

# Contribution de l'analyse moléculaire à l'investigation d'un cluster de Covid-19 en long séjour gériatrique

Céline Sakr<sup>1</sup>, Audrey Maurand<sup>1</sup>, Mélanie Mercier-Darty<sup>1,2</sup>, Olivier Henry<sup>3</sup>, Slim Fourati<sup>4</sup>, Jean-Winoc Decousser<sup>1,5</sup>

1- Équipe opérationnelle d'hygiène – Département de prévention, diagnostic et traitement des infections – 2- Plateforme de caractérisation des génomes microbiens par séquençage à haut débit – 3- Département médico-universitaire de gériatrie – 4- Laboratoire de virologie – Département de prévention, diagnostic et traitement des infections

Hôpitaux universitaires Henri-Mondor – Assistance publique-Hôpitaux de Paris – Créteil – France

5- Equipe d'accueil 7380 Dynamic – Faculté de santé – Université Paris-Est Créteil – Créteil – France

✉ **Pr Jean-Winoc Decousser** – EA 7380 Dynamic – Faculté de santé – Université Paris-Est Créteil – 8, rue du Général-Sarrail – 94000 Créteil – France – E-mail: jean-winoc.decousser@aphp.fr

## RÉSUMÉ

**Contexte et objectif.** La lutte contre la Covid-19 nécessite l'identification des événements et acteurs de transmission. L'objectif de cette étude est de montrer la complémentarité de l'approche épidémiologique classique et de l'approche moléculaire par séquençage et comparaison des souches de Sars-CoV-2. Les performances et limites de ces méthodes sont évaluées. **Méthodes.** L'investigation épidémiologique classique d'un cluster en long séjour gériatrique a été réalisée en se focalisant sur la disponibilité des données. En parallèle, les souches de Sars-CoV-2 retrouvées chez les patients ont été séquencées, caractérisées phylogénétiquement et comparées. **Résultats.** Le cluster a impliqué 29 patients et 6 soignants entre le 26/12/2020 et le 25/01/2021 dans un service de long séjour gériatrique composé de deux unités. L'étude épidémiologique a montré une contamination secondaire des soignants auprès des patients et l'indisponibilité de certaines informations nécessaires à la compréhension des transmissions (entraide des soignants entre unités, interventions des médecins de garde, plans de table des repas partagés...). Les 18 séquences analysables ont mis en évidence une souche dérivant de la souche ancestrale Wuhan, et 4 groupes de séquences identiques entre elles séparées par une à deux différences nucléotidiques. L'apparition de mutations a permis d'identifier et d'exclure des voies de transmission (deux voisins de chambre présentant deux souches différentes). Une souche très divergente (25 différences) a été identifiée en fin d'épidémie au sein du cluster, soulignant le risque de la multiplication des séries de dépistage. **Conclusion.** Une stratégie combinant une investigation épidémiologique basique et une comparaison moléculaire des souches permettrait d'économiser les forces des hygiénistes.

## MOTS-CLÉS

Épidémie – Sars-CoV-2 – Séquençage – Comparaison – Gériatrie – Soins longue durée.

## ABSTRACT

### **Contribution of molecular analysis to the investigation of a Covid-19 cluster in a long-stay geriatric unit**

**Background and objective.** The fight against Covid-19 requires the identification of transmission-related events and actors. This study aimed to demonstrate the complementarity of conventional epidemiological approaches and modern approaches involving the sequencing and comparison of Sars-CoV-2 strains. These methods are assessed for limitations and performance. **Methods.** A conventional epidemiological investigation of a cluster was undertaken in a long-stay geriatric unit, focusing on data availability. In parallel, the Sars-CoV-2 strains detected in the patients were sequenced, characterised from a phylogenetic point of view and compared. **Findings.** The cluster identified between 26 December 2020 and 25 January 2021 concerned 29 patients and 6 carers, all from the same long-stay geriatric unit that counted two wards. The epidemiological study showed a secondary contamination of staff caring for patients and the lack of availability of certain data to explain transmission modes (staff from different wards helping each other, intervention of on-call physicians, dining room seating plans...). The 18 analysable sequences revealed the presence of a strain derived from the ancestral Wuhan strain, and four groups of identical sequences separated by one to two nucleotide differences. The apparition of mutations allowed the identification and exclusion of two transmission routes (two patients from the same room presenting with two different strains). One highly divergent strain (25 differences) was identified within the cluster towards the end of the outbreak, thus highlighting the risk inherent to the multiplication of screening series. **Conclusions.** A strategy combining basic epidemiological investigations with molecular comparison of strains could help spare the energy of infection control staff.

## KEYWORDS

Disease outbreak – Sars-Cov-2 – Sequencing – Comparison – Geriatrics – Long-term care.

