NNT/NL : 0000AIXM0000/000ED000

THÈSE DE DOCTORAT

Soutenue à Aix-Marseille Université

Le 30 juin 2022 par

Lanceï Kaba

Place de la diversité bactérienne et de sa dynamique dans la surveillance épidémiologique

Discipline
Biologie Santé

Spécialité
Maladies Infectieuses

École doctorale
ED 62 - Sciences de la Vie et de la Santé

Laboratoire/Partenaires de recherche
Institut Hospitalo-Universitaire, Méditerranée Infection
Vecteurs – Infections Tropicales et Méditerranéennes (VITROME)
Institut de Recherche pour le Développement (IRD)
Assistance Publique - Hôpitaux de Marseille (AP-HM)

Composition du jury

Raymond RUIMY Université de Nice	Rapporteur
Alpha Kabinet KEITA Université de Conakry (UGAN)	Rapporteur
Florence FENOLLAR Aix Marseille Université	Présidente du jury
Hervé CHAUDET Aix-Marseille Université	Directeur de thèse
Philippe COLSON Aix Marseille Université	Invité

Affidavit

Je soussigné, Mr Lanceï KABA, déclare par la présente que le travail présenté dans ce manuscrit est mon propre travail, réalisé sous la direction scientifique du Dr Hervé CHAUDET, dans le respect des principes d'honnêteté, d'intégrité et de responsabilité inhérents à la mission de recherche. Les travaux de recherche et la rédaction de ce manuscrit ont été réalisés dans le respect à la fois de la charte nationale de déontologie des métiers de la recherche et de la charte d'Aix-Marseille Université relative à la lutte contre le plagiat.

Ce travail n'a pas été précédemment soumis en France ou à l'étranger dans une version identique ou similaire à un organisme examinateur.

Fait à Marseille, le 10 février 2022

Liste des publications et collaborations

1) **Kaba, L.**; Giraud-Gatineau,A.; Jimeno, M.-T.; Rolain, J.-M.; Colson, P.; Raoult, D.; Chaudet a, H. Consequences of the COVID-19 Outbreak Lockdown on Non-viral Infectious Agents as Reported by a Laboratory-Based Surveillance System at the IHU Méditerranée Infection, Marseille, France. *J. Clin. Med.* 2021, *10*, 3210. https://doi.org/10.3390/jcm10153210
2) Giraud-Gatineau A, **Kaba L**, Boschi C, Devaux D, Casalta JP, Gautret P, Chaudet H, Colson P, Raoult D. Control of common viral epidemics but not of SARS-CoV-2 through the application of hygiene and distancing measures. In Press in Journal of Clinical Virology (April 15, 2022). https://doi.org/10.1016/j.jcv.2022.105163
3) **Lanceï Kaba**, Audrey Giraud-Gatineau, Philippe Colson, Pierre-Edouard Fournier, Didier Raoult and Hervé Chaudet. Influence of infection origin, type of sampling and weather factors on the periodicity of some infectious pathogens in Marseille university hospitals, France. (En cours de preparation)
4) **Lanceï Kaba**, Hervé Chaudet and Philippe COLSON. Diversity study in bacteriology: A Review (En cours de preparation)
5) **Lanceï Kaba**, Philippe Colson, Hervé Chaudet and Didier Raoult. Analysis of bacterial diversity in clinical microbiology laboratories affiliated to the epidemiological surveillance system of the PACA region (PACASurvE) from 2013 to 2019, France. (En cours de preparation)
6) Colson P, Giraud-Gatineau A, Fournier PE, Ninove L, Zandotti C, Jimeno MT, Boschi C, Luciani L, **Kaba L**, Parola P, Ranque S, Rolain JM, Gautret P, Drancourt M, Lagier JC, La Scola B, Chaudet H, Raoult D. Epidemiological surveillance of respiratory viral infections at IHU Méditerranée Infection and its application to SARS-CoV-2. DOI: https://doi.org/10.35088/gpgh-wq98. In the process of submission (2021)
7) Youssouf, S.; Moise, M.; Soraya, M.; Cheick, G.O.; **Lanceï, K.**; Ghiles, G.; Thibault, M.; Jean-Louis, M.; Tu Anh, T.; Pierre, C.; Anne, F.; Joana, V. A Non-Invasive Neonatal Signature Predicts Later Development of Atopic Diseases. *J. Clin. Med*. 2022, 11, 2749. https://doi.org/10.3390/jcm11102749

Avant-propos

Le format de présentation de cette thèse correspond à une recommandation de la spécialité Maladies Infectieuses et Microbiologie, à l'intérieur du Master des Sciences dela Vie et de la Santé, qui dépend de l'Ecole Doctorale des Sciences de la Vie de Marseille. Le candidat est amené à respecter des règles qui lui sont imposées et qui comportent un format de thèse utilisé dans le Nord de l'Europe et qui permet un meilleur rangement queles thèses traditionnelles. Par ailleurs, la partie introduction et bibliographie estremplacée par une revue envoyée dans un journal afin de permettre une évaluation extérieure de la qualité de la revue et de permettre à l'étudiant de commencer le plus tôtpossible une bibliographie exhaustive sur le domaine de cette thèse.

Par ailleurs, la thèse est présentée sur article publié, accepté ou soumis associé d'un bref commentaire donnant le sens général du travail. Cette forme de présentation a paru plus en adéquation avec les exigences de la compétition internationale et permet de se concentrer sur des travaux qui bénéficieront d'une diffusion internationale.

Professeur Didier RAOULT.

Remerciements

Je tiens tout d'abord a remercié Dieu de m'avoir donné la vie, la santé et la force pour réaliser cette thèse tant souhaitée dans l'accomplissement de ma carrière professionnelle et exprimé ma gratitude à l'endroit de tous ceux qui, de près ou de loin, ont contribué tant soit peu à l'élaboration et à l'amélioration de ce travail de thèse.

Mes sincères remerciements et reconnaissances vont à l'endroit de mon directeur de thèse, Docteur **Hervé CHAUDET**, pour m'avoir donné l'opportunité de réaliser cette thèse en acceptant de m'encadrer. Votre soutien constant, votre disponibilité, votre confiance, votre rigueur scientifique, vos précieux conseils et vos réelles qualités humaines et scientifiques ont été indispensables dans l'élaboration d'un tel travail.

Mes remerciements s'adressent également au Professeur **Philippe COLSON** co-directeur de cette thèse pour le suivi, l'apport scientifique et l'accompagnement dans la réalisation de cette thèse.

Je remercie vivement mon frère jumeau **Lansana KABA** pour sa contribution personnelle au financement de cette thèse qui m'a permis de rallier Marseille et prendre en charge ma première année de thèse avant de bénéficier tout autre financement.

Je voudrais exprimer mes remerciements à Monsieur **Abdoulaye Yéro BALDE** ancien Ministère de l'Enseignement Supérieur et de la Recherche Scientifique de la République de Guinée pour son sens d'écoute, qui a prêté une oreille attentive à mon cris de cœur pour le financement de cette thèse après l'avoir débuté sur fonds propre.

Au Service de Coopération et d'Action Culturelle (SCAC) de l'Ambassade de France en Guinée pour son soutien indéfectible dans le financement des deux dernières années de ma thèse.

Je remercie le Docteur **Alpha Kabinet KEITA** et le Professeur **Raymond RIUMY** pour avoir accepté d'être les rapporteurs et lu ce travail malgré leur occupation.

Je remercie le Professeur **Florence FENOLLAR** d'avoir accepté de présider ce jury de soutenance.

Je remercie le Professeur **Didier RAOULT** et son équipe, particulièrement l'UMR VITROME d'avoir accepté de m'accueillir au sein de l'IHU et sa contribution pour l'avancement de mes projets de thèse lors des « WIP (Work In Progress) ».

Je remercie particulièrement la direction générale de l'Institut supérieur des sciences et de médecine vétérinaire (ISSMV) de Dalaba et tous les enseignants chercheurs dudit Institut pour leur soutien, plus spécifiquement à l'ancien directeur général le Professeur **Youssouf SIDIME** et au directeur général adjoint chargé de la recherche le Professeur **Alpha Oumar Sily DIALLO** pour leur soutien indéfectible, conseil et accompagnement.

J'exprime mes sentiments de gratitude au Docteur **Mohamed Lamine KEITA** pour son aide à la recherche d'un directeur de thèse et au Docteur **Audrey GIRAUD-GATINEAU** pour sa disponibilité et ses remarques constructives tout au long de ce travail.

Je tiens à remercier les collègues avec qui j'ai passé de bons moments ensemble, notamment Docteur **Adama Zan DIARRA**, Docteur **Sokhna NDONGO**, Docteur **Hamadou OUMAROU HAMA**, Docteur **David LUPANDE MWENEBITU**, Docteur **Ousmane Oumou DIALLO**, **Cheick Oumar GUINDO**, **Abdou PADANE**, **Salimatou DIAKITE**, **Aminata CAMARA**, **Fatima Zouina MEKHALIF**, **Lamine OUEDRAOGO** et **Mohamed Lamine KEITA** pour leur amitié et encouragement.

J'aimerais affectueusement remercier mes parents feu **Fodéba KABA** et Feue **Fanta KABA**, pour leur affection et soutien inconditionnels pendant mon cursus scolaire et sans qui rien n'aurait pu avoir jour. Merci pour vos prières et merci à mes frères et sœurs particulièrement à mon frère jumeau **Lansana KABA** pour leur soutien, encouragement et amour.

Enfin, je ne saurais terminer sans remercier ma modeste famille, à savoir ma chère épouse **Safiatou TRAORE** et mes chers enfants **Ciré KABA**, **Baba KABA**, **Youssouf KABA** et **Aïssatou KABA**, qui m'ont donné la force et l'énergie nécessaire à relever ce défi par leur soutien et encouragement. Je tiens à leur signifier mon affection et mon amour.

Table des matières

Liste des abréviations

AP-HM – Assistance Publique - Hôpitaux de Marseille

ANOVA – Analysis of Variance

ANOSIM – Analysis of Similarities

IHU MI – Institut Hospitalo-Universitaire Méditerranée Infection

MALDI-TOF – Matrix Assisted Laser Desorption-Ionization – Time Of Flight

MIDaS – Méditerranée Infection Data Warehousing and Surveillance

PACA – région Provence-Alpes Côte d'Azur (actuelle région du sud)

PACASurvE - PACA Surveillance Epidemiological System

SARS-CoV-2 – Severe Acute Respiratory Syndrome Corona Virus 2

VITROME – Vecteurs-Infections Tropicales et Méditerranéennes

Résumé

L'émergence ou la réémergence des épidémies dans le monde constitue une préoccupation majeure en santé publique et fait de la surveillance épidémiologique une priorité pour l'ensemble des Pays. La surveillance joue un rôle de premier plan dans la prévention et le contrôle des maladies infectieuses. Les indicateurs habituels de surveillance (taux d'incidence, de mortalité, …) donnent une vision indépendante des maladies infectieuses sous surveillance, alors que celles-ci sont de plus en plus envisagées dans le cadre d'un « écosystème » microbiologique humain. Une évolution de la surveillance épidémiologique en complétant ces indicateurs, s'impose. Cette amélioration passe par une meilleure compréhension de la co-occurrence des agents infectieux et de leur dynamique spatio-temporelle. Les objectifs de cette thèse s'inscrivent dans ce cadre pour (i) faire un état des lieux de l'étude de la diversité bactérienne, (ii) évaluer l'impact des mesures de confinement contre le SARS-CoV2 sur la population bactérienne, (iii) comprendre la diversité des espèces bactériennes identifiées par les laboratoires de microbiologie, maillon fondamentale de la surveillance épidémiologique, ainsi que sa dynamique et (iv) identifier la corrélation entre des agents bactériens avec certains facteurs météorologique, en vue de compléter les indicateurs de la surveillance épidémiologique par des indices de diversité. A cet effet, nous avons évaluer la diversité bactérienne dans les laboratoires de microbiologie affiliés au système de surveillance épidémiologique de la région Provence-Alpes-Côte d'Azur (PACASurvE) de l'Institut Hospitalo-Universitaire (IHU) Méditerranée Infection de Marseille. La diversité alpha particulièrement la richesse spécifique, l'abondance relative, l'indice de Shannon et l'indice de Simpson, et la diversité bêta ont été les paramètres évalués. Les tests d'ANOVA et de TUKEY ont été appliqués pour comparer les moyennes de la richesse et de l'abondance observées. La statistique ANOSIM a été utilisée pour comparer la moyenne des dissemblances classées entre les groupes à la moyenne des dissimilarités classées au sein des groupes. Nous avons également effectué une analyse de corrélation du top 15 des agents infectieux avec les facteurs météorologiques notamment la température, la précipitation, l'humidité, le vent et le changement de pression, pour une détection de périodicité, facilitant également la surveillance et des stratégies de prévention. Nos résultats ont montré une répartition très différente de la diversité bactérienne selon le type de laboratoire et selon les départements, en plus le mix bactérien augmentait au fil du temps. Le confinement a conduit à la diminution de la fréquence de certaines espèces tel que *Escherichia coli*, en même temps, d'autres espèces n'ont subi aucun effet, elles sont restées stable voir augmenter en fréquence. Certaines espèces ont été corrélées à un ou plusieurs facteurs météorologiques. Vue les résultats obtenus, les indices de diversité évalués pourront utiliser comme complément des indicateurs dans les systèmes de surveillance épidémiologique.

Mots clés : Diversité, bactérie, richesse spécifique, indice de Shannon, indice de Simpson, surveillance épidémiologique.

Abstract

The emergence or re-emergence of epidemics around the world is a major public health concern and makes epidemiological surveillance a priority for all countries. Surveillance plays a key role in the prevention and control of infectious diseases. The usual surveillance indicators (incidence rate, mortality rate, etc.) provide an independent view of the infectious diseases under surveillance, whereas these are increasingly considered within the framework of a human microbiological "ecosystem". An evolution of epidemiological surveillance by supplementing these indicators is essential. This improvement requires a better understanding of the co-occurrence of infectious agents and their spatio-temporal dynamics. The objectives of this thesis fall within this framework to (i) make an inventory of the study of bacterial diversity, (ii) assess the impact of lockdown measures against SARS-CoV2 on the bacterial population, (iii) understand the diversity of bacterial species identified by microbiology laboratories, a fundamental link in epidemiological surveillance, as well as its dynamics and (iv) identify the correlation between bacterial agents with certain meteorological factors, in order to complete the indicators of epidemiological surveillance using diversity indices. To this end, we have evaluated the bacterial diversity in the microbiology laboratories affiliated to the epidemiological surveillance system of the Provence-Alpes-Côte d'Azur region (PACASurvE) of the Institut Hospitalo-Universitaire (IHU) Méditerranée Infection de Marseille. The alpha diversity, particularly the specific richness, the relative abundance, the Shannon index and the Simpson index, and the beta diversity were the parameters evaluated. ANOVA and TUKEY tests were applied to compare the averages of observed richness and abundance. The ANOSIM statistic was used to compare the mean of ranked dissimilarities between groups to the mean of ranked dissimilarities within groups. We also performed a correlation analysis of the top 15 infectious agents with meteorological factors including temperature, precipitation, humidity, wind and pressure change, for periodicity detection, also facilitating monitoring and strategies. of prevention. Our results showed a very different distribution of bacterial diversity according to the type of laboratory and according to the departments, in addition the bacterial mix increased over time. The lockdown led to the decrease in the frequency of certain species such as Escherichia coli, at the same time, other species suffered no effect, they remained stable or even increased in frequency. Some species have been correlated with one or more meteorological factors. Given the results obtained, the diversity indices evaluated can be used as a complement to indicators in epidemiological surveillance systems.

Keywords: Diversity, bacteria, species richness, Shannon index, Simpson index, epidemiological surveillance

Introduction

L'émergence ou la réémergence des épidémies dans le monde constitue une préoccupation majeure en santé publique. Une maladie infectieuse émergente est un phénomène infectieux – ou présumé infectieux – inattendu (en référence à ses propriétés intrinsèques ou aux connaissances de sa biologie), touchant l'homme, l'animal ou les deux (1). La lutte contre les maladies infectieuses passe par le diagnostic précoce de la survenue de nouvelles infections, la veille épidémiologique, le développement d'un réseau global d'information, d'alerte, et d'intervention sur le terrain (2). Dans ce contexte la surveillance épidémiologique des maladies infectieuses représente un préalable indispensable à l'élaboration, l'évaluation, le contrôle, la lutte et la prévention, ce qui fait d'elle une priorité pour l'ensemble des Pays. D'où la nécessité de développer et de renforcer les systèmes de surveillance pour une détection précoce et même pour des infections à faible signaux.

Les maladies infectieuses ont longtemps représenté la principale cause de mortalité dans le monde. Le risque épidémique touche cependant d'abord les pays en voie de développement. Elles y sont responsables de 43 % du total des décès, contre 1 % dans les pays industrialisés (3). De nouvelles maladies infectieuses émergent ou ré-émergent apparaissent de façon permanente, tant dans les pays en développement que dans les pays industrialisés (4). En mars 2003, survenait une épidémie du syndrome respiratoire aigu sévère (SRAS) en Chine dans la province Guangdong, où l'OMS a reçu la notification de 7919 cas probables de SRAS de la part de 31 pays. Parmi ces malades, 662 (8,3 %) sont décédés et 3 984 (50,3 %) sont considérés comme guéris à la date du 20 mai 2003 (5). L'émergence de la fièvre hémorragique à virus Ebola en Afrique de l'Ouest en 2014 avait aussi suscité beaucoup de question sur le diagnostic et la surveillance des maladies en Afrique (6,7). Tout récemment, l'apparue de la pandémie du SARS-CoV-2 en fin décembre 2019 à Wuhan en Chine a entrainé plus de 508 000 décès confirmés dans plus de 200 pays à la date du 1er juillet 2020 (8) dans le monde avec des conséquences sanitaires, économiques et sociales énormes. Ces épidémies illustrent parfaitement l'émergence des maladies infectieuses et la nécessité d'améliorer nos systèmes de veille et de surveillance épidémiologique et de développer une collaboration étroite entre les pays en matière de gestion des événements de santé à potentiel épidémique.

Ces grandes épidémies ou pandémies ont été généralement d'origine virale. En revanche, plusieurs épidémies à échelle locale ou régionale sont dues également à des infections bactériennes, comme les épidémies de choléra en Afrique subsaharienne

(9,10). Ces bactéries ont cependant, une très grande capacité d'adaptation à leur environnement animal ou humain. En particulier, la modification des flores commensales, déterminant majeur de la genèse de notre système immunitaire, relevant de modifications globales de notre mode de vie (y compris alimentaire), pourrait jouer un rôle important, dans l'avenir, sur l'évolution des maladies infectieuses émergentes. L'émergence des maladies infectieuses à potentiel épidémique d'origine bactérienne, constitue une menace majeure et croissante en santé publique. La majorité des agents pathogènes impliqués dans les événements de maladies infectieuses émergentes sont des bactéries ou des rickettsies (54,3%) (11). L'autre aspect important dans l'émergence des maladies infectieuses est la biodiversité des pathogènes.

La biodiversité peut jouer un double rôle dans l'émergence et la transmission des maladies infectieuses. D'une part, une biodiversité élevée peut fournir une plus grande source potentielle de nouveaux agents pathogènes, mais d'autre part, la biodiversité peut réduire davantage la transmission d'agents pathogènes. Elle englobe la diversité des gènes, des espèces et des écosystèmes (12). La connaissance de la diversité de ces agents permettrait de détecter à temps et de minimiser le risque de propagation et de faciliter la mise en place des mesures de prévention, ainsi que la prise en charge adéquate des cas.

Par conséquent, plusieurs systèmes de surveillance des maladies sont mis en place dans le monde dont l'objectif principal est la détection précoce de toute menace d'épidémie et la prise de disposition idoines contre l'évolution et la propagation des éventuelles épidémies. Pour améliorer les systèmes de surveillance épidémiologique et en les rendant plus sensibles pour favoriser la détection des échappements ou des signaux faibles, particulièrement pour des infections bactériennes, il est impératif de développer ou d'introduire d'autres indicateurs outre que ceux habituels pour une meilleure efficacité dans la détection précoce des épidémies. Cela passe nécessairement par l'amélioration ou l'introduction de nouveaux outils de diagnostic.

Ainsi, l'introduction du MALDI-TOF (Matrix Assisted Laser Desorption-Ionization – Time Of Flight) dans l'identification des espèces bactériennes par les laboratoires de microbiologie clinique (13) a élargi la palette des espèces pouvant être directement identifiées par rapport aux méthodes basées sur des kits phénotypiques ou génotypiques. De fait, la diversité et la dynamique apparente des espèces bactériennes s'est largement accrue dans ces laboratoires. Cette diversité dans les laboratoires de microbiologie reste très peu documentée et peu exploitée dans le cadre de la surveillance épidémiologique des maladies infectieuses. En effet, les indicateurs

habituels de surveillance (taux d'incidence, de mortalité, …) donnent une vision indépendante des maladies infectieuses sous surveillance, alors que celles-ci sont de plus en plus envisagées dans le cadre d'un « écosystème » microbiologique humain.
Ces constatations justifient de poursuivre et d'adapter la surveillance, ainsi que la recherche de nouvelles connaissances et de nouveaux outils. Une évolution de la surveillance épidémiologique en complétant ces indicateurs est donc envisageable. Cette amélioration passe par une meilleure compréhension de la co-occurrence des agents infectieux et de leur dynamique spatio-temporelle. Les objectifs de cette thèse s'inscrivent dans ce cadre pour (i) faire un état des lieux de l'étude de la diversité bactérienne, (ii) évaluer l'impact des mesures de confinement contre le SARS-CoV-2 sur la population bactérienne, (iii) comprendre la diversité des espèces bactériennes identifiées par les laboratoires de microbiologie, maillon fondamentale de la surveillance épidémiologique, ainsi que sa dynamique et (iv) identifier la corrélation entre des agents bactériens avec certains facteurs météorologiques, en vue de compléter les indicateurs de la surveillance épidémiologique par des indices de diversité.
Ainsi, pour atteindre ces objectifs, ce manuscrit va s'articuler sur quatre chapitres :

Chapitre I : Etat des lieux des études de diversité bactérienne

Une revue exhaustive de la littérature sur les études de diversité bactérienne a été effectuée dans ce chapitre sur une période des cinq dernières années de 2015 à 2019. Nous nous sommes intéressés aux articles publiés en anglais ou en français qui ressortent de façon pertinente les indices de mesure de la diversité. Le moteur de recherche PubMed a servi de cadre pour la recherche des articles en s'appuyant sur des critères de sélection basés sur des mots clés, ce qui nous a permis de retenir 46 articles après lecture du résumé et du texte intégrale.

Chapitre II : Analyse de la diversité bactérienne et observée dans les laboratoires de microbiologie clinique du système de surveillance PACASurvE, 2013 – 2019

Les laboratoires de microbiologie constituent de nos jours un maillon important et intournable des services de santé dans le diagnostic des maladies infectieuses ou non infectieuses. En plus des techniques d'identification des agents pathogènes déjà existantes utilisées dans ces laboratoires, des nouvelles techniques d'identification sont en perpétuelle développement pour la détection des espèces rares difficilement détectables voir inconnues. La mise au point de ces techniques nouvelles permet d'élargir le champ de diversité des agents infectieux par identification de nouvelles

espèces ou d'espèces rares d'une part, et d'autre part, par une identification en nombre importante des espèces déjà connues. Cette diversité bactérienne dans les laboratoires de microbiologie reste peu documentée. Dans ce chapitre, nous évaluons la diversité bactérienne dans les laboratoires de microbiologie affiliés au système de surveillance PACASurvE de l'IHU Méditerranée Infection de Marseille.

Chapitre III : Etude d'impact du confinement contre le Covid 19 sur la diversité des agents infectieux non-viraux

La déclaration des premiers cas de Covid 19 en fin décembre 2019 à Wuhan en Chine a suscité d'intérêt et focaliser la recherche et l'attention des chercheurs sur tous les aspects d'étude d'impact de cette épidémie. Pour arrêter la propagation du virus et endiguer la pandémie, plusieurs mesures restrictives ont été prises dans le monde dont le confinement. Cette mesure de confinement orientée spécifiquement contre le SARS-CoV2, a eu sans nul doute influencé d'autres infections virales et bactériennes. L'absence de contact entre les populations, la fermeture des écoles, l'arrêt des voyages nationaux et internationaux et le lavage des mains ont impacté certains agents infectieux en termes de transmission mais aussi de l'évolution de leur population. C'est dans cette logique que nous avons jugé utile de regarder l'évolution de la diversité des espèces non-virales et de la résistance aux antibiotiques des espèces identifiées par le système de surveillance du laboratoire de l'IHU de Marseille.

Chapitre IV : Etude de saisonnalité de quelques espèces bactériennes dans la région PACA

Pour une gestion efficace des flambées ou des épidémies, la détection précoce des cas éventuels à l'aide des systèmes de surveillance s'avère incontournable. Cependant, la seule surveillance épidémiologique n'est pas suffisante, il serait important de coupler avec des connaissances de saisonnalités des différentes espèces dans le but d'identifier la corrélation entre des agents bactériens avec certains facteurs météorologiques. Bien que la saisonnalité de certaines espèces soit bien connue, surtout celle d'infections d'origine virale, mais cette saisonnalité connue des agents infectieux concerne les formes d'expression clinique des infections. Par ailleurs, la saisonnalité des agents infectieux et non la forme clinique restent encore mal connue. Cet état de fait nous a pousser à approfondir les recherches sur la périodicité de certains agents bactériens en lien avec des facteurs météorologiques dans la région PACA.

CHAPITRE I

Etat des lieux de l'étude de la diversité bactérienne de 2015 - 2019

Préambule

La perturbation de l'écosystème se traduit par la modification des communautés des différentes espèces. Cette variabilité au sein d'une espèce ou d'une communauté en termes de fréquence est qualifiée de diversité. Autrement dit, la diversité est la variabilité des organismes vivants de toute origine (14), c'est-à-dire la gamme de types d'organismes très différents et leur abondance relative dans une communauté (15,16). La diversité microbienne, ce n'est pas seulement la diversité en termes de nombre d'espèces qui existent, c'est aussi la diversité des propriétés des souches, à l'intérieur d'une espèce. Elle se réfère sans équivoque à la diversité biologique à trois niveaux : au sein des espèces (génétique), du nombre d'espèces (espèces) et de la communauté (écologique) (17,18). Le terme diversité d'espèces comprend deux composantes ; la première composante est le nombre total d'espèces présentes que l'on peut qualifier de richesse en espèces. En d'autres termes, il fait référence à la variation quantitative entre les espèces. La deuxième composante est la distribution des individus parmi ces espèces, appelée uniformité ou équité.

Les effets du changement climatique (19,20) et les actions anthropiques (21) ont impacté et continu d'impacter les écosystèmes et les communautés des micro et macroorganismes qui y vivent. Ce changement favorise l'émergence ou la réémergence d'agents responsables d'épidémies. Ces actions ont conduit à la réalisation de nombreuses études de diversité biologique, les plus connues sont celles réalisées sur le macroorganisme notamment les végétaux (22), les animaux aquatiques (23), terrestres et chez l'homme. Également, la diversité des microorganismes telles que les bactéries est largement abordée dans plusieurs études chez l'homme et tout comme chez les animaux avec des méthodes d'évaluation différentes ou similaires selon les cas et les objectifs. Les études sur la diversité bactérienne diffèrent les unes des autres par la manière d'obtenir des isolats, le mode de caractérisation, les méthodes de regroupement utilisées pour le regroupement ou l'identification, le niveau de similarité ou de distance utilisé pour définir une espèce ou un biotype et les mesures de diversité utilisées. Une gamme variée d'indice de diversité ont été proposées et utilisées par les écologistes et les mathématiciens. Les indices couramment utilisés dans la mesure de diversité ont été les indices de Shannon et de Simpson. La diversité est mesurée par une méthode de caractérisation ne peut pas être facilement comparée à celle mesurée par d'autres modes de caractérisation. Pour les études, il faut définir des critères pour un indice de diversité idéal (24).

Cette revue vise à capitaliser les différentes méthodes d'évaluation de la diversité (paramètres évalués) en bactériologie afin d'identifier les différents indices d'évaluations potentiels susceptibles d'être utilisés comme indicateur dans la surveillance épidémiologique.

REVUE DE LA LITTERATURE

Article 1

Article 1: Diversity study in bacteriology: A review

Lanceï KABA, Hervé CHAUDET, Philippe COLSON

En cours de correction

TITLE DE PAGE

Type of article: a review

Full-length title: Diversity study in bacteriology

Author list : Lanceï KABA[1,2,3], Hervé CHAUDET[1,2], Philippe COLSON[1,2]

[1] IHU Méditerranée Infection, 19-21 boulevard Jean Moulin, 13005 Marseille, France;

[2] Aix Marseille Univ., Institut de Recherche pour le Développement (IRD), Assistance

[3] Institut Supérieur des Sciences et de médecine vétérinaire (ISSMV) de Dalaba, Guinée ;

*** Corresponding author:** Hervé CHAUDET, IHU Méditerranée Infection, 19-21 boulevard Jean Moulin, 13005 Marseille, France. E-mail: herve.chaudet@gmail.com

Abstract

The disturbance of the ecosystem results in the modification of the communities of different species living there. This modification of the diversity for infectious agents would lead to the emergence of new epidemics. The main objective is to make a systematic synthesis of bacterial diversity assessment methods published in the literature. We performed this systematic review on articles deemed relevant available on PubMed from 2015 to 2019 related to bacterial diversity. The keywords used for the search were: bacterial AND (diversity OR biodiversity). From a total of 70 articles identified after reading the title and abstract, only 46 articles were selected based on their relevance. Of the 46 articles, 25 (54.3%) focused on humans and 21 (45.7%) on animals. The diversities assessed were 100% (46/46) alpha diversity, 87.0% (40/46) beta diversity and 91.3% (42/46) genetic diversity. The most frequently calculated were richness 91.3% (42/46), Shannon index 84.8% (39/46) and relative abundance 78.6% (36/46). Phylogenetic diversity was assessed by 43.5% (20/42) of the articles and taxonomic diversity by 28.6% (13/42). The species identification method used was the molecular biology method focused on the sequencing of 16S rRNA or rDNA. Forty-six of 46 articles (100%) used the molecular method. The different measures of bacterial diversity by bacteriology laboratories are basically the evaluation of the local diversity observed and the dissimilarity between samples. It would be important to extend this study to other aspects of diversity such as indicator species and dark diversity.

