Typage et comparaisons de souches : rôle et implications

BARON Sophie

MCU-PH

sophie.baron.2@univ-amu.fr

OBJECTIFS DU COURS

1. **Formuler la bonne question**
   Savoir ce que le typage peut apporter, et ce qu'il ne peut pas dire.
2. **Connaitre les methodes**
   Antibiotype, PFGE, MLST, MALDI-TOF, WGS.
3. **Lire une figure**
   Minimum spanning tree, dendrogramme, matrice de SNP.
4. **Manier les seuils**
   Comprendre pourquoi il n'existe pas de seuil universel.
5. **Decider**
   Traduire un resultat de typage en mesure d'hygiene.

Plan du cours

Pré-requis : concept d'espèce bactérienne/ clonalité bactérienne

Méthodes phénotypiques

Méthodes génomiques

- Multi-Locus Sequence Typing
- Séquençage complet
  - Principe et outil du séquençage génomique
  - Principe de l'assemblage
  - Méthode comparaison 2 à 2 : k-mers
  - cgMLST
  - SNPs
  - Les bons outils pour les bonnes questions
- Interpréter les résultats
  - Minimum Spanning Tree
  - Matrice de Distance
  - Quel seuil ?

Cas pratiques

CAS CLINIQUE

Réanimation polyvalente, 16 lits. En six semaines, le laboratoire signale **cinq patients** porteurs d'entérobactéries productrices de carbapénémase (EPC).

- P1, P2, P3, P5 : *Klebsiella pneumoniae* NDM-1, dépistage rectal
- P4 : *Enterobacter cloacae* complex NDM-1, hémoculture
- Antibiogrammes superposables pour les cinq isolats
- P1 et P5 n'ont jamais partagé la même chambre ni la même équipe

Les questions qui se posent

Faut-il déclarer une épidémie, arrêter les admissions et dépister l'ensemble des contacts ?

Et sur quel argument ?

Le typage répond à trois questions, pas plus

Est-ce bien le même microorganisme ?

Identifier l'espèce ne suffit pas : deux *K. pneumoniae* peuvent être aussi différentes que deux inconnus dans la rue.

Y a-t-il eu transmission ?

Exclure une parenté est simple et fiable. La démontrer est toujours probabiliste.

Où est le réservoir ?

Patient, soignant, siphon, endoscope, solution : le typage relie les isolats humains et environnementaux.

Une quatrième question lui échappe totalement : **qui a contaminé qui, et par quel geste**. Cela reste le travail de l'enquête épidémiologique.

Pré-requis : Concept d'espèce bactérienne

- **Espèce bactérienne basée sur le phénotype = Taxospecies (phenospecies)**

  Ensemble d'individus ayant en commun la majorité de leurs caractères, et donc phénotypiquement plus ou moins semblables.

Choix des caractères :

- Coloration de Gram
- Morphologie
- Mobilité
- Capacité à sporuler
- Température de croissance
- Besoins nutritionnels
- Mode respiratoire
- Capacité de photosynthèse
- Utilisation des différentes sources de carbone ou d'azote
- …

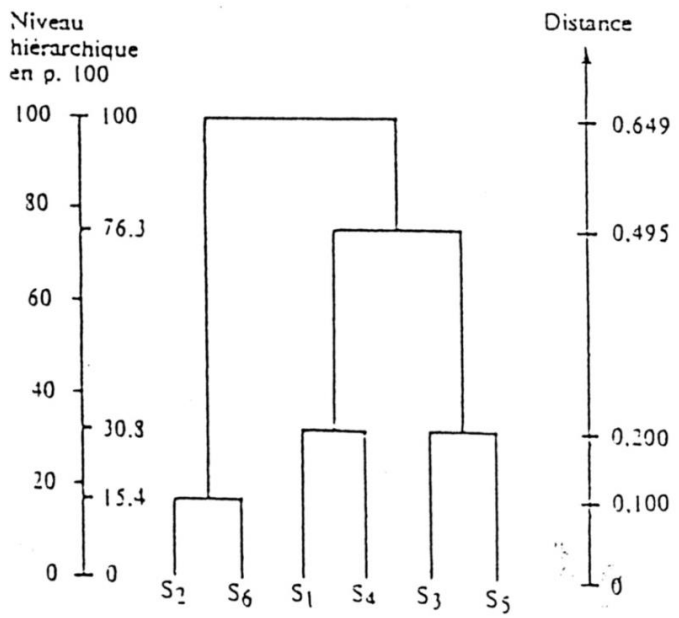


Galerie API 20E d'E.coli

Codage

Souches	Réponses aux tests T1 à T20 0: caractère négatif; 1: caractère positif
S1	001 001 011 111 101 000 00
S2	111 011 000 011 101 010 00
S3	001 001 011 101 011 000 10
S4	001 100 011 111 101 000 00
S5	010 001 011 101 011 000 10
S6	011 011 000 011 101 010 00

Dendrogramme



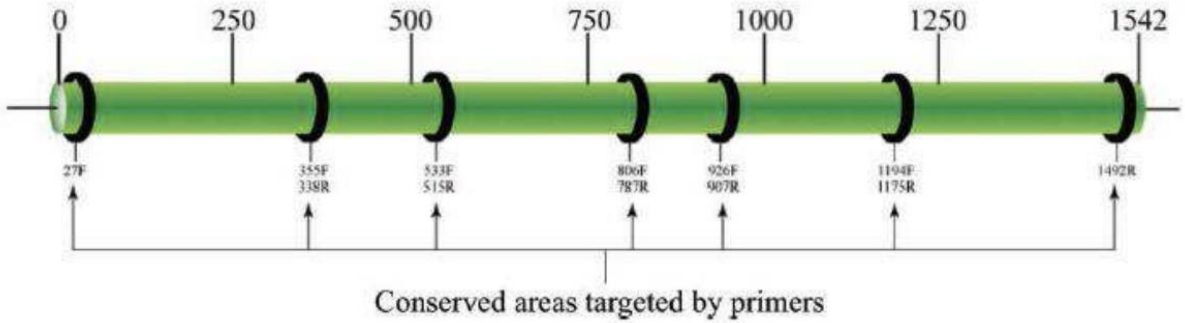
Pré-requis : Concept d'espèce bactérienne

- **La genospecies**: ensemble d'individus complètement ou partiellement identiques d'un point de vue génétique.

Définit en fonction :

- **GC%** ou coefficient de CHARGAFF
- La séquence de **l'ARN ribosomal 16S**
- Le taux d'**hybridation ADN/ADN (DDH)**

Structure ARNr 16S

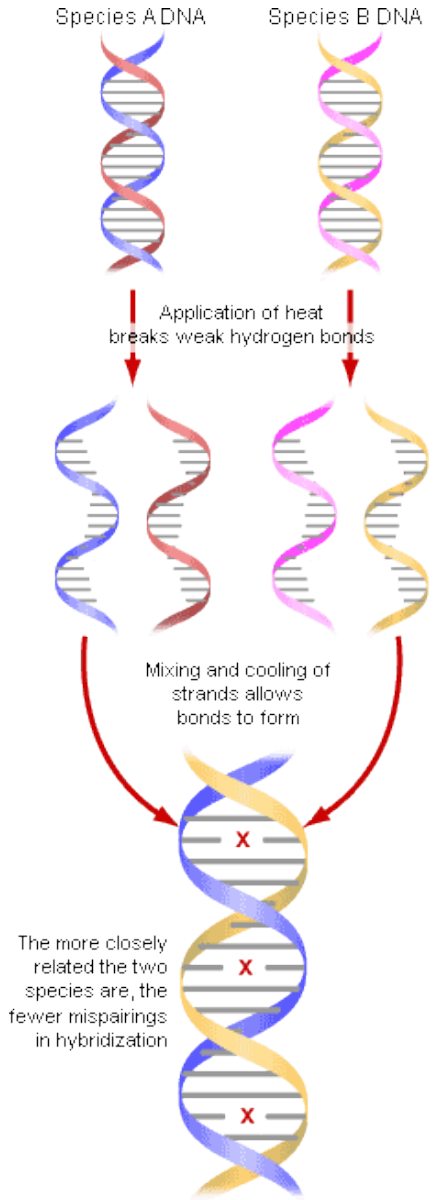


Même espèce =

GC% identique

ARNr 16S similarité >97%

Hybridation ADN/ADN > 70%



Pré-requis : Concept d'espèce bactérienne

Division asexuée par mitose

MAIS

Plasticité génome bactérien

Mutations chromosomiques, réarrangements

Perte de matériel génétique

Acquisition de matériel génétique :

