



Principes, intérêts et mise en place pratique en routine du séquençage de nouvelle génération au CHU Henri Mondor (Créteil)

Dr Jean-Winoc DECOUSSER ^{1,2}



1. Equipe Opérationnelle d'Hygiène - Laboratoire de Bactériologie et d'Hygiène, CHU Henri Mondor, Assistance Publique - Hôpitaux de Paris, Créteil, France
2. EA 7380 Dynamyc Université Paris-Est Créteil (UPEC), Ecole nationale vétérinaire d'Alfort (EnvA), Faculté de Médecine de Créteil, 8 rue du Général Sarrail, 94010 Créteil, FRANCE.



DÉCLARATION DE LIENS D'INTÉRÊTS

Conférencier: Jean-Winoc, DECOUSSER, CRETEIL

Je n'ai pas de lien d'intérêt potentiel à déclarer

Objectifs: exemple de mise en place en routine de la comparaison des souches bactériennes par séquençage de nouvelle génération

Plan:

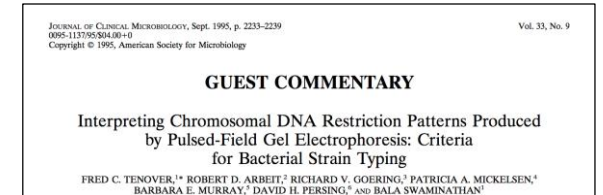
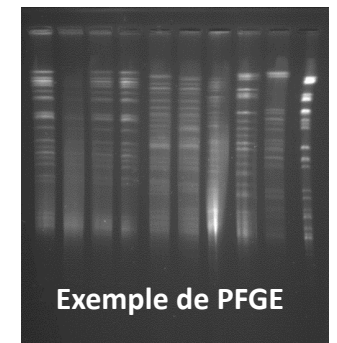
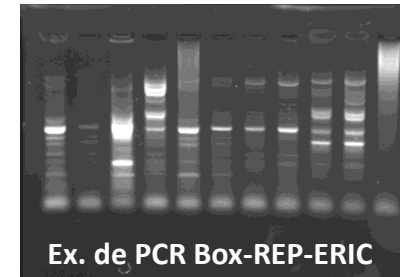
- **Contexte**
- **Situation antérieure**
- **Situation actuelle**
- **Introduction du séquençage de nouvelle génération (1) – principe et théorie**
- **Introduction du séquençage de nouvelle génération (2) – challenges et réalité**
- **Organisation locale**
 - pré-analytique
 - Analytique
 - Post analytiques
- **Exemples**
- **Discussion et perspectives**

Contexte

- **Nécessité de comparer des souches afin:**
 - D'objectiver un phénomène épidémique
 - D'identifier **une source / un réservoir** bactérien
 - D'évaluer **l'efficacité des mesures de prévention**
- **GH Henri MONDOR:** Politique de dépistage des BHRe « ambitieuse »
 - Antécédents d'hospitalisation **ou de séjours > 30j** à l'étranger dans les **2 années** précédentes
- **Comparaison de souches: conséquences sur la prise en charge du phénomène épidémique éventuel**
 - « Les précautions standard et les mesures éventuellement mises en place (PCC) n'ont pas contrôlé la transmission croisée : **au moins un cas secondaire a été identifié** au cours des dépistages réalisés chez les contacts d'un patient porteur. »
 - 1« Activer le plan local de maîtrise d'une épidémie.
 2. **Arrêter les transferts des porteurs et des contacts** encore hospitalisés.
 3. **Arrêter temporairement les admissions** le temps d'organiser 3 secteurs distincts.



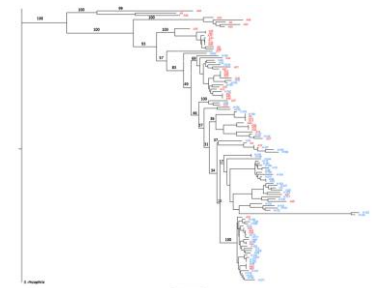
Comparaison de 2 souches: Situation antérieure



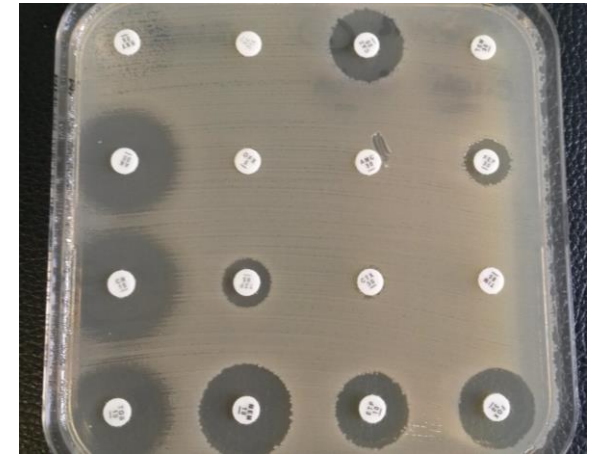
- **Antibiotype:** comparaison des profils de résistance
 - Pas de coût supplémentaire, mais performances discutables
- **Techniques par PCR (RAPD, ERIC-PCR): *fingerprinting***
 - Rapides, peu onéreuses mais peu reproductibles (à refaire à chaque nouvelle souche)
 - Interprétation visuelle
- **Gold Standard: analyse des profils de macro-restriction en champ pulsé (PFGE)**
 - Coût (matériel, réactif X 2 enzymes, main d'oeuvre)
 - Critères d'interprétation visuels « historiques » ou nécessitant un logiciel dédié
 - Délai: 5 j et une révélation finale sans contrôle intermédiaire.
- Typage par **séquençage d'un ou plusieurs gènes** et comparaison
(*Spa typing*, **Multilocus Sequence Typing / MLST – Combinaison d'allèles**