Keywords: bacterial diversity; bacterial species; diversity measures; bacteria.

1. Introduction

The disturbance of the ecosystem results in the modification of the communities of the various species living there [1]. This change leads to the emergence or re-emergence of species responsible for epidemics. Diversity represents the variability within a species or community in terms of both number and frequency. Numerous studies have made it possible to inventory thousands of species using the cultural method [2–4] before the implementation of molecular methods which nowadays remains the reference method of description of the species. However, many non-culturable species remained unknown at that time. The cultivable bacteria represented a tiny part of the total bacterial population present [5]. However, it is important to continue work on cultivable and non-cultivable bacteria from different environments. With the advent of molecular biology [6], the genetic method has allowed not only to highlight thousands of new species but also to overcome the difficulties of identification of uncultivable species. Regarding microorganisms, in natural ecosystems, they are very numerous, while there are several thousand microbial species not yet described [7]. The diversity of bacteria in the soil [8] is enormous and soil bacterial communities can vary considerably in structure. For example, one gram of soil or sediment can contain 10^{10} bacteria, in pure seawater, the number of bacteria is about 10^{6} per ml [9], [10]. These microorganisms are an important cause of morbidity and mortality in developing countries, even epidemics, as in developed countries. For infections of bacterial origin, many infections are caused by species of bacteria that are present in most people without causing disease. Often, these bacteria become pathogenic when they are found in a place in the body where they should not be present, or in abnormally high numbers during a decline in the immune system. Or when it is a subtype (strain) of bacteria that consistently makes you sick (e.g. *enterohemorrhagic* strains of *Escherichia coli*). Another important aspect in the study of bacterial species diversity is the knowledge of the nature of the ecosystem. Well-organized communities with a certain level of diversity are stable [7], [11]. If some kind of

stress is introduced into this community, the stability may collapse and the diversity will change [12]. Nowadays, many indicators exist for the epidemiological surveillance of infectious agents. However, none of them use diversity for the detection of possible epidemics. Thus, a thorough understanding of diversity detection and quantification methods could be of paramount importance in detecting an outbreak or epidemic based on diversity data.

In this review, we aim to provide an overview of commonly used methods for assessing bacterial diversity and for identifying bacterial species.

2. Materials and Methods

2.1. Search strategy and selection criteria

Our literature search strategy was limited to articles published and available on PubMed from 01/01/2015 to 31/12/2019 related to bacterial diversity. Inclusion criteria included any human or animal study of bacterial diversity that detailed methods for measuring diversity and was published in either English or French. Initially, a search based on keywords in the title and abstract was performed and subsequently articles that did not meet our inclusion criteria for defining eligible articles were excluded. Eligible articles for systematic reviews were read and analyzed to retain only relevant articles, and a flow diagram is attached as an additional figure (Figure 1). In the search, articles that did not mention species identification methods and biodiversity assessment methods were excluded after reading the abstract. After reading the full text of the various eligible articles, only those describing species identification techniques and diversity assessment methods were retained. The bibliography management software Zotero and Rayyan QCRI [13] were used to search and select bibliographic references. Data collected included authors and publication dates, study location, types of diversity studied (alpha, beta, genetic, and taxonomic), bacterial identification methods, nature and size of samples analyzed, population and species studied. Our unfiltered PubMed search equation is as follows: (((bacterial AND (diversity OR biodiversity)) NOT (soil)) NOT (water)) NOT (air). Les

éléments de filtre appliqués ont été : 2015-2019 ; Free full text ; Journal de revue ; Humans and animals ; English and French, ainsi l'équation avec les éléments du filtre était : (((bacterial AND (diversity OR biodiversity)) NOT (soil)) NOT (water)) NOT (air) AND ((ffrft[Filter]) AND (journalarticle[Filter]) AND (animal[Filter] OR humans[Filter]) AND (english[Filter] OR french[Filter]) AND (2015:2019[pdat])). All these steps are summarized in Table 1.

2.2. Statistical analysis

A descriptive analysis of the selected articles was done in terms of number and frequency based on the variables of interest.

3. Results

3.1. Documentary research

The raw number of articles obtained without filtering resulted in 54327 identified references. Applying this led to 9878 items. A total of 46 articles were retained after excluding 24 of the 70 eligible articles that were found to be irrelevant after reading the full text. Figure 1 details the methodology of the literature search, following the PRISMA recommendations.

3.2. Characteristics of the selected items

Out of 46 articles selected according to the inclusion criteria and the research objective, 25 (54.3%) related to humans and 21 (45.7%) to animals. In relation to the countries where the studies were carried out, 13 countries were concerned, of which the USA ranks first with 15 articles out of 46, i.e., 32.6% of the articles selected, followed by China 11 (23.9%). The details are presented in figure 2. However, more than half of the articles were published in 2019 or 24 (52.2%), followed by 2018 with 9 articles (19.6%), 8 (17.4%) in 2017, 4 (8.7%) in 2016 and 1 (2.2%) in 2015 (Figure 3). All papers (100%) used the molecular method of 16S rRNA region sequencing for bacterial species identification (Table 2), as it remains the best technique applicable to all bacteria, whether cultivable or not [14]. Hundreds of bacterial species were

identified throughout the papers.

3.3. Identified diversity assessment indicators.

Bacterial diversity was measured considering different types of diversity (Table 3) which are among others: alpha diversity 100% (46/46), beta diversity 87.0% (40/46) and genetic diversity 42 (91.3%). However, 87.0% (40/46) assessed both alpha and beta diversity [8], [15]–[17]. The diversity observed locally is often called "alpha diversity" [18], it is measured within a well-defined sample or site [19] and beta diversity, or spatial turnover of species, is the association between alpha diversity and gamma diversity, either multiplicative (beta = gamma / alpha) or additive (beta = gamma - alpha) [18], [20]. One of the indicators of beta diversity is the indicator species 5.0% (2/40) which is rarely determined. For alpha diversity, the indices used were specific and estimated richness, relative abundance, Shannon, Simpson, inverse Simpson, equitability, and rarefaction curve. The most frequently calculated were richness 91.3% (42/46), Shannon index 84.8% (39/46) and relative abundance 78.6% (36/46). As for beta diversity, three distance matrices were calculated namely matrices based on Bray-Curtis dissimilarity distance 67.5% (27/40), weighted and unweighted UniFrac distances 27.5% (11/40), Jaccard distance 7.5% (3/40) and Sorensen distances 5.0% (2/40). Similarities or dissimilarities were visualized either by principal coordinate analysis (PCoA) 65.0% (26/40) or by Non-metric multidimensional scaling (NMDS) 5.0% (2/40). Phylogenetic diversity 43.5% (20/42) and taxonomic diversity 28.6% (13/42) were assessed by some papers and others assessed both 19.6% (9/42). Phylogenetic diversity is a genetic diversity that measures the total length of the branches of a phylogenetic tree, is the component that describes how different species are from each other [21]. However, taxonomic diversity considers phylogenetic information. It corresponds, according to the definition provided by Clarke and Warwinck in 1995, to the average length of the path, in the hierarchical classification, between two organisms chosen

143 randomly in a community and thus considers all taxonomic levels (species but also genera,
144 families and orders).

145 **Discussion**

146 In this review, we focused on tools and metrics for assessing bacterial biodiversity that are used
147 by researchers today. The study of diversity is necessary to understand the structures,
148 distribution, and evolution of bacterial species communities, which can play an important role
149 in ecosystem functioning. Our results showed that alpha diversity is the most commonly
150 assessed by all studies [19], [22]–[24], specifically species richness and frequency (relative
151 abundance), followed by Shannon and Simpson diversity as indicated by Chao [25]. Also, Bray-
152 Curtis dissimilarity and principal coordinate analysis (PCoA) were the primary metrics for
153 assessing and visualizing beta diversity assessed [26], [27]. PCoA and PCA are used to observe
154 patterns in samples as reported in a study by Chao [28]. Compared to the nature of diversity
155 studied, the diversity of genes (phylogenetic diversity and taxonomic diversity) has remained
156 the most widely discussed in most articles because the technique most currently used in
157 microbiological research is based on identification. polymorphisms (sequencing) of the
158 bacterial gene encoding 16S rRNA or rDNA [29]. This technique is the most widely used
159 because it allows (i) to identify bacterial species that are generally non-cultivable or difficult to
160 cultivate, (ii) to identify a large number of taxa simultaneously, (iii) to detect previously
161 unknown bacterial taxa in the human microbiota and to assign them taxonomically to an order
162 or family with a high degree of accuracy, and (iv) to estimate the relative abundance of
163 individual taxa, providing quantitative information that is very important for understanding the
164 structure of the microbiota [30]. Although this molecular technique has many advantages, it
165 also has limitations [31]. A combination of the classical (cultural) and molecular methods would
166 give a better approach to the identification of bacterial species [32], [33]

167

Prospects

We plan to develop new indicators based on bacterial diversity indices to complement existing epidemiological surveillance indicators. All the diversity assessment approaches available today have advantages and limitations, although none of them offers complete access to the extremely large and complex bacterial world. These new molecular biology methods, which are constantly advancing, have provided powerful and important confirmation of previous phenotypic and genotypic diversity studies on bacteria. The combination of different methods is always the best way to have a better understanding of the diversity, phylogeny, ecosystem, evolution, and taxonomy of many living organisms.

Conclusion

The different measures of bacterial diversity by bacteriology laboratories are basically the evaluation of the local diversity observed and the dissimilarity between samples. It would be important to extend this study to other aspects of diversity such as indicator species and dark diversity.

References

[1] C. Aubertin, « La biodiversité : une notion en quête de stabilité », in *Représenter la nature ? ONG et biodiversité*, Marseille: IRD Éditions, 2013, p. 99-122.

[2] T. Ito, T. Sekizuka, N. Kishi, A. Yamashita, et M. Kuroda, « Conventional culture methods with commercially available media unveil the presence of novel culturable bacteria », *Gut Microbes*, vol. 10, nº 1, p. 77-91, 2019, doi: 10.1080/19490976.2018.1491265.

[3] N. F. Montalvo, J. Davis, J. Vicente, R. Pittiglio, J. Ravel, et R. T. Hill, « Integration of culture-based and molecular analysis of a complex sponge-associated bacterial community », *PloS One*, vol. 9, nº 3, p. e90517, 2014, doi: 10.1371/journal.pone.0090517.

192 [4] S.-H. Yoon, N. W. Choi, et S.-R. Yun, « Detecting bacterial growth in continuous
193 ambulatory peritoneal dialysis effluent using two culture methods », *Korean J. Intern.*
194 *Med.*, vol. 25, nº 1, p. 82-85, mars 2010, doi: 10.3904/kjim.2010.25.1.82.
195 [5] S. J. Giovannoni, T. B. Britschgi, C. L. Moyer, et K. G. Field, « Genetic diversity in
196 Sargasso Sea bacterioplankton », *Nature*, vol. 345, nº 6270, p. 60-63, mai 1990, doi:
197 10.1038/345060a0.
198 [6] E. M. Bik *et al.*, « Molecular analysis of the bacterial microbiota in the human stomach »,
199 *Proc. Natl. Acad. Sci.*, vol. 103, nº 3, p. 732-737, janv. 2006, doi:
200 10.1073/pnas.0506655103.
201 [7] Md. Fakruddin, K. S. B. Mannan, et S. Andrews, « Viable but Nonculturable Bacteria:
202 Food Safety and Public Health Perspective », *ISRN Microbiol.*, vol. 2013, p. 1-6, 2013,
203 doi: 10.1155/2013/703813.
204 [8] K. E. Walters et J. B. H. Martiny, « Alpha-, beta-, and gamma-diversity of bacteria varies
205 across habitats », *PLOS ONE*, vol. 15, nº 9, p. e0233872, sept. 2020, doi:
206 10.1371/journal.pone.0233872.
207 [9] R. Daniel, « The metagenomics of soil », *Nat. Rev. Microbiol.*, vol. 3, nº 6, p. 470-478,
208 juin 2005, doi: 10.1038/nrmicro1160.
209 [10] V. Torsvik, J. Goksøyr, et F. L. Daae, « High diversity in DNA of soil bacteria », *Appl.*
210 *Environ. Microbiol.*, vol. 56, nº 3, p. 782-787, mars 1990.
211 [11] A. C. Yannarell et E. W. Triplett, « Geographic and environmental sources of variation in
212 lake bacterial community composition », *Appl. Environ. Microbiol.*, vol. 71, nº 1, p.
213 227-239, janv. 2005, doi: 10.1128/AEM.71.1.227-239.2005.
214 [12] Y. Guan, J. Jia, L. Wu, X. Xue, G. Zhang, et Z. Wang, « Analysis of Bacterial Community
215 Characteristics, Abundance of Antibiotics and Antibiotic Resistance Genes Along a
216 Pollution Gradient of Ba River in Xi'an, China », *Front. Microbiol.*, vol. 9, p. 3191, 2018,

doi: 10.3389/fmicb.2018.03191.

[13] M. Ouzzani, H. Hammady, Z. Fedorowicz, et A. Elmagarmid, « Rayyan-a web and mobile app for systematic reviews », *Syst. Rev.*, vol. 5, nº 1, p. 210, déc. 2016, doi: 10.1186/s13643-016-0384-4.

[14] C. Will *et al.*, « Horizon-specific bacterial community composition of German grassland soils, as revealed by pyrosequencing-based analysis of 16S rRNA genes », *Appl. Environ. Microbiol.*, vol. 76, nº 20, p. 6751-6759, oct. 2010, doi: 10.1128/AEM.01063-10.

[15] C.-H. Chen *et al.*, « Bacterial diversity among four healthcare-associated institutes in Taiwan », *Sci. Rep.*, vol. 7, nº 1, p. 8230, 15 2017, doi: 10.1038/s41598-017-08679-3.

[16] S. R. Longford *et al.*, « Comparisons of diversity of bacterial communities associated with three sessile marine eukaryotes », *Aquat. Microb. Ecol.*, vol. 48, nº 3, p. 217-229, août 2007, doi: 10.3354/ame048217.

[17] A. Prehn-Kristensen *et al.*, « Reduced microbiome alpha diversity in young patients with ADHD », *PloS One*, vol. 13, nº 7, p. e0200728, 2018, doi: 10.1371/journal.pone.0200728.

[18] M. Pärtel, R. Szava-Kovats, et M. Zobel, « Dark diversity: shedding light on absent species », *Trends Ecol. Evol.*, vol. 26, nº 3, p. 124-128, mars 2011, doi: 10.1016/j.tree.2010.12.004.

[19] E. Marcon, « Mesure de la biodiversitsé ». 2018, [En ligne]. Disponible sur: https://hal-agroparistech.archives-ouvertes.fr/cel-01205813.

[20] H. Tuomisto, « A diversity of beta diversities: straightening up a concept gone awry. Part 1. Defining beta diversity as a function of alpha and gamma diversity », *Ecography*, vol. 33, nº 1, p. 2-22, févr. 2010, doi: 10.1111/j.1600-0587.2009.05880.x.

[21] M. A. K. Sydenham, S. R. Moe, D. N. Stanescu-Yadav, Ø. Totland, et K. Eldegard, « The effects of habitat management on the species, phylogenetic and functional diversity of bees are modified by the environmental context », *Ecol. Evol.*, vol. 6, nº 4, p. 961-973,

févr. 2016, doi: 10.1002/ece3.1963.

[22] T. B. Ault *et al.*, « Uterine and vaginal bacterial community diversity prior to artificial insemination between pregnant and nonpregnant postpartum cows1 », *J. Anim. Sci.*, vol. 97, nº 10, p. 4298-4304, oct. 2019, doi: 10.1093/jas/skz210.

[23] M. J. LaMonte *et al.*, « Composition and diversity of the subgingival microbiome and its relationship with age in postmenopausal women: an epidemiologic investigation », *BMC Oral Health*, vol. 19, nº 1, p. 246, 13 2019, doi: 10.1186/s12903-019-0906-2.

[24] M. Tal, J. S. Weese, D. E. Gomez, M. Hesta, J. M. Steiner, et A. Verbrugghe, « Bacterial fecal microbiota is only minimally affected by a standardized weight loss plan in obese cats », *BMC Vet. Res.*, vol. 16, nº 1, p. 112, avr. 2020, doi: 10.1186/s12917-020-02318-2.

[25] A. Chao, « NONPARAMETRIC-ESTIMATION OF THE NUMBER OF CLASSES IN A POPULATION », 1984.

[26] E. Karlsson *et al.*, « Airborne microbial biodiversity and seasonality in Northern and Southern Sweden », *PeerJ*, vol. 8, p. e8424, 2020, doi: 10.7717/peerj.8424.

[27] J. Wang *et al.*, « Plant functional traits regulate soil bacterial diversity across temperate deserts », *Sci. Total Environ.*, vol. 715, p. 136976, mai 2020, doi: 10.1016/j.scitotenv.2020.136976.

[28] C. Liang *et al.*, « Diversity and enterotype in gut bacterial community of adults in Taiwan », *BMC Genomics*, vol. 18, nº Suppl 1, p. 932, 25 2017, doi: 10.1186/s12864-016-3261-6.

[29] A. Ticinesi *et al.*, « The impact of intestinal microbiota on bio-medical research: definitions, techniques and physiology of a “new frontier” », *Acta Bio-Medica Atenei Parm.*, vol. 89, nº 9-S, p. 52-59, 17 2018, doi: 10.23750/abm.v89i9-S.7906.

[30] M. Ventura, F. Turroni, C. Canchaya, E. E. Vaughan, P. W. O’Toole, et D. van Sinderen, « Microbial diversity in the human intestine and novel insights from metagenomics »,

Front. Biosci. Landmark Ed., vol. 14, p. 3214-3221, janv. 2009, doi: 10.2741/3445.

[31] J. F. Siqueira et I. N. Rôças, « The Oral Microbiota in Health and Disease: An Overview of Molecular Findings », *Methods Mol. Biol. Clifton NJ*, vol. 1537, p. 127-138, 2017, doi: 10.1007/978-1-4939-6685-1_7.

[32] D. S. Pontes, C. I. Lima-Bittencourt, E. Chartone-Souza, et A. M. Amaral Nascimento, « Molecular approaches: advantages and artifacts in assessing bacterial diversity », *J. Ind. Microbiol. Biotechnol.*, vol. 34, nº 7, p. 463-473, juill. 2007, doi: 10.1007/s10295-007-0219-3.

[33] A. Santiago *et al.*, « Processing faecal samples: a step forward for standards in microbial community analysis », *BMC Microbiol.*, vol. 14, nº 1, p. 112, mai 2014, doi: 10.1186/1471-2180-14-112.

[34] A. Acharya, T. Chen, Y. Chan, R. M. Watt, L. Jin, et N. Mattheos, « Species-Level Salivary Microbial Indicators of Well-Resolved Periodontitis: A Preliminary Investigation », *Front. Cell. Infect. Microbiol.*, vol. 9, p. 347, 2019, doi: 10.3389/fcimb.2019.00347.

[35] B. Adhikari, S. W. Kim, et Y. M. Kwon, « Characterization of Microbiota Associated with Digesta and Mucosa in Different Regions of Gastrointestinal Tract of Nursery Pigs », *Int. J. Mol. Sci.*, vol. 20, nº 7, avr. 2019, doi: 10.3390/ijms20071630.

[36] A. C. Freitas, A. Bocking, J. E. Hill, D. M. Money, et VOGUE Research Group, « Increased richness and diversity of the vaginal microbiota and spontaneous preterm birth », *Microbiome*, vol. 6, nº 1, p. 117, juin 2018, doi: 10.1186/s40168-018-0502-8.

[37] M. Aira, M. Pérez-Losada, et J. Domínguez, « Diversity, structure and sources of bacterial communities in earthworm cocoons », *Sci. Rep.*, vol. 8, nº 1, p. 6632, avr. 2018, doi: 10.1038/s41598-018-25081-9.

[38] I. Allali *et al.*, « A comparison of sequencing platforms and bioinformatics pipelines for compositional analysis of the gut microbiome », *BMC Microbiol.*, vol. 17, nº 1, p. 194,

292 sept. 2017, doi: 10.1186/s12866-017-1101-8.
293 [39] C. Angebault, A. Ghozlane, S. Volant, F. Botterel, C. d'Enfert, et M.-E. Bougnoux,
294 « Combined bacterial and fungal intestinal microbiota analyses: Impact of storage
295 conditions and DNA extraction protocols », *PloS One*, vol. 13, nº 8, p. e0201174, 2018,
296 doi: 10.1371/journal.pone.0201174.
297 [40] M. Bili *et al.*, « Bacterial Community Diversity Harboured by Interacting Species », *PloS*
298 *One*, vol. 11, nº 6, p. e0155392, 2016, doi: 10.1371/journal.pone.0155392.
299 [41] M. Caputo *et al.*, « Bacterial community structure and effects of picornavirus infection on
300 the anterior nares microbiome in early childhood », *BMC Microbiol.*, vol. 19, nº 1, p. 1,
301 janv. 2019, doi: 10.1186/s12866-018-1372-8.
302 [42] J. Chopyk *et al.*, « Presence of pathogenic Escherichia coli is correlated with bacterial
303 community diversity and composition on pre-harvest cattle hides », *Microbiome*, vol. 4,
304 p. 9, mars 2016, doi: 10.1186/s40168-016-0155-4.
305 [43] J. B. Clayton *et al.*, « Bacterial community structure and function distinguish gut sites in
306 captive red-shanked doucs (Pygathrix nemaeus) », *Am. J. Primatol.*, vol. 81, nº 10-11, p.
307 e22977, oct. 2019, doi: 10.1002/ajp.22977.
308 [44] L. Delhalle *et al.*, « Exploring the Bacterial Diversity of Belgian Steak Tartare Using
309 Metagenetics and Quantitative Real-Time PCR Analysis », *J. Food Prot.*, vol. 79, nº 2, p.
310 220-229, févr. 2016, doi: 10.4315/0362-028X.JFP-15-185.
311 [45] Z. Gao *et al.*, « Microbiota of Inflammatory Bowel Disease Models », *Annu. Int. Conf.*
312 *IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Annu. Int. Conf.*, vol. 2018, p.
313 2374-2377, juill. 2018, doi: 10.1109/EMBC.2018.8512848.
314 [46] R. N. Gibson, *Oceanography and Marine Biology, An Annual Review, Volume 39: An*
315 *Annual Review:* CRC Press, 2001.
316 [47] J.-J. Godon, P. Arulazhagan, J.-P. Steyer, et J. Hamelin, « Vertebrate bacterial gut diversity:

size also matters », *BMC Ecol.*, vol. 16, p. 12, mars 2016, doi: 10.1186/s12898-016-0071-2.

[48] M. Kolasa, R. Ścibior, M. A. Mazur, D. Kubisz, K. Dudek, et Ł. Kajtoch, « How Hosts Taxonomy, Trophy, and Endosymbionts Shape Microbiome Diversity in Beetles », *Microb. Ecol.*, vol. 78, nº 4, p. 995-1013, nov. 2019, doi: 10.1007/s00248-019-01358-y.

[49] X. Liu *et al.*, « Alterations of gastric mucosal microbiota across different stomach microhabitats in a cohort of 276 patients with gastric cancer », *EBioMedicine*, vol. 40, p. 336-348, févr. 2019, doi: 10.1016/j.ebiom.2018.12.034.

[50] Q. Liu *et al.*, « The Diversity of the Endobiotic Bacterial Communities in the Four Jellyfish Species », *Pol. J. Microbiol.*, vol. 68, nº 4, p. 465-476, déc. 2019, doi: 10.33073/pjm-2019-046.

[51] T. Lu *et al.*, « Altered Gut Microbiota Diversity and Composition in Chronic Urticaria », *Dis. Markers*, vol. 2019, p. 6417471, 2019, doi: 10.1155/2019/6417471.

[52] B. Ma *et al.*, « Altered Gut Microbiota in Chinese Children With Autism Spectrum Disorders », *Front. Cell. Infect. Microbiol.*, vol. 9, p. 40, 2019, doi: 10.3389/fcimb.2019.00040.

[53] G. Maskarinec *et al.*, « Fecal Microbial Diversity and Structure Are Associated with Diet Quality in the Multiethnic Cohort Adiposity Phenotype Study », *J. Nutr.*, vol. 149, nº 9, p. 1575-1584, sept. 2019, doi: 10.1093/jn/nxz065.

[54] V. Mattei *et al.*, « Evaluation of Methods for the Extraction of Microbial DNA From Vaginal Swabs Used for Microbiome Studies », *Front. Cell. Infect. Microbiol.*, vol. 9, p. 197, 2019, doi: 10.3389/fcimb.2019.00197.

[55] T. G. McDaneld, L. A. Kuehn, et J. W. Keele, « Microbiome of the upper nasal cavity of beef calves prior to weaning12 », *J. Anim. Sci.*, vol. 97, nº 6, p. 2368-2375, mai 2019, doi: 10.1093/jas/skz119.

342 [56] R. J. Pandit *et al.*, « Microbial diversity and community composition of caecal microbiota
343 in commercial and indigenous Indian chickens determined using 16s rDNA amplicon
344 sequencing », *Microbiome*, vol. 6, nº 1, p. 115, juin 2018, doi: 10.1186/s40168-018-0501-
345 9.
346 [57] A. N. Sarangi, A. Goel, A. Singh, A. Sasi, et R. Aggarwal, « Faecal bacterial microbiota
347 in patients with cirrhosis and the effect of lactulose administration », *BMC Gastroenterol.*,
348 vol. 17, nº 1, p. 125, nov. 2017, doi: 10.1186/s12876-017-0683-9.
349 [58] A. Sarkar, M. Stoneking, et M. R. Nandineni, « Unraveling the human salivary
350 microbiome diversity in Indian populations », *PloS One*, vol. 12, nº 9, p. e0184515, 2017,
351 doi: 10.1371/journal.pone.0184515.
352 [59] J. H. Savage *et al.*, « A prospective microbiome-wide association study of food
353 sensitization and food allergy in early childhood », *Allergy*, vol. 73, nº 1, p. 145-152, janv.
354 2018, doi: 10.1111/all.13232.
355 [60] J. Selvin *et al.*, « Culture-dependent and metagenomic analysis of lesser horseshoe bats'
356 gut microbiome revealing unique bacterial diversity and signatures of potential human
357 pathogens », *Microb. Pathog.*, vol. 137, p. 103675, déc. 2019, doi:
358 10.1016/j.micpath.2019.103675.
359 [61] J. M. Shikany *et al.*, « Association of dietary patterns with the gut microbiota in older,
360 community-dwelling men », *Am. J. Clin. Nutr.*, vol. 110, nº 4, p. 1003-1014, oct. 2019,
361 doi: 10.1093/ajcn/nqz174.
362 [62] T. A. Suzuki *et al.*, « Host genetic determinants of the gut microbiota of wild mice », *Mol.
363 Ecol.*, vol. 28, nº 13, p. 3197-3207, juill. 2019, doi: 10.1111/mec.15139.
364 [63] A. Ticinesi *et al.*, « Gut microbiota composition is associated with polypharmacy in
365 elderly hospitalized patients », *Sci. Rep.*, vol. 7, nº 1, p. 11102, sept. 2017, doi:
366 10.1038/s41598-017-10734-y.

[64] G. N. Tzanetakis *et al.*, « Comparison of Bacterial Community Composition of Primary and Persistent Endodontic Infections Using Pyrosequencing », *J. Endod.*, vol. 41, nº 8, p. 1226-1233, août 2015, doi: 10.1016/j.joen.2015.03.010.

[65] A. K. Vasquez *et al.*, « The microbiome of Escherichia coli and culture-negative nonsevere clinical mastitis: Characterization and associations with linear score and milk production », *J. Dairy Sci.*, vol. 102, nº 1, p. 578-594, janv. 2019, doi: 10.3168/jds.2018-15062.

[66] B. Wagner Mackenzie *et al.*, « Longitudinal study of the bacterial and fungal microbiota in the human sinuses reveals seasonal and annual changes in diversity », *Sci. Rep.*, vol. 9, nº 1, p. 17416, nov. 2019, doi: 10.1038/s41598-019-53975-9.

[67] R. E. Walker *et al.*, « Nasal microbial composition and chronic otitis media with effusion: A case-control study », *PloS One*, vol. 14, nº 2, p. e0212473, 2019, doi: 10.1371/journal.pone.0212473.

[68] Q. Wang *et al.*, « Oral Microbiome in Patients with Oesophageal Squamous Cell Carcinoma », *Sci. Rep.*, vol. 9, nº 1, p. 19055, déc. 2019, doi: 10.1038/s41598-019-55667-w.

[69] S. Wei *et al.*, « Short- and long-term impacts of azithromycin treatment on the gut microbiota in children: A double-blind, randomized, placebo-controlled trial », *EBioMedicine*, vol. 38, p. 265-272, déc. 2018, doi: 10.1016/j.ebiom.2018.11.035.

[70] L. Yang *et al.*, « Helicobacter pylori Infection Aggravates Dysbiosis of Gut Microbiome in Children With Gastritis », *Front. Cell. Infect. Microbiol.*, vol. 9, p. 375, 2019, doi: 10.3389/fcimb.2019.00375.

[71] X. Yang *et al.*, « The normal vaginal and uterine bacterial microbiome in giant pandas (Ailuropoda melanoleuca) », *Microbiol. Res.*, vol. 199, p. 1-9, juin 2017, doi: 10.1016/j.micres.2017.01.003.

[72] L. Yin *et al.*, « Association Between Gut Bacterial Diversity and Mortality in Septic Shock Patients: A Cohort Study », *Med. Sci. Monit. Int. Med. J. Exp. Clin. Res.*, vol. 25, p. 7376-7382, oct. 2019, doi: 10.12659/MSM.916808.