## Introduction

Depuis la fin de l'année 2019, le virus Sars-CoV-2<sup>1</sup> s'est répandu en Chine puis sur l'ensemble de la planète. Lors des différentes vagues de cette pandémie qui a touché la France à partir de mars 2020, les établissements prenant en charge les personnes âgées ont fait face à de nombreux foyers de contamination (clusters) impliquant à la fois des patients et des soignants : les établissements hébergeant les personnes âgées dépendantes et les établissements hospitaliers de moyen et de long séjour gériatrique (soins de longue durée, soins de suite et de réadaptation) ont ainsi été particulièrement touchés. Ces unités de soins hébergent des patients vulnérables et fragiles, souvent dépendants et parfois déambulants. La présence de troubles cognitifs rend difficile la détection des symptômes et le respect de mesures barrières au sein d'un lieu de vie collectif [1]. Différents acteurs interviennent pour la prise en charge de ces patients : personnel soignant avec une prédominance d'aides-soignants, diététiciens, ergothérapeutes, animateurs... La charge de travail est souvent élevée, notamment la nuit et le week-end [2,3]. La lutte contre la propagation du virus dans ces unités nécessite la parfaite connaissance des modes de transmission et l'identification des événements et acteurs de cette transmission, ceci afin de mettre en place les mesures nécessaires à la maîtrise de la diffusion sans atteinte inutile à la qualité de vie des patients. Cette identification des modes de transmission se base sur deux approches complémentaires : l'approche épidémiologique, qui vise à identifier les cas positifs ainsi que leur répartition dans le temps et dans l'espace afin d'émettre des hypothèses sur la transmission virale, et l'approche par analyse phylogénique et comparaison moléculaire des séquences virales [4]. Cette dernière peut néanmoins devenir une alternative de première intention à l'approche épidémiologique : cette stratégie, dite « *sequence first*<sup>2</sup> », est basée sur l'analyse et la comparaison systématique sans a priori des souches virales [5]. L'objectif de cette étude est d'évaluer la complémentarité de ces deux approches dans le cadre de l'investigation d'un cluster en long séjour gériatrique. Les critères évalués sont la disponibilité des données, l'exactitude des informations recueillies mais aussi le temps et le coût de leur collecte.

## Matériels et méthodes

### Approche épidémiologique classique

#### Description temporo-spatiale de l'épidémie

L'épidémie s'est déroulée dans une unité de long séjour gériatrique qui hébergeait 33 patients fragiles présentant des pathologies chroniques évolutives nécessitant des soins médicaux et une surveillance 24 h/24. Ce service est formé de deux ailes, d'un local infirmier,

de deux salles à manger communes aux résidents et d'un local lave-bassin. La première aile est composée de 7 chambres individuelles et de 5 chambres doubles, la seconde, de 7 chambres individuelles et de 6 chambres doubles. L'étude épidémiologique de ce cluster a consisté à retracer rétrospectivement le parcours de 26 patients dans le service afin de détecter le cas index et les événements de transmission au cours du temps. Un tableau synoptique a été établi, représentant pour chaque patient et soignant, symptomatique ou non, la durée d'incubation et de contagiosité, la date des premiers symptômes et la date de prélèvement. La courbe épidémique a été tracée. La répartition géographique des patients infectés a été reportée.

### Identification des paramètres et acteurs impliqués dans la transmission du virus

Les déterminants individuels, collectifs, organisationnels et fonctionnels impliqués dans la transmission du virus au sein des collectivités de personnes âgées ont été recherchés dans la littérature. Une recherche de ces déterminants au sein des bases de données disponibles dans l'établissement a été réalisée à l'aide de l'examen des dossiers médicaux électronique et physique des patients, ainsi qu'au travers d'entretiens informels avec les différents acteurs, dont les cadres responsables du service. Les données ont été considérées comme « non disponibles » si elles n'existaient pas, si leur localisation n'était pas établie, ou si plusieurs demandes successives auprès de la personne ou de la structure en disposant restaient sans réponse.

### Approche d'analyse et de comparaison phylogénétique

#### Comparaison des souches virales responsables d'infection par séquençage du génome complet

Les isolats de l'épidémie ont été entièrement séquencés et comparés à l'aide de la reconstruction phylogénique moléculaire. Le séquençage des souches issues des prélèvements nasopharyngés a été réalisé par la technique Illumina® COVIDSeq™ Test sur automate NextSeq® 500 Sequencing System (Illumina Inc., San Diego, États-Unis). Les données ont été analysées avec le pipeline Dragen COVIDSeq, comprenant un alignement sur référence NC\_0456512, Wuhan-Hu-1. La lignée est identifiée via le programme Pangolin<sup>3</sup> [6] (*open source*). Les séquences ont été assemblées en génome complet et comparées à l'aide du programme Muscle<sup>4</sup> du logiciel MEGA<sup>5</sup> X (Penn-

3- *Phylogenetic assignment of named global outbreak lineages*: affectation phylogénétique de la lignée d'épidémie mondiale citée (<https://virological.org/t/pangolin-web-application-release/482>).

4- *Multiple sequence comparison by log-expectation*: comparaison de plusieurs séquences par espérance logarithmique.

5- *Molecular evolutionary genetics analysis*: analyse de génétique évolutive moléculaire.

1- *Severe acute respiratory syndrome coronavirus 2*, coronavirus 2 du syndrome respiratoire aigu sévère.

2- Séquençer en premier.

sylvania State University, State College, États-Unis) [7]. Cette analyse a été réalisée avec l'aide technique de l'ingénieure hospitalière de la plateforme qui est en partie intégrée à l'équipe opérationnelle d'hygiène (EOH).

## Résultats

### Description temporo-spatiale de l'épidémie

Entre le 26 décembre 2020 et le 25 janvier 2021, 29 patients ont été considérés comme infectés par le virus : 26 patients ont été testés positifs au Sars-CoV-2 par test RT-PCR<sup>6</sup> et trois autres patients ont été testés négatifs par RT-PCR mais ont été considérés comme associés au cluster selon des critères cliniques et temporo-spatiaux (Tableau I). Parmi les soignants du service, six ont été testés positifs de façon concomitante au cluster des

patients (quatre aides-soignants et deux infirmières). Le chronogramme regroupant les 26 patients et les six soignants du service plaide en faveur d'une contamination secondaire des soignants auprès des patients (Figure 1). La courbe épidémique représente l'apparition des nouveaux cas symptomatiques et asymptomatiques au cours du temps, détectés lors de la réalisation des campagnes de dépistage (les 28 décembre et 4, 11, et 18 janvier) (Figure 2).