Maiden 1998)

- Données reproductibles, non ambiguës, transférables mais représente 0.2% du génome
- Plus cher, plus long
- Plutôt de la phylogénie que de la comparaison de souche (fonction de l'épidémiologie bactérienne)



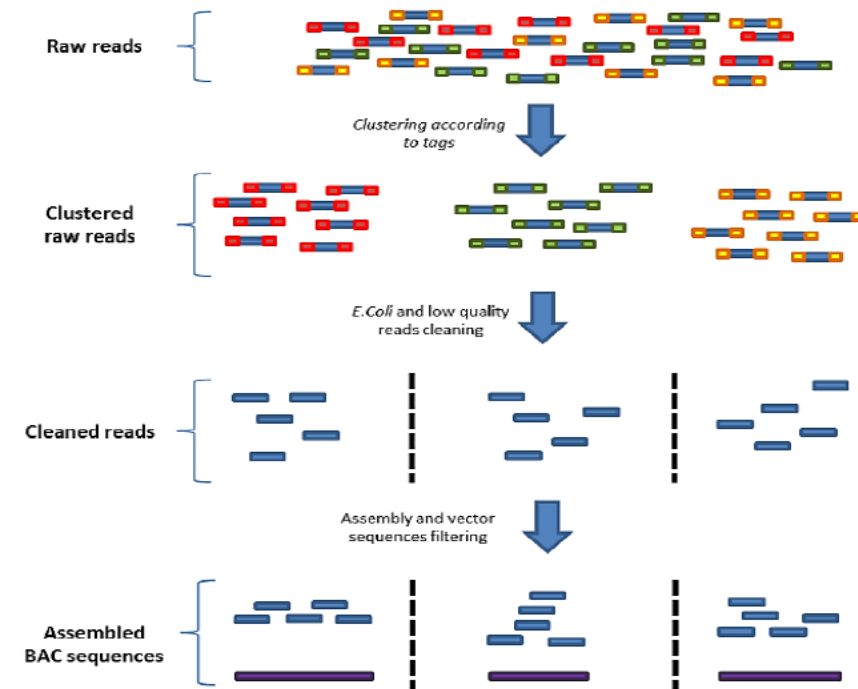
Situation antérieure



- **Problème supplémentaire : épidémie de plasmides**
 - Cas (notamment) des Entérobactéries Productrices de Carbapénémase (EPC) de la famille des **OXA-48**
 - Support génétique sur des plasmides très mobiles au sein des différentes entérobactéries du tube digestif d'un même patient.
 - Transmission inter- espèces masquant une transmission croisée.
 - **Technique de référence:**
 - Extraction des plasmides et électroporation dans une souche de référence
 - Conjugaison dans souche de référence puis criblage sur milieu sélectif
 - Comparaison des tailles de plasmides + marquage par sonde froide du gène de résistance recherché.....

Introduction du séquençage de nouvelle génération (2) – challenges et réalité

- **Qualité des résultats bruts à valider**
 - Validation biologique
 - Expertise de l'hygiéniste qui doit avoir un œil critique
- **Nettoyage** des millions de séquences obtenues (« Trimming »)
 - biais
- **Assemblage** selon différentes approches / logiciels sur *références* ou *de novo*
 - biais
- **Interprétation des résultats**
 - Organisation des Génomes
 - Plasmides?
- **Comparaison des génomes**
 - Base/base: recherche des Single Nucleotide Polymorphism
 - Autres approches (*core genome* MLST)
 - Critères d'interprétation?



NGS à l'hôpital Henri-Mondor

- **Plateforme** de Séquençage de Nouvelle Génération sur site gérée par un microbiologiste d'un département intégré de **»Prévention, diagnostic et traitement des infections«** (Dr Christophe RODRIGUEZ - Virologue)
- **Ingénieur Plateforme** (Mélanie DARTY)
- Formation ESCMID Université de Groningen
- Familiarisation avec les génomes complets dans le cadre d'activités de recherche
- AHU- thésard immergé dans la **bioinformatique** au sein du LabGeM (Dr Guilhem ROYER)