[73] G. Yu *et al.*, « The effect of cigarette smoking on the oral and nasal microbiota », *Microbiome*, vol. 5, n° 1, p. 3, janv. 2017, doi: 10.1186/s40168-016-0226-6.

[74] Z. Zhang, S. Jiao, X. Li, et M. Li, « Bacterial and fungal gut communities of Agrilus mali at different developmental stages and fed different diets », *Sci. Rep.*, vol. 8, n° 1, p. 15634, oct. 2018, doi: 10.1038/s41598-018-34127-x.

[75] G. Zhao *et al.*, « Gut Microbiome of Chinese Forest Musk Deer Examined across Gender and Age », *BioMed Res. Int.*, vol. 2019, p. 9291216, 2019, doi: 10.1155/2019/9291216.

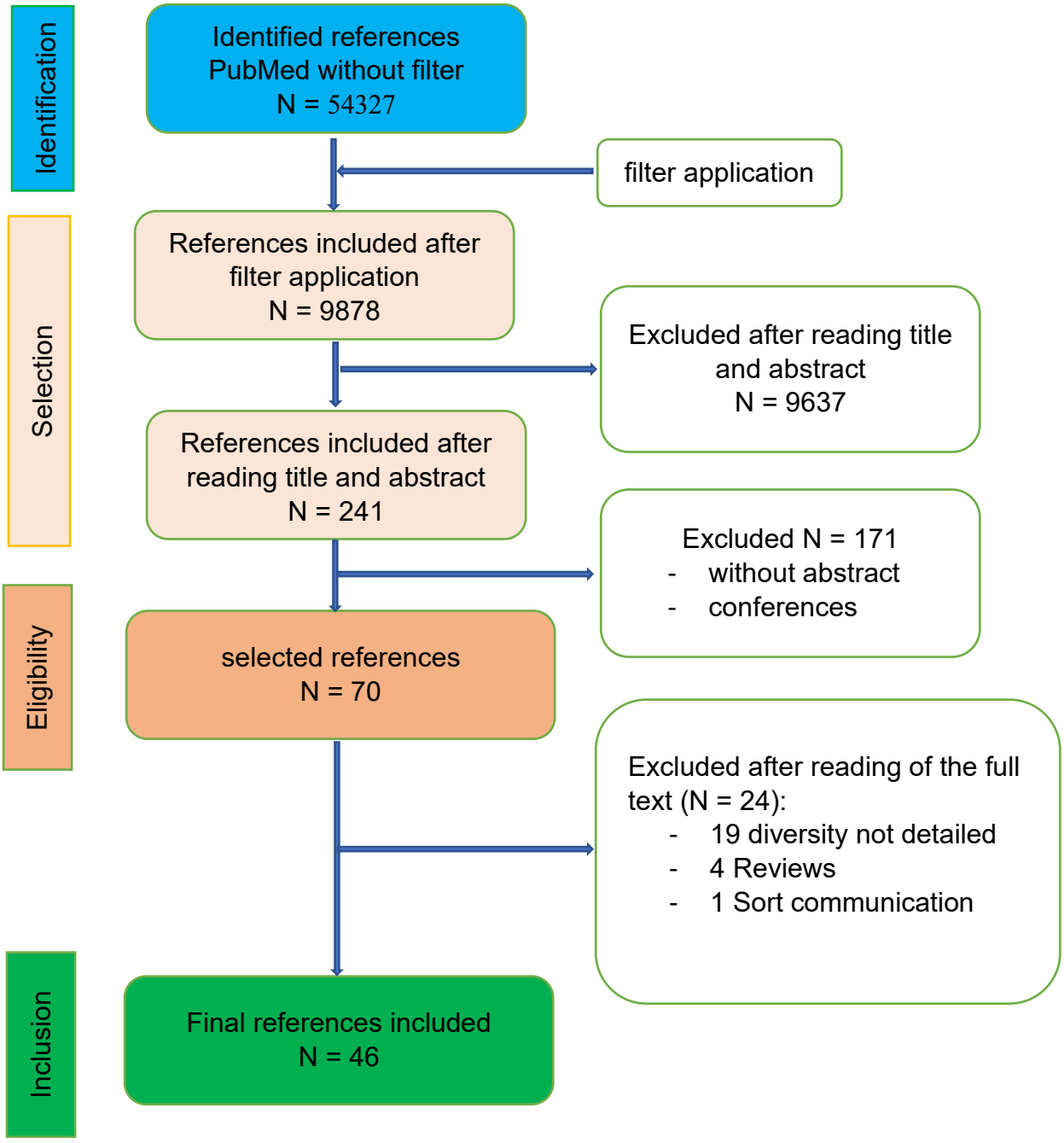


Figure 1: Flowchart of the selection process for articles on diversity indicators for synthesis

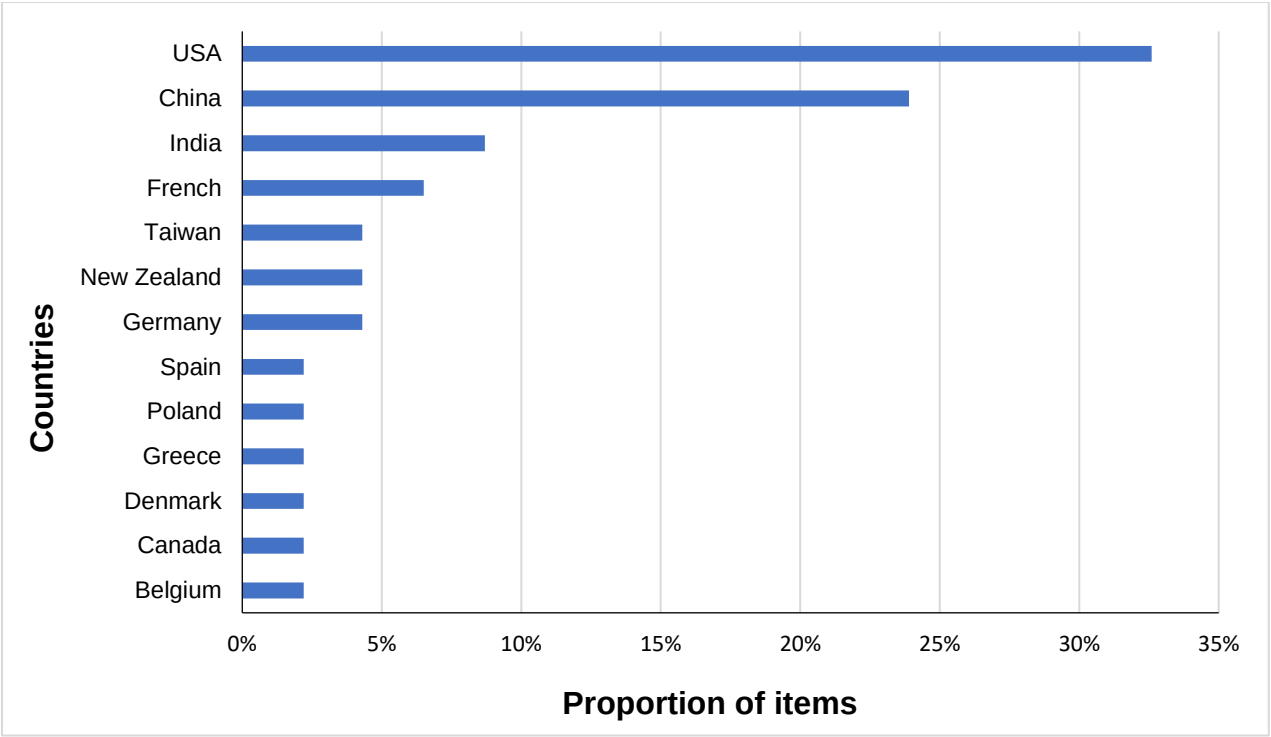


Figure 2: Distribution of selected items by country

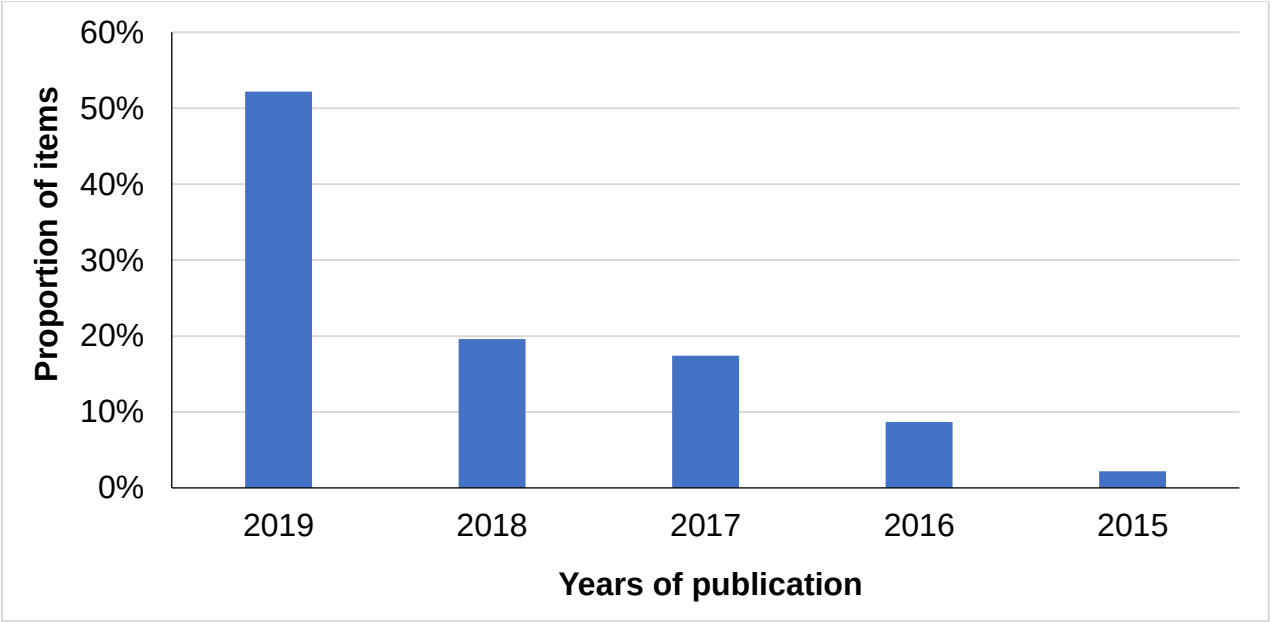


Figure 3: Proportion of articles selected by year of publication.

475 Table 1: Summary of search and selection criteria

Step 1: Topic Definition of the subject of the review	PICO method: - Population: Bacterial - Intervention: measurement of bacterial diversity - Comparator: none - Results: to know the indicators of evaluation of the bacterial diversity
Step 2: Parameters Definition of the parameters	Parameters according to STARLITE: - Sampling strategy: Selective search - Type of studies: Filtering of "Journal Articles" subcategories in PubMed - Approach: Electronic search - Range of years: choice to limit ourselves from January 1, 2015 to December 31, 2019 - Limits: English and French - Inclusion: The study must present types of diversity or indices for measuring diversity. The study must be an article. - Exclusion: the language of the study is not English or French or articles that do not present data on bacterial diversity. - Terms used without filter: (((bacterial AND (diversity OR biodiversity)) NOT (soil)) NOT (water)) NOT (air). With filter: (((bacterial

	AND (diversity OR biodiversity)) NOT (soil)) NOT (water)) NOT (air) AND ((ffrft[Filter]) AND (journalarticle[Filter]) AND (animal[Filter] OR humans[Filter]) AND (english[Filter] OR french[Filter]) AND (2015:2019[pdat])) - Electronic source: PubMed only
Step 3: Literature review Conducting the literature review according to the research parameters	The literature review was performed on PubMed on May 31, 2020. A total of 46 included articles were identified.
Step 4: Filtering Read the title and the summary to make a first sort according to the parameters of the search	The first filter was performed based on the title and abstract of the articles. We excluded articles that seemed less detailed, or short papers.
Step 5: Eligibility Get the complete texts of the articles. Saving in pdf Reading of articles and justification of excluded articles.	At this stage, we have excluded 24 articles including 19 articles which do not detail the measures of diversity, 4 reviews of the literature and a short communication.
Step 6: Inclusion Inclusion of the articles for their analysis	Finally, a total of 46 articles were included for analysis.
Presentation of the results	For the publication of this project, Figure 1 was used to document the steps of the literature review. The following text was used to describe the parameters of

	the search, following the STARLITE recommendations: «First, we listed journal articles presenting indices for measuring bacterial diversity. The literature search was performed on PubMed. Articles included had to be written in English or French, and published between January 1, 2015 and December 31, 2019. A selection of articles over the past five years allowed us to include recent studies that are more representative of current practice. The following search strategy was used: (((bacterial AND (diversity OR biodiversity)) NOT (soil)) NOT (water)) NOT (air). Articles presenting diversity types and diversity indices were selected based on title and abstract. From the full texts of the selected articles, literature reviews and a short communication were excluded to keep only journal articles. Studies where diversity measures are well explained were retained, but articles with poorly detailed diversity measures were excluded. »

476

477 PICO: Population, Intervention, Comparators, Outcomes

478 STARLITE: Standard for Reporting Literature searches

479

480 Table 2: Description of the articles included in the review.

N°	First author, Year [ref]	Countries	Identification methods (Sequencing)	Samples (n)	Population	Species
1	Acharya 2019 [34]	China	16S rRNA	Saliva (n=35)	Human	Human
2	Adhikari 2016 [35]	USA	16S rRNA	Digesta mucosal (n=10)	Animal	pigs
3	Freitas 2018 [36]	Canada	16S rRNA	Vaginal swabs (n=47)	Human	Human
4	Aira 2018 [37]	Spain	16S rRNA	Earthworm cocoons	Animal	earthworm
5	Allali 2017 [38]	USA	16S rRNA	Chicken cecum (n=36)	Animal	Chicken
6	Angebault 2018 [39]	French	16S rRNA	Fecal (n=5)	Human	Human
7	Ault 2019 [22]	USA	16S rRNA	Uterine and Vaginal (n=68)	Animal	Cow
8	Bili 2016 [40]	French	16S rRNA	Fly (n=25)	Animal	Fly
9	Caputo 2019 [41]	Germany	16S rRNA	Nasal swabs (n=76)	Human	Human
10	Chen 2017 [15]	Taiwan	16S rRNA	Environment (n=203)	Human	Human
11	Chopyk 2016 [42]	USA	16S rRNA	Feces (n=576)	Animal	Cattle
12	Clayton 2019 [43]	USA	16S rDNA	Feces (n=6)	Animal	Bird
13	Delhalle 2016 [44]	Belgium	16S rRNA	Steak tartare (n=58)	Animal	Beef
14	Gao 2018 [45]	USA	16S rRNA	Feces (n=13)	Animal	Rats
15	Gibson 2019 [46]	USA	16S rRNA	Feces (n=100)	Animal	Rhinoceros
16	Godon 2016 [47]	French	16S rRNA	Feces (n=189)	Animal	Mammals, birds, and reptiles
17	Kolasa 2019 [48]	Poland	16S rRNA	Beetles (n=24)	Animal	Beetles
18	Liang 2017 [28]	Taiwan	16S rRNA	Fecal (n=181)	Human	Human
19	Liu X 2019 [49]	China	16S rRNA	Stomach tissues (n=276)	Human	Human
20	Liu Q 2019 [50]	China	16S rRNA	Jellyfish (n=4)	Animal	Jellyfish
21	Lu 2019 [51]	China	16S rRNA	Fecal (n=10)	Human	Human
22	Ma 2019 [52]	China	16S rRNA	Fecal	Human	Human

				(n=45)		
23	Maskarinec 2019 [53]	USA	16S rRNA	Fecal (n=1735)	Human	Human
24	Mattei 2019 [54]	USA	16S rDNA	Vaginal swabs (n=5)	Human	Human
25	McDaneld 2019 [55]	USA	16S rRNA	Nasal swabs (n=1614)	Animal	Beef
26	Pandit 2018 [56]	India	16S rRNA	Feces (n=20)	Animal	Chicken
27	Prehn-Kristensen 2018 [17]	Germany	16S rRNA	Fecal (n=31)	Human	Human
28	Sarangi 2017 [57]	India	16S rRNA	Fecal (n= 53)	Human	Human
29	Sarkar 2017 [58]	India	16S rRNA	Saliva (n=92)	Human	Human
30	Savage 2018 [59]	USA	16S rRNA	Fecal (n=216)	Human	Human
31	Selvin 2019 [60]	India	16S rRNA	Gut (n= 3)	Animal	Bats
32	Shikany 2019 [61]	USA	16S rRNA	Gut (n=517)	Human	Human
33	Suzuki 2019 [62]	USA	16S rRNA	Gut (n=50)	Animal	Mouse
34	Ticinesi 2017 [63]	Italy	16S rRNA	Fecal (n=76)	Human	Human
35	Tzanetakis 2015 [64]	Greece	16S rRNA	Teeth (n=48)	Human	Human
36	Vasquez 2019 [65]	USA	16S rRNA	Milk	Animal	Cow
37	Wagner 2019 [66]	New Zealand	16S rRNA	Nasal swabs (n=4)	Human	Human
38	Walker 2019 [67]	New Zealand	16S rRNA	Nasal swabs (n=178)	Human	Human
39	Wang 2019 [68]	China	16S rRNA	Oral (n=41)	Human	Human
40	Wei 2018 [69]	Denmark	16S rRNA	Fecal (n=116)	Human	Human
41	Yang L 2019 [70]	China	16S rRNA	Fecal (n=154)	Human	Human
42	Yang X 2017 [71]	China	16S rRNA	Vaginal swabs (n=11)	Animal	Panda
43	Yin 2019 [72]	China	16S rDNA	Fecal (n=150)	Human	Human
44	Yu 2017 [73]	USA	16S rRNA	Plaque nasal swabs (n=43)	Human	Human
45	Zhang 2018 [74]	China	16S rRNA	Gut	Animal	Insect
46	Zhao 2019 [75]	China	16S rRNA	Fecal (n=20)	Animal	Deer

481

482 Table 3: Identified indicators for measuring diversity.

Types of diversity n (%)	Evaluation Indicators	Frequency of articles using these indicators (%)
Alpha 46 (100)	Richness (observed and estimate)	42 (91.3)
	Relative abundance	36 (78.6)
	Shannon index	39 (84.8)
	Simpson and invsimpson	15 (32.6)
	Evenness	4 (8.7)
	Rarefaction curve	11 (23.9)
Beta 40 (87.0)	Bray-Curtis dissimilarity distances	27 (67.5)
	Weighted and unweighted UniFrac distances	11 (27.5)
	Jaccard distances	3 (7.5)
	Sorensen distances	2 (5.0)
	Indicator species	2 (5.0)
	Visualization PCoA	26 (65.0)
	Visualization NMDS	2 (5.0)
Genetic 42 (91.3)	Taxonomic	13 (28.6)
	Phylogenetic	20 (43.5)
	Taxonomic and phylogenetic	9 (19.6)

483

484

485

CHAPITRE II

Dynamique de la diversité bactérienne dans la région Provence-Alpes-Côte d'Azur (PACA)

Préambule

Les maladies infectieuses représentent un problème majeur de santé publique dans le monde, où environ 43% des décès dans les pays les plus pauvres sont dus aux agents infectieux (25). En 2017, le groupe d'étude Global Burden of Disease (GBD) a indiqué que plus de 50 agents pathogènes étaient responsables de 6,2 millions de décès dans le monde et 18,6% de la mortalité étaient imputables aux maladies transmissibles (26). Une gamme très variée de ces maladies transmissibles relève d'une grande diversité des agents infectieux bactériens et viraux.

L'étude de la diversité est restée longtemps orienté sur la diversité des macro-organismes (végétaux et animaux). Cependant, les laboratoires de microbiologie des hôpitaux et les cliniques publics et privés mettent en évidence un nombre important de microbes dont la diversité des agents identifiés est très peu documentée (13). Un autre aspect important de l'étude de la diversité des espèces bactériennes est la nature de l'écosystème. Les communautés bien organisées qui présentent un certain niveau de diversité sont stables (15,27). Si une sorte de stress est introduit dans cette communauté, la stabilité peut s'effondrer et la diversité changera.

Une première analyse du système de biosurveillance basée sur les données de surveillance hebdomadaire de 210 laboratoires de microbiologie dans la région Provence-Alpes-Côte d'Azur (région PACA) en 2016, effectuée par Michael Huart (28) a mis en évidence 673 espèces bactériennes sur 611 espèces surveillées par le système, du 01 juillet 2013 au 20 mars 2016. Le système a détecté 62 nouvelles espèces bactériennes. Cependant, aucune étude de la variabilité des espèces dans les laboratoires n'a été menée sur les données de surveillance en microbiologie clinique des laboratoires membres de ce système de biosurveillance. Alors nous nous sommes interrogés est-il possible de compléter la surveillance épidémiologique par de nouveaux indicateurs basés sur l'évolution temporelle de la diversité ou du mix bactérien ?

Sur la base de ces études préliminaires, nous abordons dans cette thèse, une analyse de diversité des espèces bactériennes identifiées dans les laboratoires de microbiologie clinique de la région PACA, en vue d'une identification de nouveaux indicateurs potentiels de surveillance épidémiologique basés sur des indices de diversité.

Article 2

Article 2: Analysis of bacterial diversity in clinical microbiology laboratories affiliated to the epidemiological surveillance system of the PACA region (PACASurvE) from 2013 to 2019, France

Lanceï KABA, Philippe COLSON, Hervé CHAUDET and Didier Raoult

En cours de correction

TITLE PAGE

Analysis of bacterial diversity in clinical microbiology laboratories affiliated to the epidemiological surveillance system of the PACA region (PACASurvE) from 2013 to 2019, France

Author list: Lanceï KABA[1,2,3], Hervé CHAUDET[1,2], Audrey GIRAUD-GATINEAU[1,2], Philippe COLSON[1,2]

[1] IHU Méditerranée Infection, 19-21 boulevard Jean Moulin, 13005 Marseille, France ;

[2] Aix Marseille Univ., Institut de Recherche pour le Développement (IRD), Assistance

[3] Institut Supérieur des Sciences et de médecine vétérinaire (ISSMV) de Dalaba, Guinée ;

*** Corresponding author:** Hervé CHAUDET, IHU Méditerranée Infection, 19-21 boulevard Jean Moulin, 13005 Marseille, France. E-mail: herve.chaudet@gmail.com

Abstract

The study of diversity is important for understanding the modification of the ecosystem. The aim of this work was to assess and compare the bacterial diversity of microbiology laboratories in the PACA region from 2013 to 2019. To this end, the laboratories were organized into two types either in a hospital center (CH) or in a medical biology laboratory (LBM), then we determined the observed richness and the relative abundance of each species before comparing the medians. For the analysis of the similarity, the laboratories were grouped into three groups according to the Bray-Curtis distance and the similarity was evaluated by the ANOSIM test. A total of 765 species were identified with 1 112 201 bacterial isolates. LBMs had higher diversity in richness ($Avg_{(LBM)}$=257.5(222.5-265.8); $Avg_{(CH)}$ =140(80-233); p-value =0.04) and in abundance ($Mean_{(LBM)}$=74617(54003-104623); $Avg_{(CH)}$ =9909(2959-16097); p-value = 0.002) that hospital centers. However, the Shannon (p-value=0.55) and Simpson (p-value=0.27) indices were similar between these laboratory types. Moreover, the laboratory groups were distinct from one another in terms of richness ($Avg_{(G1)}$=247.4±9.2 : $Avg_{(G2)}$=478±16.8 ; $Avg_{(G3)}$=99.8±11.2 ; p-value=$1.3*10^{-11}$), but also in terms of

abundance ($Mean_{(G1)}$=53473±11884.6 ; $Mean_{(G2)}$=163247.3±21698.3 ; $Mean_{(G3)}$=7310,8±1752.7 ; p-value=0.0001). The similarity study showed that the groups were dissimilar to each other (R=0.57; p-value = 0.001). Identification of specific and common species would be necessary to characterize the distribution of species by type and laboratory group.

Key words: Bacterial diversity, Similarity, Shannon index, Simpson index, PACA

I. Introduction

Infectious diseases represent a major public health problem in the world. They are one of the main causes of morbidity and mortality, including in developed countries. A study by the Global burden group in 2017, indicates a mortality of 18.6% related to communicable diseases [1]. In France, they represent 5% of all-cause mortality according to Santé Publique France [2]. The epidemiology of bacterial infections remains to be better monitored in France.

Several epidemiological surveillance systems exist today in France that collaborate with microbiology laboratories. These surveillance systems either conduct syndromic surveillance or surveillance of well-known infectious agents that pose a major public health threat or both. In this paper, we look at the changes in the bacterial mix through a study of the diversity of bacterial species identified by microbiology laboratories in the PACA region.

Biodiversity being a variability of living organisms of any origin [3], i.e. the range of very different types of organisms and their relative abundance in a community [4, 5]. Microbial diversity unequivocally refers to biological diversity at three levels: within species (genetic), species number (species) and community (ecological) [6, 7]. Diversity studies are important for increasing understanding of the diversity of genetic resources and understanding the distribution of organisms, identifying differences in diversity associated with disruptive management, understanding the regulation of biodiversity, and understanding the consequences of biodiversity. The term species diversity has two components; the first component is the total number of species present, which can

be referred to as species richness. The second component is the distribution of individuals among these species, called evenness or equity.

The study of diversity could therefore be used to monitor successions and the effects of disturbances, but also to detect the emergence of new pathogens likely to trigger a potential epidemic or to identify unknown agents. The epidemiological surveillance system of the Provence Alpes-Côte d'Azur region (PACASurvE) is a system that has been in operation since 2013 [8] and collaborates with more than 300 laboratories which are hospitals or medical biology laboratories, aiming to specifically monitor infections of bacterial origin [9, 10], accordingly, the system is designed to issue alarms if an outbreak is detected or if a single case of a rare but serious infectious disease or an unknown infectious agent is discovered.

Today, the diversity of infectious agents identified by epidemiological surveillance systems has not been studied at all. In addition, knowledge of the diversity of species and their specificity according to geographical areas is interesting for understanding their evolution over time.

The objective of this paper was to describe bacterial diversity in microbiology laboratories with the usual tools of quantitative ecology to understand how this diversity would contribute to epidemiological surveillance indicators for early detection of potential epidemics.

II. Methods

1. Study area and laboratories

The Provence-Alpes-Côte d'Azur (PACA) region is a region in southeastern France, north of Corsica and east-northeast of the Occitanie region. It is bordered to the north by the Auvergne-Rhône-Alpes region and to the west by the Occitanie region, with the Rhône River marking the regional border. PACA is made up of six departments from the former provinces of Provence and Dauphiné (Figure 1).

This study is a longitudinal retrospective study on the bacterial diversity identified in the laboratories of epidemiological surveillance of the PACA region from 01/07/2013 to 25/11/2019. Laboratories send bacterial species identification results weekly by email, which are collected and uploaded to the server of the MIDaS database of the Marseille Institut Hospitalo-Universitaite (IHU). The laboratories have been categorized into hospital centers (CH) and medical biology laboratories (LBM).

The composition of the bacterial communities identified by the laboratories was evaluated using diversity indices: specific richness, relative abundance, Shannon's diversity index and Simpson's index [11]. The Shannon index takes into account the richness and uniformity of species [12]. Bacterial diversity in terms of species richness (number of species observed), relative abundance (number of isolates of each species), Shannon and Simpson diversity to determine alpha diversity and dissimilarity between laboratory groups based on the Bray-Curtis distance matrix for beta diversity. This similarity was visualized via the method of non-metric multidimensional positioning (NMDS) and principal coordinate analysis (PCoA).

The Shannon index (evenness) was used to characterize the specific richness and distribution of isolates within these species. The Shannon index is a positive real number generally between 0 and 5 but having no maximum value in theory. The higher the value of the index, the better the isolates are well distributed, and therefore the greater the diversity.

The Simpson index was used to determine the concentration of species, as well as the abundance of each species [13]. The Simpson's index (D) is a representation of the probability that two individuals, in the same sample and chosen at random, belong to the same species. The range of Simpson's index is from 0 to 1, but for an easier interpretation we have determined the complement of Simpson's index (1-D). The index value also oscillates between 0 and 1, the higher the value, the greater the diversity of the sample.

2. Data analysis

For the analysis of species richness and relative abundance, we applied the method of Chao1 [14] and the concentration of isolates for each species by the Simpson index. Indicator species were identified using the multipatt function of the "indicspecies" package [15]. The dissimilarity (or similarity) analysis was performed by grouping the laboratories into three groups according to their similarity based on the Bray-Curtis distance [16, 17] and the groups were visualized using the ordination method which is an NMDS method [18].

3. Statistical analysis

Wilcoxon test was used as a test to compare the medians by type of laboratory. As for the constituted groups, we used the ANOVA and TUKEY tests to compare means of specific richness and the Shapiro and Bartlett tests to check normality and homogeneity respectively. However, the means of the relative abundances of the isolates were compared by the Kruskal Wallis test and multiple comparison with the bonferroni method using Student's pairwise function. The similarity of the groups was tested by the ANOSIM (Analysis of similarity) function of the "vegan" package. The ANOSIM statistic compares the mean of ranked dissimilarities between groups to the mean of ranked dissimilarities within groups. An R-value close to "1.0" suggests dissimilarity between groups, while an R-value close to "0" suggests a uniform distribution of high and low ranks within and between groups. R values less than "0" suggest that dissimilarities are greater within groups than between groups [19].

Analyzes were performed using Rstudio software [20] with the "vegan" and "iNEXT" packages. The differences were considered significant below the threshold $\alpha = 0.05$.