### Disponibilité des paramètres impliqués dans la transmission du virus

Plusieurs facteurs et acteurs ont probablement joué un rôle dans la transmission virale mais la disponibilité de ces paramètres dans les bases de données des résidents n'est pas toujours évidente. Concernant les patients, nous avons pu retrouver leurs caractéristiques individuelles, leur date d'entrée et de sortie du service, la date de leur test RT-PCR positif, leur niveau de dépendance,

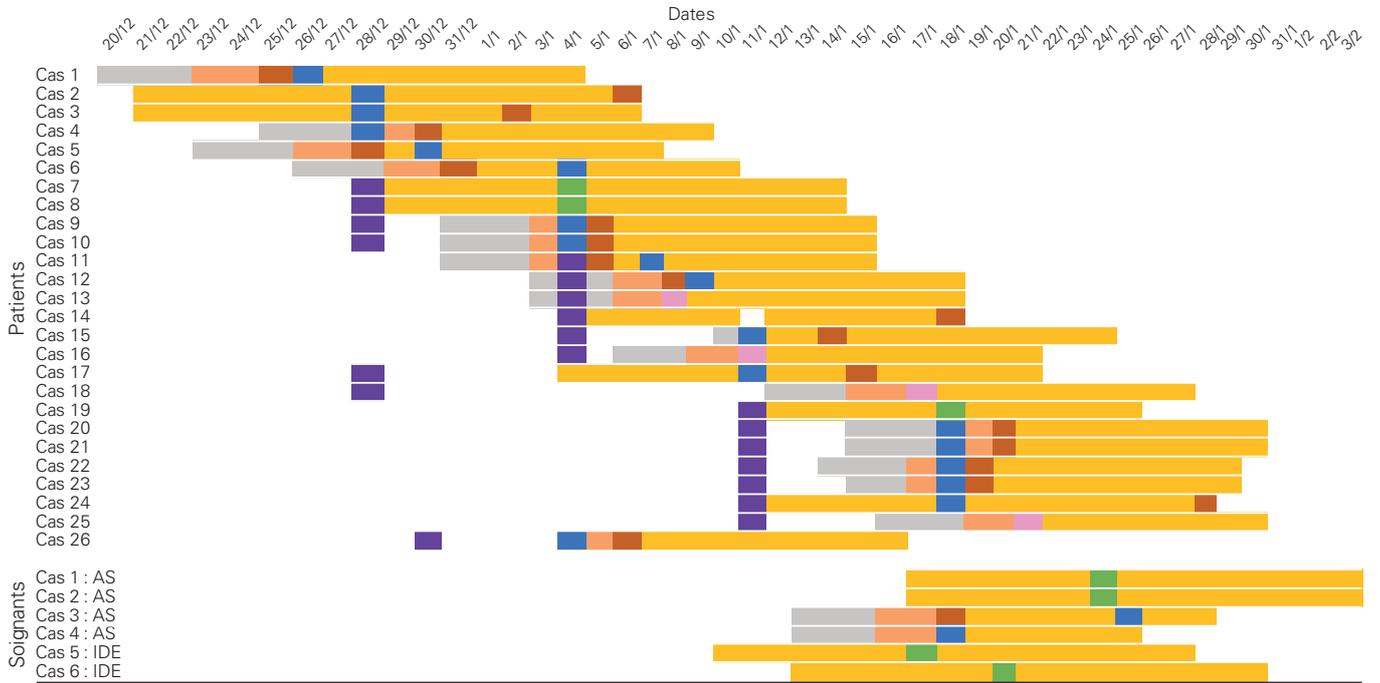
6- Reverse transcription-polymerase chain reaction : transcriptase inverse-réaction de polymérisation en chaîne.

**Tableau I – Résultats des tests RT-PCR et du typage des virus des patients dans le cadre d'un cluster en long séjour gériatrique.**

Patient	Sexe	Date du test PCR	Résultat du test PCR	Résultat du typage
Cas 1	F	26/12/2020	Positive	SEQi
Cas 2	F	28/12/2020	Positive	Séquencé mais critère de qualité insuffisant
Cas 3	F	28/12/2020	Positive	SEQi
Cas 4	M	28/12/2020	Positive	SEQi+ c1372t
Cas 5	F	30/12/2020	Positive	SEQi
Cas 6	M	04/01/2021	Positive	SEQi
Cas 7	F	04/01/2021	Positive	SEQi+g2180a
Cas 8	M	04/01/2021	Positive	SEQi
Cas 9	M	04/01/2021	Positive	SEQi+c22347t
Cas 10	F	04/01/2021	Positive	SEQi+ c1372t
Cas 11	F	07/01/2021	Positive	Non séquencé non typé
Cas 12	F	09/01/2021	Positive	Non séquencé non typé
Cas 13	M	08/01/2021	Positive	Non séquencé non typé
Cas 14	F	11/01/2021	Positive	SEQi+c22347t
Cas 15	M	11/01/2021	Positive	Séquencé mais critère de qualité insuffisant
Cas 16	F	11/01/2021	Positive	SEQi+c1299t+a17566g
Cas 17	F	11/01/2021	Positive	SEQi+ c1372t
Cas 18	M	17/01/2021	Positive	SEQi+ c1372t
Cas 19	M	18/01/2021	Positive	SEQi+ c1299t + c15911t
Cas 20	F	18/01/2021	Positive	Séquencé mais critère de qualité insuffisant
Cas 21	M	18/01/2021	Positive	SEQi+ c1372t+c13848t
Cas 22	M	18/01/2021	Positive	SEQi+ c1372t+ c13848t
Cas 23	F	18/01/2021	Positive	Séquencé mais critère de qualité insuffisant
Cas 24	F	18/01/2021	Positive	SEQi+g2180a+a28428g
Cas 25	F	21/01/2021	Positive	SEQi+25 SNP
Cas 26	F	04/01/2021	Positive	Non séquencé non typé
Cas 27	F	25/01/2021	Négative	Non fait ( PCR négative)
Cas 28	M	04/01/2021	Négative	Non fait ( PCR négative)
Cas 29	F	11/01/2021	Négative	Non fait ( PCR négative)