 ESCMID EUROPEAN SOCIETY OF CLINICAL MICROBIOLOGY AND INFECTIOUS DISEASES

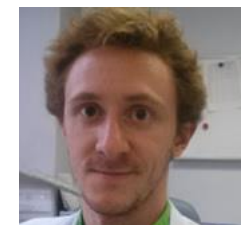
ESCMID Postgraduate
Technical Workshop

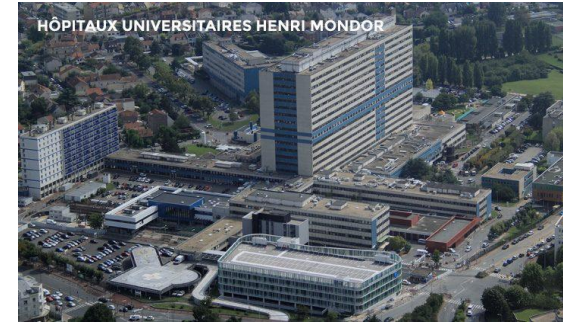
**Capacity-building
Workshop: Rapid NGS for
Characterization and
Typing of Resistant
Gram-Negative Bacilli**

**Groningen, Netherlands
7 – 9 October 2015**



Infection • Antimicrobials • Modelling • Evolution





Organisation locale - Pré analytique

- Pré analytique: **validation des demandes** au sein de l'EOH
 - Investigation EOH
 - Suspicion de transmission croisée
 - Conséquence sur la prise en charge des patients / implication du matériel
 - Nécessité d'avoir des résultats **rapidement** pour prendre des décisions vs envoi aux CNR
 - Situation apparemment maîtrisée mais un « nouveau » cas apparaît: efficacité des mesures de prévention mises en place?
 - **Réflexions / stratégies** sur les actions en fonction des résultats
- Et non pas « pour voir »...
- **Choix de la profondeur de séquençage / nombre de souches à passer / taille du génome / degré d'urgence:**
 - Type d'automates MiSeq, NextSeq[®] (Illumina)
 - Technologie (150 ou 300 *paired end*)
 - 150 bp / Miseq = 9 à 10 échantillons
 - 300 bp / Miseq : 24 E. coli

Organisation locale - analytique

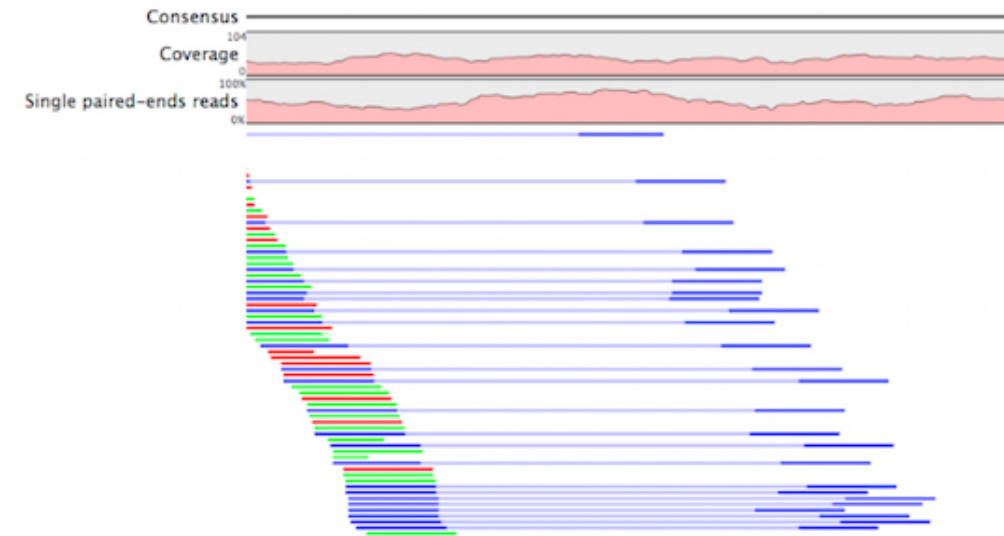
- Préparation des extraits d'ADN (3 automates QiaSymphony®)
 - Vérification de la qualité des ADN
 - Préparation des banques
 - Vérification de la qualité des banques
 - Séquençage
- *contrôles intermédiaires et finaux*





Organisation locale – post analytique

- Les problèmes commencent....
- **Contrôles de la qualité du séquençage +++**
 - 1. Qualité globale du « run »**
 - Nombre de clusters obtenus + qualité de toutes les bases (MiSeq 2 X 300 bp, objectif 44 000 000 reads)
 - 2. Qualité des « reads »:**
 - « Q score > 30 » supérieur à 70% sur une fenêtre de 20 nucléotides
 - 3. Qualité de l'assemblage**
 - Nombre de « contigs » (entre 200 et 500, Maximum 1000)
 - Profondeur de **100 X** théorique en moyenne, minimum **30X**, soit 1 million de « reads » pour une souche de *E. coli* qui fait 5 mégabases
- Souche de référence séquencée en Sanger (*E. coli* S88) séquencée à chaque fois en NGS:
 - bilan annuel < 5 à 6 SNPs