Acquisition de plasmides

Acquisition d'ilots génomiques au niveau du chromosome

Changements **stables ou transitoires**

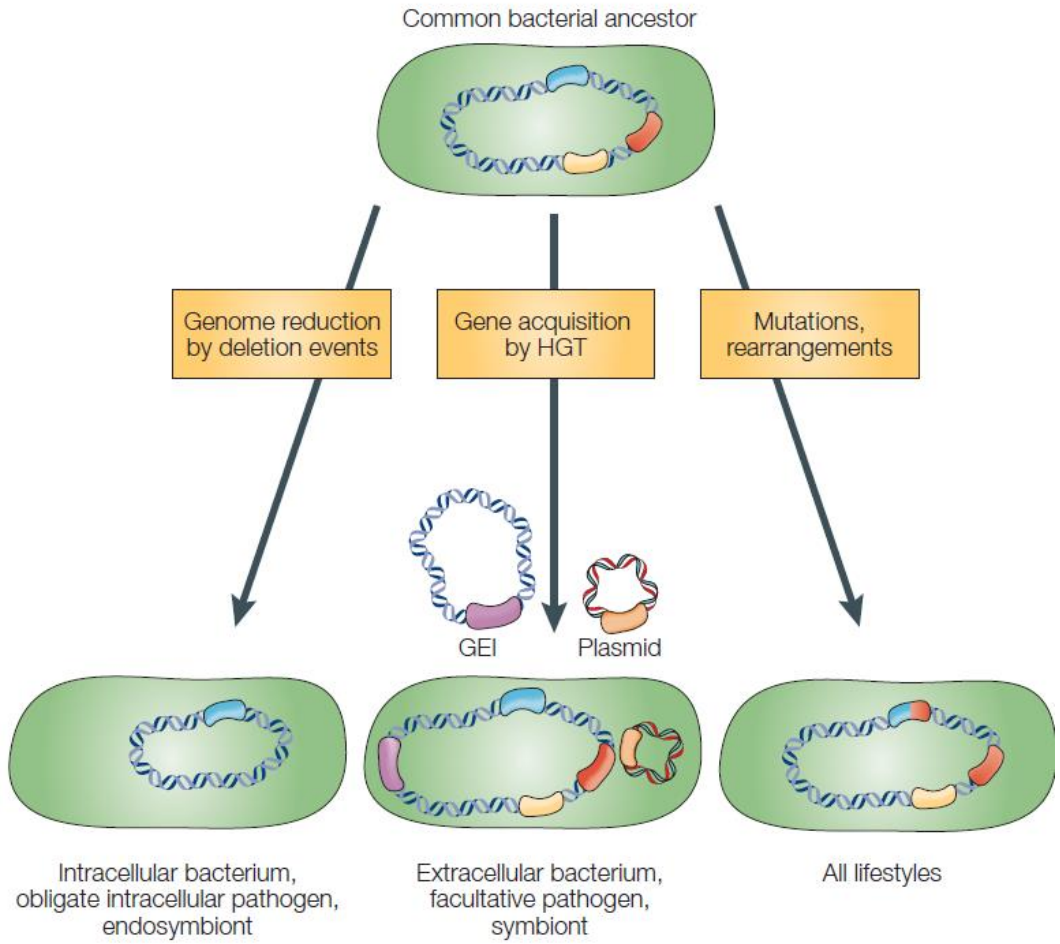


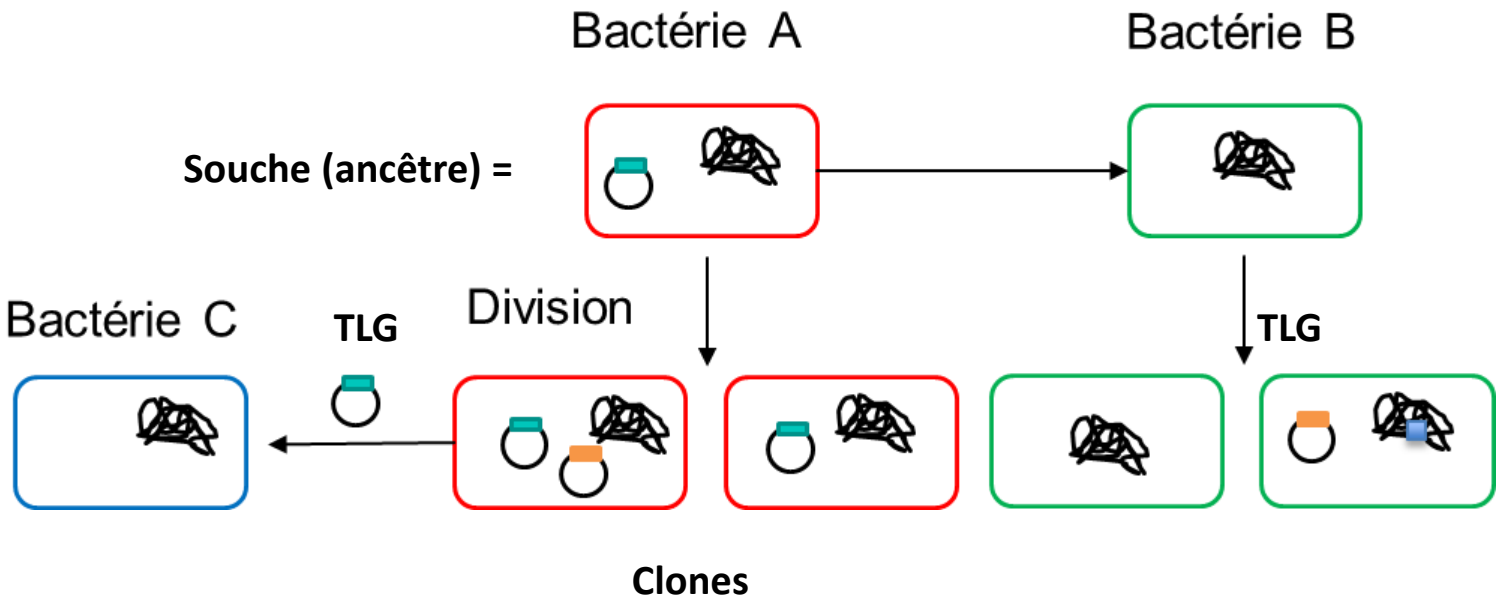
Figure 1 | **Evolution of bacterial variants by acquisition and loss of genetic information.**

TLG : Transfert latéral de gène ≈ reproduction sexuée

Dobrindt et. al, Nature Review Microbiology, 2004

Comment définit-on la clonalité bactérienne ?

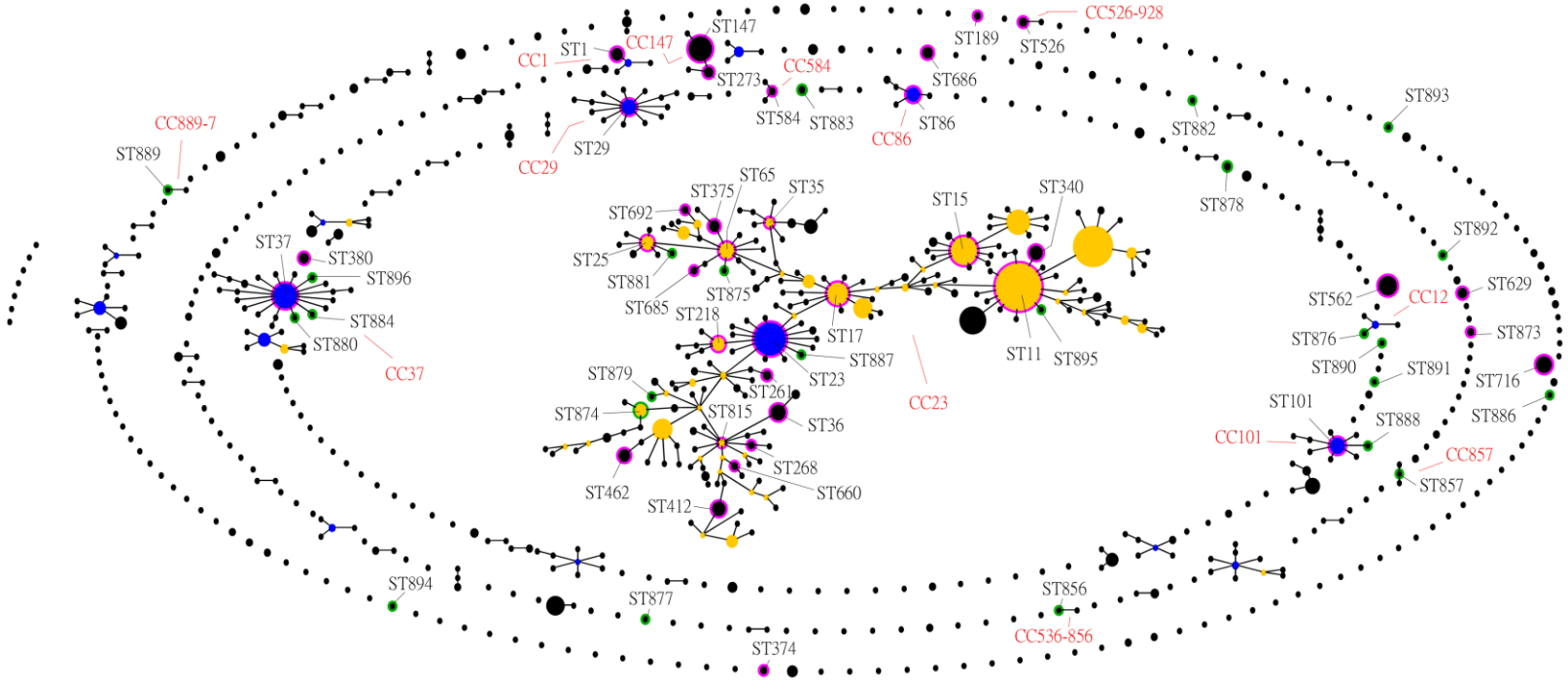
- La clonalité bactérienne désigne un ensemble de bactéries **génétiquement identiques, descendantes d'une seule cellule ancestrale**, formant ainsi un groupe **homogène** au sein d'une espèce.
- Bactéries clonales = lignée cellulaire complètement ou partiellement identique au sens génétique et issue d'une même souche (ancêtre)



Comment définit-on la clonalité bactérienne ?

- Différences **phénotypiques** au sein d'une même espèce bactérienne : **clone particulier induisant une pathologie ou présentant des différences phénotypiques.**

- Différences de **pathogénicité** : souches/clones pathologiques vs portage asymptomatique de nombreuses bactéries (*S. aureus, Vibrio….)*



Snapshot de la population de *Klebsiella pneumoniae*

Qi Wang, Plos One, 2013

Pré-requis : Six mots à ne pas confondre

Terme	Exemple	Définition
Espèce	*Klebsiella pneumoniae*	Niveau de l'identification (MALDI-TOF). Ne dit rien de la parenté.
Souche	L'isolat d'un patient donné	Une culture pure issue d'un prélèvement, à un moment donné.
Clone	Isolats issus d'un ancêtre commun récent	Notion relative : « récent » dépend de l'échelle.
ST / lignée	ST147, ST307, ST258…	Groupe défini par le MLST. Échelle de la décennie et du continent.
Clone à haut risque	ST147 NDM, ST258 KPC	Lignée mondialement diffusée, associée à la multirésistance.
Cluster	Isolats sous un seuil de parenté	Notion opérationnelle, définie par l'analyse, pas par la nature.

FONDAMENTAUX

Comment juger une méthode de typage

Pouvoir discriminant

Sépare-t-elle deux souches souches non apparentées ? apparentées ?

Typabilité

Donne-t-elle un résultat pour pour tous les isolats ?

Reproductibilité

Même résultat demain, dans dans un autre labo ?

Portabilité

Le résultat est-il comparable comparable entre centres ? ?

Une cinquième qualité, souvent oubliée : la **concordance épidémiologique**. Une méthode parfaite au laboratoire mais qui contredit systématiquement l'enquête de terrain n'a aucune valeur.

AVANT LE SÉQUENÇAGE

Quarante ans de typage en une frise

1980s
Ribotypage,
plasmidotypage

1993
PFGE
standardisée

1998
MLST
7 gènes

2000s
MLVA, *spa*,
rep-PCR

2015+
WGS : cgMLST
et SNP

Chaque méthode a été la référence de son époque. Toutes répondaient à la même question avec une résolution croissante.

Quelles sont les méthodes utilisées pour détecter une clonalité ?

METHODES PHENOTYPIQUES

- Méthodes **phénotypiques** basée sur les caractères :
  - Biochimiques (classification en biotypes ou biovars)
  - Antigéniques (classification en sérotypes ou sérovars)
  - De sensibilité aux antibiotiques (classification en antibiotypes)
  - De sensibilité aux bactériophages (classification en lysotypes ou lysovars)…

- **Ex: L'antibiotype**

Avantages

- Disponible pour tout isolat, sans surcoût
- Premier signal d'alerte d'un regroupement
- Un profil différent oriente vers des souches différentes

Limites

- Un même clone peut perdre ou gagner une résistance en quelques jours
- Des clones sans lien peuvent converger vers le même profil
- Le profil dépend du panel testé et de la lecture
- Pouvoir discriminant très faible chez les BHRe

Même logique pour le sérotypage et le lysotypage : utiles pour classer, incapables de trancher une transmission croisée.

Quelles sont les méthodes utilisées pour détecter une clonalité ?

