

Suite à l'appel d'offre de Santé Publique France pour la nomination des CNR pour le mandat 2017 – 2021, le CNR de la peste et autres yersiniose a été reconduit pour cette période. Le Pr Elisabeth Carniel qui dirigeait le CNR depuis 1990 est partie prendre la direction du Centre Pasteur du Cameroun. C'est à présent le Dr Anne-Sophie Le Guern qui dirige le CNR.

## RETOUR D'EXPERIENCE SUR LES PCR DE DIAGNOSTIC

Plusieurs laboratoires membres du Réseau National de Surveillance des *Yersinia* entéropathogènes nous ont informé utiliser des PCR multiplex (comme le système BD MAX™) pour la détection des entérobactéries avant d'effectuer la culture.

Nous aimerions avoir un retour des laboratoires utilisant le screening par PCR pour connaître les réactifs utilisés, les cibles détectant les *Yersinia* et la corrélation que vous observez entre PCR et culture.

Nous attirons votre attention sur le fait que la cible utilisée par le BD MAX™ pour identifier des *Yersinia* ne permet pas de détecter les *Yersinia pseudotuberculosis*. La PCR de screening est utile pour orienter le diagnostic mais ne permet pas de s'affranchir de l'étape de culture.

---

***DANS LE PROCHAIN FASCICULE,  
« L'INVESTIGATION EPIDEMIOLOGIQUE DE CAS GROUPES AU CNR DE  
LA PESTE ET AUTRES YERSINIOSES »***

---



**CENTRE NATIONAL DE REFERENCE DE LA PESTE ET AUTRES YERSINIOSES**

**INSTITUT PASTEUR**

**UNITE DES *YERSINIA***

**28, RUE DU DOCTEUR ROUX**

**75724 PARIS CEDEX 15 (France)**

**☎ 01 40 61 37 67 📠 01 45 68 89 54**

**Site web : <http://www.pasteur.fr>**

**CONTACT: Cyril SAVIN. E-mail : [cyril.savin@pasteur.fr](mailto:cyril.savin@pasteur.fr)**