Organisation locale – post analytique

- Quelle méthode de comparaison des souches?
- Principe: *core genome MLST (cgMLST)*
 - Repose sur un ensemble de gènes conservés au sein d'une espèce ou d'espèces proches
- Utilisation d'un logiciel commercial: SeqSphere® (RidomSeqsphere)
 - Licence payante (<http://www.ridom.de/seqsphere/>)
 - Approche standardisée
 - Mise à jour
 - Intervention si bug
 - Publications
 - **Performances** (reproductibilité): 20 *S. aureus* / 5 laboratoires : 3 erreurs d'attribution d'allèle sur 183 927, soit 0.0016% (Mellmann et al. J Clin Microbiol 2017)
 - **Critères d'interprétation**

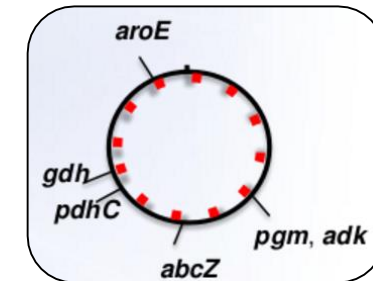
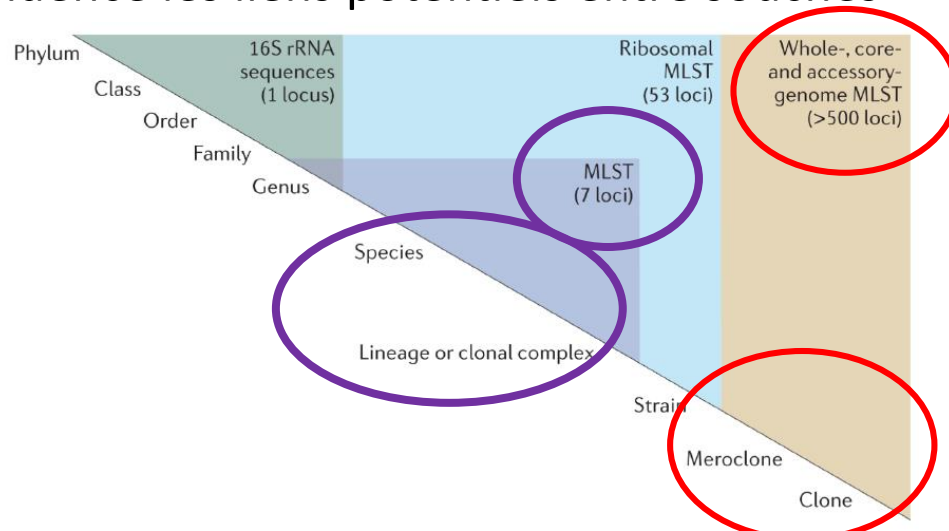
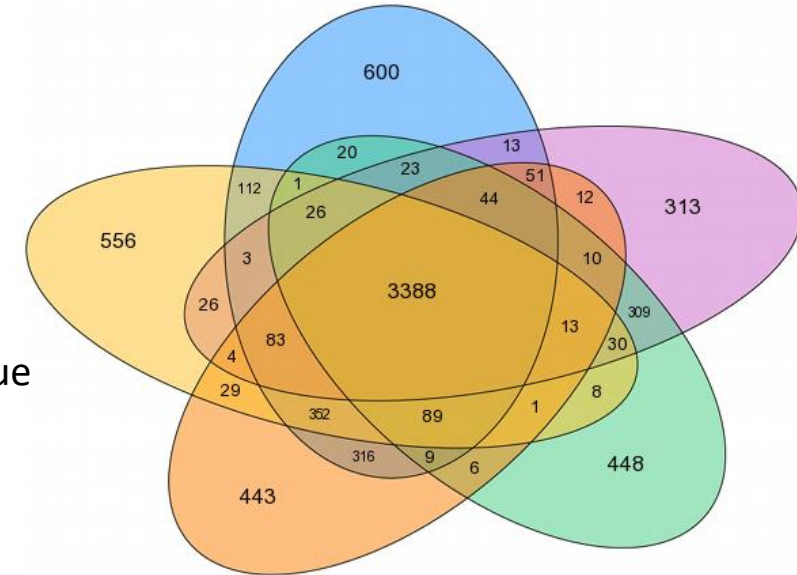




Organisation locale – post analytique

- **cg MLST = super MLST discriminant +++**

- Les gènes communs = le **core genome** (cf. diagramme de Venn)
- Gènes codant uniquement
- Attribution d'allèle à chaque gène du cgMLST:
 - Nombre d'allèles identiques vs nombre d'allèles différents distance allélique
- Matrice de distance et *spanning-tree* pour mettre en évidence les liens potentiels entre souches