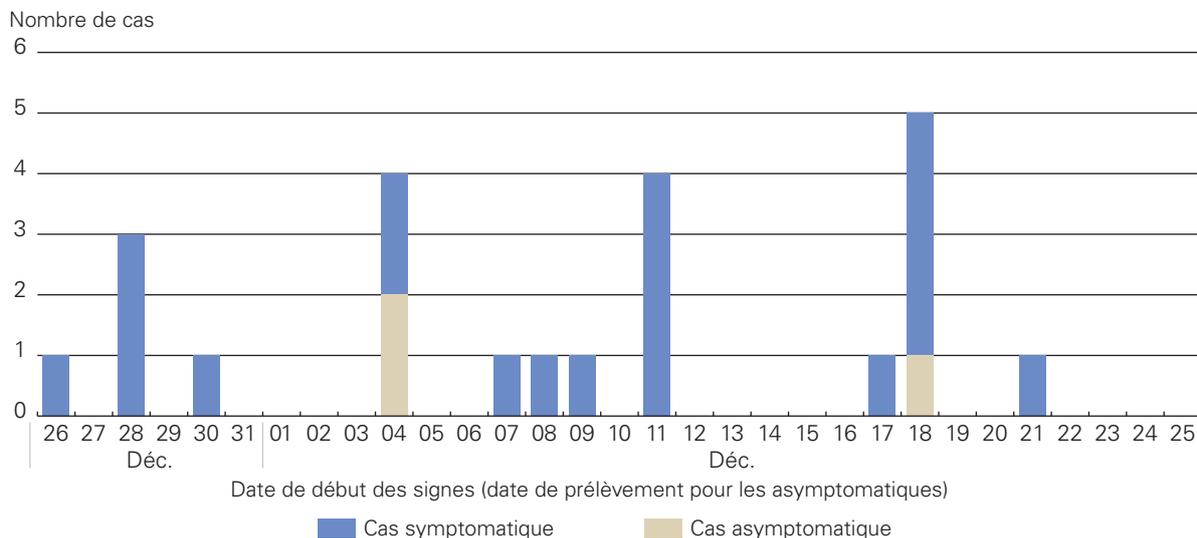
ARN : Acide ribonucléique ; RT-PCR : Reverse transcription-polymerase chain reaction, transcriptase inverse-réaction de polymérisation en chaîne ; SEQi : séquence index avec 25 SNP comparé avec la souche Wuhan-Hu-1 ; SNP : single nucleotide polymorphism, polymorphisme d'un seul nucléotide ; mutation.

**Figure 1 – Tableau synoptique et chronogramme des 26 patients et des 6 soignants positifs au Sars-CoV-2 dans le cadre d'un cluster en long séjour gériatrique.**



AS : aide-soignant(e) ; IDE : infirmier(e) diplômé(e) d'État ; RT-PCR : *Reverse transcription-polymerase chain reaction*, transcriptase inverse-réaction de polymérisation en chaîne ; Sars-CoV-2 : *Severe acute respiratory syndrome coronavirus 2*, coronavirus 2 du syndrome respiratoire aigu sévère.

**Figure 2 – Courbe épidémique représentant la détection des cas symptomatiques et asymptomatiques au cours du temps dans le cadre d'un cluster en long séjour gériatrique.**



la présence de troubles cognitifs ainsi que leur type, et la présence de déambulation à l'intérieur ou à l'extérieur du service. Concernant les soignants, leurs horaires de travail (médecins, infirmières, aides-soignants, psychologues, diététiciennes, kinésithérapeutes et intérimaires...) ainsi que leurs mobilités entre services ont été recherchés; les données concernant la mobilité des soignants entre les unités, services et étages dans le cadre d'une entraide ne sont pas tracées. Les données concernant les intérimaires et les remplacements sont plus difficiles à collecter car les tableaux obtenus correspondent au planning prévisionnel et ne sont pas corrigés en fonction des présences effectives et des absences non programmées. Les données concernant les visites de diététicienne et de kinésithérapeute dans les chambres des patients ont été recueillies mais celles des consultations d'ergothérapeute et de psychologue n'étaient pas disponibles. Le week-end, il n'y avait pas de traçabilité des interventions du médecin de garde dans les différents secteurs. Concernant le dépistage, le personnel médical était moins observant quant aux recommandations et à la réalisation des tests, avec une inconnue quant à l'exhaustivité des tests demandés et effectivement réalisés. Concernant les visiteurs, la traçabilité a été retrouvée, à l'exception de la réalisation et des résultats de leur dépistage avant les visites. Les données relatives à la géographie du service, la présence d'une salle à manger commune, l'emplacement des chambres, leur type (simple ou double) ainsi que la présence d'un voisin de chambre infecté ont été relevés. La participation à des activités communes où les patients auraient été regroupés n'était pas tracée. Il était possible de retrouver en recherchant patient par patient les consultations internes