METHODES PHENOTYPIQUES

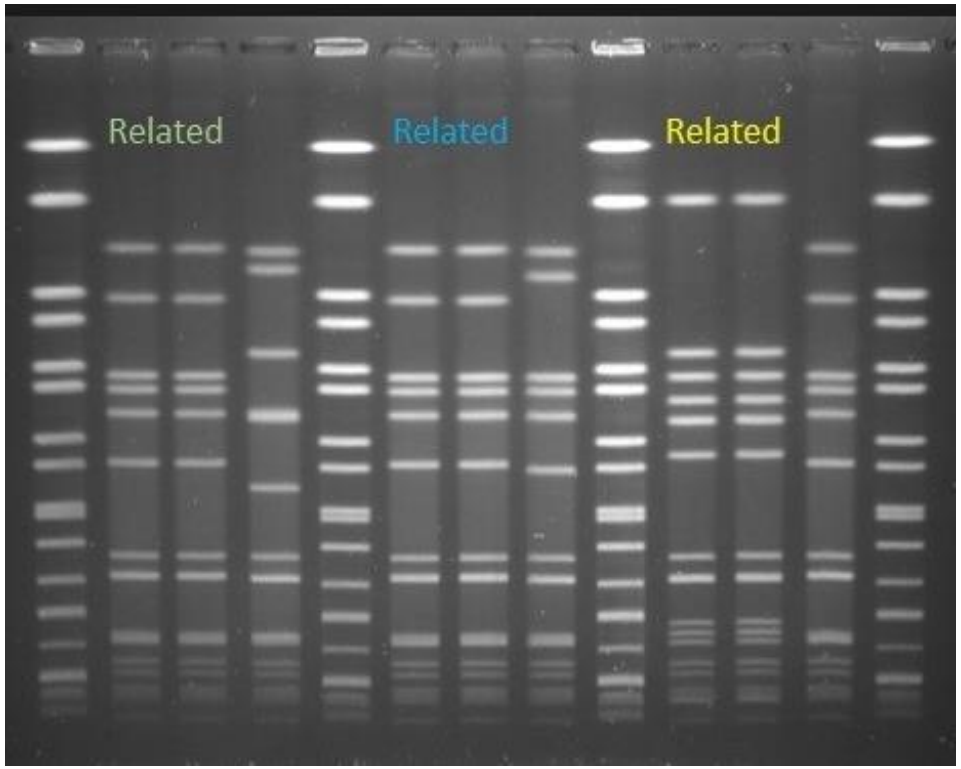
- Méthodes **phénotypiques moléculaires** :
  - Electrophorèse multilocus des isoenzymes (MLEE)
  - Polymorphisme de longueur des fragments de restriction (RFLP)
  - **Electrophorèse en champs pulsé (PFGE)**
  - Amplification aléatoire d'ADN polymorphe (RAPD)…

Niveau local ++
Fastidieux

- PFGE : la référence historique

L'ADN entier est coupé par une enzyme rare puis séparé en champ pulsé. Chaque souche donne un profil de bandes, comparé à l'œil ou par dendrogramme.

- Excellent pouvoir discriminant pour son époque
- Mais : 3 à 4 jours, technicien expérimenté
- Comparaison inter-laboratoires difficile
- Une seule mutation peut créer ou supprimer une bande



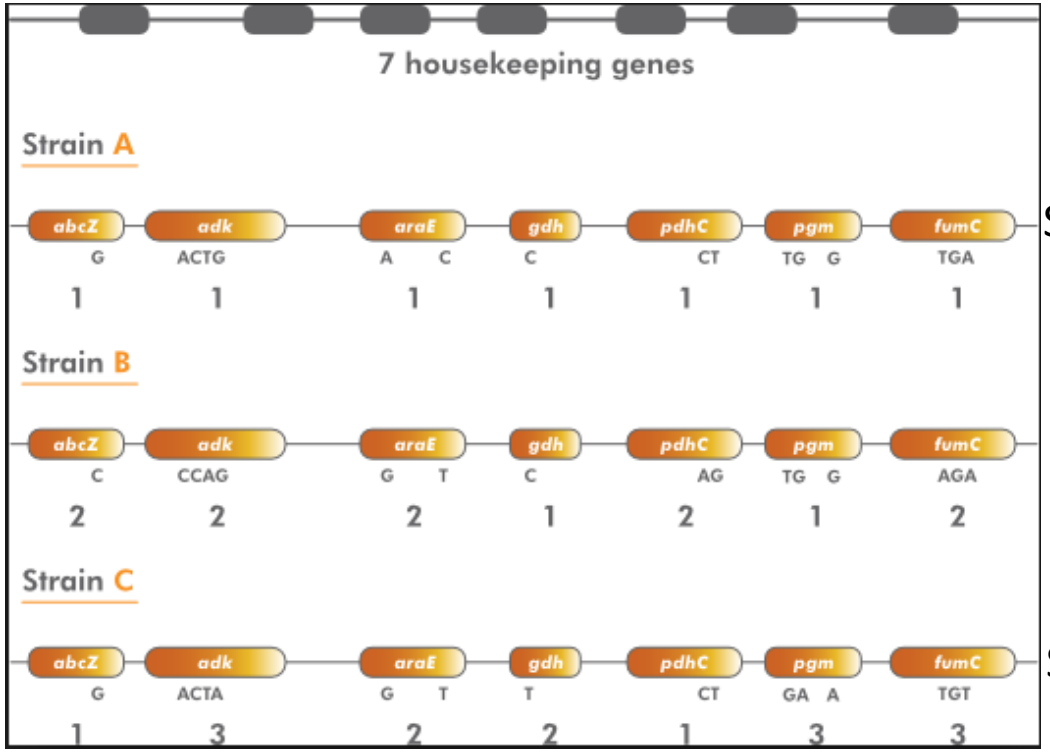
Quelles sont les méthodes utilisées pour détecter une clonalité ?

METHODES GENOMIQUES

3 grands types d'approches :

- **Approche gène-par-gène**, dite *Multi-Locus Sequence Typing* (MLST) et ses dérivés : analyse de **marqueurs génétiques prédéfinis**, typiquement les gènes en copie unique dans les génomes.
- Alignement et comparaison à des **séquences de référence** contre lesquelles les génomes étudiés seront alignés et comparés.
- **Comparaison deux à deux :** comparaison du contenu en 'mots' c'est-à-dire en séquences nucléotidiques d'une longueur 'k' choisie, appelés **k-mères**.

MLST = Multi Locus Sequence Typing



- Composée de plusieurs **gènes de ménage (7)** (*présent dans tous les génomes d'une même espèce)*
- Identification de **toutes** les mutations *(SNP = single nucleotide polymorphism)* de chaque gene
- Chaque **variant** du gène est codé par un allele
- Combinaison des 7 allèles = 1 séquenço-type (*ST = Sequence Type)*

MLST de *K. pneumoniae* ST307 et ST147

ST	gapA	infB	mdh	pgi	phoE	rpoB	tonB
307	4	1	2	52	1	1	7

ST	gapA	infB	mdh	pgi	phoE	rpoB	tonB
147	3	4	6	1	7	4	38

Avantage

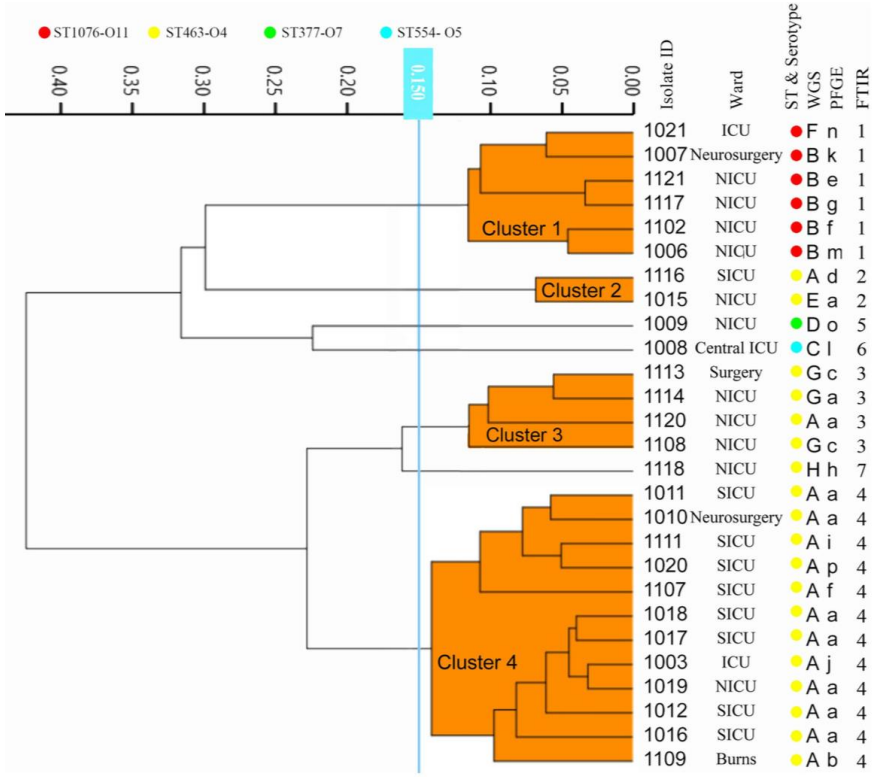
Un ST est un identifiant mondial et pérenne. Il permet de parler de *clones à haut risque* entre hôpitaux, entre pays, entre publications.

La limite

Sept gènes, soit environ **0,1 %** du génome. L'échelle est celle de la décennie, pas celle d'une épidémie de six semaines.

Le MALDI-TOF identifie l'espèce.
Il ne type pas la souche.

IR Biotyper



m was obtained automatically by FT-IR. The lineages of IRBT, WGS-based typing, PFGE, STs, and serotypes are given for each isolate.

Des applications de typage MALDI existent pour quelques situations (clones spécifiques), mais elles ne sont ni généralisables ni suffisantes pour une investigation.

Méthode basée sur la **technologie IR**

Etudie le profil des sucres, lipides et protéines des bactéries

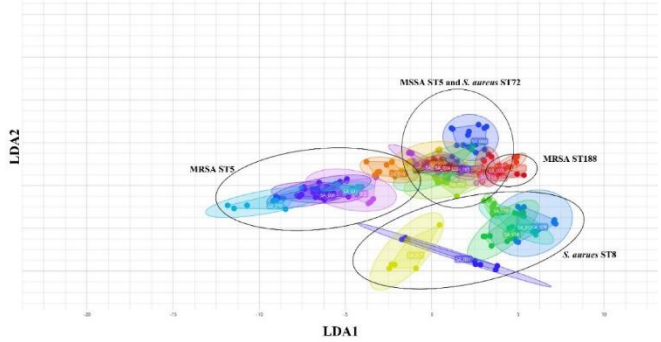
Méthode typage de souche couplé au MALDI-TOF

Permet l'analyse de scénarios épidémiques « en temps réel »

<u>Application</u> :

sérotypage pour *Streptococcus, Legionelle, Salmonella* et *Listeria*

Comparaison de profils pour les autres espèces



Récapitulatif

Méthode	Cible	Discrimination	Portabilité	Place en 2026
Antibiotype	Phénotype	Très faible	Faible	Signal d'alerte
Sérotype, spa	Antigènes, 1 locus	Faible à moyen	Bonne	Surveillance ciblée
PFGE	Génome entier, bandes	Bon	Médiocre	Quasi abandonnée
MLST	7 gènes	Moyen	Excellente	Nomenclature des clones
MLVA	Répétitions	Bon	Moyenne	Quelques espèces
WGS	Tout le génome	Maximal, réglable	Excellente	Référence actuelle

Le séquençage permet :

- Une résolution maximale
- Une comparaison des résultats comparables d'un centre à l'autre et d'une année sur l'autre.

LE SÉQUENÇAGE COMPLET

Ce que le WGS change

D'un profil à une distance

On ne compare plus des images : on compte des différences. Le résultat est un nombre.

Une seule analyse, plusieurs réponses

Espèce, ST, cgMLST, gènes de résistance, plasmides, virulence : tout sort du même fichier.

Une mémoire cumulative

Un génome séquencé aujourd'hui sera comparable à celui d'un patient admis dans trois ans.

MAIS il faut toujours une **question épidémiologique précise** et une **collection d'isolats bien choisie**.
Séquencer au hasard produit des données, pas des réponses.

LE SÉQUENÇAGE COMPLET:

Le circuit

1	2	3	4	5
Isolat	**Extraction**	**Séquençage**	**Analyse**	**Restitution**
Culture pure, une colonie	ADN génomique	Illumina ou Nanopore	QC, assemblage, distances	Avec l'équipe d'hygiène
1-2 j	**2-4 h**	**1-2 j**	**2-24 h**	**réunion**

Délai réaliste : 4 à 10 jours entre la décision et le rendu commenté, hors urgence organisée.