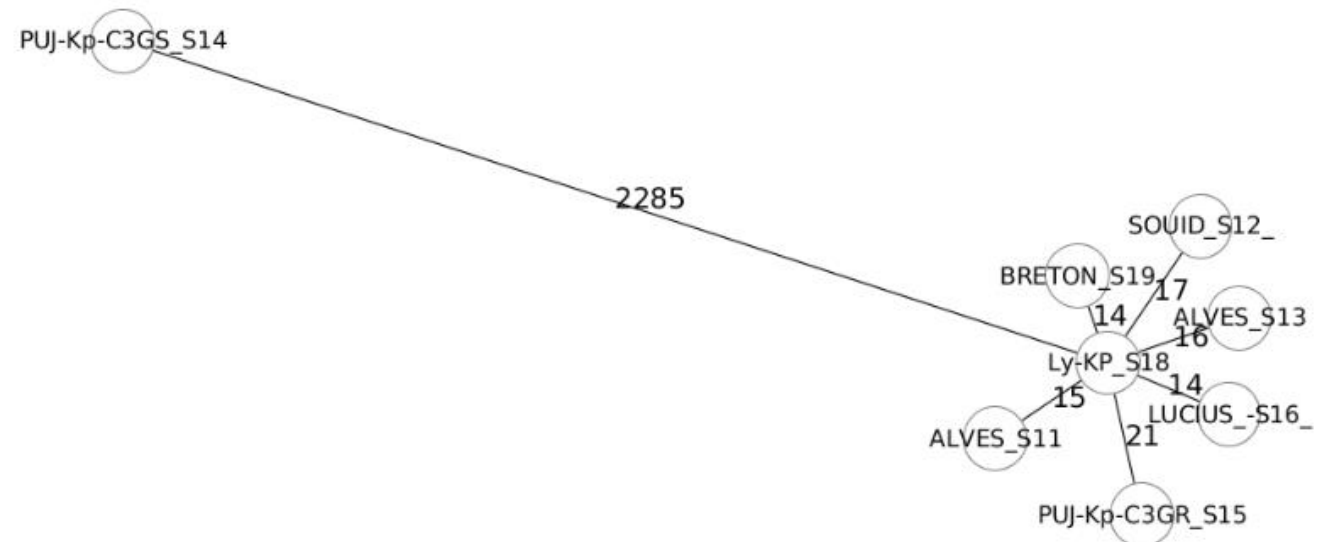
Critères d'interprétation

- **Quantitatifs et publiés**
- Affinés avec le temps et les publications
- **Entérocoques** (ERV *E. faecium*, De Been et al. J Clin Microbiol 2015)
 - 1423 gènes cibles (sur 2860)
 - Lien épidémiologique avéré si 0 -20 allèles de différence
 - Lien épidémiologique possible si 21-40 allèles
 - Pas de lien si > 40 allèles
- *Listeria monocytogenes* (*Ruppitsch et al* , J Clin Microbiol 2015, *Halbedel et al. J Clin Microbiol 2018*)
 - 1701 gènes cibles, seuil fixé à 10 allèles pour différencier 2 souches de clusters différents
- SARM: >6, *Pseudomonas aeruginosa*: > 14, *E. coli*: > 10 (Mellman et al., J Clin Microbiol, 2016) , *Legionella pneumophila*: > 4 (Motan-Gilad et al., Euro Surveill 2015)

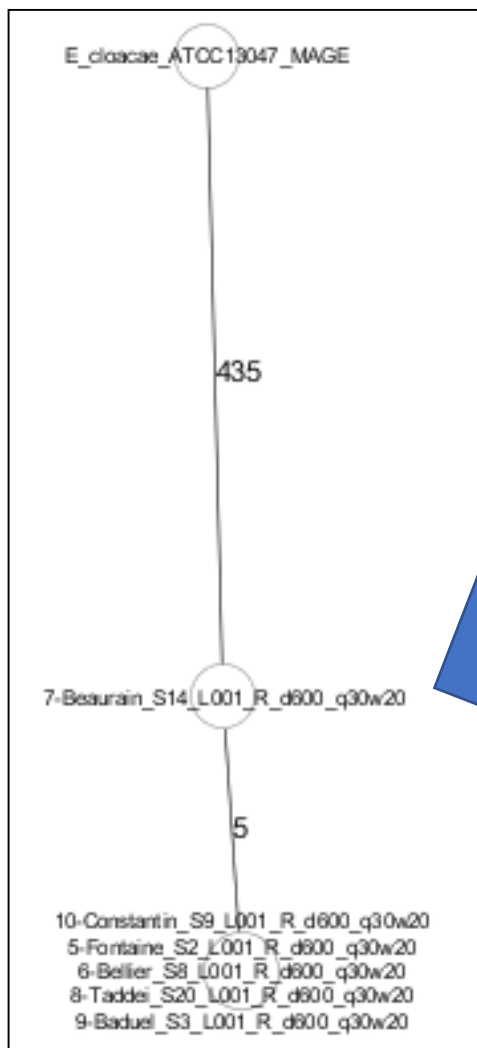
Exemple: transmission d'EPC lors d'actes d'endoscopie invasive (Cholangio-Pancréatographie Rétrograde Endoscopique -CPRE)

- 07/09/2017 Découverte fortuite d'une BHRé (bactériémie *K. pneumoniae* OX-48) chez un patient sans facteur de risque
- CPRE le 30/08/2017 – endoscope 22
- Le 29/08/2017 CPRE avec le même endoscope n° 22 chez un patient connu *K. pneumoniae* OXA-48
- Dépistage des patients ayant bénéficié d'une CPRE avec l'endoscope 22 après le 29/08
- 4 patients suivants positifs avec *K. pneumoniae* OXA-48 ou *E. coli* OXA-48
- Prélèvements de l'endoscope n° 22: négatifs