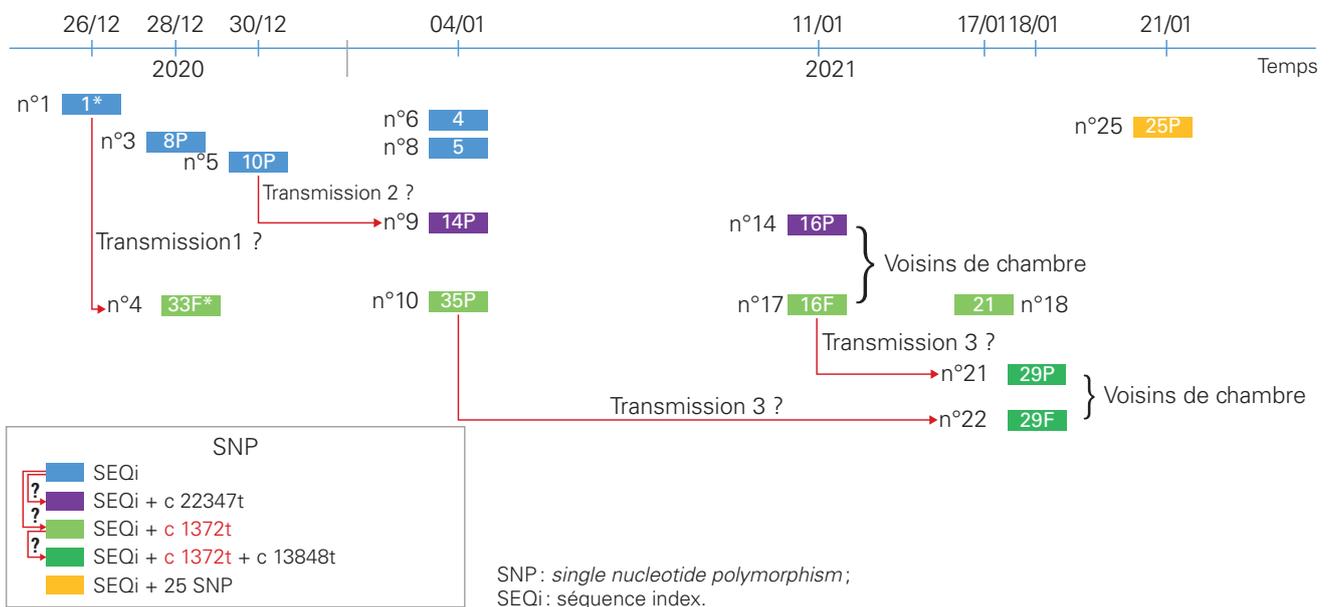
et externes, mais les conditions de ces sorties du service n'étaient pas précisées et auraient nécessité une recherche beaucoup plus fine (opérateurs hors médecin signataire, brancardiers) voire impossible (ambulanciers, personnes rencontrées sur un site extrahospitalier...).

### Séquençage et comparaison des souches impliquées dans l'épidémie

Parmi les prélèvements des 26 patients testés positifs, 22 ont pu être séquencés et 18 ont atteint les critères de qualité permettant une interprétation des séquences. Les souches impliquées étaient toutes proches de la souche de référence de Wuhan-Hu-1 (NC\_045512.2), avec 25 mutations (SNP<sup>7</sup>) de différence par rapport à la souche historique. Les résultats du séquençage viral ont également montré la présence de quatre groupes de patients qui partageaient la même séquence Sars-CoV-2 et qui présentaient un ou deux SNP de différence entre les groupes. De plus un patient a été infecté par une souche très divergente (25 SNP par rapport à la souche du patient n° 1). Il est intéressant de noter que ce patient a été dépisté tardivement par rapport aux autres événements de transmission, lors de la dernière série hebdomadaire de dépistage des patients exposés. Un groupe de six patients a été infecté par un virus avec la même séquence, ainsi que deux groupes de deux patients et un groupe de quatre patients. Une analyse génétique couplée d'une analyse temporelle et spatiale a été réalisée afin d'appréhender le dynamisme de la diffusion (Figure 3). Les prélèvements

7- Single nucleotide polymorphisms, polymorphisme d'un seul nucléotide: variation d'une seule paire de bases du génome entre individus d'une même espèce (d'après Wikipédia).

Figure 3 – Données génétiques, temporelles et spatiales des patients infectés dans le cadre d'un cluster en long séjour gériatrique.