Conséquence : les mesures barrières sont prises avant le résultat. Le typage confirme, oriente ou allège ; il ne déclenche pas.

LE SÉQUENÇAGE COMPLET

Deux technologies, trois pièges

Lectures courtes (Illumina)

Très précises à la base près. Standard pour cgMLST et SNP. Assemblage fragmenté : mal adapté à la reconstruction des plasmides.

Lectures longues (Nanopore)

Reconstituent les plasmides entiers et les structures répétées. Délai court. Indispensable dès qu'on suspecte une diffusion de plasmide.

Trois causes classiques de résultat ininterprétable

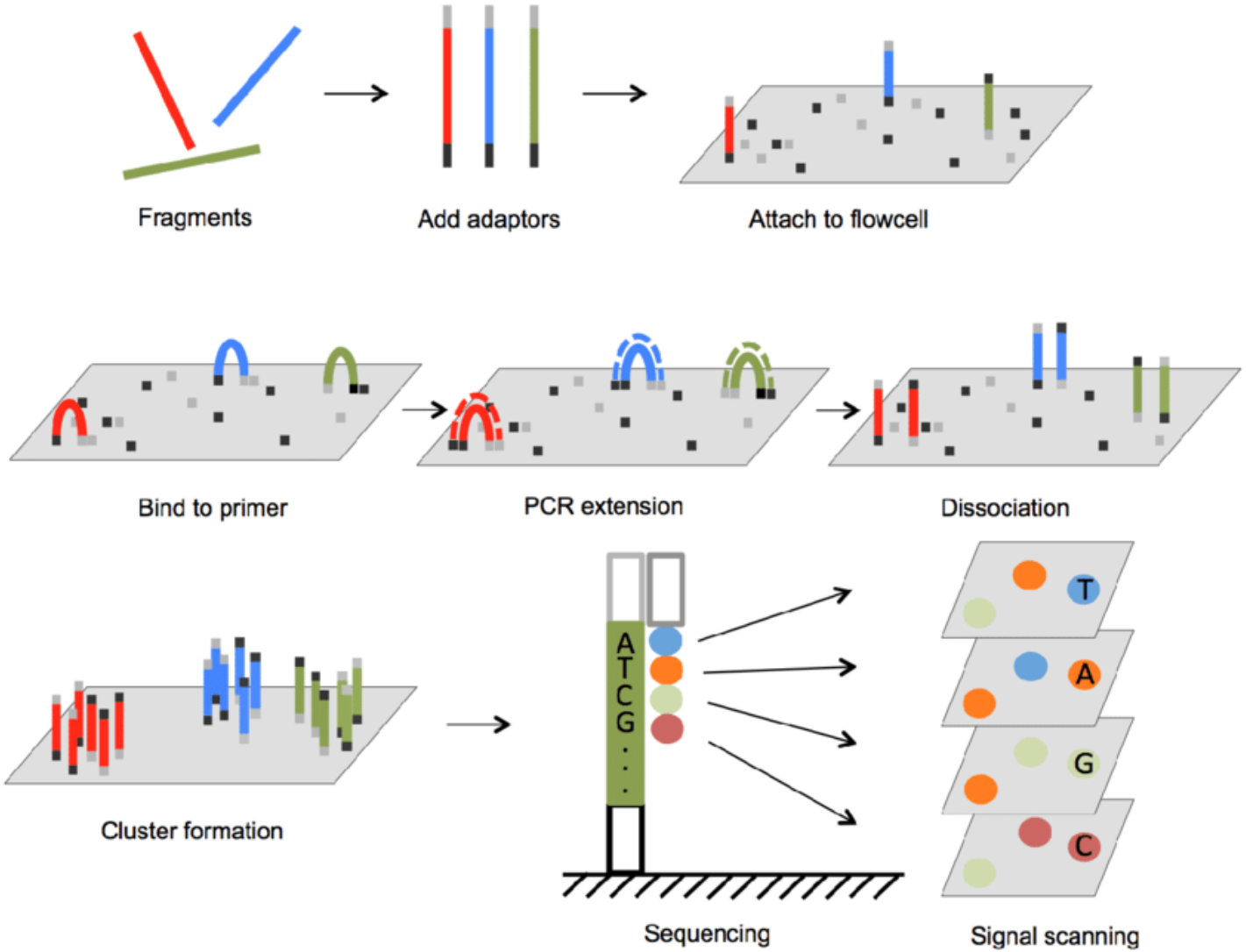
- **Culture non pure** : deux clones dans le même tube, le génome est un mélange
- **Profondeur insuffisante** : en dessous de 30×, les allèles manquants faussent la distance
- **Pipelines différents** : deux laboratoires, deux schémas, des chiffres non comparables

Séquençage par synthèse (Illumina)

Librairie pour fragmenter et amplifier les brins d'ADN (≃1,5 jours)

Plusieurs projets par cassette





Séquençage molécule unique en temps réel (Oxford, Pacbio)

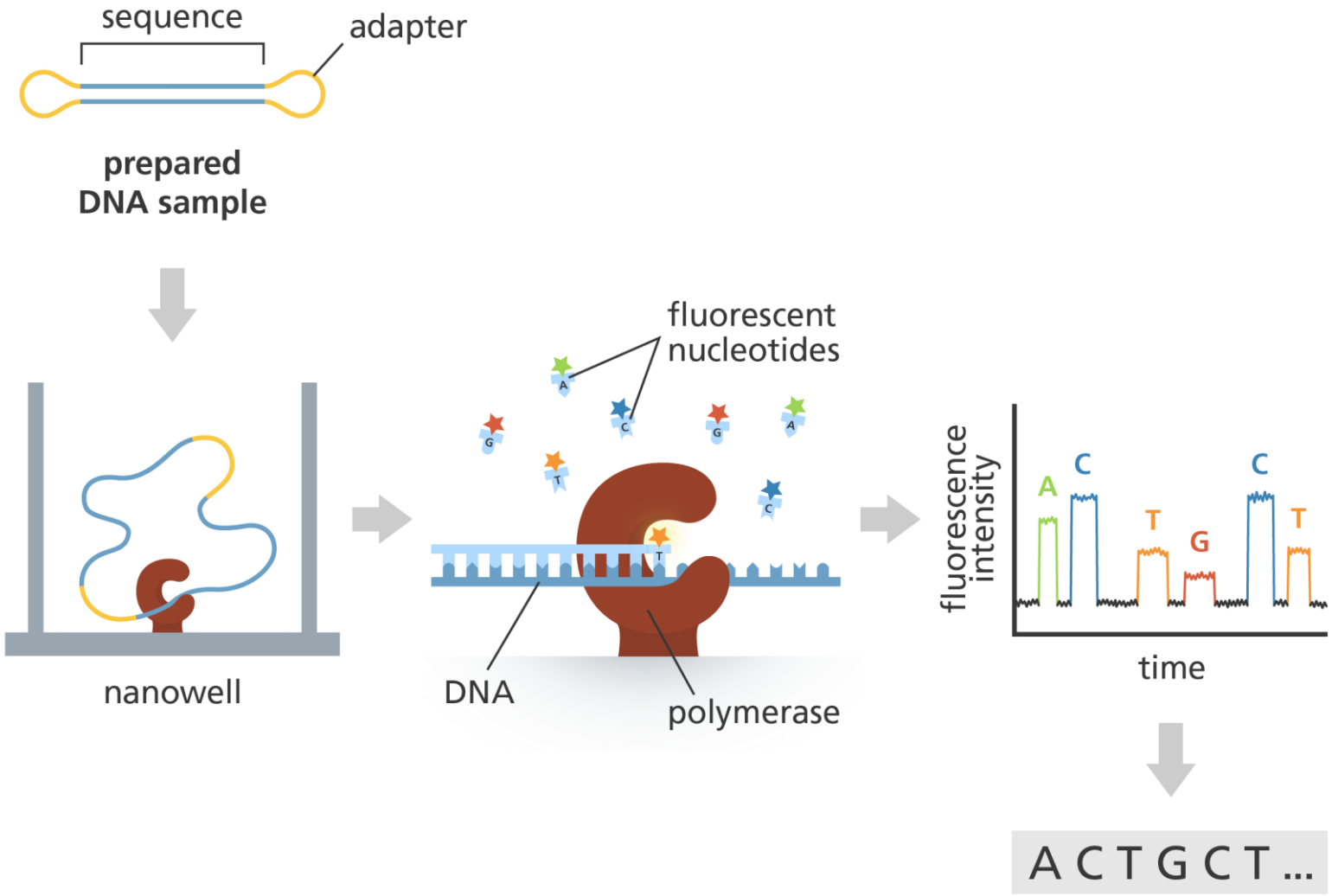
Préparation de l'échantillon rapide (≃1h)

Projet par projet

Plusieurs projets possibles sur la même flowcell

MAIS risque de contamination



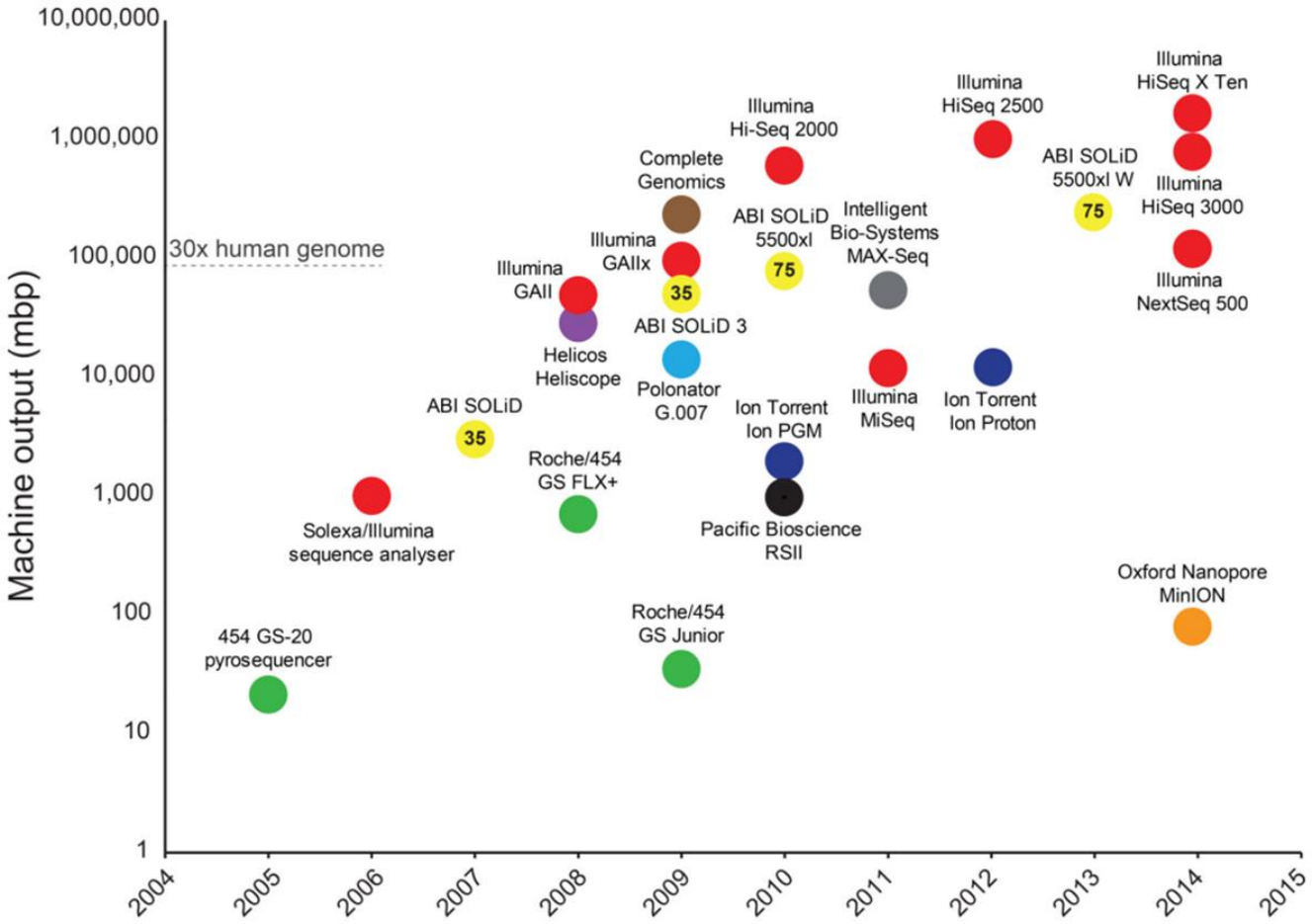


Technologies actuelles de séquençage

		Next generation sequencing	
	Méthode Sanger	**Séquençage par synthèse (Illumina)**	**Séquençage molécule unique en temps réel (Oxford, Pacbio)**
Avantages	Standardisation Qualité de séquençage	Simple, évolutif, haut rendement Possibilité de standardisation	Rapide, Equipement peu couteux (starter pack 1,000$)
Inconvénients	Cout au nucléotide Nombre de gènes limité	Equipement coûteux	Faible rendement des séquences de haute qualité
Délai	2-3 jours	4-5 jours	1-2 jours
Analyse des données	Rapide	+/- long	Inclus en « real-time » pour une partie des données +/- long