➤ Intérêt d'une **méthode quantitative** pour prendre des **dispositions contraignantes** (séquestration définitive d'un endoscope)

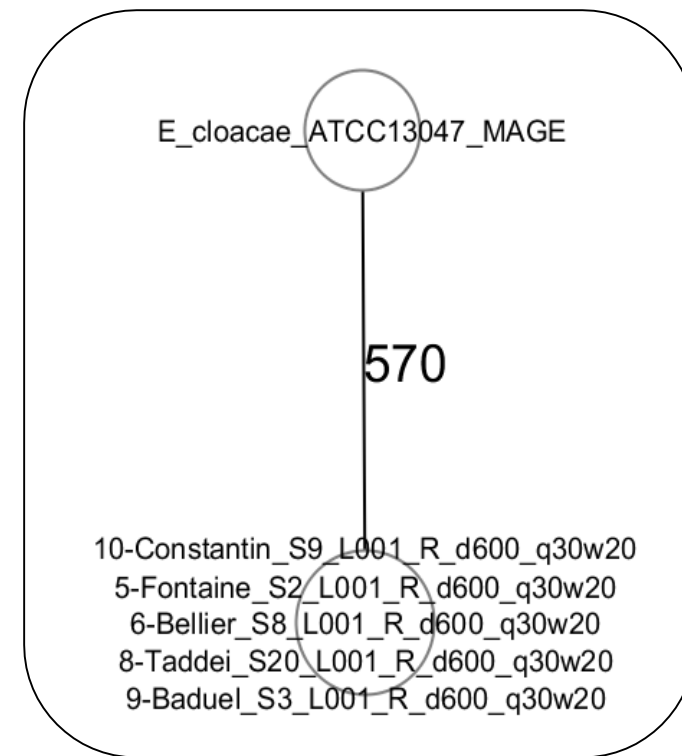


Exemple: épidémie d'*E. cloacae* BLSE



- **Variabilité de l'espèce *E. cloacae***
 - "seulement" **590 gènes** composent le schéma *cgMLST* (493989 bases)
- 1 souche avec résultats discordants
 - *qualité insuffisante du séquençage du génome*

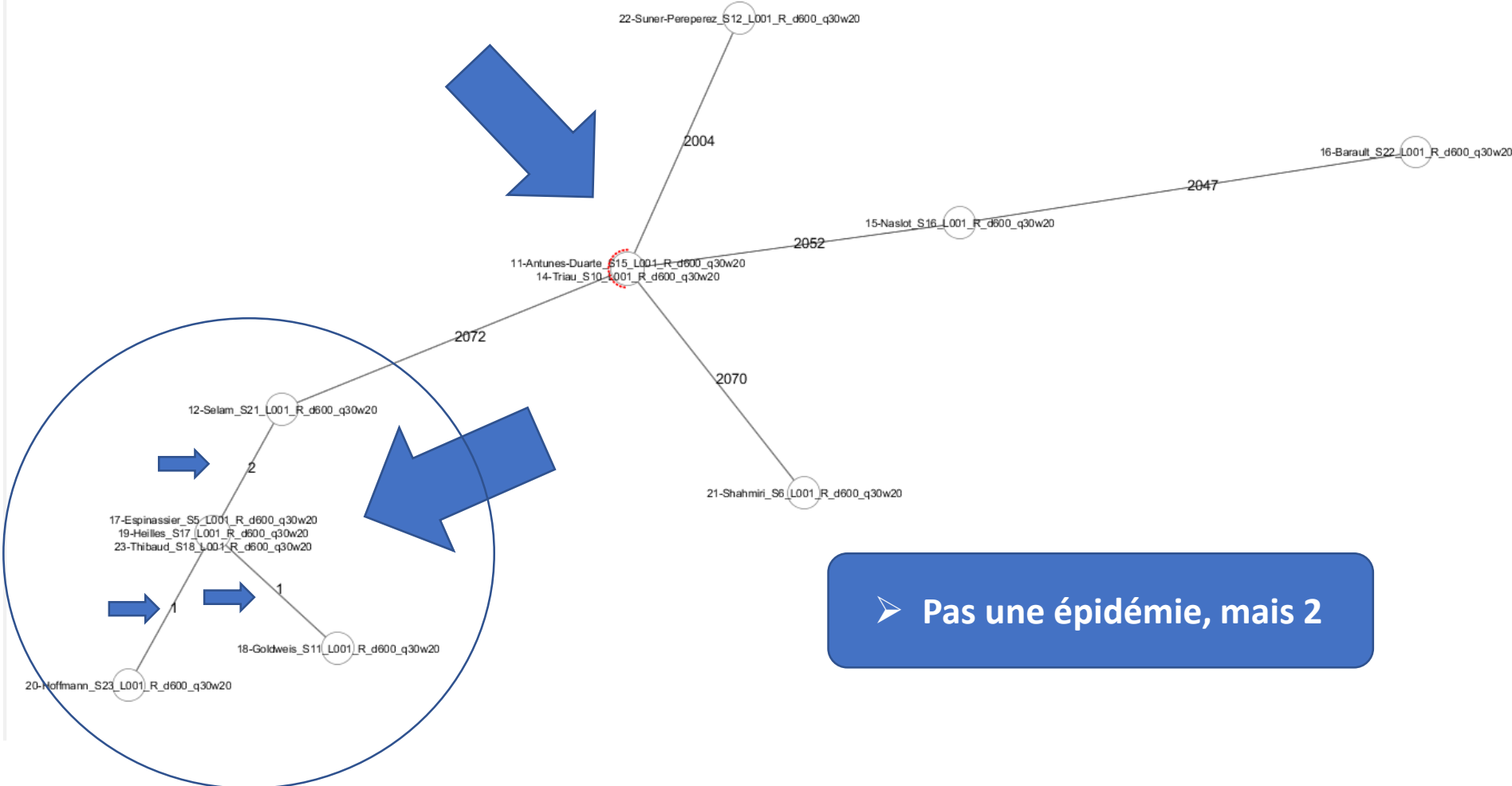
Performances dépendantes de l'espèce et de la qualité du séquençage