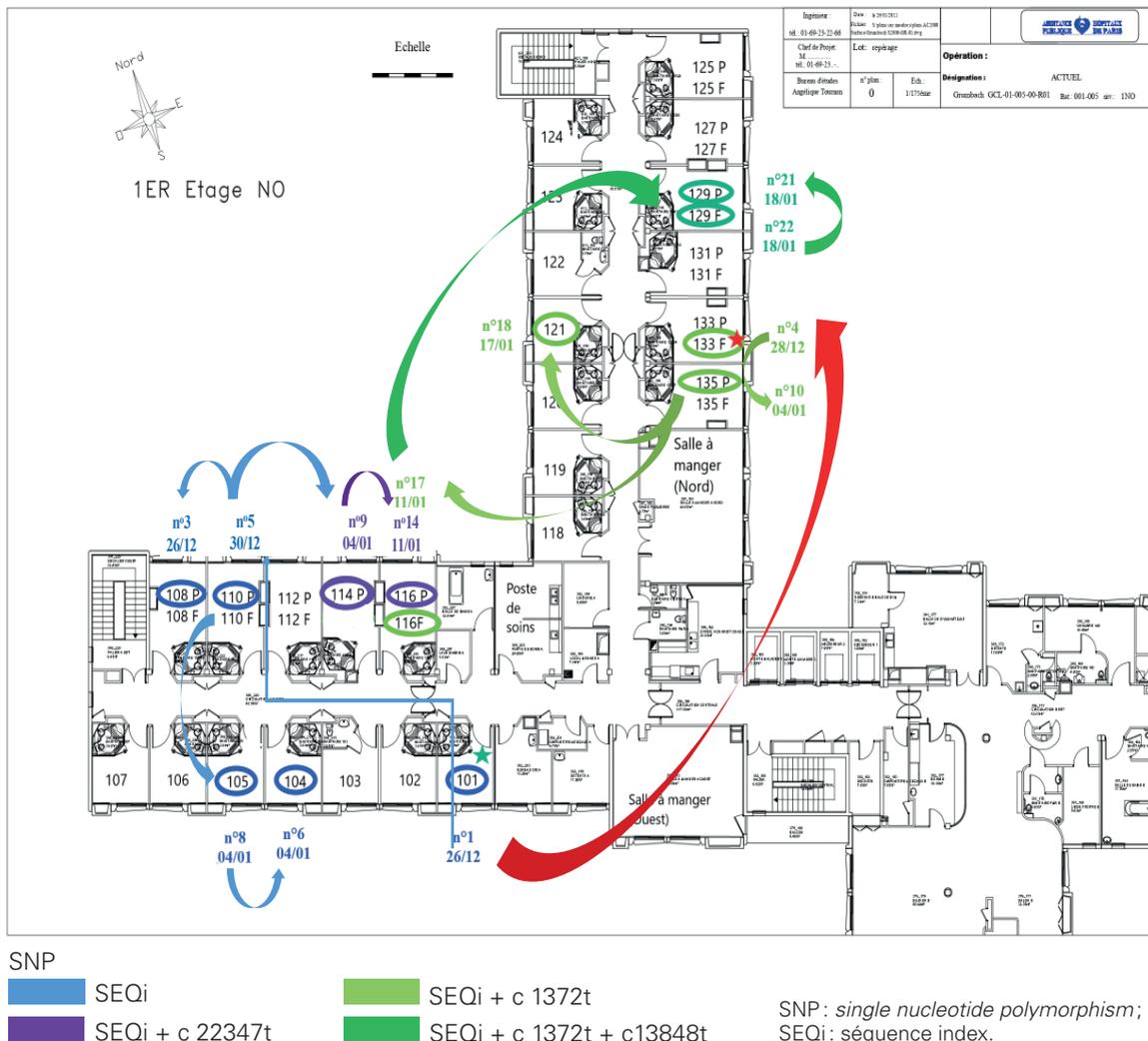
des six soignants testés positifs en fin d'épidémie n'ont pas été pris en considération dans les analyses moléculaires. Le cas n° 1 a été considéré comme étant le cas index à l'origine de l'épidémie puisqu'il s'agissait d'un résident qui sortait deux à trois fois par semaine pour une dialyse. Ce cas portait la même séquence que les cas n° 3, 5, 6 et 8. Cette série de cas pourrait être à l'origine d'une transmission aux cas n° 9 et 14, qui partageaient la même séquence présentant une mutation (C22347T) par rapport à la souche index. Par ailleurs, elle pourrait aussi être responsable de la transmission aux cas n° 4, 10, 17 et 18 qui partagent une mutation différente de celle des cas n° 9 et 14. La comparaison moléculaire des cas n° 21 et 22 a révélé la même mutation que celle retrouvée chez les cas n° 4, 10, 17 et 18 (C1372T) accompagnée d'une mutation supplémentaire (C13848T). Ces deux patients étaient voisins de chambre et ont partagé une exposition commune ou ont échangé la même souche. L'apparition de cette nouvelle mutation (C13848T) atteste que cette souche a évolué très probablement à partir d'une des souches portant

déjà la mutation C1372T; il est en effet très peu probable que deux mutations apparaissent en même temps à partir de la souche index. L'analyse et les enseignements tirés de l'évolution des souches sont très intéressants car ils ajoutent une notion de temporalité aux informations sur la transmission du virus. Concernant les cas n° 14 et 17, voisins de chambre, l'enquête épidémiologique standard aurait suggéré à l'évidence que ces deux patients partageaient la même souche virale. L'analyse moléculaire a pourtant montré que chacune présentait une mutation différente par rapport au génome index. Il est intéressant de noter que, sans l'apport des techniques moléculaires, la transmission entre les cas n° 14 et 17 aurait été attribuée à une contamination entre voisins de chambre (Figure 4).

### Discussion

La pandémie de Covid-19 a constitué un événement exceptionnel par son ampleur, sa durée et sa mortalité notamment au sein des établissements de SLD. L'implication d'un nouveau virus dont les modes de

Figure 4 – Totalité des transmissions au sein du service 1A (long séjour gériatrique).



transmission font encore aujourd'hui débat, l'échec des mesures de prévention basiques au sein de ces établissements et la surcharge de travail phénoménale des EOH nous ont amenés à comparer deux stratégies.