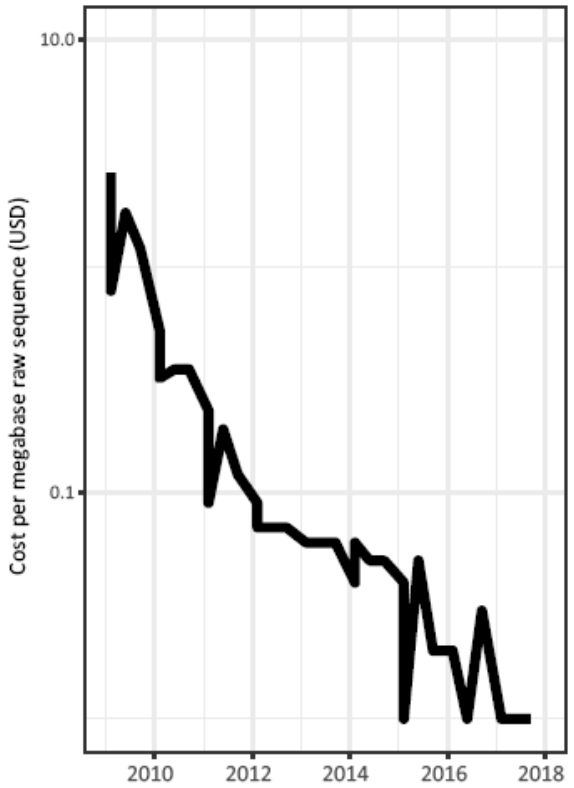
Séquençage Génome Entier : Performance et Coût

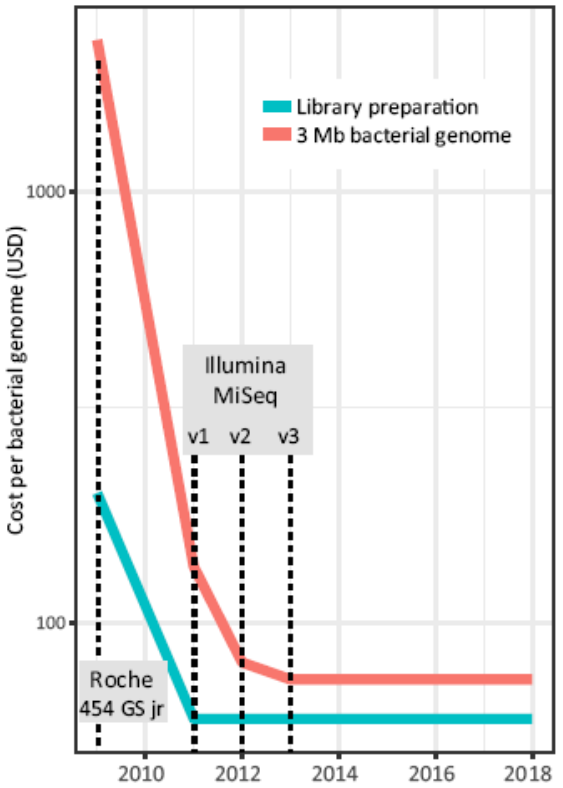
Next Generation Sequencing



Reuter et al. Mol Cell. 2015

Decrease of cost

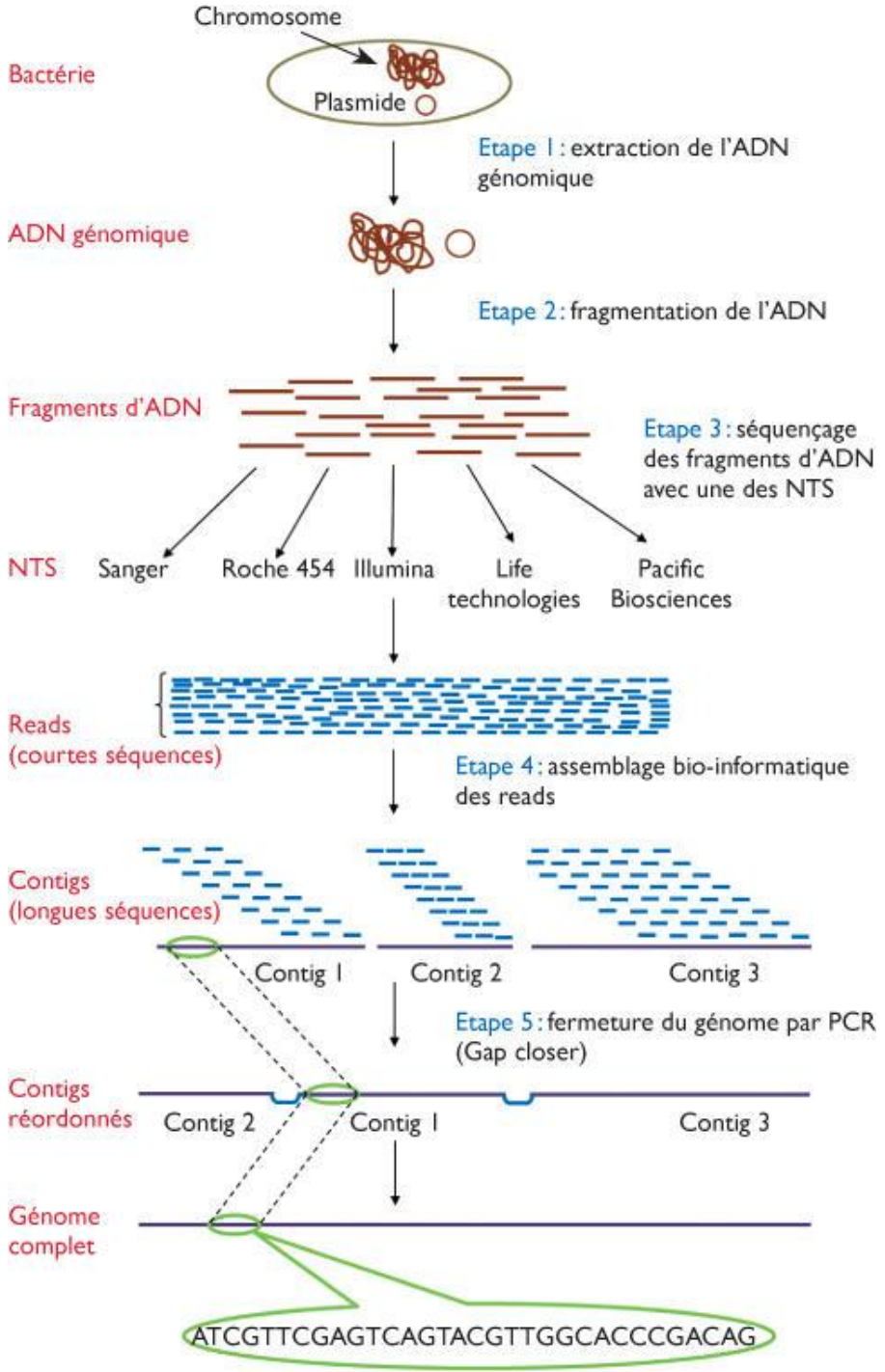




Balloux et al. Trends Microbiol 2018

Séquençage génome entier

Principe



Comparaison de génomes

gene par gene

cgMLST

On compte les **allèles différents** sur 1 500 à 3 000 gènes du génome cœur.

base par base

SNP

On aligne sur une référence proche et on compte les **positions nucléotidiques** qui diffèrent.

sans reference

k-mer

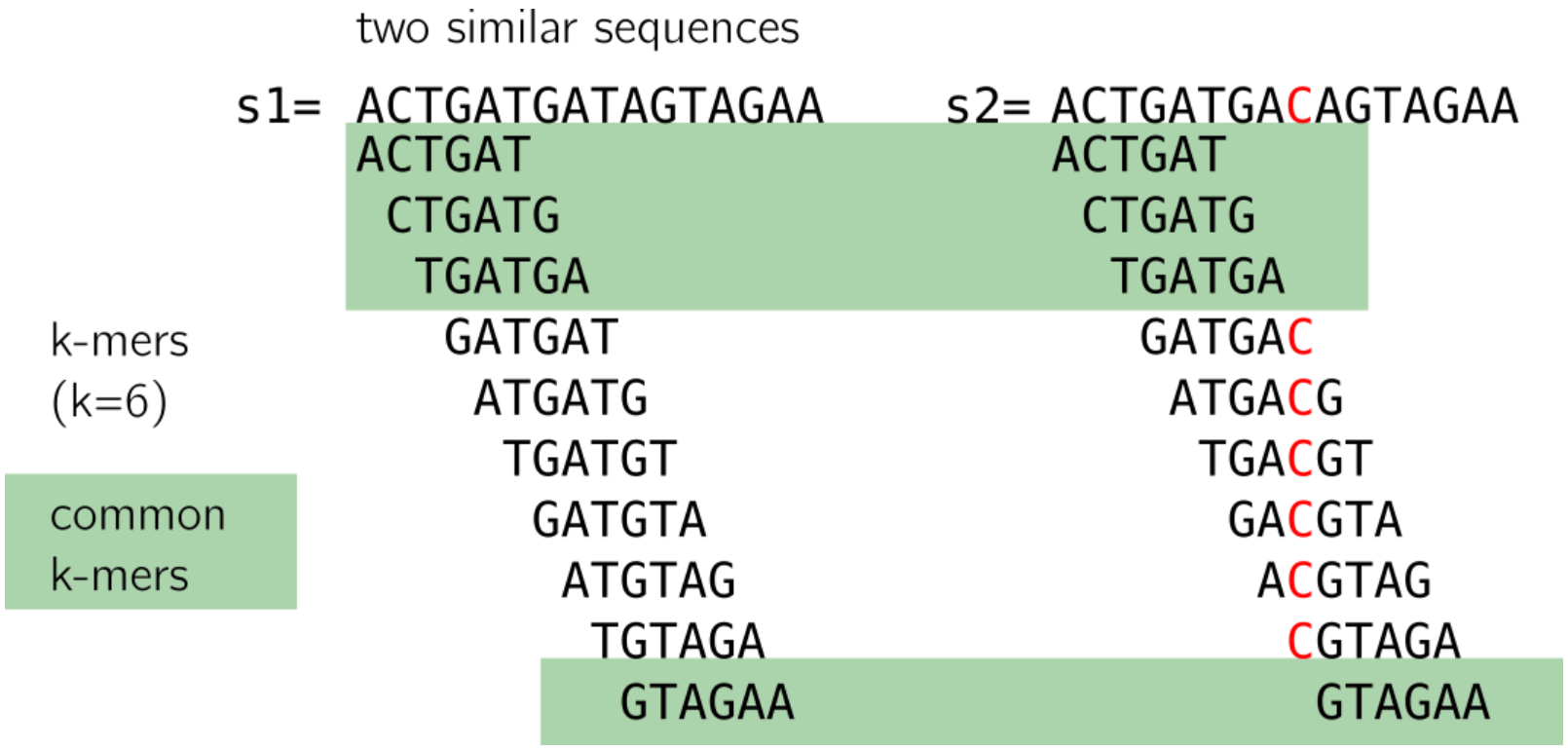
On compare le **contenu global** en fragments de longueur k, sans marqueurs prédéfinis.