III. Results

1. Characterization of the laboratory groups of the PACASurvE system

The PACASurvE surveillance system laboratories are located in the six (6) departments of the PACA region with a very variable number of sites per grouping. BIOESTEREL laboratory alone is spread over 83 sites, the best represented in the region, followed by LABOSUD with 58 sites. All the general hospital centers are represented by only one site except the center of MARTIGUES. Among the six (6) departments, the Bouches-du-Rhône department has the largest number of laboratories (60%) (15/25) and the 40% are distributed among the other departments. The volume of activity carried out, i.e., the number of samples analyzed in these laboratories, is highly variable by laboratory group. APHM group with only four (4) sites has 36.9% of activity volume, almost triple the volume of activity of BIOESTEREL (12.5%) for being represented by 83 sites. We also note very low volumes of activity, less than 1% in 40% (10/25) of the laboratories, all general hospitals. Private medical laboratories (LBMP) have more or less a larger volume of activity (table 1). Overall, the average activity was 238853 ± 90674 samples analyzed

2. Alpha diversity

- **Species richness and relative abundance of bacterial isolates**

A total of 1,112,201 bacterial isolates were obtained corresponding to 765 bacterial species. In terms of number of species, the APHM laboratory was richer than all the laboratories with 562 species, followed by the BIOESTEREL laboratory with 498 species. As for the relative abundance of isolates, the APHM laboratory also showed a high diversity with 16.12% contrary to the RSABRAN laboratory which recorded the lowest diversity in abundance with 0.12% (table 2). Five laboratories (APHM, BIOESTEREL, NICE, LABAZUR_NICE and LABOSUD) out of 25 provided more than 100,000 isolates, i.e. 63.5% of the overall relative abundance.

Regarding the observed species richness and relative abundance by laboratory type, we observed the medians in richness of 140 (80 – 233) and 257.5 (222.5 – 265.8) (p-value = 0.04) for hospital centers (CH) and medical biology laboratories (LBM) and the median abundances of 9909 (2959 – 16097) for CH and 74617 (54003 – 104623) for LBM (p-value = 0.002). Medical biology laboratories turn out to be more diversified (Figure 2 (A) and (B)) in terms of richness and abundance than those of hospitals, although the number of LBMs is half that of CHs.

The second aspect of diversity addressed in this study, which concerns the distribution of isolates between species, was shown for each of the laboratories by the diversity of Shannon. Thus, according to this diversity, the laboratories of APHM (Shannon = 3.27) and STJOSEPH (Shannon = 3.03) proved to be more diversified than the other laboratories. Although the STJOSEPH laboratory was 2 times less species rich than the BIOESTEREL laboratory (Shannon = 2.40), but its Shannon index indicates a better species frequency distribution, which means that it is more diverse than the BIOESTEREL laboratory. On the other hand, the least diverse was the RSABRAN laboratory (Shannon = 1.62). However, this Shannon diversity is dependent on the sample size. The high diversity observed in the APHM laboratory by the Shannon index was confirmed by the Simpson index (Simpson = 0.91) and the least was the HYERES laboratory (Simpson = 0.62).

The analysis by type of laboratory of the Shannon and Simpson indices (Figure 2 (C) and (D)) shows that there is no significant difference in diversity according to the Shannon index (p-value = 0.5) between hospital laboratories and medical laboratories and compared to the Simpson's index (p-value = 0.3). For Simpson's diversity, a site is more diverse if the complement of its index is closer to 1 and this illustrates a more important combination of species and an equitable distribution of isolates between species. In this logic, the APHM site (Simpson = 0.91) is the most diversified and the HYERES site (Simpson = 0.62) the least diversified. The three (3) most frequent specific species were presented in Figure 3.

Looking at the distribution of specific richness and relative abundance by laboratory grouping which are, among others, university hospital centers (CHU), city laboratories (LABOVILLE), mixed laboratories (Mix), general hospital centers (CHG) and the Armed Forces Training Hospital (HIA), the CHUs and the LABOVILLEs show greater diversity in species and abundance than the other laboratory structures (Figures 4 and 5). The two laboratories that showed more diversity in species and in abundance were the laboratories of the CHU of APHM with 562 species observed for 179,307 bacterial isolates, i.e., 16.12% relative abundance and the group of city laboratories of BIOESTEREL with 498 species and 178069 bacterial isolates. The laboratories of the general hospital centers generally identified less species although representative in terms of number (13 laboratories out of 25). This analysis allowed us to observe 88 species specific to the APHM laboratory group, 67 species specific to the BIOESTEREL laboratory and 12 species specific to the Nice University Hospital. However, seven (7) laboratories had no specific species and eight (8) had only one specific species (Figure 6). Looking at the dynamics of diversity over time, we found that the abundance of identified species increased from year to year (Figure 7) and new species were identified.

The spatial analysis of the species was carried out considering the department where the laboratories were located. PACA region has six (6) departments including Bouches-du-Rhône (BdR), Var (VAR), Alpes-Maritimes (AM), Alpes-de-Haute-Provence (AHP), Vaucluse (VAU) and Hautes-Alpes (HA). The results showed a strong identification of species in the department of Bouches-du-Rhône, with 663 species among which 150 have been identified only in this department, followed by the department of Var with 551 species. The department of Vaucluse recorded fewer identifications (91 species) than all other departments (Figure 8), however, the department of Hautes-Alpes does not have a laboratory included in the surveillance system.

3. Beta diversity

- **Similarity between laboratory groups**

Based on the measurement of the Bray-Curtis distance, three groups of laboratories were formed (Figure 9) according to their similarity, including group 1 composed of 10 laboratories, group 2 of 3 laboratories and group 3 of 12 laboratories. The average richness in the groups were respectively 247.4 ± 9.2, 478 ± 16.8 and 99.8 ± 11.2. The means of specific richness in these three groups differ significantly from one group to another (p-value = 1.3*10-11). Group 2 shows a high diversity in species compared to the other groups.

With regard to the average abundances, we observed averages of 53473 ± 11884.6, 163247.3 ± 21698.3 and 7310.8 ± 1752.7 respectively for groups I, II and III. Also, these abundance means differ significantly (p-value = 0.0001) from one group to another and group 2 was the richest in abundance. Visualization of the clusters using non-metric multidimensional positioning (NMDS) and principal coordinate analysis allowed us to construct a cluster three map of laboratories (Figure 10), in which the more similar two laboratories are in terms of abundance or biomass, the closer they are to each other on the map.

The average of the ranked dissimilarities between groups to the average of the ranked dissimilarities within groups (Figure 11), indicates a value of R = 0.6 and p-value = 0.001. Under the null hypothesis, the ranges of (ranked) dissimilarities within groups are equal, or at least very similar, so the p-value shows dissimilarity within groups.

The other important aspect of the study of diversity is the knowledge of indicator species, in our case species that are specific to each laboratory, in other words species that are only identified by one laboratory. Indicator species are species used as ecological indicators of community or habitat types, environmental conditions, or environmental changes.

IV. Discussion

In this present study, we evaluate and compare bacterial diversity from microbiology laboratories affiliated with the PACA region bacteria surveillance system (PACASurvE) [8, 9]. The study of diversity in microbiology laboratories would improve the early detection of possible epidemics in the areas of origin of the samples based on the richness and abundance of the identified species but also of rare species with epidemic potential. The medical biology laboratories (LBM) have shown to be richer in species and abundance than the hospital centers (CH) (Figure 2). This could be explained by the greater number and the proximity of these LBMs to the populations. However, considering the species richness and uniformity of the species and the concentration of isolates, these two types of laboratories were similar (Figure 2). The grouping of laboratories on the dissimilarity of Bray-Curtis showed that group 2 composed of only three laboratories from APHM, NICE (hospital centers) and BIOESTEREL (medical biology laboratory) had the highest diversity in terms of richness in species just as in abundance. This richer diversity of this group would be due to a more intense level of activity in these laboratories, certainly related to their location in denser geographical areas on the one hand and on the other hand by their equipment in detection and identification material of last generation. Dissimilarity (or similarity) between groups (Figure 9) grouped the laboratories into three groups sharing roughly the same species.

Previous studies of bacterial diversity were performed on different parts of the human body including the oral microbiome of the elderly where the results showed that alpha diversity did not exist significantly different between carious and non-caries patients [21], the gut microbiota [22] where the observed specific richness decreases significantly with stool firmness ($p=0.0007$).

In our study, alpha diversity varied significantly between laboratories, however, considering both types of laboratories, hospitals (CH) and medical laboratories (LBM) did not show a significant difference in Shannon diversity ($p = 0.55$) and Simpson diversity ($p = 0.27$). Ten species out of 765 were observed by all laboratories (*Aerococcus urinae*, *Citrobacter freundii*, *Citrobacter koseri*,

244 *Staphylococcus aureus*, *Staphylococcus epidermidis*, *Staphylococcus hominis*, *Streptococcus*
245 *anginosus*, *Streptococcus gallolyticus* and *Streptococcus pneumoniae*). In contrast, some species
246 were identified by only one laboratory (Table 3 in Supplementary Data). Among these specific
247 species considered as indicator species [23], the APHM laboratories alone detected 88 species,
248 followed by the group of BIOESTEREL laboratories with 67 species. This could be justified by the
249 volume of activity carried out in these two laboratories. The analysis of beta diversity by the method
250 of study of similarity between the groups formed based on Bray-Curtis's distance and the analysis
251 of principal coordinates (PCoA), showed a significant dissimilarity between the groups.

252 The strength of this study lies in the fact that several diversity assessment methods were applied to
253 better understand the diversity in the laboratories. Our work was able to highlight for the first time
254 a significant and specific bacterial diversity in microbiology laboratories and a possible introduction
255 of diversity indices to complete the indicators of epidemiological surveillance systems.

256 It is also noted certain weaknesses related to the restriction only on infections of bacterial origin. It
257 goes without saying that it should also be extended to infections of viral origin which lead to
258 epidemics or pandemics with more serious forms.

259 In short, molecular biology laboratories identified more species and isolates than hospital center
260 laboratories. However, taken individually, the bacterial diversity was richer in the laboratories of
261 the APHM hospital center. Group 2 made up of the APHM, NICE and BIOESTEREL laboratories
262 identified more species than the other two groups. These results may guide the choice of inclusion
263 of laboratories in a future program of epidemiological surveillance based on diversity. Diversity
264 indices such as species richness, relative abundance, Shannon index and Simpson index could be
265 used to complement epidemiological monitoring indicators.

266 However, not all aspects could be addressed in this thesis, so it would be imperative for future
267 research to extend these studies to infections of viral origin and to edapho-climatic factors in the
268 areas from which patients come. Within the framework of cooperation, strengthen the partnership

with researchers from southern countries for active research to improve the surveillance systems in these countries where epidemics impose a heavy price on the population deprived of adequate healthcare structures and faced with extreme poverty.

V. References

1. Global Burden of Disease. Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980–2017: a systematic analysis for the Global Burden of Disease Study 2017. The Lancet. 2018;392:1736–88.

2. Lefèvre H, Pavillon G, Le Toullec A, Péquignot F, Jougla E. Mortalité par maladies infectieuses en France. Situation actuelle et tendances évolutives. Surveillance nationale des maladies infectieuses. 2003.

3. Whittaker RH. EVOLUTION AND MEASUREMENT OF SPECIES DIVERSITY. TAXON. 1972;21:213–51.

4. Fakruddin Md, Mannan KSB, Andrews S. Viable but Nonculturable Bacteria: Food Safety and Public Health Perspective. ISRN Microbiol. 2013;2013:1–6.

5. Torsvik V, Goksøyr J, Daae FL. High diversity in DNA of soil bacteria. Appl Environ Microbiol. 1990;56:782–7.

6. Harpole W. Neutral Theory of Species Diversity. Nat Educ Knowl. 2010;1:31.

7. Torsvik V, Daae FL, Sandaa R-A, Øvreås L. Novel techniques for analysing microbial diversity in natural and perturbed environments. J Biotechnol. 1998;64:53–62.

8. Huart M, Bedubourg G, Abat C, Colson P, Rolain JM, Chaudet H, et al. Implementation and Initial Analysis of a Laboratory-Based Weekly Biosurveillance System, Provence-Alpes-Côte d'Azur, France. Emerg Infect Dis. 2017;23:582–9.

291 9. Abat C, Chaudet H, Colson P, Rolain J-M, Raoult D. Real-Time Microbiology Laboratory
292 Surveillance System to Detect Abnormal Events and Emerging Infections, Marseille, France. Emerg
293 Infect Dis. 2015;21:1302–10.

294 10. Hulth A, Andrews N, Ethelberg S, Dreesman J, Faensen D, van Pelt W, et al. Practical usage of
295 computer-supported outbreak detection in five European countries. Euro Surveill Bull Eur Sur Mal
296 Transm Eur Commun Dis Bull. 2010;15.

297 11. Do TT, Delaney S, Walsh F. 16S rRNA gene based bacterial community structure of wastewater
298 treatment plant effluents. FEMS Microbiol Lett. 2019;366.

299 12. Hollenbeck JP, Ripple WJ. ASPEN AND CONIFER HETEROGENEITY EFFECTS ON BIRD
300 DIVERSITY IN THE NORTHERN YELLOWSTONE ECOSYSTEM. West North Am Nat.
301 2007;67:92–101.

302 13. Marcon E. Mesure de la biodiversitsé. 2018.

303 14. Chao A. Nonparametric Estimation of the Number of Classes in a Population. Scand J Stat.
304 1984;11:265–70.

305 15. Dufrêne M, Legendre P. SPECIES ASSEMBLAGES AND INDICATOR SPECIES:THE NEED
306 FOR A FLEXIBLE ASYMMETRICAL APPROACH. Ecol Monogr. 1997;67:345–66.

307 16. Ferrier S, Manion G, Elith J, Richardson K. Using Generalized Dissimilarity Modelling to
308 Analyse and Predict Patterns of Beta Diversity in Regional Biodiversity Assessment. Divers Distrib.
309 2007;13:252–64.

310 17. Gillett DJ, Mazor RD, Norton SB. Selecting comparator sites for ecological causal assessment
311 based on expected biological similarity. Freshw Sci. 2019;38:554–65.

312 18. Grall J, Coïc N. Synthèse des méthodes d'évaluation de la qualité du benthos en milieu côtier.

2006.

19. Clarke K, Warwick R. Clarke KR, Warwick RM.Change in Marine Communities: An Approach to Statistical Analysis and Interpretation. Primer-E Ltd: Plymouth, UK. 2001.

20. R Core Team. A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/. 2021.

21. Jiang Q, Liu J, Chen L, Gan N, Yang D. The Oral Microbiome in the Elderly With Dental Caries and Health. Front Cell Infect Microbiol. 2019;8:442.

22. Vandeputte D, Falony G, Vieira-Silva S, Tito RY, Joossens M, Raes J. Stool consistency is strongly associated with gut microbiota richness and composition, enterotypes and bacterial growth rates. Gut. 2016;65:57–62.

23. De Cáceres M, Legendre P, Moretti M. Improving indicator species analysis by combining groups of sites. Oikos. 2010;119:1674–84.

334 *Table 2: Specific richness and relative abundance of isolates observed in PACASurvE laboratory clusters,*

335 *2013 - 2019, PACA region*

N°	Laboratories	Type of laboratories	Estimated richness (Chao1±sd)	Observed species richness	Number of bacterial isolates	Relative abundance (%)
1	APHM	CH	665±26	562	179307	16.12
2	BIOESTEREL	LBM	630±32	498	178069	16.01
3	NICE	CH	449±22	374	132366	11.90
4	LABAZUR_N	LBM	309±12	279	115426	10.38
5	LABOSUD	LBM	354±26	274	101022	9.08
6	LABAZUR_P	LBM	217±9	194	75035	6.75
7	CERBALLIANCE	LBM	255±10	232	74198	6.67
8	ALPHABIO	LBM	343±20	282	64536	5.80
9	STJOSEPH	CH	325±18	273	34902	3.14
10	BARLA	LBM	153±15	120	22404	2.01
11	AIX	CH	355±34	245	20684	1.86
12	CASAMANCE	LBM	318±24	241	20449	1.84
13	LAVERAN	CH	299±20	233	16097	1.45
14	FREJUS	CH	295±25	221	12381	1.11
15	MARTIGUES	CH	187±16	145	10643	0.96
16	TOULON	CH	188±20	137	10208	0.92
17	GAP	CH	249±40	147	9909	0.89
18	AUBAGNE	CH	91±8	80	9806	0.88
19	SALON	CH	233±35	140	9577	0.86
20	HYERES	CH	94±12	71	3491	0.31
21	LACIOTAT	CH	72±5	64	2959	0.27
22	DIGNE	CH	162±20	114	2930	0.26
23	ARLES	CH	81±21	51	2409	0.22
24	PERTUIS	CH	141±22	91	2090	0.19
25	RSABRAN	CH	48±7	38	1303	0.12

336 *CH = Hospital Center; LBM = Laboratory of Medical Biology*

337

338

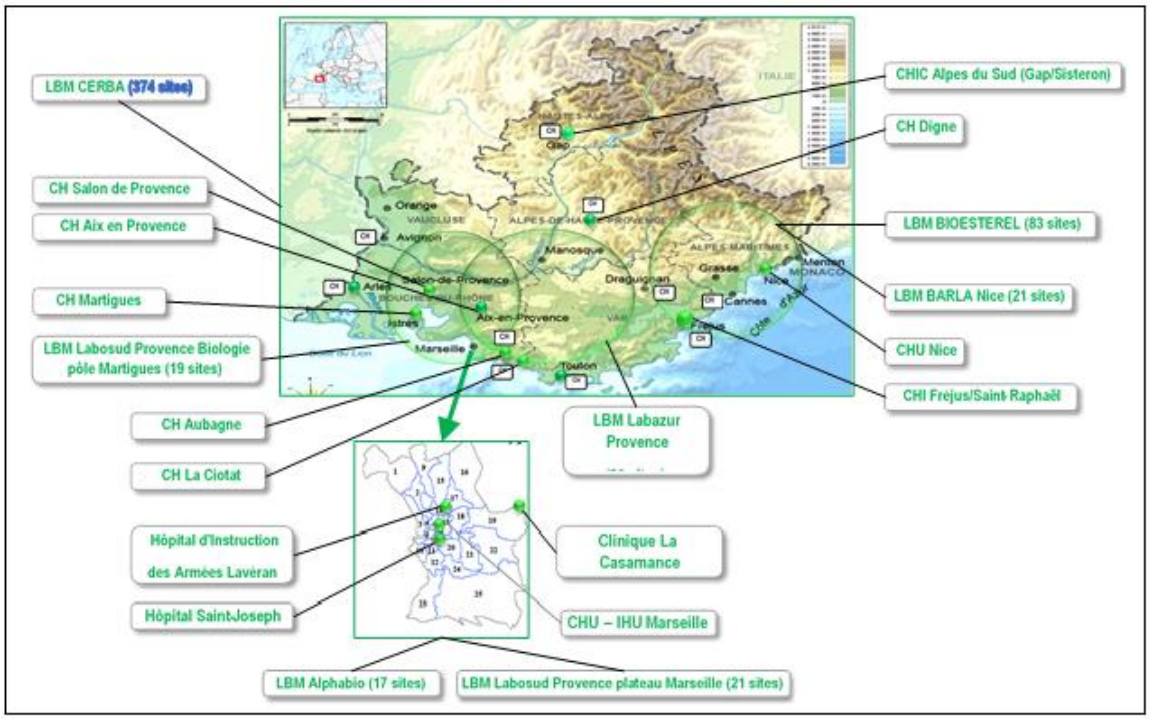


Figure 1: Participation in the PACA network of epidemiological surveillance of infections based on data from microbiology laboratories

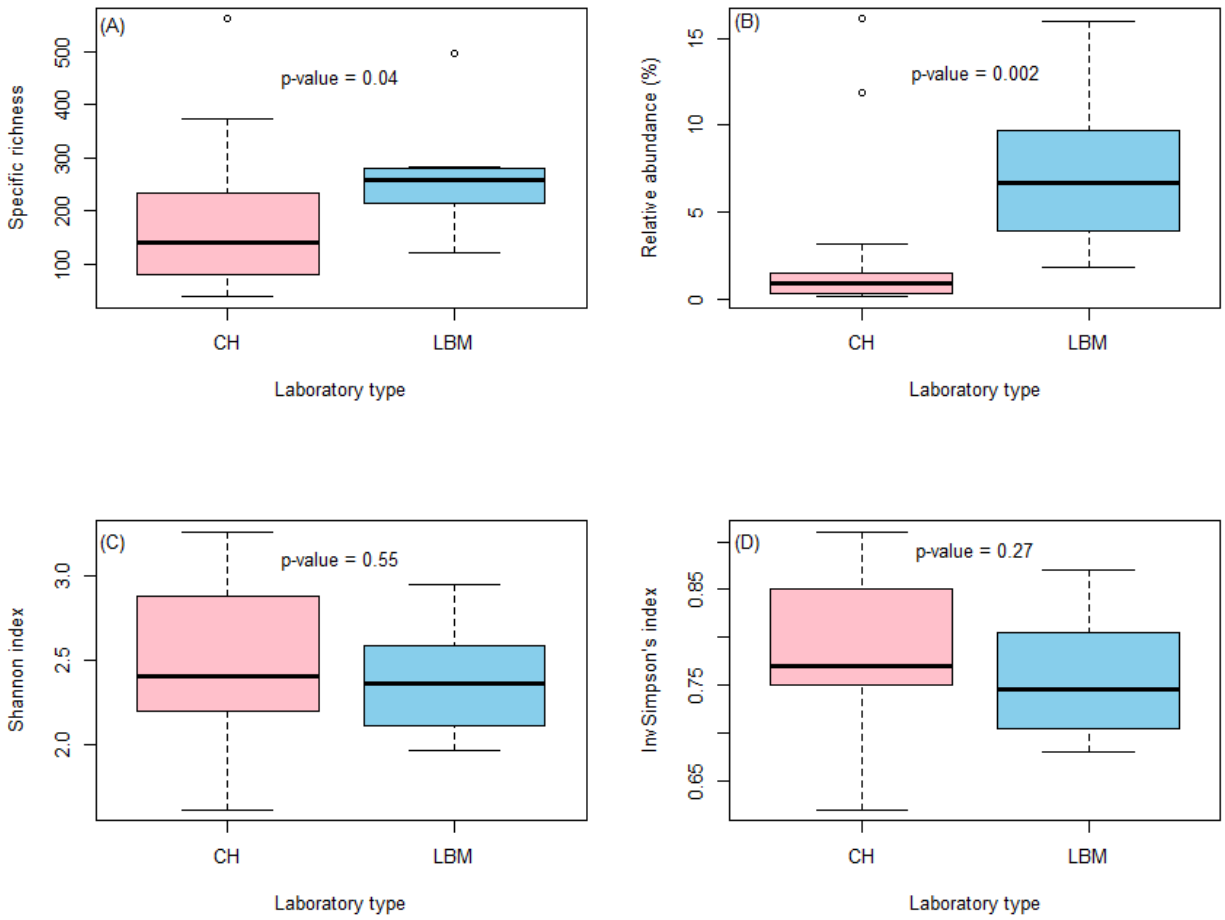


Figure 2: Descriptive analysis of species richness (A), relative abundance (B), Shannon index (C) and Inverse Simpson's index (D) laboratory type.