### L'investigation épidémiologique

L'investigation épidémiologique historique « fine » est capable de mettre en évidence divers paramètres liés aux résidents (statut clinique, déambulation...), aux personnels soignants (intervenant transversaux), aux organisations de soins (absences, remplacements, aides la nuit et les week-ends) ou aux visiteurs qui jouent un rôle important dans la transmission virale. Certains éléments de traçabilité existent, mais sont difficiles d'accès donc coûteux en temps (planning des interventions des intérimaires) ou ne sont pas présentés sous un format permettant un accès rapide et transversal des données (exemple du listing des interventions du médecin de garde la nuit et le week-end). La formalisation et la traçabilité de l'ensemble des informations au sein d'un outil informatique unique facilement manipulable pourrait constituer une source de données intéressante. Ces entrepôts de données sont aujourd'hui disponibles mais de consultation complexe, non compatible avec l'analyse d'un phénomène épidémique durant quelques semaines. L'identification fine de l'intervention des soignants au sein des différents secteurs pourrait bénéficier d'un traçage par puces RFID<sup>8</sup> [8].

### L'approche *sequence first*

L'approche *sequence first* appliquée à la caractérisation et à la comparaison des souches par séquençage de leur génome [9] a fait l'objet de nombreuses publications très positives dans le domaine des épidémies de bactéries, qu'elles soient à l'échelon d'un hôpital, d'un groupement d'hôpitaux ou encore d'une région ou d'un pays entier [10]. Leur contribution dans le cadre des épidémies virales « banales » par opposition aux investigations spécifiques et ponctuelles (transmission de virus hémotogène) reste encore à préciser [11]. Cette approche a permis d'objectiver plusieurs événements de transmission au sein du service (cas n° 4, 10, 17 et 18) ou, au contraire, d'en écarter alors qu'ils étaient « épidémiologiquement évidents » comme une transmission suspectée entre voisins de chambre (patients n° 14 et 17). La succession dans le temps d'événements de transmission est aussi établie par l'accumulation de mutations (souches des patients n° 4, 10, 17 et 18 évoluant vers la souche des patients n° 21 et 22 par acquisition d'une mutation supplémentaire). La survenue de mutations différentes, certaines s'accumulant avec le temps, permet d'exclure certaines transmissions qui correspondraient à l'acquisition d'une nouvelle mutation et surtout à la réver-

sion d'une autre, ce qui est statistiquement hautement improbable (souche des patients n° 9 et 14 dérivant de la souche initiale mais ne pouvant provenir de la souche ayant déjà muté comme celle des patients n° 4, 10, 17, 19, 21 et 22). De plus, cette approche a permis d'objectiver un risque lié à la multiplication des séries de dépistages transversaux à distance de l'événement initial de contamination : en effet, la dernière souche de patient identifiée (cas n° 25) était génétiquement très éloignée des autres (25 différences). Or ce cas tardif contaminé en dehors de la chaîne de transmission étudiée peut remettre en question une gestion pourtant efficace du cluster initial, et prolonger les séries de dépistage au risque de détecter de nombreux cas non épidémiologiquement reliés. Ceci est particulièrement vrai en l'absence de l'analyse moléculaire systématique et lorsqu'il y a une forte circulation communautaire de l'agent pathogène impliqué. Cependant, cette approche présente également des limites : elle nécessite tout d'abord l'obtention de spécimens biologiques en quantité et en qualité suffisante. Ainsi sur les 29 patients concernés par l'épidémie, trois ont été testés négatifs par RT-PCR, quatre ne présentaient pas les critères de qualité suffisants pour être typés, et quatre séquences n'étaient pas exploitables. Néanmoins, l'objectif du travail n'était pas d'explorer entièrement le cluster mais d'évaluer les contributions respectives et les limites des deux approches. La disponibilité de la technologie sur le plateau technique du groupe hospitalier a été à l'évidence un avantage majeur pour tester l'approche du séquençage. Le coût, de quelques dizaines d'euros, n'est pas rédhibitoire ; le délai d'obtention de la séquence est d'environ une semaine à dix jours. Néanmoins la volonté des autorités de tutelle d'augmenter et d'accélérer le séquençage des prélèvements positifs de RT-PCR afin de suivre au plus près l'apparition et la diffusion des variants va dans le sens d'une démocratisation de cette technique et de la disponibilité des séquences [12]. Au-delà de la problématique technique, l'interprétation des données de séquençage nécessite des compétences en bio-informatique. Cette limite à la démocratisation de l'approche par séquençage peut être contournée par la mise en place de collaborations avec des professionnels ayant des compétences en bio-informatique ou le développement de ces compétences au sein des EOH. Nous avons décidé d'intégrer dans notre EOH un ingénieur hospitalier à mi-temps disposant de ces compétences afin de gérer l'ensemble des problématiques bio-informatiques liées à la comparaison des souches virales, bactériennes et fongiques. La sensibilité de la méthode de comparaison des génomes dépend quant à elle de la variabilité et de la diversité des souches circulantes. Lors de la deuxième vague de l'épidémie, la diversité des variants et leur évolution individuelle, notamment sous la pression immunitaire, a permis de disposer de données utilisables pour étudier des clusters [11]. Le taux de mutation, évalué à une ou deux mutations par mois,

8- *Radio frequency identification*: identification par radio-fréquence ou radio-identification.

est compatible avec les trois mutations qui sont apparues chez les souches infectant les patients de notre cluster [12]. Dans le cas présent, il aurait été intéressant de disposer des séquences s'étant diffusées dans le reste de l'hôpital, d'une part, et dans la région en dehors de notre hôpital, d'autre part, afin de s'assurer que leur diversité était suffisante pour avoir une spécificité et une sensibilité acceptables de la méthode [12, 13]. Une déclinaison prospective encore plus large de l'approche *sequence first* à partir de l'ensemble des séquences obtenues pour un service, un bâtiment ou un hôpital permettrait aux EOH de ne s'investir que dans des événements patents et établis de transmission croisée, à l'instar de ce qui a été fait pour certaines bactéries multirésistantes [2].