Ces approches sont complémentaires et s'enchaînent :

- k-mer pour dégrossir
- cgMLST pour cadrer
- SNP pour trancher.

Comparaison deux à deux : k-mers

- **k-mers :** sous-chaîne de longueur *k* d'une chaîne de caractères plus grande.
- Pour un génome : extraction de toutes les sous-séquences de longueur fixe (k) de chaque génome,
- Comparaison de la fréquence et de la composition de ces k-mères entre eux.



Comparaison de génomes

cgMLST : le MLST étendu à tout le génome cœur

Le principe du MLST, appliqué non plus à 7 gènes mais à l'ensemble des gènes présents chez presque tous les isolats de l'espèce. Chaque variant reçoit un numéro d'allèle ; la distance entre deux souches est le **nombre de loci discordants**.

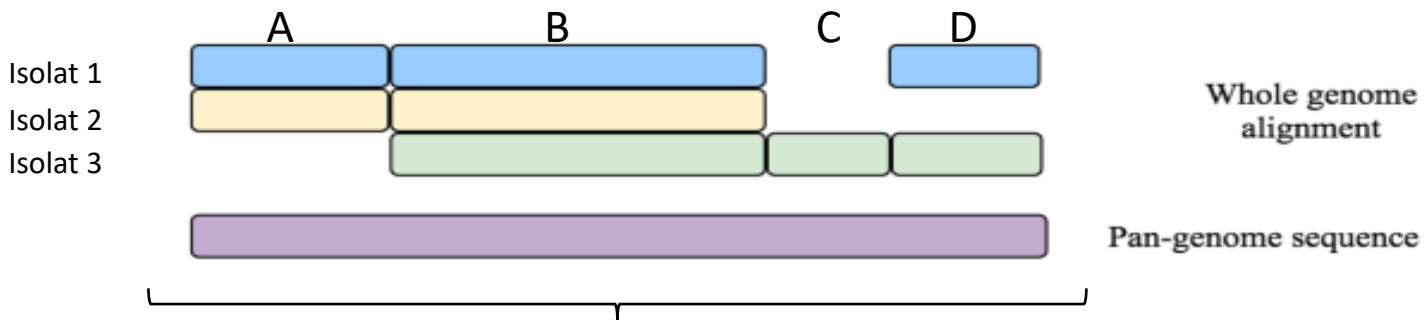
Schémas publics : **BIGSdb-Pasteur**, **EnteroBase**, **PubMLST**, **cgMLST.org**. Un numéro d'allèle veut dire la même chose dans tous les laboratoires qui utilisent le même schéma.

Atouts

Standardisé, portable, tolérant aux données manquantes, pas besoin de génome de référence.

Limites

Un gène avec 1 ou 40 mutations compte pour **1 allèle**. Ignore le génome accessoire, donc les plasmides. N'existe pas pour toutes les espèces.



Pangenome = **core genome (B)** + gènes accessoires (**AD**) + gènes spécifiques (**C**)

Klebsiella pneumoniae/variicola/quasipneumoniae cgMLST

Show Targets

Name	K. pneumoniae sensu lato cgMLST
Version	1.0
Seed Genome	NTUH-K2044 (NC_012731.1, 15-JUN-2016)
Genus	Klebsiella
Species	pneumoniae/variicola/quasipneumoniae
Locus Count	2,358
Complex Type Distance	15
Complex Type Count	19,661
Last Change	Sep 19, 2025, 4:28 PM
Curators	J. Rossen, Department of Medical Microbiology, University of Groningen, the Netherlands
	D. Harmsen, University Münster, Germany
Publications	**Weber RE, Pietsch M, Frühauf A, Pfeifer Y, Martin M, Luft D, Gatermann S, Pfennigwerth N, Kaase M, Werner G, and Fuchs S.** IS26-Mediated Transfer of blaNDM-1 as the Main Route of Resistance Transmission During a Polyclonal, Multispecies Outbreak in a German Hospital. *Front Microbiol.* 2019, **10**: 2817 [PubMed 31921015]
	Piazza A, Comandatore F, Romeri F, Brilli M, Dichirico B, Ridolfo A, Antona C, Bandi C, Gismondo MR, and Rimoldi SG. Identification of blaVIM-1 Gene in ST307 and ST661 Klebsiella pneumoniae Clones in Italy: Old Acquaintances for New Combinations. *Microb. Drug Resist.* 2019, **25**: 787-790 [PubMed 30589602]

Comparaison de génomes

SNP : une résolution maximale mais pour moins de comparabilité

Avantages

- Distingue des isolats que le cgMLST ne sépare pas
- Permet de dater et d'orienter la transmission
- Sans marqueurs prédéfinis : aucun biais de sélection

Limites

- Dépend du **génome de référence** choisi
- Dépend du filtrage des régions répétées et recombinées
- Deux analyses, deux références : deux nombres différents

Un nombre de SNP n'a de sens qu'accompagné de la référence utilisée, du pipeline et des filtres appliqués. Doit être renseigné dans le compte rendu.

Comparaison de génomes

Quelle approche pour quelle question ?

Votre question	Approche	Pourquoi
Exclure une parenté	cgMLST	Rapide, standardisé, portable
Confirmer un cluster	cgMLST puis SNP	Le SNP affine ce que le cgMLST a regroupé
Dater ou orienter la transmission	SNP + phylogénie	Nécessite des dates de prélèvement fiables
Espèce rare ou sans schéma	k-mer (Mash)	Aucun marqueur prédéfini requis
Suspicion de plasmide partagé	Lectures longues + analyse plasmidique	Le génome cœur ne répondra pas
Situer dans l'épidémiologie mondiale	MLST (ST) + cgMLST	Le ST reste le langage commun

Interpréter un résultat de comparaison de génomes

Le minimum spanning tree



Chaque cercle est un isolat, chaque trait porte le **nombre d'allèles différents**. La longueur du trait ne veut rien dire : seul le chiffre compte.

Interpréter un résultat de comparaison de génomes

La matrice de distances

	P1	P2	P3	P5	Q1	Q2
P1	–	3	2	5	418	421
P2	3	–	4	6	420	423
P3	2	4	–	5	417	420
P5	5	6	5	–	419	422
Q1	418	420	417	419	–	9
Q2	421	423	420	422	9	–

Doit figurer sur le compte-rendu

Le même jeu de données que l'arbre précédent, sans mise en forme.
Deux blocs apparaissent : P1-P2-P3-P5 (2 à 6 allèles) et Q1-Q2 (9 allèles), séparés par plus de 400.

> La matrice est plus sûre que l'arbre : elle montre **toutes** les distances, là où l'arbre n'en affiche qu'une sélection.

Quel seuil ?

Ordres de grandeur souvent rapportés

Espèce	Schéma cgMLST	Cluster évoqué si	Remarque
S. aureus	~1 860 loci	≤ 24 allèles	Schéma ancien, largement validé
L. monocytogenes	~1 750 loci	≤ 7 allèles	Seuil utilisé pour les types de cluster européens
C. difficile	~2 270 loci	≤ 6 allèles	Espèce peu variable
K. pneumoniae	630 ou 2 360 loci	≤ 5 à 10 allèles	Dépend fortement du schéma choisi
E. coli	~2 500 loci	≤ 20 à 26 allèles	Espèce plus diversifiée
Enterobacter cloacae cx	~2 700 loci	≤ 10 à 15 allèles	Schémas récents, moins de recul

Ces chiffres sont des **repères de lecture**, pas des règles de décision.
Ils ne valent que pour un schéma donné, un pipeline donné et une épidémie de courte durée.

Quel seuil ?

Contents lists available at ScienceDirect

Clinical Microbiology and Infection

journal homepage: www.clinicalmicrobiologyandinfection.com

Review

Whole genome sequencing options for bacterial strain typing and epidemiologic analysis based on single nucleotide polymorphism versus gene-by-gene–based approaches

A.C. Schürch [1], S. Arredondo-Alonso [1], R.J.L. Willems [1], R.V. Goering [2, *]

[1] Department of Medical Microbiology, University Medical Center, Utrecht, The Netherlands
[2] Department of Medical Microbiology and Immunology, Creighton University School of Medicine, Omaha, NE, USA

Table 1
Examples of relatedness criteria for wg/cgMLST and SNP typing schemes of representative clinically relevant bacteria

Organism	Relatedness threshold[a]		References
	wg/cgMLST (allele) SNPs		
Acinetobacter baumannii	≤8	≤3	[25,26]
Brucella spp.	Epidemiologic validation in progress[b]		http://www.applied-maths.com/applications/wgmlst
Campylobacter coli, C. jejuni	≤14	≤15	[27,28]
Cronobacter spp.	Epidemiologic validation in progress[b]		http://www.applied-maths.com/applications/wgmlst
Clostridium difficile	Epidemiologic validation in progress[b]	≤4	[29], http://www.cgmlst.org/ncs, http://www.applied-maths.com/applications/wgmlst
Enterococcus faecium	≤20	≤16	[30]
Enterococcus raffinosus	Epidemiologic validation in progress[b]		http://www.applied-maths.com/applications/wgmlst
Escherichia coli	≤10	≤10	[31,32], https://enterobase.warwick.ac.uk/
Francisella tularensis	≤1	≤2	[33,34]
Klebsiella oxytoca	Epidemiologic validation in progress[b]		http://www.applied-maths.com/applications/wgmlst
Klebsiella pneumonia	≤10	≤18	[35,36]
Legionella pneumophila	≤4	≤15	[37]
Listeria monocytogenes	≤10	≤3	[38,39]
Mycobacterium abscessus		≤30	[40]
Mycobacterium tuberculosis	≤12	≤12	[41]
Neisseria gonorrhoeae	Epidemiologic validation in progress[b]	≤14	[42], http://www.applied-maths.com/applications/wgmlst
Neisseria meningitidis	Epidemiologic validation in progress[b]		http://www.cgmlst.org/ncs
Pseudomonas aeruginosa	≤14	≤37	[31,43]
Salmonella dublin	Epidemiologic validation in progress[b]	≤13	[44], https://enterobase.warwick.ac.uk/
Salmonella enterica	Epidemiologic validation in progress[b]	≤4	[45], http://www.cgmlst.org/ncs, http://www.applied-maths.com/applications/wgmlst, https://enterobase.warwick.ac.uk/
Salmonella typhimurium	Epidemiologic validation in progress[b]	≤2	[46], https://enterobase.warwick.ac.uk/
Staphylococcus aureus	≤24	≤15	[47,48]
Streptococcus suis		≤21	[49]
Vibrio parahaemolyticus	≤10		[50]
Yersinia spp.	0		[51]

cg, core genome; MLST, multilocus sequence typing; SNP, single nucleotide polymorphism; wg, whole genome.
[a] Data often represent single studies that can be used to begin formulation of species-specific interpretation criteria. Thus, these data should be coupled with newly published similar studies to ensure that resulting values are not atypical and can be generally applied.
[b] Proposed wg/cgMLST schemes are available online (http://www.cgmlst.org/ncs, http://www.applied-maths.com/applications/wgmlst, https://enterobase.warwick.ac.uk/) but as yet have not been epidemiologically validated.