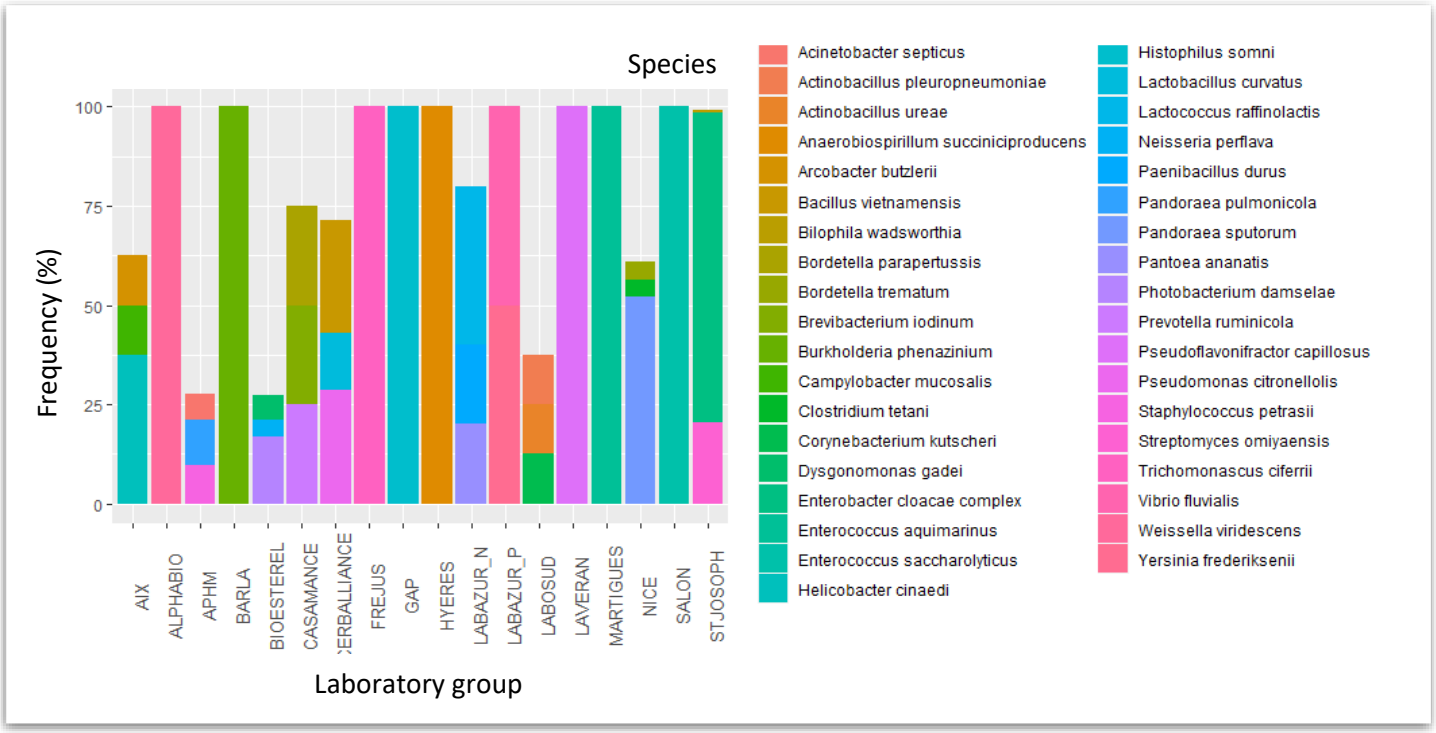


Figure 3: Diagram of the three most frequent specific species by laboratory

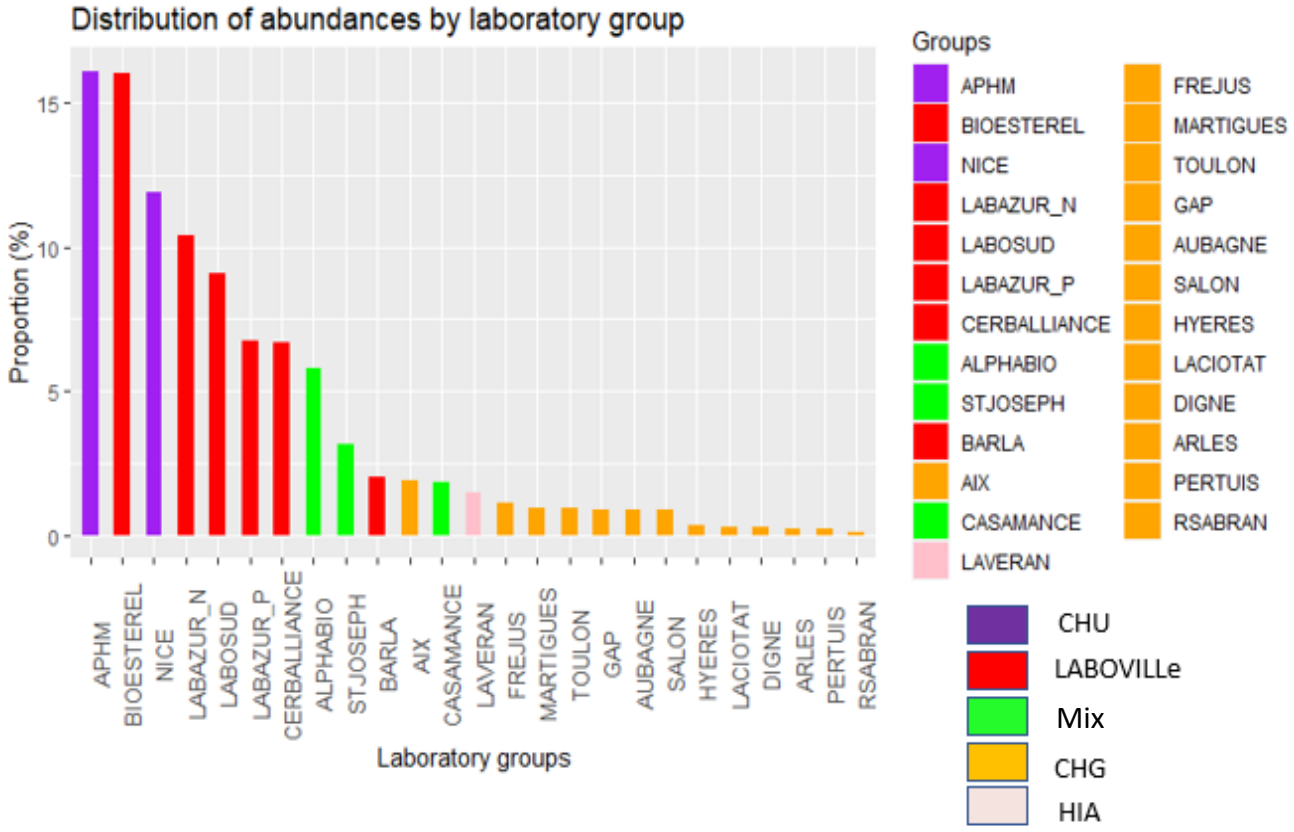


Figure 4: Distribution of richness by laboratory group

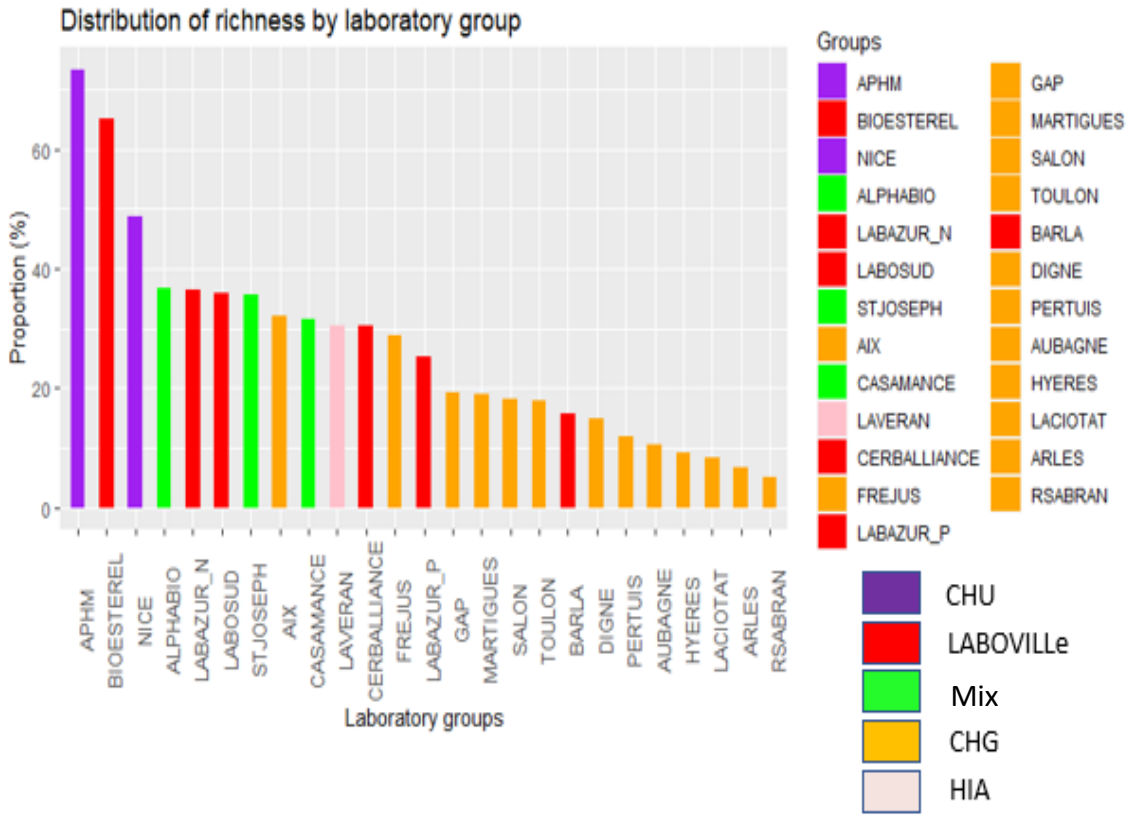


Figure 5: Distribution of abundance by laboratory group

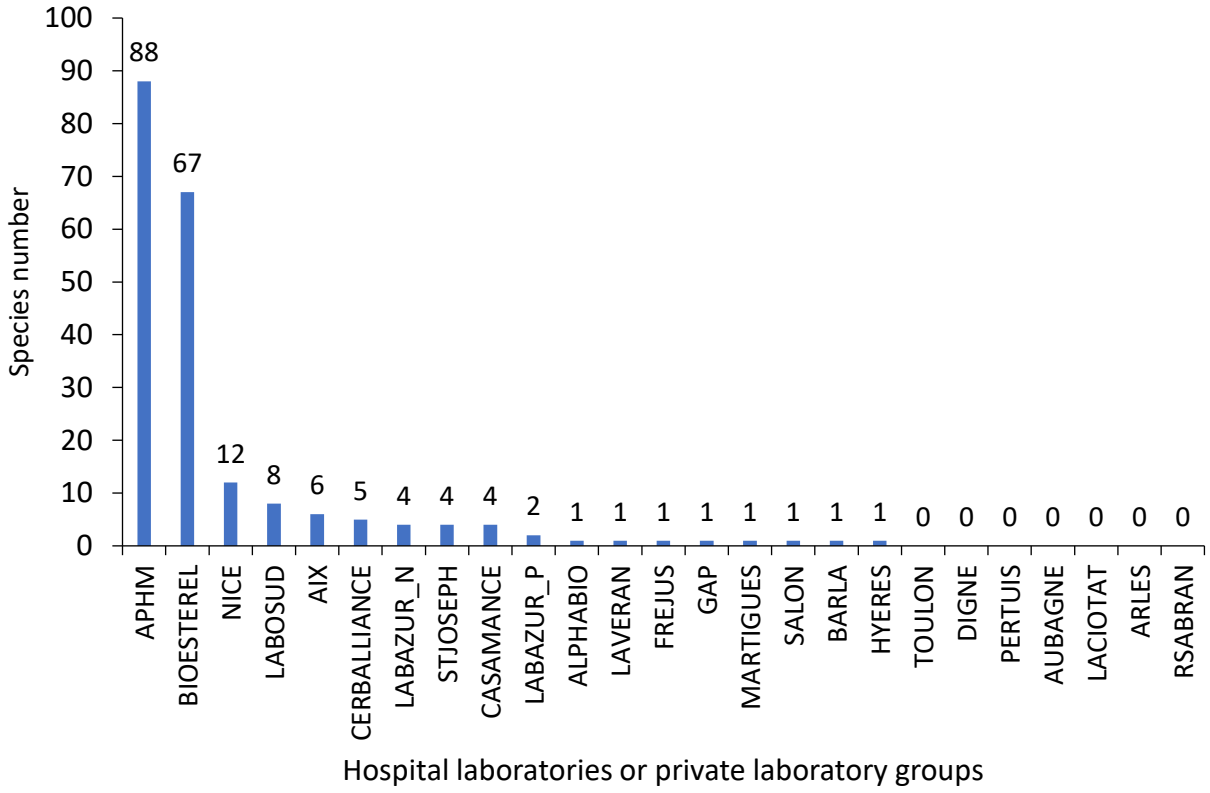


Figure 6: Number of bacterial species specific to laboratory groups

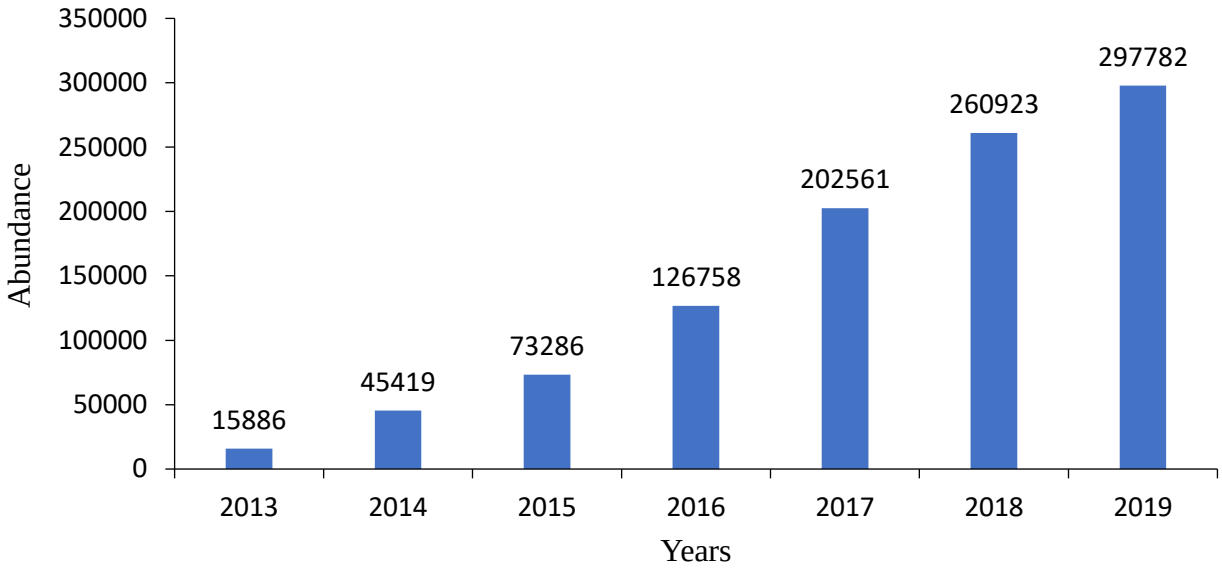


Figure 7: Evolution of observed species abundance per year

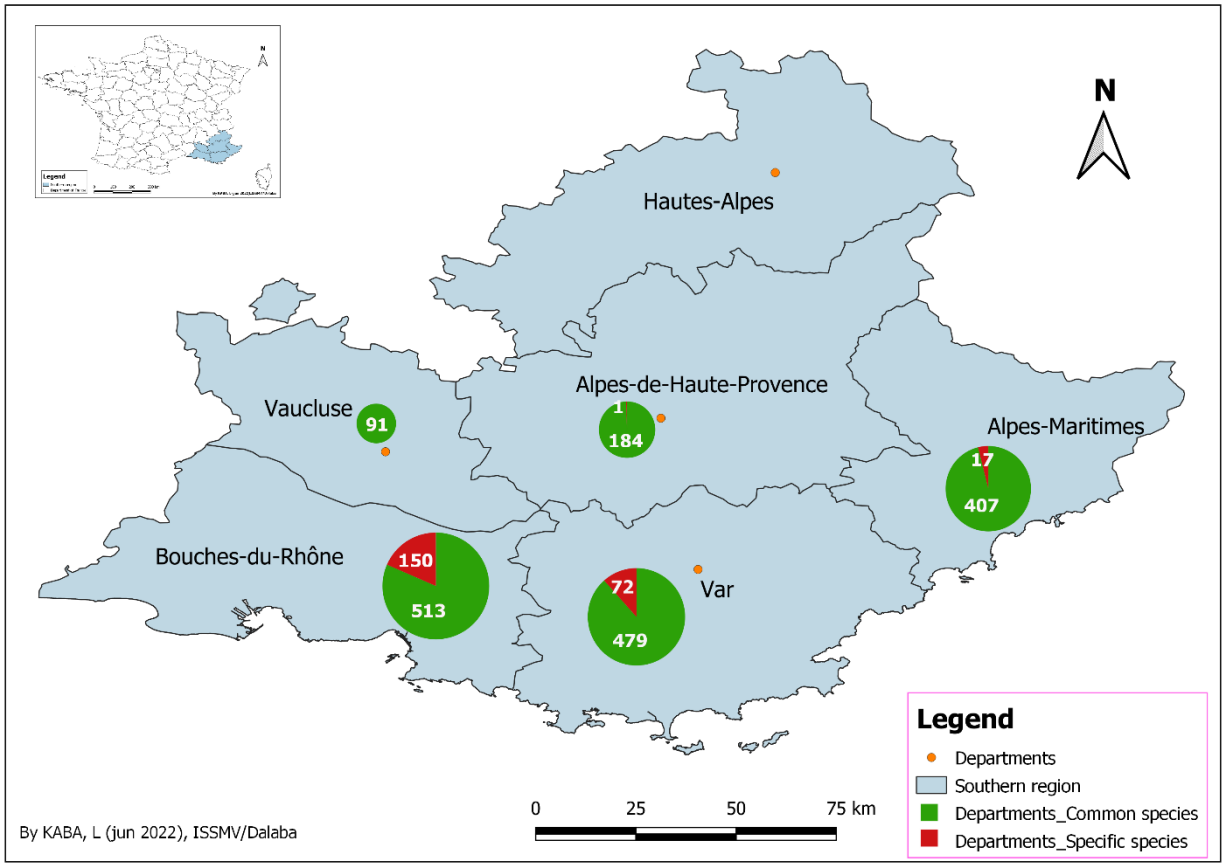


Figure 8: Distribution of species identified from 2013 to 2019 by department

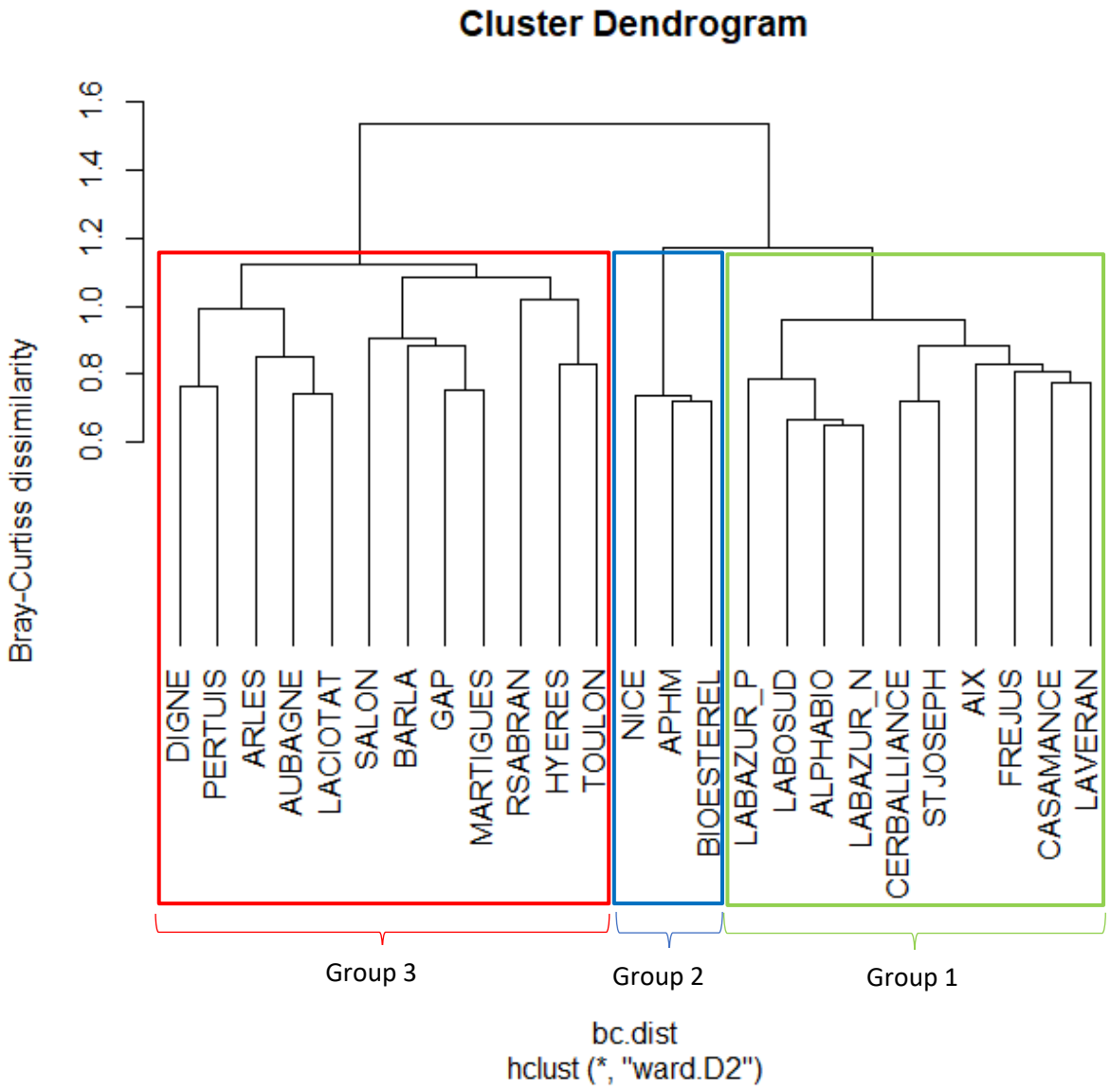


Figure 9: Clustering dendrogram of laboratory groups

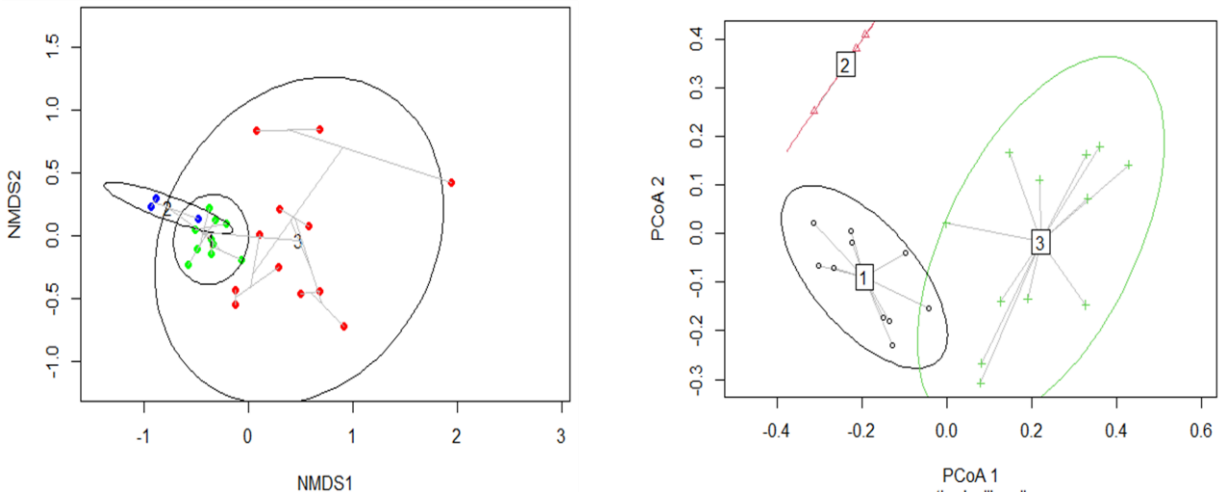


Figure 10: Map of clustering of laboratories into three groups based on similarity (NMDS and PCoA)

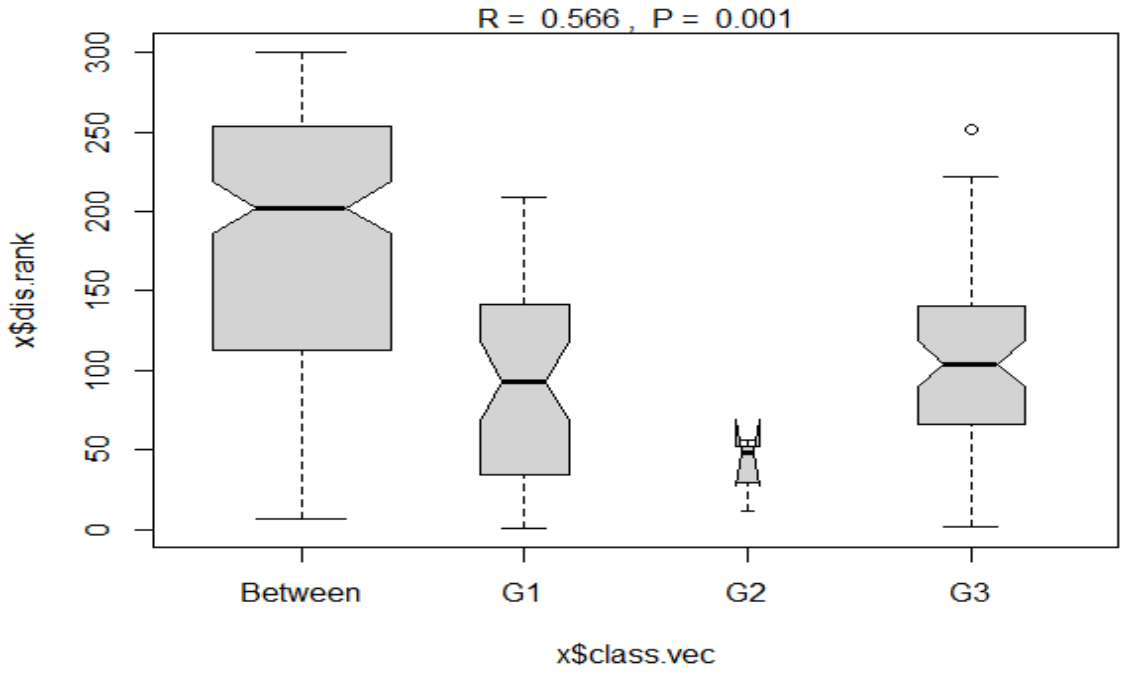


Figure 11: Dissimilarity between and within groups

495 **SUPPLEMENTARY MATERIELS**

496 Table 3: Species detected only by a single laboratory

BIOESTEREL (67 species)	Freq	**BIOESTEREL (67 species)**	**Freq**
Photobacterium damselae	25	*Actinomyces bowdenii*	1
Dysgonomonas gadei	9	*Advenella kashmirensis*	1
Neisseria perflava	6	*Aerococcus christensenii*	1
Citrobacter gillenii	5	*Aeromonas encheleia*	1
Edwardsiella tarda	5	*Alistipes indistinctus*	1
Flavobacterium lindanitolerans	4	*Arthrobacter woluwensis*	1
Peptoniphilus tyrrelliae	4	*Avibacterium endocarditidis*	1
Staphylococcus delphini	4	*Bacteroides clarus*	1
Streptococcus halichoeri	4	*Bacteroides finegoldii*	1
Actinomyces funkei	3	*Blastomonas natatoria*	1
Bacteroides coagulans	3	*Clostridium symbiosum*	1
Neisseria animaloris	3	*Corynebacterium matruchotii*	1
Paenibacillus glucanolyticus	3	*Curtobacterium luteum*	1
Vibrio harveyi	3	*Clostridium symbiosum*	1
Vibrio mytili	3	*Corynebacterium matruchotii*	1
Actinomyces canis	2	*Curtobacterium luteum*	1
Fusobacterium canifelinum	2	*Enterococcus devriesei*	1
Gallibacterium anatis	2	*Enterococcus phoeniculicola*	1
Haemophilus paraphrohaemolyticus	2	*Glutamicibacter creatinolyticus*	1
Murdochiella asaccharolytica	2	*Gordonia rubripertincta*	1
Peptoniphilus coxii	2	*Haemophilus parasuis*	1
Pseudomonas cichorii	2	*Lactobacillus kitasatonis*	1
Pseudomonas taetrolens	2	*Lactobacillus saerimneri*	1
Streptococcus didelphis	2	*Lysinibacillus boronitolerans*	1
Vibrio furnissii	2	*Oceanobacillus profundus*	1
Vibrio xuii	2	*Paenibacillus xylanilyticus*	1
Achromobacter spanius	1	*Paraburkholderia fungorum*	1
Acidaminococcus fermentans	1	*Peptoniphilus olsenii*	1
Acidiphilium acidophilum	1	*Porphyromonas uenonis*	1

497

498 Table 3 continued: Species detected only by a single laboratory

BIOESTEREL (65 species)	**Number**	**BIOESTEREL (65 species)**	**Number**
Propionimicrobium lymphophilum	1	*Lactobacillus saerimneri*	1
Providencia rustigianii	1	*Lysinibacillus boronitolerans*	1
Pseudochrobactrum asaccharolyticum	1	*Oceanobacillus profundus*	1
Pseudomonas agarici	1	*Paenibacillus xylanilyticus*	1
Pseudomonas fragi	1	*Paraburkholderia fungorum*	1
Pseudomonas savastanoi	1	*Peptoniphilus olsenii*	1
Providencia rustigianii	1	*Porphyromonas uenonis*	1
Pseudochrobactrum asaccharolyticum	1	*Propionimicrobium lymphophilum*	1
Pseudomonas agarici	1	**STJOSOPH (4 species)**	
Pseudomonas fragi	1	*Enterobacter cloacae complex*	95
Pseudomonas savastanoi	1	*Streptomyces omiyaensis*	25
Pseudomonas straminea	1	*Bilophila wadsworthia*	1
Pseudomonas synxantha	1	*Enterococcus dispar*	1
Shewanella baltica	1	**CASAMANCE (4 species)**	
Staphylococcus lutrae	1	*Bordetella parapertussis*	1
Stenotrophomonas rhizophila	1	*Brevibacterium iodinum*	1
Vibrio ponticus	1	*Prevotella ruminicola*	1
Pseudomonas straminea	1	*Streptococcus pluranimalium*	1
Pseudomonas synxantha	1	**ALPHABIO (1 species)**	
Shewanella baltica	1	*Weissella viridescens*	2
Staphylococcus lutrae	1	**LAVERAN (1 species)**	
Stenotrophomonas rhizophila	1	*Pseudoflavonifractor capillosus*	2
Vibrio ponticus	1	**FREJUS (1 species)**	
Enterococcus devriesei	1	*Trichomonascus ciferrii*	1
Enterococcus phoeniculicola	1	**GAP (1 species)**	
Glutamicibacter creatinolyticus	1	*Histophilus somni*	1
Gordonia rubripertincta	1	**MARTIGUES (1 species)**	
Haemophilus parasuis	1	*Enterococcus aquimarinus*	1
Lactobacillus kitasatonis	1		

499

500

501 Table 3 continued: Species detected only by a single laboratory

APHM (88 species)	Number	APHM (88 species)	Number
Pandoraea pulmonicola	30	*Porphyromonas endodontalis*	3
Staphylococcus petrasii	26	*Prevotella conceptionensis*	3
Acinetobacter septicus	18	*Bacillus weihenstephanensis*	2
Atopobium parvulum	16	*Brevibacterium massiliense*	2
Actinomyces ihumii	9	*Cellulosimicrobium cellulans*	2
Mycobacterium xenopi	6	*Desulfovibrio desulfuricans*	2
Pseudomonas massiliensis	6	*Flavonifractor plautii*	2
Corynebacterium ihumii	5	*Ignavigranum ruoffiae*	2
Pantoea eucrina	5	*Kosakonia cowanii*	2
Actinomyces grossensis	4	*Microbacterium kitamiense*	2
Clostridium bifermentans	4	*Mycobacterium bovis*	2
Clostridium subterminale	4	*Mycobacterium europaeum*	2
Cupriavidus gilardii	4	*Mycobacterium indicus pranii*	2
Delftia tsuruhatensis	4	*Mycobacterium mageritense*	2
Janibacter hoylei	4	*Mycobacterium simiae*	2
Leptotrichia trevisanii	4	*Roseomonas genomospecies 5*	2
Macrococcus caseolyticus	4	*Streptococcus castoreus*	2
Mycobacterium abscessus subsp. bolletii	4	*Acidovorax temperans*	1
Mycobacterium chimaera	4	*Aeromonas bestiarum*	1
Bacillus vallismortis	3	*Akkermansia muciniphila*	1
Corynebacterium lascolaensis	3	*Anaerococcus lactolyticus*	1
Corynebacterium ureicelerivorans	3	*Bacillus firmus*	1
Cupriavidus respiraculi	3	*Bacillus marisflavi*	1
Microbacterium oxydans	3	*Brachyspira pilosicoli*	1
Mycobacterium kansasii	3	*Brevibacillus agri*	1
Nocardia cyriacigeorgica	3	*Buttiauxella gaviniae*	1
Peptoniphilus grossensis	3	*Butyricimonas phoceencis*	1
Peptoniphilus sp. EL1	3	*Butyricimonas virosa*	1
Campylobacter sputorum	1	*Chlamydia abortus*	1
Carnobacterium divergens	1	*Citrobacter murliniae*	1

502

503 Table 3 continued and end: Species detected only by a single laboratory

APHM (88 species)	Number	CERBALLIANCE (5 species)	Number
Clostridium cadaveris	1	*Pseudomonas citronellolis*	2
Clostridium celerecrescens	1	*Bacillus vietnamensis*	2
Corynebacterium pilosum	1	*Lactobacillus curvatus*	1
Eggerthia catenaformis	1	*Listeria innocua*	1
Janibacter sanguinis	1	*Moraxella atlantae*	1
Kytococcus sedentarius	1	**LABAZUR_N (4 species)**	
Leuconostoc citreum	1	*Lactococcus raffinolactis*	2
Lysinibacillus massiliensis	1	*Paenibacillus durus*	1
Massilia timonae	1	*Pantoea ananatis*	1
Mycobacterium avium complex (MAC)	1	*Alicyclobacillus acidoterrestris*	1
Mycobacterium colombiense	1	**NICE (12 species)**	
Mycobacterium lentiflavum	1	*Pandoraea sputorum*	12
Neisseria canis	1	*Bordetella trematum*	1
Nocardia otitidiscaviarum	1	*Clostridium tetani*	1
Paenibacillus provencensis	1	*Erwinia persicina*	1
Pantoea annanatis	1	*Paenibacillus lactis*	1
Peptostreptococcus stomatis	1	*Prevotella dentalis*	1
Prevotella massiliensis	1	*Prevotella pallens*	1
Prevotella oulorum	1	*Pseudomonas libanensis*	1
Pseudomonas kuykendallii	1	*Rhodococcus erythropolis*	1
Pseudomonas rhodesiae	1	*Stenotrophomonas acidaminiphila*	1
Rhodococcus equi	1	*Streptococcus sobrinus*	1
Rhodococcus rhodochrous	1	*Yersinia kristensenii*	1
Rothia terrae	1	**SALON (1 species)**	
Sphingobacterium thalpophilum	1	*Enterococcus saccharolyticus*	2
Sporolactobacillus laevolacticus	1	**LABAZUR_P (2 species)**	
Sporosarcina luteola	1	*Vibrio fluvialis*	1
Yersinia intermedia	1	*Yersinia frederiksenii*	1

504

CHAPITRE III

Etude d'impact du confinement et des mesures barrières appliqués contre la pandémie du Covid 19 sur des agents infectieux outre que le SARS-CoV-2

Préambule

L'émergence du nouveau coronavirus en fin décembre 2019 à Wuhan précisément dans la province de Hubei en Chine et sa propagation rapide à l'échelle mondiale a conduit à des prises de décisions générales visant à endiguer la transmission du virus. L'épidémie de départ a été déclaré comme une pandémie le 11 mars 2020 par l'OMS (29). Ce nouveau virus nommé SARS-CoV 2 (Severe Acute Respiratory Syndrome CoronaVirus 2) par le Comité international de taxonomie des virus (ICTV) (30) et la maladie quant à elle, a été nommé COVID 19 par l'OMS en février 2020 (31),ce virus est un virus appartenant à la famille des *Coronaviridae*. A la date du 11 octobre 2020, l'OMS affirma qu'à l'échelle mondiale il y avait plus de 37 millions de cas confirmés de COVID-19 et 1 million de décès (32). La France a franchi la barre de 20000 décès le 20 avril 2020 dont 12 513 en milieu hospitalier et 7 552 dans les établissements médico-sociaux (33). Les mesures prises contre le virus du Covid 19 a bouleversé le comportement et les habitudes des populations telles que la fermeture des écoles, annulations des voyages nationaux et internationaux, la fermeture des commerces, le lavage des mains et la distanciation sociale, etc. Ces dispositions ont impacté fortement la vie des populations de tous les pays.