Spanning tree après retrait de la souche « aberrante »

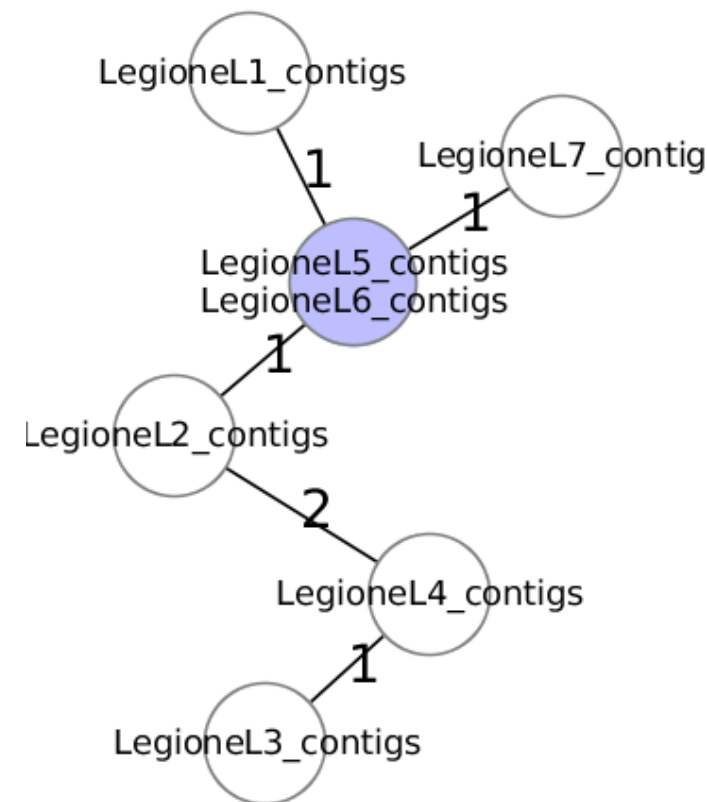
Epidémie de *K. pneumoniae* BLSE en réanimation chirurgicale cardiaque?

cgMLST 2544 gènes cibles



Exemple: cas de légionellose nosocomiale

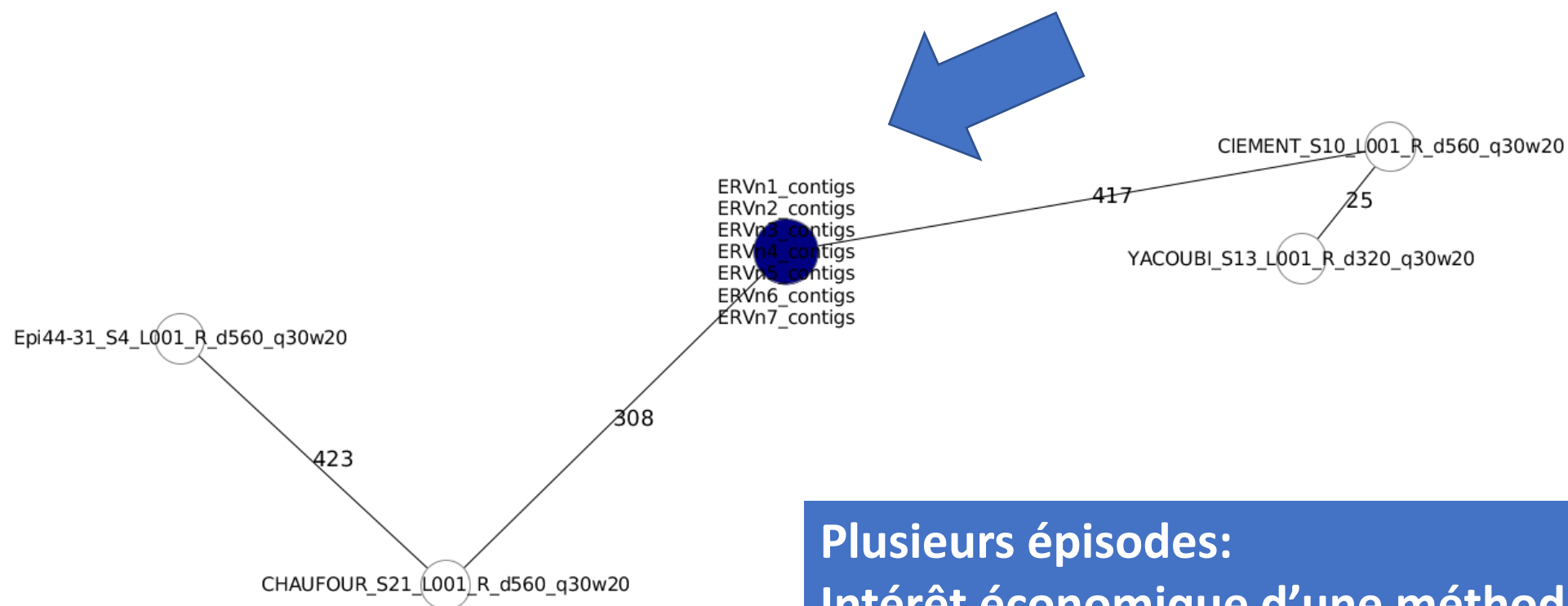
- Cas de Légionellose nosocomiale avérée en avril 2018
- Enquête environnementales: 6 souches dont 5 dans l'eau froide
- Seuil: 4 différences alléliques / 1283 gènes
- En attente du retour du CNR
- Attention aux recombinaisons: toujours confronter aux données épidémiologiques
 - Épidémie de Légionnelles, 2 clusters avec 40 différences alléliques, finalement correspondant à une recombinaison
(Petzold et al. Euro Surveill 2017)