Quel seuil ?

Six paramètres qui déplacent le seuil

Durée de l'épidémie

Plus elle dure, plus les isolats liés divergent. Une diffusion sur huit ans peut dépasser 350 SNP.

Vitesse d'évolution

Variable selon l'espèce. Un phénotype hypermutateur peut multiplier le taux de mutation par plus de cent.

Pression de sélection

Antibiotiques, stress immunitaire, biofilm de siphon : l'environnement accélère ou fige l'évolution.

Portage prolongé

Un patient colonisé depuis des mois héberge une population diversifiée ; l'isolat séquencé n'en est qu'un échantillon.

Site de prélèvement

Selles, plaie, sang : des sous-populations différentes du même clone chez un même patient.

Nombre d'isolats par patient

Un seul isolat par patient sous-estime systématiquement la diversité réelle.

Quel seuil ?

La démarche « ad hoc » : un seuil par investigation

1

Dater la fenêtre

Combien de temps sépare le premier et le dernier prélèvement ?

2

Estimer la divergence attendue

Vitesse d'évolution de l'espèce ×durée, plutôt qu'un chiffre importé d'un article.

3

Ancrer sur un témoin

Inclure des isolats non épidémiques du même ST, même période, même région.

4

Confronter au terrain

Le seuil retenu doit être cohérent avec les liens connus, sinon il est faux.

Le point 3 est le plus négligé : sans isolats témoins, on ne sait pas si 10 allèles c'est proche ou banal dans cette lignée.

INTERPRÉTER ET DÉCIDER

Ce n'est pas toujours le clone qui circule !

Transmission clonale

La **bactérie** passe d'un patient à l'autre. Le cgMLST la voit : même espèce, même ST, distance faible.

Transmission plasmidique

Le **gène de résistance** passe d'une bactérie à l'autre, parfois entre espèces. Le cgMLST ne la voit **pas du tout** : il n'analyse que le génome cœur.

Signal d'alerte : **la même carbapénémase dans des espèces différentes**, dans la même unité. Il faut alors des lectures longues et une comparaison des plasmides, pas un cgMLST.

CAS CLINIQUE

Résultats

P1, P2, P3

K. pneumoniae ST147, **1 à 4 allèles** entre eux. Cluster confirmé.

P5

K. pneumoniae **ST307**, plus de 400 allèles des autres. Aucun lien clonal.

P4

E. cloacae complex. Espèce différente : la comparaison de génome cœur n'a pas de sens.

L'analyse complémentaire en lectures longues

P4 et P5 portent le **même plasmide IncX3 porteur de *bla*NDM-1**, quasi identique à celui des trois isolats ST147. Il ne s'agit pas d'une, mais de **deux épidémies superposées** : une diffusion clonale et une diffusion de plasmide.

CAS CLINIQUE

Quelle interpretation ?

Une transmission croisée

Les mesures ciblées sur le seul « clone ST147 » étaient insuffisantes : il fallait traiter l'unité comme un réservoir global.

Un évènement isolé

Aucun lien clonal entre P5 et les autres : inutile de remonter la chaîne de soins entre eux. Deux jours d'enquête économisés.

Les limites

Pas d’information sur le sens de la transmission, ni le véhicule, ni s'il existe un porteur non dépisté. L'enquête de terrain reste entière.

Le typage n’est qu’un complément à l’enquête et ne le remplace en aucun cas

INTERPRÉTER ET DÉCIDER

Une conclusion toujours basée sur plusieurs éléments

Génomique

Les souches sont-elles compatibles avec un ancêtre commun récent ?

Spatiale

Les patients ont-ils partagé un lieu, un matériel, une équipe ?

Temporelle

La chronologie rend-elle la transmission plausible ?

> Deux points sur trois : une hypothèse. Trois points concordants : une conclusion. Un seul point génomique : un chiffre, rien de plus.

INTERPRÉTER ET DÉCIDER

Question	Réponse opérationnelle
Qui décide de typer ?	L'équipe d'hygiène et le microbiologiste, ensemble, sur une question écrite.
Quels isolats envoyer ?	Tous les cas, plus des témoins non épidémiques, plus les prélèvements d'environnement pertinents.
Quand ?	Dès le 2e ou 3e cas groupé, sans attendre que l'épidémie soit installée.
Où ?	Laboratoire local équipé, plateforme régionale, ou Centre national de référence selon l'espèce.
Que demander dans le compte rendu ?	Espèce, ST, schéma et pipeline utilisés, matrice complète, témoins inclus, interprétation datée.
Et après ?	Conserver les isolats et les génomes. La prochaine alerte se lira à la lumière de celle-ci.

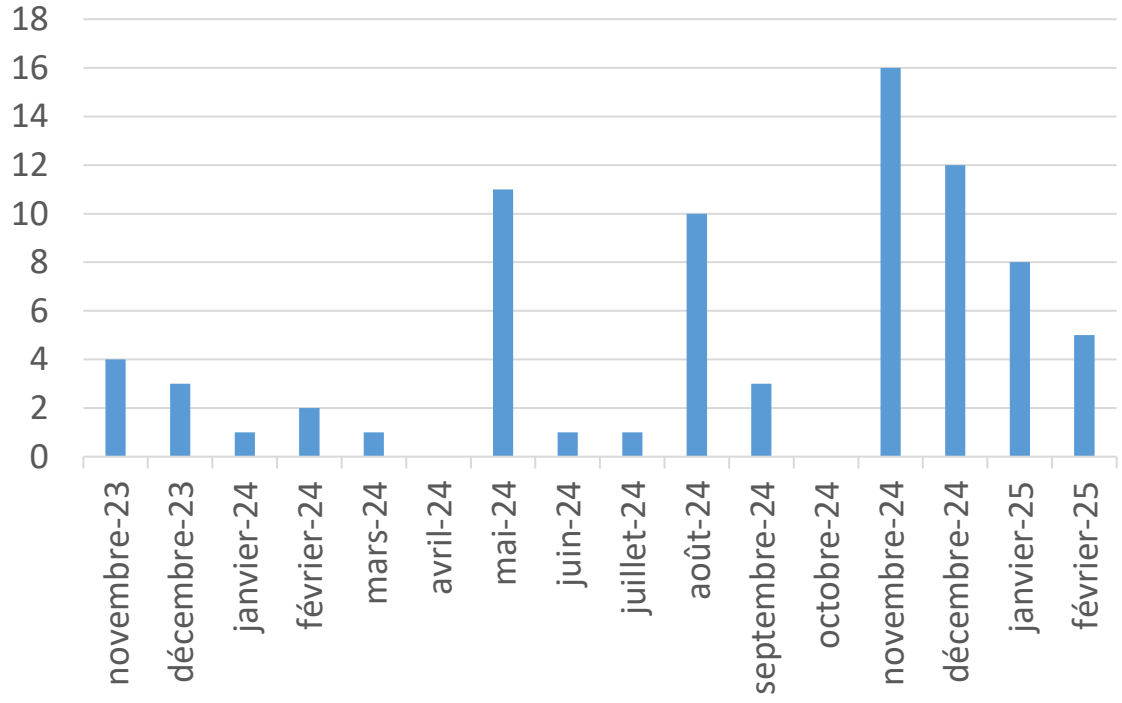
Cas pratique n°2

Contexte : 3 décès chez des nouveau-nés hospitalisés en soin intensif ayant présentés des chocs septiques fulminants à *Klebsiella pneumoniae* sur une période de 3 mois.

	QA1547	QA1548	QA2345
Identification (Hybridation ADN-ADN par TYGS) Score DDH Référence	*Klebsiella quasipneumoniae* subsp. *similipneumoniae* 92.8 % GCA_000613225	*Klebsiella pneumoniae* 92.9% GCA_000742135	*K. pneumoniae* 93% ATCC13883
MLST	Nouveau (*19bb)	5079	792
Analyse des SNPs (SNP-dists) - nombre de gènes (core-genome)	Non applicable	Non applicable	Non applicable
Locus O	O3/O3a	O1ab	O3/O3a
Locus K	**K1**	KL64 (K64)	KL25 (K25)
Gènes de virulence	Aucun	Aucun	**Yersiniabactine** **Colibactine** ***rmpA2***
Score de virulence	0	0	**2**
Gènes de résistance	OKP-B-36	SHV-1	SHV-1
Phénotype de résistance	Phénotype sauvage	Phénotype sauvage	Phénotype sauvage

Cas pratique n°3

- Identification de cas d'ERG au sein de 3 services de médecine conventionnelle d'un même établissement.
- Ces 3 services sont contigus et situés sur un **même étage** avec un personnel paramédical pouvant ponctuellement intervenir d'une unité à l'autre.



Répartition des cas au sein des 3 unités en fonction du temps

Détail de l'épidémie par Unité : souches

Résultats de l'analyse des souches après séquençage génome entier

Souche	U1	U2	U3	Total
VanA (ST1469)	- 3 patients -Acquisition nosocomial e : 3/3 patients	- 8 patients - 1 Porteur connu : Patient 13 - Acquisition nosocomiale : 7/8 patients	–	11
VanA (ST483)	–	–	- 1 patient - 1 Porteur connu : Patient 21	1
VanB (ST152)	-4 patients - 1 Porteur connu : Patient 2 -Acquisition nosocomiale : 3/4 patients	- 1 patient - 1 Porteur connu : Patient 14	- 6 patients - Acquisition nosocomiale : 6/6 patients	11
Total	7	9	7	23

Epidémie associée aux soins suspectée : décision d'une investigation

Enquête environnementale

Méthodes :

- Décembre 2024 : 84 prélèvements répartis dans les 3 unités
- Laboratoire Hygiène : culture + antibiogramme + PCR
- Souches séquencées pour comparaison génomique

Résultats :

- U1 : écouvillon d'une bassine (chambre 83) était positif à **ERG VanB ST152.**
- U2 : écouvillon d'un ordinateur du bureau infirmier était positif à **ERG VanA ST1469.**
- U3 : écouvillon d'une bassine (chambre 11) était positif à **ERG VanB ST152,** écouvillon d'un brassard de tension partagé entre les patients (positionné dans le couloir) était positif à ERG **VanA ST1469.**

Unité	Chambre simple	Chambre double
U1	11	8
U2	2	12
U3	18	3

Description des unités

Chaque chambre est équipée de sanitaires comprenant un lavabo, un WC, une douche, ainsi qu'une bassine destinée à la toilette des patients alités

Résumé chronologique



Figure 8. Répartition chronologique des chambres ayant accueilli des patients porteurs de souches cliniques d'ERG présentant ≤ 20 SNP de différence avec la souche environnementale (tensiomètre, bureau infirmier).

Les patients peuvent occuper plusieurs chambres différentes durant leur séjour.