Cependant, d'autres épidémies précédentes à conséquences désastreuses avaient également marqué l'histoire de la santé publique (34), telle que la pandémie de la grippe espagnole de 1918-1919 qui a tué plus de 20 à 40 millions de personnes dans le monde entier (35). En revanche, les mesures de confinement appliquées contre le Covid 19 a concerné la meilleure partie des pays touchés et ce n'est qu'une première dans l'histoire de la lutte contre les pandémies. Ces mesures ont entrainé des conséquences tant sur le plan économique, social et sanitaire (36–38). Les différentes études d'impact du confinement se sont focalisées principalement sur ces aspects économiques, sociaux et sanitaires. Cependant, l'absence de contact entre les populations, des voyages nationaux et internationaux et le lavage des mains auraient impacté directement ou indirectement la communauté de certains agents infectieux. Nous avons jugé nécessaire de regarder dans ce chapitre les effets du confinement sur l'évolution de la diversité des agents infectieux non-viraux avec l'évolution à la résistance aux antibiotiques (article 3) et le contrôle des épidémies virales courantes mais pas SARS-CoV2 (article 4) rapporté par le système de surveillance épidémiologique des laboratoires basé à l'IHU Méditerranée Infection de Marseille.

Article 3

Article 3: Consequences of the COVID-19 Outbreak Lockdown on Non-viral Infectious Agents as Reported by a Laboratory-Based Surveillance System at the IHU Méditerranée Infection, Marseille, France. *J. Clin. Med.* 2021, *10*, 3210. https://doi.org/10.3390/jcm10153210

Kaba, L.; Giraud-Gatineau, A.; Jimeno, M.-T.; Rolain, J.-M.; Colson, P.; Raoult, D.; Chaudet a, H.
Publié dans le *Journal of Clinical Medicine* (juillet 2021)

Journal of
Clinical Medicine

MDPI

Article

Consequences of the COVID-19 Outbreak Lockdown on Non-Viral Infectious Agents as Reported by a Laboratory-Based Surveillance System at the IHU Méditerranée Infection, Marseille, France

Lanceï Kaba [1,2,3,†], Audrey Giraud-Gatineau [1,2,4,5,†], Marie-Thérèse Jimeno [5], Jean-Marc Rolain [1,5,6], Philippe Colson [1,5,6], Didier Raoult [1,5,6] and Hervé Chaudet [1,2,5,*]

1 IHU Méditerranée Infection, 19-21 Boulevard Jean Moulin, 13005 Marseille, France; lancekaba@yahoo.fr (L.K.); audrey.giraud.gatineau@gmail.com (A.G.-G.); jean-marc.rolain@univ-amu.fr (J.-M.R.); philippe.colson@univ-amu.fr (P.C.); didier.raoult@gmail.com (D.R.)
2 Aix Marseille Université, Institut de Recherche pour le Développement (IRD), Assistance Publique-Hôpitaux de Marseille (AP-HM), Service de Santé des Armées (SSA), VITROME, 13005 Marseille, France
3 Institut Supérieur des Sciences et de Médecine Vétérinaire (ISSMV) de Dalaba, Dalaba BP 09, Guinea
4 French Armed Forces Center for Epidemiology and Public Health (CESPA), Service de Santé des Armées (SSA), 13014 Marseille, France
5 Assistance Publique-Hôpitaux de Marseille (AP-HM), 13005 Marseille, France; marie-therese.jimeno@ap-hm.fr
6 Aix-Marseille Université, Institut de Recherche pour le Développement (IRD), Assistance Publique-Hôpitaux de Marseille (AP-HM), MEPHI, 27 Boulevard Jean Moulin, 13005 Marseille, France
* Correspondence: herve.chaudet@gmail.com; Tel.: +33-413-732-401; Fax: +33-413-732-402
† The authors have contributed equally.

Citation: Kaba, L.; Giraud-Gatineau, A.; Jimeno, M.-T.; Rolain, J.-M.; Colson, P.; Raoult, D.; Chaudet a, H. Consequences of the COVID-19 Outbreak Lockdown on Non-viral Infectious Agents as Reported by a Laboratory-Based Surveillance System at the IHU Méditerranée Infection, Marseille, France. *J. Clin. Med.* **2021**, *10*, 3210. https://doi.org/10.3390/jcm10153210

Academic Editors: Philippe Parola, Corneliu Petru Popescu, Michela Sabbatucci and Gregory A. Hand

Received: 27 April 2021
Accepted: 14 June 2021
Published: 21 July 2021

Abstract: The objective of this paper is to describe the surveillance system MIDaS and to show how this system has been used for evaluating the consequences of the French COVID-19 lockdown on the bacterial mix of AP-HM and the antibiotic resistance. MIDas is a kind of surveillance activity hub, allowing the automatic construction of surveillance control boards. We investigated the diversity and resistance of bacterial agents from respiratory, blood, and urine samples during the lockdown period (from week 12 to 35 of 2020), using the same period of years from 2017 to 2019 as control. Taking into account the drop in patient recruitment, several species have exhibited significant changes in their relative abundance (either increasing or decreasing) with changes up to 9%. The changes were more important for respiratory and urine samples than for blood samples. The relative abundance in respiratory samples for the whole studied period was higher during the lockdown. A significant increase in the percentage of wild phenotypes during the lockdown was observed for several species. The use of the MIDaS syndromic collection and surveillance system made it possible to efficiently detect, analyze, and follow changes of the microbiological population as during the lockdown period.

Keywords: syndromic surveillance; clinical microbiology laboratory; epidemiology; lockdown; COVID-19; diversity; wild

1. Introduction

The Hospital University Institute Méditerranée Infection (IHU-MI) hosts the clinical microbiology and virology laboratory for all four of the public university hospitals of Marseille (AP-HM) that perform the diagnosis of infectious agents including bacteria, microscopic fungi, parasites, and viruses. Since 2013, it has implemented and improved five syndromic epidemiological surveillance sub-systems that use its laboratory results from

a single collection and an analysis system named MIDaS (Mediterranée Infection Data Warehousing and Surveillance). Besides traditional surveillance based on patients' clinical diagnoses of notifiable infectious diseases, syndromic surveillance uses data about analysis requests from clinicians' prescriptions and laboratory's tests results, as well as other laboratory markers, through innovative approaches.

In December 2019, Wuhan in the Hubei province became the epicenter of the spread of a new emerging pathogen called SARS-CoV-2. It spreads rapidly to other continents, a pandemic being declared officially in March 2020 (https://www.who.int/emergencies/diseases/novel-coronavirus-2019/interactive-timeline/, accessed 27 May 2021). For controlling the rapid spread of this virus, the French government implemented various measures including the closure of schools, cultural centers, and socialization places such as bars and restaurants [1,2], before finally promulgating a total lockdown one week later, from 16 March 2020 [3] to 11 May 2020 (week 12 to 19). Many European, American, and Asian countries have made the same choice with a more or less strict lockdowns [4]. Several studies have already assessed the effectiveness and impact of the diverse consequences of lockdowns in various fields such as the medical, economic, or sociological fields [5–8]. They showed that lockdown measures led to major alterations of patient cares in hospital settings, especially those with chronic non-communicable diseases such as hypertension, diabetes, mental depression, etc. [9]. In the UK, a 71% decrease in blood counts was reported in the first four weeks of containment, and 57% fewer patients were sent for specialist hematology review [10]. However, to our knowledge, the microbiological impact of lockdown measures has not been yet studied.

The objective of this paper is to describe the current organization of the surveillance system and to show how it has been used for evaluating the consequences of the French lockdown on bacterial identifications in the AP-HM, as well as its influence on antibiotic resistance.

2. Materials and Methods

2.1. MIDaS, an Epidemiological Hub

The MIDaS system can be considered as a surveillance activity hub, hosting data coming from the hospital information system, the analysis automata, and several other sources such as the taxonomy database maintained by NCBI/GenBank (https://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi, accessed on 27 May 2021) and the demographical data coming from the Institut national de la statistique et des études économiques (https://www.insee.fr/fr/statistiques, accessed on 27 May 2021). Before their analysis, data are preprocessed depending on their nature (bacterial identification, molecular biology, antiobiotic resistance, etc.) for specific aggregations or expert processing, as the identification of resistance phenotype by expert rules. Raw data and preprocessing results are then systematically analyzed and presented by statistical automata in search of statistical aberrations within time series [11] (e.g., a significant increase suggesting a possible health concern or a disease outbreak). All significant increases are automatically documented by the system, including the sampling profiling, the search for atypical antibiograms, and the mapping of cases. Analysis results are dynamically presented using tailored control boards on an internal dedicated website, and are discussed each week during a staff meeting that is attended by biologists, clinicians, and epidemiologists. During this staff meeting, in silico investigations (notably additional comparisons and cross-referencing of data) can be used. Furthermore, epidemiological investigations can be initiated if the investigation confirms the alarm and, if required, an epidemiological alert may be simultaneously transmitted to the health institutions concerned including the Regional Health Agency (Agence Régionale de Santé, ARS, Marseille, France) and the Infection Control Committee (Comité de Lutte contre les Infections nosocomiales, CLIN, Marseille, France). The surveillance results are also disseminated weekly through the IHU Méditerranée Infection website As a surveillance activity hub, this system also suggests

the weekly enrichment of our microbial strain collection (CSUR) [12], and supports the quality control of laboratory activities in the search for deviations in laboratory processes. The overall structure of MIDaS is presented in Figure 1.

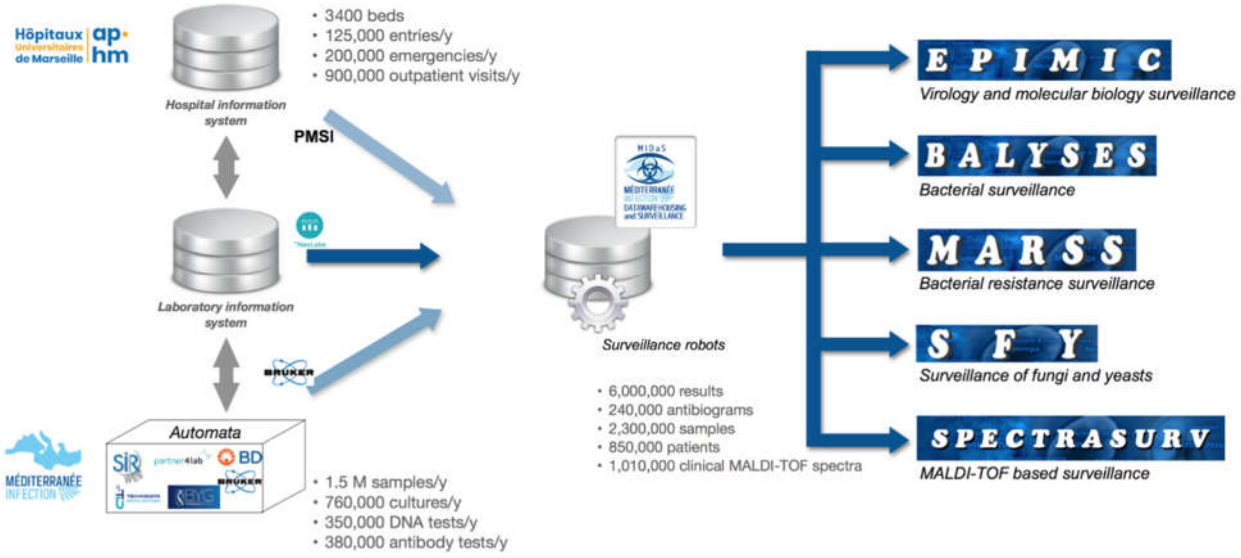


Figure 1. Structure of Méditerranée Infection Data warehousing and Surveillance (MIDaS).

2.2. MIDaS Data Collection

The main role of MIDaS is to gather surveillance-related data from the hospital information system and the laboratory information system by weekly extraction–transform–load processes. The AP-HM consists of four public university hospitals: Timone (1307 beds), Conception (767 beds), North Hospital (793 beds), and South Hospital (421 beds). It has approximately 125,000 admissions and 1 million consultations per year. The clinical microbiology and virology laboratory performs approximately 8 million tests per year. Tests results, as well as patients and specimen information, are collected from the laboratory information system. Other data come from the hospital information systems, such as the hospital medico-economic data (Programme de Médicalisation des Systèmes d'Information, PMSI) used for patients' death status in order to study death-associated infections. Data from other systems are also collected, such as spectra files generated by the Matrix Assisted Laser Desorption IonizationTime of Flight (MALDI–TOF) mass spectrometry instruments used for bacterial and fungal routine identification. MIDaS is therefore a data warehouse that groups together microbiological analysis results (sample number, requesting unit, sample date, type of analysis, antibiotic susceptibility test results, and antibiotic resistance phenotype) and patient information (anonymized patient identifier, age, gender, postal code of residence, anonymized identifier of hospital stay, date of hospitalization, length of stay, and death). Six million microbiological results are stored in this data warehouse, representing 240,000 antibiotic susceptibility tests, 2,300,000 samples, 850,000 patients, and nearly 1 million MALDI–TOF clinical spectra (with more than 3 million for spectra being produced for research purposes).

2.3. MIDaS Domain-Specific Monitoring Systems

Five domain-specific monitoring sub-systems are connected to the data warehouse for producing fully automated dashboards. Historically, EPIMIC (for EPIdemiological surveillance and alert based on MICrobiological data) is the first surveillance system that was implemented (in 2002) in our laboratory to allow monitoring of the weekly counts of clinical specimens sent by clinicians, diagnosis tests performed, and diagnosis results [13]. It has been later updated in 2013 for its integration into MIDaS. Since 2013, bacteria have

been more comprehensively monitored by BALYSES (Bacterial real-time Laboratory-based Surveillance System), while SFY (Surveillance of Fungi and Yeasts) has focused on microscopic fungi and yeasts and MARSS (Marseille Antibiotic Resistance Surveillance System) has monitored antibiotic resistance patterns [14]. In addition, MALDI–TOF spectra have been used as an additional tool to support surveillance and are analyzed by the SpectraSurv system (for MALDI–TOF based surveillance) [15].

2.4. Data for the Lockdown Analysis

In this study, we were particularly interested by the bacterial agents identified in respiratory, blood, and urine samples during the lockdown period (that is, from week 12 (mid-march) to week 35 (end of august) 2020). We used the same period of years from 2017 to 2019 as control.

2.4.1. Hospital Activities

A preliminary descriptive analysis based on BALYSES automatic control boards allowed us to define three a priori periods according to the evolution of the laboratory activity during 2020: a lockdown phase (weeks 12–19), a restoration phase (weeks 20–24) and a post-lockdown phase (weeks 25–35) (Figure 2).

Figure 2. Follow-up of patients with at least one bacterial identification at IHU.

2.4.2. Bacterial and Fungal Community

The bacterial community was studied in terms of species richness and abundance for the 3 most frequent samples: urine, respiratory, and blood samples. The specific richness represents the total number of species present in a sample and the relative abundance (or relative frequency) indicates the frequency of a species.

2.4.3. Evolution of Antibiotic Resistance

We used the percentage of non-resistant (wild) isolates for each species monitored by the surveillance system without differentiating the resistance phenotypes, taking into account the nosocomial or community origin of the isolate.

2.5. Statistical Analysis

For further investigations of the surveillance system results, the log-linear model, and the Fisher and Chi2 tests for point comparisons, were used to evaluate the evolution of diversity and antimicrobial resistance, with a statistical significance threshold of 0.05 [16]. All statistical processes were done using software R version 4.0.3 (https://www.R-project.org, accessed on 27 May 2021.

3. Results

3.1. Hospital Activities

The follow-up of the laboratory bacterial identification activity (Figure 2) showed a drop in the moving average of the number of patients during the lockdown period from 641.5 patients on 11 March to 412.5 patients on 13 May 2020. After the end of the lockdown, the activity level gradually returned to the normal, from 412.5 patients on 13 May to 512.5 patients on 17 June and then to 611.25 patients on 15 July 2020.

3.2. Bacterial and Fungal Community

From weeks 12 to 35 in 2017–2020, a total of 349 bacterial and fungal species were identified from 30,918 identifications including 24,946 from urine samples (186 distinct species), 4555 from respiratory samples (230 distinct species), and 1417 from blood samples (111 distinct species). The top twenty species alone represent 87.4% (27,037/30,918) of the total number of identifications. While the relative abundance in respiratory samples for the whole studied period was higher in 2020, it decreased for urine samples and was constant for blood samples (Figure 3A). However, while the species richness was constant over time in respiratory and urine samples it decreased in blood samples (Figure 3B).

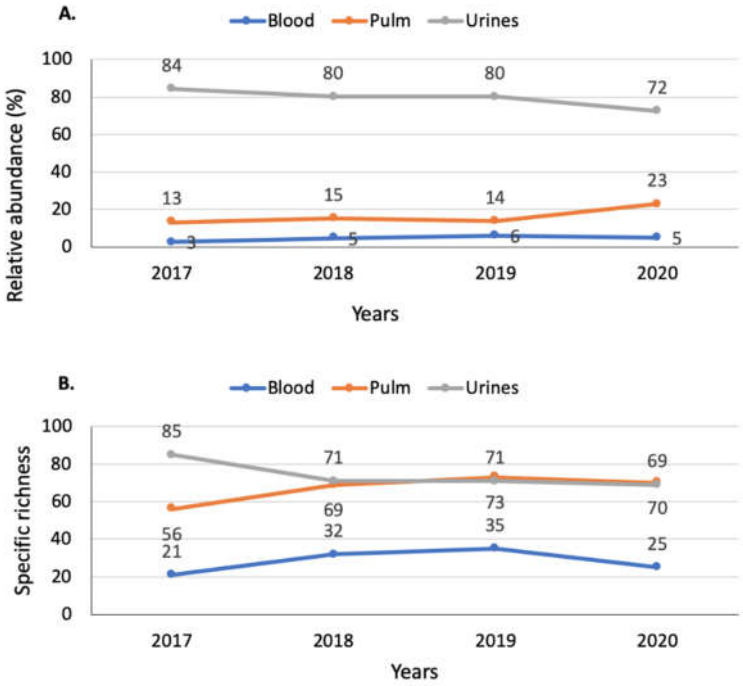


Figure 3. Evolution of the relative abundance (**A**) and the specific richness (**B**) from weeks 12 to 35, 2017 to 2020, at Assistance Publique—Hôpitaux de Marseille, Marseille, France.

When comparing diversities between 2020 and 2017–2019 for the pooled three kinds of samples (urine, respiratory, and blood samples), we found a significant variation in the relative frequency of nine species out of the top twenty (45%) during the lockdown period, and four species during the restoration and post-lockdown periods (although not for the same species) (Table 1, Figure 4). The species that significantly decreased during the lockdown were *Escherichia coli* (39.3% to 28.6%, p-value < 2.2×10^{-16}), *Klebsiella oxytoca* (1.5% to 0.8%, p-value = 0.02), and *Haemophilus influenzae* (1.2% to 0.7%, p-value = 0.02). There was a significant increase of *Candida albicans*, *Staphylococcus epidermidis*, *Enterobacter cloacae*, *Staphylococcus haemolyticus*, *Enterobacter aerogenes*, and *Candida glabrata* (Table 1). *S. epidermidis* and *C. albicans* species increased during all three time periods. Conversely, *E. coli* significally decreased. *Citrobacter*

koseri experienced a significant decrease only during the restoration period, and *Staphylococcus aureus* experienced significant growth during the post-lockdown (Table 1).

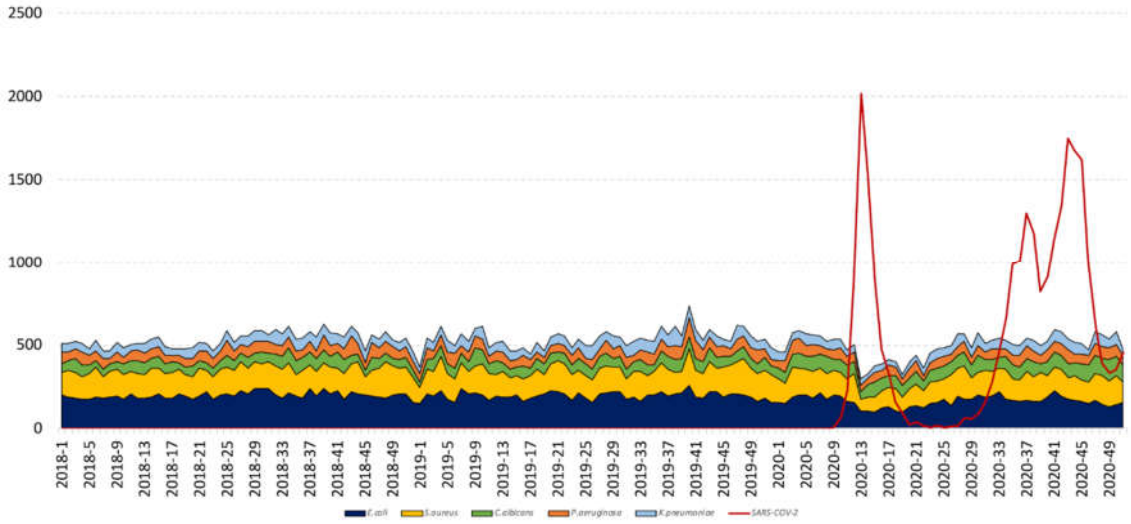


Figure 4. Weekly incidence of the five most identified species in our institute and SARS-CoV-2 from 2018 to 2020, Marseille, France.

Table 1. Evolution of the relative frequency (abundance) during the lockdown period 2020 vs. 2017–2019 for the three types of sampling, Marseille.

N	Species	During Lockdown (Weeks 12–19)				During Restoration (Weeks 20–24)				During Post-Lockdown (Weeks 25–35)			
		2017-19 %	2020 %	*p*-Value	Evol	2017-19 %	2020 %	*p*-Value	Evol	2017-19 %	2020 %	*p*-Value	Evol
1	*E. coli*	39.3	28.6	$<2.2 \times 10^{-16}$	↘	37.8	32.2	8.2×10^{-5}	↘	37.2	35.0	0.02	↘
2	*K. pneumoniae*	8.8	7.7	0.12	→	8.7	8.9	0.78	→	10.4	9.3	0.07	→
3	*E. faecalis*	6.7	7.4	0.28	→	6.3	6.3	0.94	→	6.2	5.9	0.50	→
4	*P. aeruginosa*	4.4	5.3	0.06	→	4.7	5.0	0.61	→	4.8	4.8	0.84	→
5	*C. albicans*	4.0	7.9	8.62×10^{-14}	↗	3.9	6.4	4.2×10^{-5}	↗	4.2	5.2	0.02	↗
6	*S. aureus*	3.7	4.5	0.11	→	3.7	4.9	0.05	→	3.5	4.5	0.01	↗
7	*S. epidermidis*	3.5	5.0	0.001	↗	3.3	4.7	0.02	↗	2.9	3.6	0.04	↗
8	*P. mirabilis*	2.7	2.8	0.80	→	2.6	2.9	0.54	→	3.0	3.3	0.36	→
9	*E. cloacae*	2.4	3.7	0.0005	↗	2.6	3.2	0.24	→	3.3	3.9	0.10	→
10	*S. agalactiae*	2.3	2.1	0.46	→	2.	1.7	0.47	→	2.2	2.4	0.48	→
11	*K. oxytoca*	1.5	0.8	0.02	↘	1.4	1.7	0.44	→	1.3	1.1	0.37	→
12	*E. faecium*	1.5	1.4	0.71	→	1.2	1.2	0.34	→	1.2	1.2	0.80	→
13	*H. influenzae*	1.2	0.7	0.02	↘	1.0	0.8	0.53	→	0.8	0.5	0.05	→
14	*S. haemolyticus*	1.2	2.0	0.004	↗	1.3	1.7	0.34	→	1.1	1.3	0.19	→
15	*C. koseri*	1.1	1.4	0.33	→	1.4	0.7	0.03	↘	1.0	1.0	0.92	→
16	*E. aerogenes*	1.1	1.6	0.04	↗	1.0	1.0	0.87	→	1.3	1.6	0.12	→
17	*S. saprophyticus*	1.0	0.8	0.49	→	1.1	1.1	0.77	→	1.0	1.2	0.51	→
18	*M. morganii*	0.8	1.2	0.09	→	1.0	0.5	0.08	→	0.9	1.0	0.60	→
19	*S. pneumoniae*	0.7	0.5	0.18	→	0.6	0.4	0.40	→	0.5	0.5	0.85	→
20	*C. glabrata*	0.6	1.1	0.01	↗	0.7	0.8	0.81	→	0.6	0.7	0.65	→
21	*Autres*	11.6	13.4	0.03	↗	13.5	14.0	0.58	→	12.5	12.2	0.62	→

↗ Significant growth; ↘ Significant decrease; → Non-significant change; Evol = Evolution.

3.2.1. Diversity in Respiratory Samples

During the lockdown and restoration periods, five species out of the top twenty recorded a significant variation in their relative frequency, and one species recorded a significant variation during the post-lockdown period (Table 2). *E. coli*, *S. pneumoniae*, and *H. influenzae* significantly decreased during the lockdown and remained stable during the next two phases. *C. albicans* is the only species that increased during the three periods. *K. pneumoniae* decreased during the restoration period, whereas species including *E. cloacae* and *S. agalactia* increased.

Table 2. Evolution of relative abundance for respiratory samples from week 12 to week 35, 2020, Marseille.

N	Species	During Lockdown (Weeks 12–19)				During Restoration (Weeks 20–24)				During Post-Lockdown (Weeks 25–S35)			
		2017-19 %	2020 %	*p*-Value	Evol	2017-19 %	2020 %	*p*-Value	Evol	2017-19 %	2020 %	*p*-Value	Evol
1	*E. coli*	7.5	3.3	0.001	↘	5.9	5.0	0.61	→	5.0	5.0	0.98	→
2	*K. pneumoniae*	5.0	4.5	0.69	→	5.7	2.2	0.02	↘	6.5	4.6	0.11	→
3	*E. faecalis*	2.5	3.1	0.51	→	1.1	3.9	0.004	↗	1.6	1.4	0.82	→
4	*P. aeruginosa*	9.0	9.6	0.69	→	10.0	11.1	0.61	→	8.6	9.4	0.57	→
5	*C. albicans*	7.4	17.6	2.1×10^{-9}	↗	9.0	13.3	$<2.2 \times 10^{-16}$	↗	8.4	13.5	0.00	↗
6	*S. aureus*	13.8	12.2	0.40	→	12.3	12.2	0.97	→	13.4	14.9	0.36	→
7	*S. epidermidis*	5.0	6.9	0.11	→	3.7	6.8	0.04	↗	3.3	5.0	0.07	→
8	*P. mirabilis*	0.9	0.6	0.76	→	0.9	1.8	0.31	→	0.8	1.6	0.12	→
9	*E. cloacae*	2.6	4.7	0.03	↗	3.1	3.2	0.95	→	5.3	5.0	0.68	→
10	*S. agalactiae*	1.0	0.4	0.36	→	0.0	1.1	0.02	↗	0.8	0.5	0.77	→
11	*K. oxytoca*	1.3	0.4	0.11	→	1.4	1.8	0.77	→	1.4	1.2	0.84	→
12	*E. faecium*	0.6	0.0	0.19	→	0.4	0.0	0.56	→	0.2	0.5	0.36	→
13	*H. influenzae*	8.7	2.9	2.5×10^{-5}	↘	7.3	3.9	0.05	→	6.1	2.8	0.00	→
14	*S. haemolyticus*	2.3	3.1	0.37	→	2.9	2.5	0.76	→	1.7	1.8	0.90	→
15	*C. koseri*	1.0	0.8	1.00	→	0.7	0.4	1.00	→	0.5	0.5	1.00	→
16	*E. aerogenes*	1.3	2.7	0.07	→	1.1	0.7	0.73	→	1.6	2.1	0.38	→
17	*S. saprophyticus*	0.0	0.0	1.00	→	0.0	0.0	1.00	→	0.0	0.0	1.00	→
18	*M. morganii*	0.5	1.4	0.06	→	0.4	0.0	0.56	→	0.8	0.5	0.77	→
19	*S. pneumoniae*	4.6	2.0	0.01	↘	4.0	2.2	0.15	→	3.7	3.0	1.48	→
20	*C. glabrata*	1.0	1.4	0.41	→	1.3	0.7	0.74	→	0.7	1.1	0.40	→
21	*Autres*	24.3	22.5	0.42	→	28.7	27.2	0.64	→	29.8	25.4	0.05	→

↗ Significant growth; ↘ Significant decrease; → Non-significant change; Evol = Evolution.

3.2.2. Diversity in Blood Samples

A significant increase in relative frequency was observed for *E. faecalis* and *S. haemolyticus* for blood samples during the lockdown, which was maintained during the post-lockdown only for *S. haemolyticus* (Table 3). No other variation in relative frequency was observed during the restoration period.

Table 3. Evolution of relative abundance for blood cultures sample from week 12 to week 35, 2020, Marseille.