Rapidité des résultats

Exemple: épidémie ERV

1 496 gènes du cgMLST



Plusieurs épisodes:
Intérêt économique d'une méthode reproductive

Discussion:

Avantages vs techniques dites « de référence »

➤ **Objectivité de l'information:**

- **Séquences:** reproductibilité + incrémentation d'une base de données qui peut évoluer sur des jours et des années
- **Quantification des différences**, à interpréter selon la variabilité génétique de l'espèce et les taux de mutation (*L. monocytogenes*: stable +++ avec $2.5 \cdot 10^{-7}$ substitutions par position par an)
 - Attention à la durée de l'épidémie
 - Publications justifiant des seuils de décisions

➤ **Délai:** \leq techniques dites de référence

➤ **Coût:** 70 à 150 euros par souche (cf. Diversilab® - rep PCR, BioMérieux)

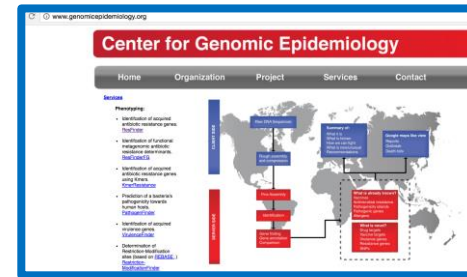
- Pas de matériel dédié
- Pas de réactif dédié
- Intérêt d'optimiser les séries (« runs »): **centralisation des analyses sur une plateforme**
- Codification: RIHN

➤ **Alternative concurrentielle** (coût, temps): utilisation des profils obtenus lors de l'identification par MALI-TOF?

Discussion: perspectives

➤ Au delà de la comparaison de souches, obtention des informations suivantes:

- Fond génétique (ST, serogroupage ...)
- Mécanismes de résistance
- Facteur de virulence
- Présence de plasmides
 - Identification des origines de répllication
 - Reconstruction « in silico » : logiciel PlasCope®



➤ **Évaluation de l'épidémiogénicité des souches?**

- **Adaptation du niveau de précautions complémentaires en fonction du contenu génétique**
- **Médecine (« hygiène ») personnalisée en fonction du micro-organisme**

➤ **Séquençage et comparaison en prospectif de certaines BMR** (Mellman et al. J Clin Microbiol 2016)

- Définition des taux de transmission de base (hors épidémie)
 - SARM 5.8%, *E. coli* MDR 2.3%
- **Adaptations des stratégies de mise en place de précautions complémentaires**
- Évaluation des conséquences

Discussion: perspectives

- Encore plus discriminant, le « **Single nucleotide polymorphism - *SNP calling*** »:
 - Basée sur le principe statistique d'apparition indépendante des substitutions
 - Mais utilisation de différents programmes successifs, dont les différents paramètres de calcul (filtres) sont des sources de variabilité
 - Problème des insertions d'éléments génétiques mobiles, transferts horizontaux, qualité +++ des génomes: expertise bioinformatique
- Encore plus rapide : **techniques par nanopore / séquenceur portable**
 - **Cout, erreurs+++ mais reconstruction de l'architecture du génome**
- **Accréditation:** problèmes des souches environnementales, de la portée d'accréditation, des CQ externes, bases de données, versions..
- **Nouveaux métiers et compétences** à l'hôpital (bioinformatique)

Conclusion

- Page qui se tourne, à l'évidence
- Quelle **codification** (disparition de la nomenclature « parallèle » dite de **Montpellier**)?
- Difficulté d'évaluer le retour sur investissement (RIHN) vs prévention
- Analyse « **accréditable** »?