## Conclusion et perspectives

Le travail réalisé a mis en évidence la difficulté d'accès ou de traçabilité de nombreuses données démographiques susceptibles de jouer un rôle important dans la diffusion d'un micro-organisme au sein d'un service. Le travail que nous avons mené plaide pour une approche séquentielle pragmatique de l'investigation au sein des clusters viraux en période pandémique par les EOH. Devant une suspicion de cluster, nous proposons de

limiter l'intervention de l'EOH à un rappel général des mesures de lutte adaptées au type de transmission croisée associé au micro-organisme identifié. En parallèle du recueil des données épidémiologiques de base pour construire les tableaux synoptiques, le chronogramme, la courbe épidémique et la répartition géographique des cas, une collecte la plus exhaustive possible de prélèvements biologiques permettra de disposer d'un matériel génétique abondant et exhaustif pour la comparaison des génomes. Dans un second temps, l'étude fine des événements de transmission ne sera réalisée que pour les cas confirmés par les résultats de séquençage, en se focalisant sur les transmissions inattendues ou inexplicables afin de mettre en évidence des contextes particuliers, donc la nécessité de mesures spécifiques. Parmi les outils d'avenir dont l'hygiéniste disposera, l'utilisation et la manipulation des données de séquençage constitueront à l'évidence un moyen d'optimisation des forces vives de l'EOH qui devra sous-traiter ou, mieux, intégrer ces compétences. La place centrale et initiale de ces outils dans le contexte particulier de la pandémie de Covid-19 au sein d'une unité de long séjour gériatrique gagnerait à être évaluée dans d'autres contextes épidémiques, notamment viraux. ■

## Références

- 1- Mehta HB, Li S, Goodwin JS. Risk factors associated with Sars-CoV-2 infections, hospitalization, and mortality among US nursing home residents. *JAMA Netw Open* 2021;4(3):e216315.
- 2- Gardner W, States D, Bagley N. The coronavirus and the risks to the elderly in long-term care. *J Aging Soc Policy* 2020;32(4-5):310-315.
- 3- Ikitimur H, Borku Uysal B, Cengiz M, et al. Determining host factors contributing to disease severity in a family cluster of 29 hospitalized Sars-CoV-2 patients: could genetic factors be relevant in the clinical course of COVID-19? *J Med Virol* 2021 Jan;93(1):357-365.
- 4- Forster P, Forster L, Renfrew C, et al. Phylogenetic network analysis of Sars-CoV-2 genomes. *Proc Natl Acad Sci USA* 2020;117(17):9241-9243.
- 5- Peacock SJ, Parkhill J, Brown NM. Changing the paradigm for hospital outbreak detection by leading with genomic surveillance of nosocomial pathogens. *Microbiology (Reading)* 2018;164(10):1213-1219.
- 6- Rambaut A, Holmes EC, O'Toole Á, et al. A dynamic nomenclature proposal for Sars-CoV-2 lineages to assist genomic epidemiology. *Nat Microbiol* 2020;5(11):1403-1407.
- 7- Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Mol Biol Evol* 2016;33(7):1870-1874.
- 8- Brouqui P, Boudjema S, Soto Aladro A, et al. New approaches to prevent healthcare-associated infection. *Clin Infect Dis* 2017;65(suppl\_1):S50-S54.
- 9- Løvestad AH, Jørgensen SB, Handal N, et al. Investigation of intra-hospital Sars-CoV-2 transmission using nanopore whole-genome sequencing. *J Hosp Infect* 2021;111:107-116.
- 10- Sundermann AJ, Babiker A, Marsh JW, et al. Outbreak of vancomycin-resistant *Enterococcus faecium* in interventional radiology: detection through whole-genome sequencing-based surveillance. *Clin Infect Dis* 2020;70(11):2336-2343.
- 11- Esper FP, Cheng YW, Adhikari TM, et al. Genomic epidemiology of Sars-CoV-2 infection during the initial pandemic wave and association with disease severity. *JAMA Netw Open* 2021;4(4):e217746.
- 12- Feder KA, Pearlowitz M, Goode A, et al. Linked clusters of Sars-CoV-2 variant B.1.351 – Maryland, January-February 2021. *MMWR Morb Mortal Wkly Rep* 2021;70(17):627-631.
- 13- Braun KM, Moreno GK, Buys A, et al. Viral sequencing reveals US healthcare personnel rarely become infected with Sars-CoV-2 through patient contact. *Clin Infect Dis* 2021;ciab281.

### Citation

Sakr C, Maurand A, Mercier-Darty M, Henry O, Fourati S, Decusser JW. Contribution de l'analyse moléculaire à l'investigation d'un cluster de Covid-19 en long séjour gériatrique. *Hygiènes* 2021;29(4):261-268.

### Historique

Reçu 16 juillet 2021 – Accepté 28 août 2021 – Publié 30 septembre 2021

**Remerciements:** Nous tenons à remercier l'ensemble du personnel soignant et administratif ayant participé à cette étude.

**Financement:** le groupe hospitalier Henri-Mondor a financé le poste de l'étudiant en Master 2 ayant permis de réaliser cette étude.

**Liens d'intérêt:** les auteurs déclarent ne pas avoir de lien d'intérêt.