N°	Species	During Lockdown (Weeks 12–19)				During Restoration (Weeks 20–24)				During Post-Lockdown (Weeks 25–S35)			
		2017-19 %	2020 %	*p*-Value	Evol	2017-19 %	2020 %	*p*-Value	Evol	2017-19 %	2020 %	*p*-Value	Evol
1	*E. coli*	5.8	1.0	0.06	→	6.8	1.2	0.08	→	3.0	4.2	0.85	→
2	*K. pneumoniae*	3.7	1.9	0.53	→	2.4	2.4	1.00	→	5.7	3.7	0.30	→
3	*E. faecalis*	2.4	7.8	0.03	↗	1.9	0.0	0.58	→	2.8	4.8	0.18	→
4	*P. aeruginosa*	2.7	4.9	0.34	→	3.4	2.4	1.00	→	6.5	4.3	0.27	→
5	*C. albicans*	9.8	3.9	0.07	→	6.3	3.6	0.57	→	3.9	2.7	0.42	→
6	*S. aureus*	11.3	8.7	0.58	→	14.1	16.9	0.58	→	12.4	13.8	0.61	→
7	*S. epidermidis*	24.1	33.0	0.07	→	23.8	32.5	0.14	→	23.4	25.0	0.66	→
8	*P. mirabilis*	1.8	0.0	0.34	→	1.0	0.0	1.00	→	1.2	2.1	0.47	→
9	*E. cloacae*	2.7	1.9	1.00	→	3.9	4.8	0.75	→	4.5	5.3	0.66	→
10	*S. agalactiae*	0.3	0.0	1.00	→	0.0	0.0	1.00	→	0.4	0.5	1.00	→
11	*K. oxytoca*	0.9	1.0	1.00	→	1.0	0.0	1.00	→	0.6	0.0	0.57	→
12	*E. faecium*	0.6	1.0	1.00	→	1.0	0.0	1.00	→	0.6	1.1	0.62	→
13	*H. influenzae*	0.3	0.0	1.00	→	0.5	1.2	0.49	→	0.0	0.0	1.00	→
14	*S. haemolyticus*	3.4	10.7	0.003	↗	2.4	6.0	0.16	→	2.8	8.5	0.001	↗
15	*C. koseri*	0.0	0.0	1.00	→	0.0	1.2	0.29	→	0.2	0.0	1.00	→
16	*E. aerogenes*	0.3	1.0	1.00	→	0.5	1.2	0.49	→	0.4	1.1	0.30	→
17	*S. saprophyticus*	0.0	0.0	1.00	→	0.5	1.2	0.49	→	0.0	0.0	1.00	→
18	*M. morganii*	0.0	1.0	1.00	→	0.0	0.0	1.00	→	0.0	0.0	1.00	→
19	*S. pneumoniae*	0.3	0.0	1.00	→	0.5	0.0	1.00	→	0.0	0.0	1.00	→
20	*C. glabrata*	0.3	0.0	1.00	→	0.5	0.0	1.00	→	0.6	0.0	0.57	→
21	*Autres*	29.3	22.3	0.17	→	29.6	25.3	0.46	→	30.3	22.9	0.05	→

↗ Significant growth; → Non-significant change; Evol = Evolution.

3.2.3. Diversity in Urine Samples

E. coli significantly decreased from 46.5% to 38.4% during the lockdown, in contrast to *C. albicans* (from 3.0% to 5.1%), *E. cloacae* (from 2.3% to 3.5%), and *C. glabrata* (from 0.6% to 1.1%) (Table 4) which significantly increased (Table 4). During the restoration period, only *C. albicans* (2.9% to 4.9%) increased and *C. koseri* (1.6% to 0.8%) decreased. No significant variation was observed for the post-lockdown period.

Table 4. Evolution of relative abundance for urine samples from week 12 to week 35, 2020, Marseille.

N°	Species	During Lockdown (Weeks 12–19)				During Restoration (Weeks 20–24)				During Post-Lockdown (Weeks 25–S35)			
		2017-19 %	2020 %	*p*-Value	Evol	2017-19 %	2020 %	*p*-Value	Evol	2017-19 %	2020 %	*p*-Value	Evol
1	*E. coli*	46.5	38.4	8.9×10^{-9}	↘	44.7	41.6	0.07	→	44.1	43.4	0.53	→
2	*K. pneumoniae*	9.7	9.2	0.48	→	9.5	11.2	0.10	→	11.3	10.7	0.43	→
3	*E. faecalis*	7.7	8.7	0.17	→	7.4	7.4	0.96	→	7.1	6.9	0.70	→
4	*P. aeruginosa*	3.6	4.0	0.51	→	3.9	3.6	0.72	→	4.2	3.8	0.43	→
5	*C. albicans*	3.0	5.1	8.3×10^{-5}	↗	2.9	4.9	0.002	↗	3.6	3.6	0.91	→
6	*S. aureus*	1.6	1.7	0.62	→	1.8	2.1	0.57	→	1.5	1.6	0.55	→
7	*S. epidermidis*	2.1	2.5	0.35	→	2.3	2.0	0.56	→	1.8	1.9	0.80	→
8	*P. mirabilis*	3.1	3.7	0.22	→	3.0	3.5	0.46	→	3.5	3.8	0.45	→

9	*E. cloacae*	2.3	3.5	0.005	↗	2.5	3.1	0.28	→	2.9	3.5	0.08	→
10	*S. agalactiae*	2.7	2.7	0.93	→	2.3	2.0	0.53	→	2.5	2.9	0.24	→
11	*K. oxytoca*	1.6	1.0	0.07	→	1.4	1.8	0.38	→	1.3	1.2	0.48	→
12	*E. faecium*	1.7	1.8	0.70	→	1.3	1.6	0.46	→	1.4	1.3	0.67	→
13	*H. influenzae*	0.0	0.0	1.00	→	0.0	0.0	1.00	→	0.0	0.0	1.00	→
14	*S. haemolyticus*	0.8	1.0	0.47	→	1.0	1.1	0.79	→	0.8	0.7	0.51	→
15	*C. koseri*	1.2	1.7	0.17	→	1.6	0.8	0.04	↘	1.1	1.1	0.95	→
16	*E. aerogenes*	1.1	1.4	0.35	→	1.0	1.0	0.99	→	1.3	1.6	0.27	→
17	*S. saprophyticus*	1.2	1.2	0.81	→	1.4	1.3	0.89	→	1.3	1.5	0.34	→
18	*M. morganii*	0.9	1.2	0.39	→	1.1	0.5	0.05	→	1.0	1.2	0.38	→
19	*S. pneumoniae*	0.1	0.0	0.61	→	0.1	0.0	1.00	→	0.1	0.0	0.59	→
20	*C. glabrata*	0.6	1.1	0.02	↗	0.6	0.8	0.43	→	0.6	0.6	0.80	→
21	*Autres*	8.5	10.2	0.03	↗	10.2	9.9	0.75	→	8.9	8.7	0.82	→

↗ Significant growth; ↘ Significant decrease; → Non-significant change, Evol = Evolutio.

3.2.4. Diversity in Intensive Care Units and Emergency Reception

In intensive care units, eleven of the twenty species (55%) had a significant change in relative abundance during the lockdown period. Of these eleven species, seven (63.6%) experienced a significant decrease in relative abundance (Table 5). However, the restoration and post-lockdown phases experienced, respectively, a decrease of 53% (7/13) and 58% (7/12) of their relative abundance. At the adult emergency unit, only 7/17 species (41.2%) showed a significant decrease in relative abundance compared to an increase for 10/17 species (58.8%) during lockdown. A sharp decrease in relative abundance was observed for almost all species with a significant change 7/8 (87.5%) during the post-lockdown period (Table 6).

Table 5. Evolution of relative abundance in intensive care units from week 12 to week 35, 2020, Marseille.

N°	Species	During Lockdown (Weeks 12–19)				During Restoration (Weeks 20–24)				During Post-Lockdown (Weeks 25–35)			
		2017-19 %	2020 %	*p*-Value	Evol	2017-19 %	2020 %	*p*-Value	Evol	2017-19 %	2020 %	*p*-Value	Evol
1	*E. coli*	10.9	8.6	0.2	→	9.2	4.6	0,01	↘	15.3	13.4	0.27	→
2	*K. pneumoniae*	13.6	7.8	0.003	↘	0	0	1	→	8.1	5.5	0.04	↘
3	*E. faecalis*	0.8	2.8	0.003	→	4.3	0	0.0001	↘	6.8	7.4	0.64	→
4	*P. aeruginosa*	6.0	0	7.30×10^{-7}	↘	10.7	17.4	0.002	↗	7.1	2.6	7.50×10^{-5}	↘
5	*C. albicans*	4.1	10.4	3.40×10^{-6}	↗	8.3	0	7.30×10^{-8}	↘	8.7	19.3	2.70×10^{-11}	↗
6	*S. aureus*	3.0	24.8	$<2.2 \times 10^{-16}$	↗	23	0.6	$<2.2 \times 10^{-16}$	↘	10.1	8.6	0.3	↘
7	*S. epidermidis*	9.5	4.3	0.001	↘	3.5	34.9	$<2.2 \times 10^{-16}$	↗	7.2	3.1	0.0003	↘
8	*P. mirabilis*	1.1	0.3	0.2	→	0.1	2.1	0.001	↗	2.4	1.6	0.25	→
9	*E. cloacae*	2.1	3.8	0.07	→	8.5	3.4	0.002	↘	5.7	7.9	0.05	→
10	*S. agalactiae*	0.8	1	0.75	→	0.2	0	1	→	0.8	0.3	0.38	→
11	*K. oxytoca*	1.4	0.3	0.09	→	1.7	0.3	0.08	→	1.2	2.8	0.01	↗
12	*E. faecium*	2.4	0.3	0.01	↘	1.4	0	0.02	↘	1	1.9	0.09	→
13	*H. influenzae*	7.7	0.3	$<4.5 \times 10^{-8}$	↘	3.6	6.4	0.03	↗	0.8	0	0.03	↘

14	*S. haemolyticus*	1.8	0	0.01	↘	0.5	1.2	0.23	→	0.5	3.6	5.90×10^{-9}	↗
15	*C. koseri*	1.1	0	0.04	↘	0.4	0	0.56	→	0.8	1.7	0.05	↗
16	*E. aerogenes*	2.5	5.6	0.003	↗	0.7	0.9	0.71	→	1.7	0.3	0.02	↘
17	*S. saprophyticus*	0	0.3	0.26	→	0	0	1	→	0.1	0	1	→
18	*M. morganii*	0.2	0.3	1	→	0	0	1	→	0.8	1.9	0.03	↗
19	*S. pneumoniae*	2.8	5.6	0.01	↗	2	0	0.01	↘	0.5	0.2	0.47	→
20	*C. glabrata*	2.0	1.3	0.32	→	0.1	1.5	0.01	↗	1	0.5	0.27	→
21	*Autres*	26.3	22.5	0.14	→	21.9	73.4	$<2.2 \times 10^{-16}$	↗	80.5	17.2	$<2.2 \times 10^{-16}$	↘

↗ Significant growth; ↘ Significant decrease; → Non-significant change, Evol = Evolutio.

Table 6. Evolution of relative abundance emergency units from week 12 to week 35, 2020, Marseille.

N°	Species	During Lockdown (Weeks 12–19)				During Restoration (Weeks 20–24)				During Post-Lockdown (Weeks 25–35)			
		2017-19	2020	*p*-Value	Evol	2017-19	2020	*p*-Value	Evol	2017-2019	2020	*p*-Value	Evol
		%	%			%	%			%	%		
1	*E. coli*	31.0	17.7	3.50×10^{-7}	↘	40.3	0.6	<2.2e-16	↘	25.0	39.0	1.10e-14	↗
2	*K. pneumoniae*	1.3	3.3	0.01	↗	9.5	0.0	3.60e-09	↘	11.0	11.5	0.64	→
3	*E. faecalis*	8.2	0.8	5.30×10^{-7}	↘	0.3	13.6	<2.2e-16	↗	2.3	2.6	0.58	→
4	*P. aeruginosa*	5.9	2.2	0.004	↘	2.8	8.0	1.90e-05	↗	5.2	6.0	0.36	→
5	*C. albicans*	7.8	0.6	3.80×10^{-7}	↘	2.2	26.0	$<2.2 \times 10^{-16}$	↗	3.7	3.5	0.82	→
6	*S. aureus*	5.5	14.4	2.70×10^{-9}	↗	7.8	16.3	3.10×10^{-6}	↗	6.1	0.7	2.70e-11	↘
7	*S. epidermidis*	8.1	14.1	0.0004	↗	7.1	0.0	4.50e-07	↘	5.7	7.5	0.06	→
8	*P. mirabilis*	1.9	0	0.008	↘	0.6	0.0	0.36	→	3.9	4.2	0.7	→
9	*E. cloacae*	0.8	0.8	1	→	0.3	7.7	$<2.2 \times 10^{-16}$	↗	4.6	1.3	9.50×10^{-6}	↘
10	*S. agalactiae*	0.2	4.7	3.10×10^{-10}	↗	1.1	0.6	0.54	→	2.3	0.1	1.40×10^{-5}	↘
11	*K. oxytoca*	1.3	2.8	0.04	↗	0.8	0.0	0.13	→	1.2	0.3	0.02	↘
12	*E. faecium*	1.2	0	0.04	↘	0.5	0.0	0.35	→	1.0	1.4	0.31	→
13	*H. influenzae*	0.7	0	0.23	→	2.3	2.4	0.98	→	1.3	0.1	0.002	↘
14	*S. haemolyticus*	2.3	4.7	0.01	↗	1.9	1.2	0.36	→	1.6	0.4	0.01	↘
15	*C. koseri*	1.3	0	0.02	↘	0.1	0.0	1	→	0.9	1.0	0.73	→
16	*E. aerogenes*	0.2	0.8	0.11	→	0.3	4.4	4.60×10^{-8}	↗	2.4	1.5	0.13	→
17	*S. saprophyticus*	0.1	3.9	2.70×10^{-10}	↗	1.4	0.0	0.03	↘	1.7	0.0	6.80×10^{-5}	↘
18	*M. morganii*	0.9	2.5	0.03	↗	0.2	0.0	1	↘	0.3	0.0	0.19	→
19	*S. pneumoniae*	1.1	0	0.06	→	1.8	0.6	0.1	→	0.6	1.0	0.21	→
20	*C. glabrata*	0.4	3.3	4.90×10^{-6}	↗	0.8	0.0	0.22	→	0.6	0.0	0.13	→
21	*Autres*	19.8	76.5	$<2.2 \times 10^{-16}$	↗	17.9	81.4	$<2.2 \times 10^{-16}$	↗	81.4	82.3	0.55	→

↗ Significant growth; ↘ Significant decrease; → Non-significant change, Evol = Evolutio.

3.3. Evolution of Antibiotic Resistance

Whatever the origin of the infection, the analysis of the evolution of bacterial antibiotic resistance showed a significant increase in the percentage of wild phenotypes during 2020 compared to the control period for *E. coli* (45.4% to 48.5%), *K. pneumoniae* (59.6% to 67.7%), *P. mirabilis* (56.4% to 64.1%), and *P. aeruginosa* (56.0% to 64.9%) (Table 3). The other species belonging to the 20 most represented species did not show any significant change.

The wild percentage for community infection significantly increased for *E. coli* and *P. aeruginosa*, whereas it decreased for *K. pneumoniae* and *P. mirabilis*. However, for nosocomial infection, this percentage significantly decreased for only *P. aeruginosa* and increased for *K. pneumoniae*.

However, regarding the infection origin (nosocomial or community), the percentage of wild phenotypes significantly decreased when the origin was nosocomial and significantly increased when the origin was community for *E. aerogenes*, *E. faecium*, *K. oxytoca*, and *M. morganii* (Table 7). *E. faecalis* presented a decreasing percentage for nosocomial infection and *E. cloacae* an increasing percentage for community infection.

Table 7. Comparison of wild percentage by origin of infection 2017–2019 vs. 2020, Marseille.

Species	Global								Origin of Infection															
	2017–2019			2020			*p*-Value	Evol *	Nosocomial								Community							
									2017–2019			2020			*p*-Value	Evol *	2017–2019			2020			*p*-Value	Evol *
	W	R	%	W	R	%			W	R	%	W	R	%			W	R	%	W	R	%		
A. baumannii	74	22	77.1	9	3	75.0	1.00	→	20	22	47.6	9	3	75	0.11	→	54	0	100	0	0	-	-	→
E. aerogenes	342	49	87.5	119	16	88.2	0.84	→	121	31	79.6	0	16	0	1.6×10^{-10}	↘	221	18	92.5	119	0	100	0.001	↗
E. cloacae	786	348	69.3	265	92	74.2	0.08	→	515	0	100	0	0	-	-	→	271	348	43.8	265	92	74.2	$<2.2 \times 10^{-16}$	↗
E. faecalis	1987	9	99.6	348	1	99.7	1.00	→	1374	1	99.9	0	1	0	0.001	↘	613	8	98.7	348	0	100	0.06	→
E. faecium	393	60	86.8	64	5	92.8	0.16	→	273	42	86.7	0	5	0	5.7×10^{-5}	↘	120	18	87.0	64	0	100	0.001	↗
E. coli	4791	5771	45.4	1315	1398	48.5	0.004	↗	0	1892	0.0	0	1398	0	1.00	→	4791	3879	55.3	1315	0	100	$<2.2 \times 10^{-16}$	↗
K. oxytoca	408	91	81.8	103	22	82.4	0.87	→	148	61	70.8	0	22	0	2.2×10^{-11}	↘	260	30	89.7	103	0	100	0.0001	↗
K. pneumoniae	1542	1047	59.6	484	231	67.7	7.7×10^{-5}	↗	1027	373	73.4	484	0	100	$<2.2 \times 10^{-16}$	↗	515	674	43.3	0	231	0	$<2.2 \times 10^{-16}$	↘
M. morganii	285	52	84.6	81	14	85.3	0.87	→	180	38	82.6	0	14	0	1.8e-10	↘	105	14	88.2	81	0	100	0.001	↗
P. mirabilis	532	411	56.4	177	99	64.1	0.02	↗	154	287	34.9	0	0	-	-	→	378	124	75.3	177	99	64.1	0.001	↘
P. aeruginosa	1343	1055	56.0	432	234	64.9	4.2×10^{-5}	↗	901	637	58.6	0	234	0	$<2.2 \times 10^{-16}$	↘	442	418	51.4	432	0	100	0.001	↗
S. marcescens	304	8	97.4	119	7	94.4	0.15	→	191	6	97.0	0	7	0	6.5×10^{-10}	↘	113	2	98.3	119	0	100	0.24	→
S. aureus	4458	486	90.2	31	2	93.9	0.77	→	1455	154	90.4	0	0	-	-	→	3003	332	90.0	31	2	93.9	0.77	→
S. agalactiae	770	43	94.7	123	5	96.1	0.51	→	253	0	100	0	0	-	-	→	517	43	92.3	123	5	96.1	0.17	→

↗ Significant growth; ↘ Significant decrease; → Non-significant change, Evol = Evolution; W = Wild; R = Resistance.

4. Discussion

The consequences of COVID-19 on the diversity of non-viral infectious agents and on the evolution of antibiotic resistance is notable.

The overall diversity analysis shows that the main changes were within pulmonary samplings, which were characterized by an increase of the number of different species identified in association with an increase of the species distribution evenness. Blood samples were also affected, albeit with less different species. No evident change can be observed for the urine samples.

As we worked on species' relative frequencies for taking in account the shrinking of hospital admissions during the lockdown, the results must be interpreted in terms of species replacements. That is, the decrease of a species or group of species' relative frequency is associated with the increase of the relative frequency of another group.

Considering the overall species' distributions, *E. coli* was characterized by a decreasing of its frequency on the whole study period, while *C. albicans* and *S. epidermidis* showed

an increase during the same period. The other species show only temporary changes, such as an increase only during the lockdown for *S. haemolyticus* or *E cloacae*, or a decrease during this same period for *K. oxytoca* or *H. influenzae*.

When considering the kind of sample we can observe that, if *E. coli* presents a homogenous global decrease, this global behavior cannot be retrieved in the sample related results. We may think that this is an example of Simpson's paradox [17]. On the contrary, *C. albicans* and *C. glabrata* exhibited a homogenous increase on all sample related results, just as on the global one.

Beside *E. coli*, the only species decreasing in respiratory samplings, and only during the lockdown, were *H. influenzae* and *S. pneumoniae*. It is interesting to observe that these two species are well known viral co-infections, frequently observed in association with the influenza virus, and capable of coexisting in the same biofilm [18]

Few alterations of the blood-related bacterial mix were observed, concerning only *E. faecalis* and *S. haemolyticus*. They are considered to be important nosocomial pathogens, but the same behavior cannot be observed for other well-known nosocomial pathogens such as *S. aureus* or *S. epidermidis*.

More frequency changes can be observed in emergency than in intensive care units. For only one species, *P. aeruginosa*, the changes are the same, i.e., a decrease during lockdown followed by an increase. For several species (*C. albicans*, *S. epidermidis*, *S. haemolyticus*, and *K. pneumoniae*) the changes are inversed between the two kinds of units. Frequency changes only for intensive care units are specific of *H. influenzae* (decrease followed by an increase) and *S. pneumoniae* (increase followed by a decrease), while changes for *E. faecalis* (decrease followed by an increase), *S. saprophyticus*, *M. morganii* (in both cases, an increase followed by a decrease), and *S. agalactiae* (increase) were observed only in emergency units. Contou and al. [19] observed the same species behaviors in their ICU, at the exception of *H. influenzae*, which decreased in our series. When considering the global frequencies of Candida albicans and glabrata, a first hypothesis explaining their increase may be the relative importance of the admissions in ICU during the lockdown and the immediate period after.

Candida albicans and *glabrata* showed a significant increase of their relative frequency, whatever the type of sample and the period studied. This behavior is found in intensive care units, which were heavily impacted during the SARS-CoV-2 epidemic, and not in emergency units, possibly explaining their significant increases during the lockdown [20,21].

The wild phenotype population has also increased in comparison with the previous three years for the twenty most identified species in our institute. *E. coli* and *P. aeuruginosa* present more frequently a wild phenotype than usually in the context of community-acquired infection. However, the susceptibility of *E. coli* to most antibiotics involved in community-acquired urinary tract infections tended to decrease before the COVID-19 pandemic [22].

In France, government responses taken to limit the spread of the virus, such as lockdown measures, probably played a role in the evolution of the identification of bacteria and fungi [2,3]. Indeed, it was recommended that individuals should stay at home and contact the emergency call center (number 15) only in the event of respiratory distress in order to avoid congesting hospital resources and to prevent the spread of the disease [23]. In addition, in order to manage patients with COVID-19, many hospital departments have been transformed to accommodate these SARS-CoV-2 positive patients, which explains this increase in the number of hospitalizations. Non-emergency hospital activities were suspended. Thus, the number of patients and ordinary hospitalization outside of COVID-19 decreased considerably during the first containment, partly explaining this decrease in some pathogens and the increase in others.

5. Conclusions

The systematic use of a surveillance system, such as the MIDaS syndromic collection and the surveillance system at IHU-MI, made it possible to detect aberrations in the epi-

demic signal, to observe and analyze unexpected increases in observed cases, to implement actions to stop the spread of a pathogen, but also to understand the underlying mechanisms of its transmission. This study shows that a such system allowed us to detect and analyze the consequences of the lockdown on the bacterial and fungal population identified within our patients. Bacterial populations usually associated with seasonal viral infection where drastically reduced, while a usual raising of infections associated with *C. albicans* and *glabrata* emerged, possibly due to the increase of patients admitted in ICU.

Author Contributions: Conceived and designed by H.C. and D.R.; L.K., A.G.-G. and M.-T.J. collected data or/and performed experiments; L.K., A.G.-G., P.C., J.-M.R., H.C. and D.R. analyzed and interpreted data; A.G.-G., L.K., P.C., D.R. and H.C. wrote the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the French Government under the "Investments for the Future" program managed by the National Agency for Research (ANR), Méditerranée-Infection 10-IAHU-03, and was also supported by the Région Provence Alpes Côte d'Azur and European funding FEDER PRIMMI (Fonds Européen de Développement Régional—Plateformes de Recherche et d'Innovation Mutualisées Méditerranée Infection), FEDER PA 0000320 PRIMMI.

Institutional Review Board Statement: The information system involved in this study and the associated data analysis was declared and approved by the Commission Nationale Informatique et Liberté (declaration number 2139516 v 0).

Informed Consent Statement: The information system involved in this study and the associated data analysis was declared and approved by the Commission Nationale Informatique et Liberté (declaration number 2139516 v 0).

Data Availability Statement: The data from our surveillance system are not available on the public domain, but anyone interested in using the data for scientific purpose is free to request permission from the corresponding author: Hervé Chaudet (herve.chaudet@gmail.com).

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Vanhems, P. SARS-CoV2 infection and primary school closure. *Euro. Surveill.* **2020**, *25*, 2000617, doi:10.2807/1560-7917.ES.2020.25.15.2000617.
2. Moatti, J.P. The French response to COVID-19: Intrinsic difficulties at the interface of science, public health, and policy. *Lancet Public Health* **2020**, *5*, e255, doi:10.1016/S2468-2667(20)30087-6.
3. Décret n°2020-260 du 16 Mars 2020 Portant Réglementation des Déplacements dans le Cadre de la Lutte Contre la Propagation du Virus COVID-19. Available online: https//www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000041728476 (accessed on 18 November 2020).
4. Thomas, H.; Boby, T.; Angrist, N.; Cameron-Blake, E.; Hallas, L.; Kira, B. Variation in Government Responses to COVID19. Version 9.0. Blavatnik School of Government Working Paper. 10 December 2020. Available online: www.bsg.ox.ac.uk/covid-tracker (accessed on 22 January 2021).
5. The Lancet Gastroenterology Hepatology. Drinking alone: COVID-19, lockdown, and alcohol-related harm. *Lancet Gastroenterol. Hepatol.* **2020**, *5*, 625.
6. Pierce, M.; Hope, H.; Ford, T.; Hatch, S.; Hotopf, M.; John, A.; Kontopantelis, E.; Webb, R.; Wessely, S.; McManus, S.; et al. Mental health before and during the COVID-19 pandemic: A longitudinal probability sample survey of the UK population. *Lancet Psychiatry* **2020**, *7*, 883–892.
7. Palazzolo, C.; Maffongelli, G.; D'Abramo, A.; Lepore, L.; Mariano, A.; Vulcano, A.; Bartoli, T.A.; Bevilacqua, N.; Giancola, M.L.; Di Rosa, E.; et al. Legionella pneumonia: Increased risk after COVID-19 lockdown? Italy, May to June 2020. *Eur. Surveill.* **2020**, *25*, 2001372.
8. Faber, M.; Ghisletta, A.; Schmidheiny, K. A lockdown index to assess the economic impact of the coronavirus. *Swiss. J. Econ. Stat.* **2020**, *156*, 11, doi:10.1186/s41937-020-00056-8.
9. Flaherty, G.T.; Hession, P.; Liew, C.H.; Lim, B.C.W.; Leong, T.K.; Lim, V.; Sulaiman, L.H. COVID-19 in adult patients with pre-existing chronic cardiac, respiratory and metabolic disease: A critical literature review with clinical recommendations. *Trop. Dis. Travel Med. Vaccines* **2020**, *6*, 16, doi:10.1186/s40794-020-00118-y.
10. Willan, J.; King, A.J.; Djebbari, F.; Turner, G.D.H.; Royston, D.J.; Pavord, S.; Collins, G.; Peniket, A. Assessing the impact of lockdown: Fresh challenges for the care of haematology patients in the COVID-19 pandemic. *Br. J. Haematol.* **2020**, *189*, e224–e227, doi:10.1111/bjh.16782.

11. Stroup, D.F.; Williamson, G.D.; Herndon, J.L.; Karon, J.M. Detection of aberrations in the occurrence of notifiable diseases surveillance data. *Stat. Med.* **1989**, *8*, 323–332.
12. Lagier, J.C.; Khelaifia, S.; Alou, M.T.; Ndongo, S.; Dione, N.; Hugon, P.; Caputo, A.; Cadoret, F.; Traore, S.I.; Seck, E.H.; et al. Culture of previously uncultured members of the human gut microbiota by culturomics. *Nat Microbiol.* **2016**, *1*, 16203.
13. Colson, P.; Rolain, J.M.; Abat, C.; Charrel, R.; Fournier, P.E.; Raoult, D. EPIMIC: A simple homemade computer program for real-time EPIdemiological surveillance and alert based on MICrobiological data. *PLoS ONE* **2015**, *10*, e0144178.
14. Abat, C.; Chaudet, H.; Colson, P.; Rolain, J.M.; Raoult, D. Real-time microbiology laboratory surveillance system to detect abnormal events and emerging infections, Marseille, France. *Emerg. Infect. Dis.* **2015**, *21*, 1302–1310.
15. Giraud-Gatineau, A.; Texier, G.; Garnotel, E.; Raoult, D.; Chaudet, H. insights into subspecies discrimination potentiality from bacteria MALDI-tof mass spectra by using data mining and diversity studies. *Front. Microbiol.* **2020**, *11*, 1931.
16. R Core Team. *A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2019. Available online: https://www.R-project.org/ (accessed on 17 May 2021).
17. Chu, K.H.; Brown, N.J.; Pelecanos, A.; Brown, A.F. Simpson's paradox: A statistician's case study. *Emerg. Med. Australas.* **2018**, *30*, 431–433, doi:10.1111/1742-6723.12943.
18. Tikhomirova, A.; Kidd, S.P. Haemophilus influenzae and Streptococcus pneumoniae: Living together in a biofilm. *Pathog. Dis. Nov.* **2013**, *69*, 114–126.
19. Contou, D.; Claudinon, A.; Pajot, O.; Micaëlo, M.; Flandre, P.L.; Dubert, M.; Cally, R.; Logre, E.; Fraissé; M.; Mentec, H.; et al. Bacterial and viral co-infections in patients with severe SARS-CoV-2 pneumonia admitted to a French ICU. *Ann. Intensive Care* **2020**, *10*, 119.
20. Baldesi, O.; Bailly, S.; Ruckly, S.; Lepape, A.; L'Heriteau, F.; Aupee, M.; Boussat, S.; Bervas, C.; Machut, A.; Berger-Carbonne, A.; et al. ICU-acquired candidaemia in France: Epidemiology and temporal trends, 2004–2013—A study from the REA-RAISIN network. *J. Infect.* **2017**, *75*, 59–67, doi:10.1016/j.jinf.2017.03.011.
21. De Pascale, G.; Tumbarello, M. Fungal infections in the ICU: Advances in treatment and diagnosis. *Curr. Opin. Crit. Care* **2015**, *21*, 421–429, doi:10.1097/MCC.0000000000000230.
22. Kim, Y.J.; Lee, J.M.; Cho, J.; Lee, J. Change in the annual antibiotic susceptibility of escherichia coli in community-onset urinary tract infection between 2008 and 2017 in a Tertiary Care Hospital in Korea. *J. Korean Med. Sci.* **2019**, *34*, e228, doi:10.3346/jkms.2019.34.e228.
23. Préparation a la Phase Epidémique de COVID-19. Ministère des Solidarités et de la Santé, 16 March 2020. Available online: https://solidarites-sante.gouv.fr/IMG/pdf/guide-covid-19-phase-epidemique-v15-16032020.pdf (accessed on 27 March 2020).