Résumé chronologique

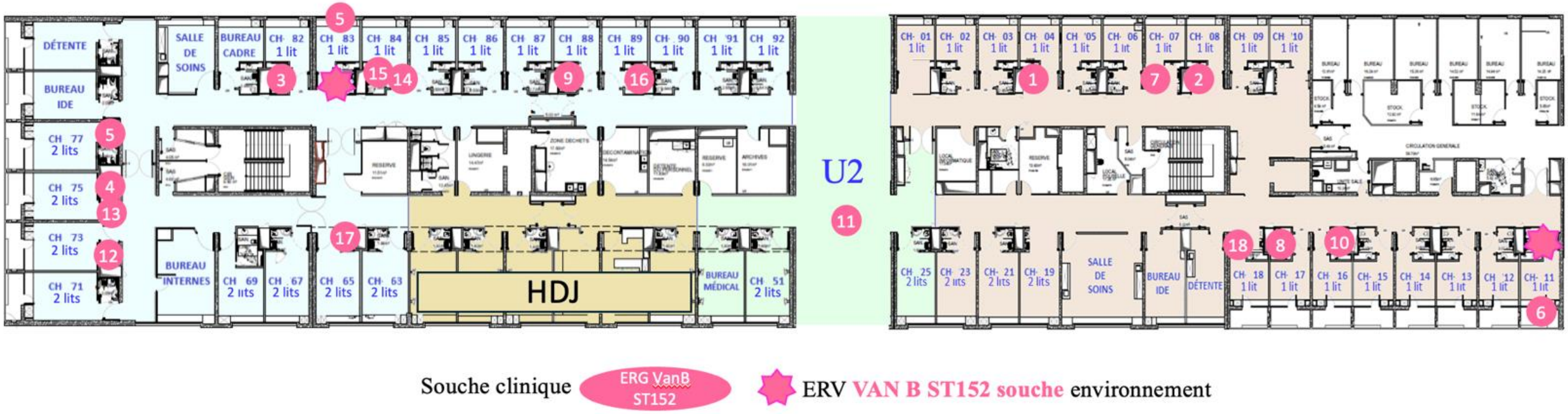


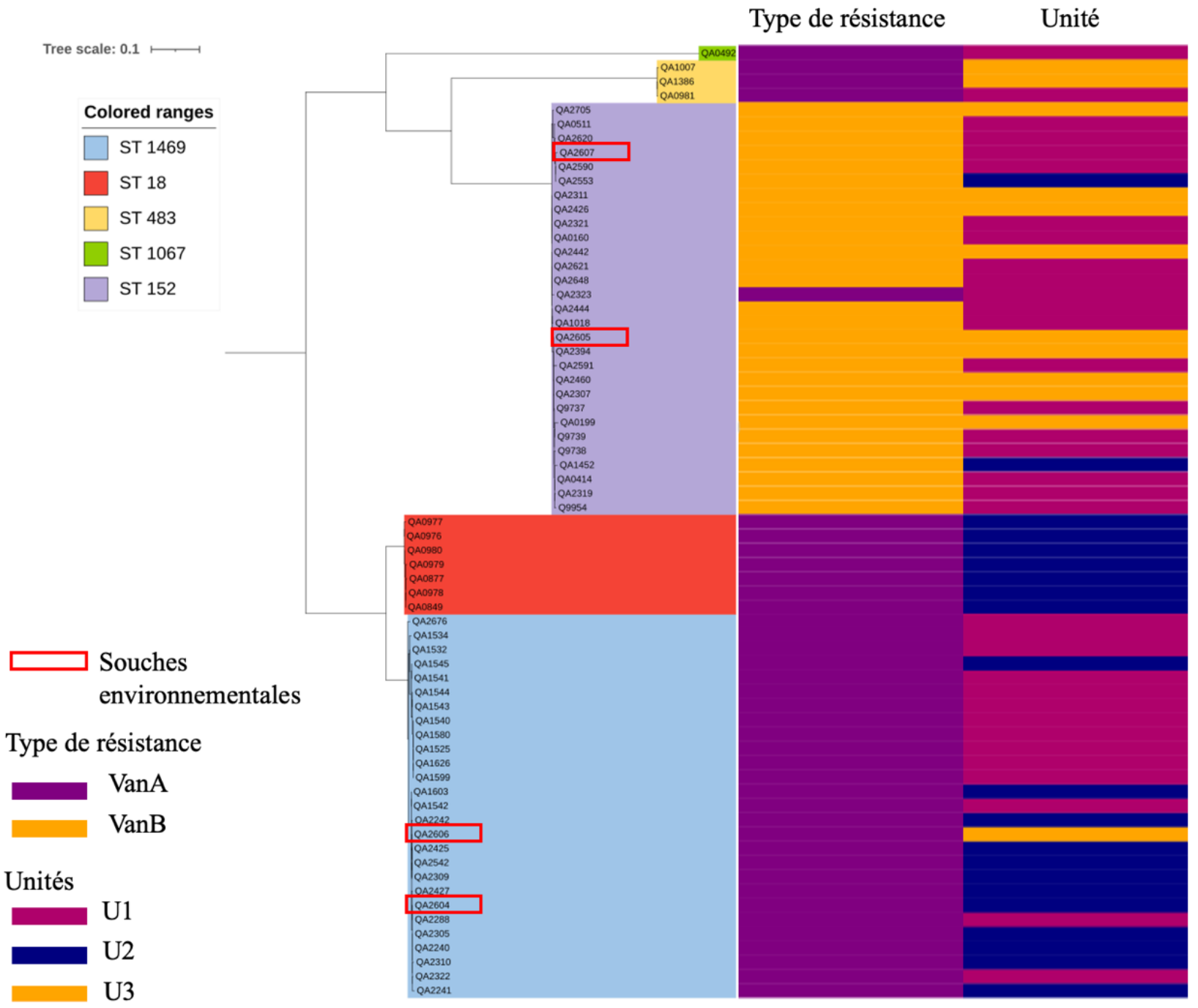
Figure 9. Répartition chronologique des chambres ayant accueilli des patients porteurs de souches cliniques d'ERG présentant ≤ 20 SNP de différence avec la souche environnementale (bassines chambres 11 U3 et 83 U1).

Les patients peuvent occuper plusieurs chambres différentes durant leur séjour.

Cas n°3

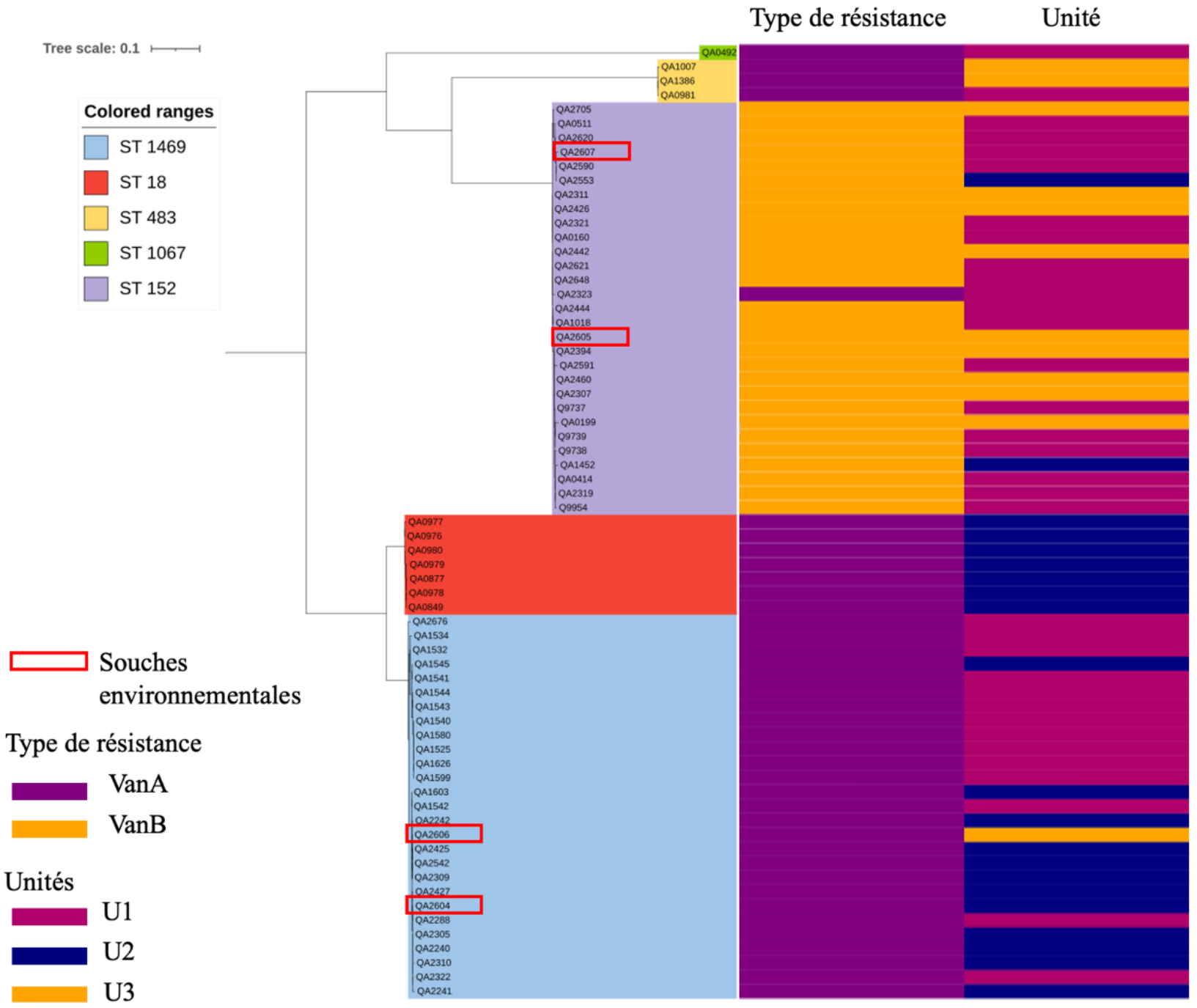
- **27 isolats VanA ST1469 :**
  - **Diversité génétique moyenne =** 18 SNPs de différence (intervalle : 0-309).

- Souche environnementale **QA2606 (tensiomètre de l'U3)** ≤20 SNPs avec 25 isolats cliniques et avec l'isolat environnemental **QA2604 (bureau infirmier de l'U2)**.



Cas n°3

- 29 souches **VanB ST152 :**
- diversité génomique moyenne = 4 SNPs (intervalle : 0 -11).
- Souche environnementale **QA2605**, (**bassine chambre U3)** ≤20 SNPs avec 27 isolats cliniques et avec l'isolat environnemental **QA2607 (bassine chambre U1)**



Cas pratique n°3

Diffusion intra- et inter-unités des ERGs

Réservoir environnemental à l'origine de la persistance des ERGs dans les trois services

- Brassard à tension
- Bureau infirmier

→ **Désinfection insuffisante du matériel**

Transmission manuportée par le personnel (non prélevé dans cette enquête)

→ **Hygiène des mains insuffisante**

Rôle de l'enquête : Educatif ++

CONCLUSION

1. Le typage **exclut** bien mais il **confirme** mal.
2. Identification, lignée (ST) et parenté fine sont trois échelles différentes.
3. Aucun seuil n'est universel : il dépend de l'espèce, du schéma, du pipeline et de la durée.
4. Sans isolats témoins, un chiffre de distance n'est pas interprétable.
5. Le génome cœur ne voit pas les plasmides : pensez-y dès que les espèces diffèrent.
6. Une conclusion solide tient sur trois points essentiels : génomique, spatial, temporel.

Remerciements

- **Equipe OHH :**
  - Paul-Rémi Petit,
  - Gwendoline Ragonnet
  - Rémi Charrel
- **Laboratoire Hygiène environnementale :**
  - Fréderique Gouriet
- **Externe en pharmacie:**
  - Marie-Alizée Borragini,
  - Nolwenn Marie
  - Manon Piat

MERCI POUR VOTRE ATTENTION