Article 4

Article 4: Control of common viral epidemics but not of SARS-CoV-2 through the application of hygiene and distancing measures. *J. Clin. Vir*.
https://doi.org/10.1016/j.jcv.2022.105163

Giraud-Gatineau, A**, Kaba L**, Boschi C, Devaux D, Casalta JP, Gautret P, Chaudet H, Colson P, Raoult D.
Accepted in Journal of Clinical Virology (April 2021)

Journal of Clinical Virology 150–151 (2022) 105163

Contents lists available at ScienceDirect

Journal of Clinical Virology

journal homepage: www.elsevier.com/locate/jcv

Control of common viral epidemics but not of SARS-CoV-2 through the application of hygiene and distancing measures

Audrey Giraud-Gatineau [a,b,c,d], Lancei Kaba [a,b,e], Céline Boschi [a,d,e], Christian Devaux [a,e,f], Jean-Paul Casalta [a,d,e], Philippe Gautret [a,b,d], Hervé Chaudet [a,b,c,d], Philippe Colson [a,d,e], Didier Raoult [a,e,*]

[a] IHU Méditerranée Infection, 19-21 boulevard Jean Moulin, 13005, Marseille, France
[b] Aix Marseille Univ, Institut de Recherche pour le Développement (IRD), Assistance Publique – Hôpitaux de Marseille (AP-HM), Service de Santé des Armées (SSA), Vecteurs – Infections Tropicales et Méditerranéennes (VITROME), 27 boulevard Jean Moulin, 13005, Marseille, France;
[c] French Armed Forces Center for Epidemiology and Public Health (CESPA), Service de Santé des Armées (SSA), camp de Sainte Marthe, BP 40026, Marseille, France
[d] Assistance Publique- Hôpitaux de Marseille (AP-HM), 264 rue Saint-Pierre, 13005, Marseille, France
[e] Aix-Marseille Univ, Institut de Recherche pour le Développement (IRD), Assistance Publique - Hôpitaux de Marseille (AP-HM), Microbes Evolution Phylogeny and Infections (MEPHI), 27 boulevard Jean Moulin, 13005, Marseille, France
[f] Centre National de la Recherche Scientifique (CNRS), Marseille, France

ARTICLE INFO

Key words:
Coronavirus
Respiratory infections
Gastrointestinal infections
SARS-CoV-2
MIDaS
Epidemic
Surveillance

ABSTRACT

Background: We systematically survey respiratory and gastrointestinal infections of viral origin in samples sent to our university hospital institute in Marseille, southern France. Here, we evaluated whether the measures implemented to fight COVID-19 had an effect on the dynamics of viral respiratory or gastrointestinal infections.
Methods: We analysed PCR performed and positive for the diagnoses of viral respiratory and gastrointestinal infections over five years (January 2017-February 2021). Data were collected from our epidemiological surveillance system (MIDaS). Dates and contents of French measures against SARS-CoV-2 were collected from: https://www.gouvernement.fr/info-coronavirus/les-actions-du-gouvernement.
Results: Over the 2017-2021 period, 990,364 analyses were carried out for respiratory infections not including SARS-CoV-2, 510,671 for SARS-CoV-2 and 27,719 for gastrointestinal infections. During winter 2020–2021, when the most restrictive lockdown measures were in place in France, a marked decrease of infections with influenza viruses (one case versus 1,839-1,850 cases during 2017-2020 cold seasons) and with the RSV (56 cases versus 988-1,196 cases during 2017-2020 cold seasons) was observed, demonstrating the relative effectiveness of these measures on their occurrence. SARS-CoV-2 incidence seemed far less affected. Rhinoviruses, parainfluenza 3 virus, and the coronavirus NL63 remained at comparable levels. Also, the norovirus winter season positivity rates decreased continuously and significantly over time from 9.3% in 2017–2018 to 2.0% in 2020–2021.
Conclusion: The measures taken to control COVID-19 were effective against lower respiratory tract infections viruses and gastroenteritis agents, but not on the agents of the common winter cold and SARS-CoV-2. This suggests that more specific measures to prevent COVID-19 and upper respiratory tract infections need to be discovered to limit the spread of this epidemic.

1. BACKGROUND

In December 2019, a new coronavirus named SARS-CoV-2 (for severe acute respiratory syndrome coronavirus 2) emerged in Wuhan, Hubei region, China. It spread rapidly to the rest of the world and was declared a pandemic in March 2020 [1]. As of February 09, 2022 402,064,265 SARS-CoV-2 cases and 5,768,927 patient deaths from COVID-19 (for coronavirus disease 2019) were reported [2]. SARS-CoV-2 variants have emerged since summer of 2020 [3,4] and have each determined an epidemic of variable intensity and duration. These variants have been revealed to be associated with differences regarding viral loads, transmissibility, and clinical severity and they have been involved in various degrees of escape to immunity elicited by vaccination or infection [5–8]. The dynamics of SARS-CoV-2 epidemics

* Corresponding author: Didier Raoult, IHU Méditerranée Infection, 19-21 boulevard Jean Moulin, 13005 Marseille, France
E-mail address: didier.raoult@gmail.com (D. Raoult).

https://doi.org/10.1016/j.jcv.2022.105163
Received 18 April 2021; Received in revised form 12 March 2022;
Available online 16 April 2022

at national and global scales proved to be unpredictable.

In order to reduce the spread of SARS-CoV-2, the French government decided to take several health and social measures. This initially involved repeated risk prevention messages on the use of protective measures including regularly hand washing with soap or alcohol-based hand gel, social distancing of 1.5 meter between individuals, and wearing a mask [9]. These measures had already been used in prevention campaigns for other viruses, particularly respiratory viruses such as influenza [10,11]. More restrictive measures on movement were also taken, with the implementation of a number of lockdowns and curfews (Decree No. 2020–260; Decree No. 2020–1310) [12,13]. Thus, in addition to the fight against COVID-19, these measures may also be effective at controlling other communicable respiratory and digestive diseases.

At the Hospital University Institute Méditerranée Infection (IHU-MI), the activity of the virology and microbiology laboratory is monitored by a collection and surveillance system known as MIDaS (for Mediterranée Infection Data Warehousing and Surveillance) [14,15]. This system enables us to monitor respiratory and digestive virus infections on a weekly basis, and has included COVID-19 since its emergence in France [16]. The objective of this study was to analyse the epidemiological curves of respiratory and gastrointestinal viruses since the emergence of SARS-CoV-2 and to evaluate if they changed under the measures implemented against COVID-19 in France by comparing them during cold seasons over the past five years.

2. MATERIALS AND METHODS

2.1. Surveillance system

Since 2003, the activity of our clinical microbiology laboratory has involved massive and unbiased monitoring of all clinical samples received for testing bacteria, viruses, parasites and microscopic fungi [16,17]. This followed recommendations from one of the authors (DR) [18] made to the French government in 2003 to set up surveillance systems of any abnormal events related to infectious diseases based on laboratory data, including through syndromic surveillance. Our laboratory is the single one to diagnose infections for all public university hospitals (Assistance Publique - Hôpitaux de Marseille (AP-HM)) in Marseille, which have a total of 3,288 beds with nearly 125,000 admissions and one million consultations per year. Our laboratory conducts approximately eight million tests every year.

Since 2013 when the IHU-MI was established, our surveillance tools have expanded further and have improved through our unique MIDaS (for Mediterranée Infection Data Warehousing and Surveillance) collection and surveillance system, which consists of five sub-systems [14]. We systematically collect all laboratory data (samples, tests, positive diagnoses) from the Nexlab laboratory management system. All microbiological analysis results (sample identification, requesting clinical department, date of sampling, analysis, result, antibiotic susceptibility testing, antibiotic resistance phenotype, bacterial co-identifications) and patient information (anonymised patient identification, age, sex, home postal code, anonymised hospital stay identification, date of stay within a department, death) are then deposited in a dedicated data warehouse. All clinical samples, tests and infectious agents are monitored on a weekly basis throughout the year. MIDaS automatically detects any aberrations in the statistical signal using the cumulative sum control chart (CUSUM) algorithm and triggers alarms [19]. These alarms are discussed during a weekly epidemiological staff meeting, which includes epidemiologists, biologists, and infectiologists.

Respiratory and gastrointestinal samples and infectious agents are some of the items surveyed. Respiratory and gastrointestinal viruses are diagnosed in our laboratory using commercial or in-house real-time Polymerase Chain Reaction (qPCR) tests and adopting a syndromic approach using multiplex or simplex tests. They included influenza A and B viruses, respiratory syncytial virus (RSV), rhinoviruses, enteroviruses, adenoviruses, metapneumovirus, endemic coronaviruses (HCoV-OC43, -NL63, -E229 and -HKU1), parainfluenza viruses 1 to 4 (HPIV1 to HPIV4) and SARS-CoV-2, over a period of time from January 2017 to February 2021. For the detection of SARS-CoV-2 RNA, we used in house qPCR procedures previously described [20]. To detect the other respiratory viruses, we used the FTD Respiratory pathogens 21 (Fast Track Diagnosis, Luxembourg), the Biofire FilmArray Respiratory panel 2 plus (BioMérieux, Marcy-l'Etoile, France), the Respiratory Multi-Well System r-gene (BioMérieux), or the GeneXpert Xpert Flu/RSV (Cepheid, Sunnyvale, CA, USA) assays [21].

Data on diagnoses of influenza A and B viruses were also collected from a private clinical microbiology and virology laboratory through the PACASurvE (for the Provence Alpes-Côte d'Azur Surveillance Epidemiological System) network that extents our surveillance system to private medical biology laboratories located in the Provence Alpes-Côte d'Azur French region that includes Marseille [22]. These diagnoses were reached by an immunochromatographic assay in 2017 and then by the GeneXpert Flu/RSV assay between 2018 and 2021.

The gastrointestinal viruses diagnosed included adenoviruses, rotaviruses, sapoviruses, noroviruses and astroviruses. The tests were performed using the Fast Track Diagnosis viral gastroenteritis pathogens assay (Fast Track Diagnosis).

2.2. Statistical analyses

In order to better understand the evolution of respiratory and gastrointestinal virus infections over time, the proportion of positive results between October and the end of February were compared for each virus for the 2017–2018, 2018–2019, 2019–2020 and 2020–2021 seasons. These evolutions were analysed using the log-linear model, and the Fisher and Chi-square tests for point comparisons with a two-tailed statistical significance threshold of 0.05. Statistical analyses were done using R version 4.1 [23].

2.3. Government measures and policies

Measures taken by the French government in the fight against the spread of SARS-CoV-2 and dates when these measures were implemented were collected from the government website (https://www.gouvernement.fr/info-coronavirus/les-actions-du-gouvernement).

3. RESULTS

3.1. Diagnoses of respiratory viral infections at IHU-MI from 2017 to 2021

Over a period of five years (January 2017 to February 2021), 990,364 analyses were performed for common respiratory viruses, with 37,915 positive results. Most of these cases were due to influenza viruses (influenza A virus, 6,544; influenza B viruses, 2,459) followed by rhinoviruses (7,379), RSVs (3,846), adenoviruses (1,991), metapneumoviruses (1,482), enteroviruses (790), HCoV HKU1 (424), HCoV NL63 (421), HCoV OC43 (227), HCoV E229 (87), HPIV3 (340), HPIV4 (68), HPIV2 (18) and HPIV4 (9) (Table 1).

Slight yearly variations were observed from 2017 to 2019 with regards to the respective prevalence of these viruses (Fig. 1). In 2017, the influenza A virus was the most frequently identified respiratory viral agent (12.1%), followed by rhinovirus (9.3%) and RSV (7.2%). In the same year, 1.7% of samples tested for influenza B virus were positive for this agent. In 2018, the rhinovirus was the most commonly diagnosed (12.6%), compared to 6.9% for RSV, 6.3% for influenza A virus and 5.2% for influenza B virus. 2019 was comparable to 2017 in terms of the ranking of respiratory viruses, although the proportions of respiratory viruses' diagnoses were higher in 2019. The intensity of the epidemic peak for each of these respiratory viruses therefore changed over the years, as did the date upon which they appeared (Fig. 1).

Table 1
Tests performed and positive for PCR detection of respiratory viruses in 2017, 2018, 2019, 2020 and 2021 at Hospital University Institute Méditerranée Infection.

	2017	2018	2019	2020	2021	Total 2017-2021	2017		2018		2019		2020		2021		Total2017-2021	
	N	N	N	N	N		N	%	N	%	N	%	N	%	N	%	N	%
Adenovirus	5,656	14,881	17,636	32,237	5,283	**75,693**	160	2.8	449	3.0	652	3.7	600	1.9	130	2.5	**1,991**	**2.6**
Common coronaviruses	2,395	3,773	8,211	32,237	5,283	**51,899**	70	2.9	110	2.9	231	2.8	998	3.1	276	5.2	**1,685**	**3.2**
HCoV 229E	2,395	3,773	8,211	11,739	5,283	**31,401**	0	0.0	0	0.0	0	0.0	84	0.7	3	0.1	**87**	**0.3**
HCoV HKU1	2,395	3,773	8,211	11,736	5,283	**31,398**	0	0.0	0	0.0	0	0.0	423	3.6	1	0.0	**424**	**1.4**
HCoV NL63	3,791	3,773	8,211	11,739	5,283	**32,797**	0	0.0	0	0.0	0	0.0	252	2.1	169	3.2	**421**	**1.3**
HCoV OC43	9,007	3,773	8,211	11,74	5,283	**38,014**	0	0.0	0	0.0	0	0.0	175	1.5	102	1.9	**277**	**0.7**
Enterovirus	4,362	8,933	17,649	32,237	5,283	**68,464**	36	0.8	164	1.8	279	1.6	307	1.0	4	0.1	**790**	**1.2**
Influenza virus	12,992	14,856	17,844	32,237	5,283	**83,212**	1,737	13.4	1,708	11.5	2,427	13.6	3,119	9.7	0	0.0	**8,991**	**10.8**
Influenza A virus	12,608	14,859	17,847	32,237	5,283	**82,834**	1,525	12.1	936	6.3	2,397	13.4	1,686	5.2	0	0.0	**6,544**	**7.9**
Influenza B virus	13,088	14,858	17,847	32,237	5,283	**83,313**	222	1.7	772	5.2	30	0.2	1,435	4.5	0	0.0	**2,459**	**3.0**
Metapneumovirus	7,654	14,75	17,622	32,237	5,283	**77,546**	230	3.0	325	2.2	445	2.5	462	1.4	20	0.4	**1,482**	**1.9**
Human parainfluenza virus	9,007	3,771	8,268	32,237	5,283	**58,566**	15	0.2	200	5.3	438	5.3	129	0.4	322	6.1	**1,104**	**1.9**
HPIV1	3,791	3,771	8,268	9,268	5,283	**30,381**	0	0.0	0	0.0	0	0.0	9	0.1	0	0.0	**9**	**0.0**
HPIV2	3,791	3,771	8,268	9,268	5,283	**30,381**	0	0.0	0	0.0	0	0.0	17	0.2	1	0.0	**18**	**0.1**
HPIV3	2,395	3,771	8,268	9,268	5,283	**28,985**	0	0.0	0	0.0	0	0.0	41	0.4	299	5.7	**340**	**1.2**
HPIV4	2,395	3,771	8,268	9,268	5,283	**28,985**	0	0.0	0	0.0	0	0.0	48	0.5	20	0.4	**68**	**0.2**
Rhinovirus	4,305	14,057	17,637	32,237	5,283	**73,519**	401	9.3	1,771	12.6	2,264	12.8	2,494	7.7	449	8.5	**7,379**	**10.0**
Respiratory syncytial virus	12,756	14,849	17,851	32,237	5,283	**82,976**	923	7.2	1,024	6.9	1,347	7.5	498	1.5	54	1.0	**3,846**	**4.6**
SARS-CoV-2	0	0	0	420,120	90,551	**510,671**	0	0	0	0	0	0	26,723	6.4	8,236	9.1	**34,959**	**6.8**

HCoV, human coronavirus; HPIV, human parainfluenza virus; SARS-CoV-2, severe acute respiratory syndrome coronavirus 2

Since February 2020, 510,671 samples have been analysed for SARS-CoV-2 and 34,959 tested positive (6.8%). Of 420,120 samples tested for SARS-CoV-2 in 2020, 6.4% (N = 26,723) were positive while in 2021, out of 90,551 samples, 9.1% were positive. The government introduced several restrictive measures in an attempt to mitigate the spread of SARS-CoV-2 and to control the epidemic as effectively as possible. A first lockdown was imposed between 17 March 2020 and 11 May 2020, recommendations have been in place on wearing masks in enclosed spaces (particularly in the workplace) since 20 July 2020, a curfew was introduced between 8:00 pm and 6:00 am between 17 October 2020 and 28 October 2020, a second lockdown took place between 29 October 2020 and 15 December 2020, and a new curfew was introduced on 16 January 2021 from 6:00 pm to 6:00 am. In addition to these actions, individual preventive measures have also been recommended, including hand washing with soap or alcohol-based hand gel, a distance of 1.5 metres between individuals and the promotion of remote working. In 2020, the proportion of positive tests dramatically decreased to 7.7% for rhinoviruses, 5.2% for influenza A virus, 4.5% for influenza B virus and 1.5% for RSV. This was also the case for the first two months of 2021, where no cases of influenza A or B were observed. In the first two months of 2021, the most frequently diagnosed virus was SARS-CoV-2 (9.1%), followed by rhinoviruses (8.5%), parainfluenza viruses (6.1%, mainly HPIV3: 5.7%) and endemic coronaviruses (5.2%, mainly HCoV NL63: 3.2%). The same results regarding influenza A and B viruses were observed from a private clinical microbiology and virology laboratory through the PACASurvE network (Table 2).

3.2. Comparison of winter seasons for respiratory viral infections

In order to avoid the Simpson effect, which is the presence of second order interactions between all factors that inverse statistical relations when data are pooled [24], we compared results during cold seasons (from October to mid-February). Over the last four such seasons, the most significant variations were observed for influenza A virus, with a positivity rate of 11.3% of the 9,819 tested samples (N= 1,106 cases) during the 2017–2018 winter season, which increased to 18.6% of the 10,973 tested samples (N= 2,042 cases) during the 2018–2019 season, dropped to 9.6% of the 11,711 tested samples (N= 1,125 cases) in 2019–2020 and accounted for 0% of the 8,786 tested samples in 2020–2021 (Fig. 1, Table 3). As of 24 February 2021, no cases of influenza A virus had been diagnosed during the 2020–2021 winter season. Influenza B virus was also absent for the 2020–2021 winter season, although this had already been observed in 2018–2019. RSV also showed a considerable decrease in the proportion of positive cases, reaching 0.6% (56 cases in 2020–2021 compared to between 9.4-10.9% (N= 988-1,196 cases) in the other three cold periods (p-value <0.001). Metapneumovirus and enterovirus had a less marked decrease (N= 21 and 9 in 2020–2021 vs N= 339 and 375 in 2019–2020, respectively; p-value < 0.001). The adenovirus positivity rate has remained relatively constant over time, at about 3% (p-value > 0.05), as was the case for endemic coronaviruses in 2017–2018 and 2018–2019. A significant decrease was nevertheless observed in 2020–2021 (p-value < 0.001). Rhinovirus exhibited a significantly higher positivity rate in 2020–2021 (12.9%) compared to 2017–2019 and 2018–2019 (9.9% and 10.9% respectively, p-value < 0.001). The positivity rate of the HPIV3 parainfluenza virus increased from 0.1% (N = 4) in 2019–2020 to 3.7% (N = 324) in 2020–2021 (p-value < 0.001).

3.3. Total gastrointestinal viral infections at IHU-MI in 2017–2021

Between 2017 and 2021, 27,719 tests were performed resulting in approximately 1,098 diagnoses of gastrointestinal infections (Table 4). Rotavirus (5.6% for 6,612 samples analysed) was the most frequently diagnosed gastrointestinal virus over the study period, followed by adenovirus (5.2% for 6,227 samples analysed) and norovirus (4.2% for 7,791 samples analysed). As was previously observed for respiratory

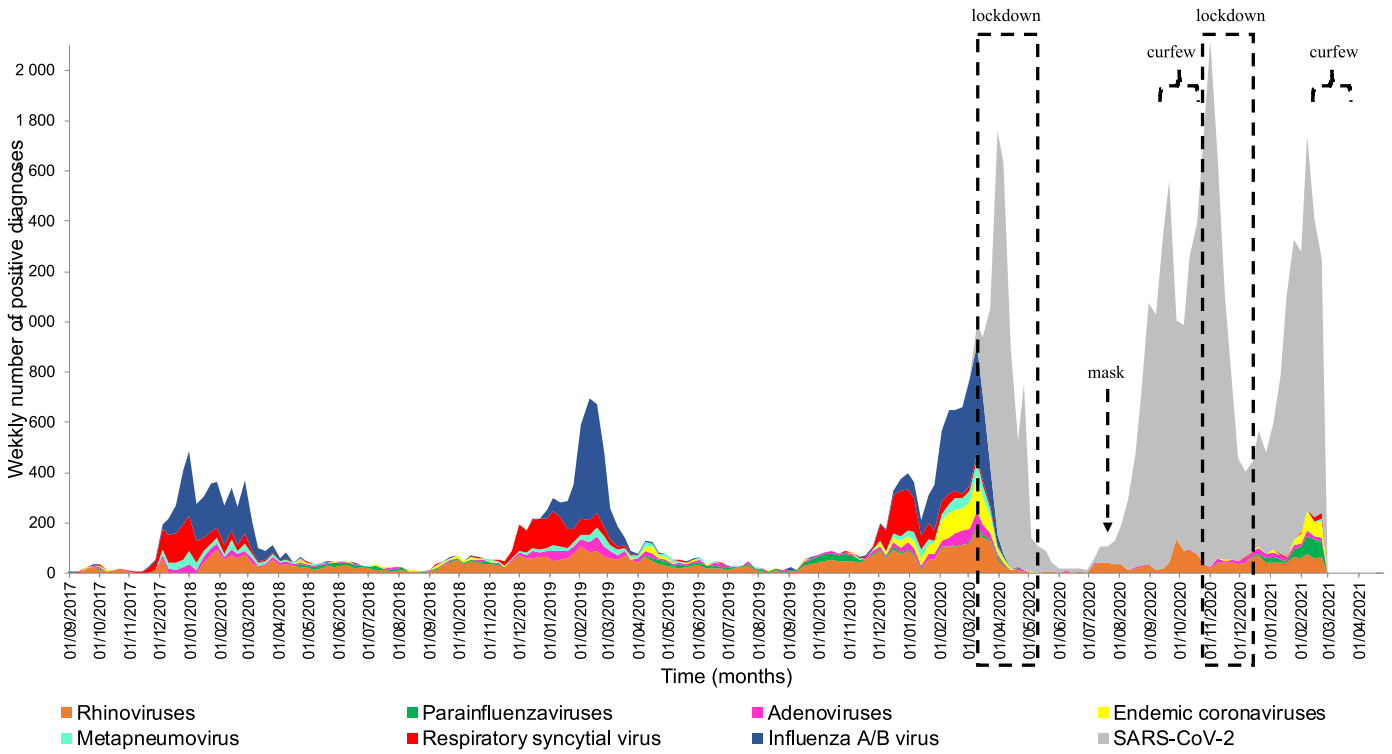


Fig. 1. – Respiratory virus infections diagnosed at Hospital University Institute Méditerranée Infection in 2017-2021. Actions taken by the government are indicated by a dotted square for lockdowns, an arrow for the obligation to wear a mask in enclosed spaces, and a brace symbol for curfews.

Table 2
– Results for diagnoses of influenza A virus and influenza B virus by year for a private clinical microbiology and virology laboratory through the PACASurvE (for Provence Alpes-Côte d'Azur Surveillance Epidemiological System) network.

Year	Number of samples tested	Influenza A virus diagnoses		Influenza B virus diagnoses	
		N	%	N	%
2017	547	67	12.2	2	0.4
2018	1,111	63	5.7	29	2.6
2019	2,410	681	28.3	32	1.3
2020	2,625	500	19.0	357	13.6
2021	106	0	0.0	1	0.9

viruses, the intensity of the epidemic peak as well as the date of its onset varied over the years (Fig. 2). In 2017, 2018 and 2020, adenovirus was the most frequently identified virus (5.0%, 7.6% and 4.1% respectively) while in 2019, rotavirus (8.6%) was the virus most commonly identified.

In the first two months of 2021, of the 255 samples analysed, rotavirus was again the most frequently identified virus (N = 12, 4.7%) followed by norovirus (N = 10, 3.9%) and adenovirus (N = 10, 3.9%).

The overall positivity rate of gastrointestinal infections decreased significantly over time during the winter seasons (Table 5). Notably, the norovirus winter season positivity rates decreased continuously and significantly over time (2017–2018: 9.3%; 2018–2019: 8.4%; 2019–2020: 5.5%; 2020–2021: 2.0%). In contrast, adenovirus and rotavirus showed stable positivity rates between 2018–2019 (3.8% and 2.4% respectively) and 2020–2021 winter seasons (3.4% and 3.6% respectively) (Fig. 2).

4. DISCUSSION

In this study, the systematic monitoring of our microbiology and virology laboratory activity has enabled us to identify changes in the epidemiology of respiratory and gastro-intestinal viral communicable diseases during the spread of a new emerging virus, SARS-CoV-2.

These data show that the epidemiology of infection with SARS-CoV-2 is not at all similar to that of other respiratory infections. As observed in other countries and in France, flu viruses have decreased dramatically [25–27]. It should be noted that the number of infections by endemic coronaviruses and rhinoviruses does not seem to be particularly affected by the preventive measures taken and may have, in common with COVID-19, modes of transmission that are different from those of influenza viruses, RSV and the other respiratory viruses studied. Curiously, in our region, a higher number of parainfluenza virus 3 (HPIV3) were observed. One of the explanations for these epidemiological figures could be that the viruses experiencing a decrease in their incidence are most often involved in pneumonia, while, conversely, the agents responsible for nasal infections and for causing colds, such as endemic coronaviruses or HPIV3, remain constant. Measures to control COVID-19 would then prevent pneumonia, and gastroenteritis. From this hypothesis, it would be interesting to study the nasal and pneumonic forms in COVID-19 patients and assess their evolution in time.

The impact of measures to control COVID-19 probably played a major role in these epidemiological changes [28]. These measures included both repeated recommendations on risk prevention measures such as hand washing with soap or alcohol-based hand gel, disinfecting surfaces, and social distancing, but also actions which were legally enforced, including wearing masks and the implementation of lockdown or curfews [9,29]. Hand washing and disinfection was probably the main factor having an impact upon the usual respiratory and gastrointestinal viral infections [30], and have been key elements of influenza prevention campaigns for several years [11]. It is not clear from the literature that lockdown measures and other social control measures have really had an impact on the spread of SARS-CoV-2 or on other respiratory infections [31]. For example, Sweden has issued very few social control measures while other countries such as France have implemented relatively strong measures without significant differences in the number of cases or mortality [32].

The lack of effectiveness of these measures on the COVID-19 epidemic raises several questions. The first is the existence of infection

Table 3
- Tests performed and positive for PCR detection of respiratory viruses, during the same cold months in 2017–2018, 2018–2019, 2019–2020 and 2020–2021

Viruses	Tests				Positive										
	2017-2018	2018-2019	2019-2020	2020-2021	2017-2018			2018-2019			2019-2020			2020-2021	
	N	N	N	N	N	%	p value	N	%	p value	N	%	p value	N	%
Adenovirus	8,876	10,831	11,687	8,786	262	3.0	0.51	416	3.8	0.006	355	3.0	0.74	274	3.1
Common coronaviruses	387	1,123	11,556	8,786	13	3.4	0.97	37	3.3	0.96	617	5.3	<0.001	292	3.3
HCoV 229E	387	1,123	6,357	8,786	0	0.0	1	0	0.0	1	31	0.5	<0.001	5	0.1
HCoV HKU1	387	1,123	6,357	8,786	0	0.0	1	0	0.0	1	236	3.7	<0.001	1	0.0
HCoV NL63	387	1,123	6,357	8,786	0	0.0	0.006	0	0.0	<0.001	114	1.8	0.49	171	1.9
HCoV OC43	657	1,123	6,357	8,786	0	0.0	0.003	0	0.0	<0.001	56	0.9	0.02	114	1.3
Enterovirus	900	10,83	11,688	8,786	13	1.4	<0.001	153	1.4	<0.001	375	3.2	<0.001	9	0.1
Influenza virus	9,819	10,973	11,711	8,786	1,839	18.7	<0.001	2,044	18.6	<0.001	1,85	15.8	<0.001	1	0.0
Influenza A virus	9,819	10,973	11,711	8,786	1,106	11.3	<0.001	2,042	18.6	<0.001	1,125	9.6	<0.001	0	0.0
Influenza B virus	9,819	10,973	11,711	8,786	743	7.6	<0.001	2	0.0	1	727	6.2	<0.001	1	0.0
Metapneumovirus	8,873	10,83	11,687	8,786	315	3.6	<0.001	258	2.4	<0.001	339	2.9	<0.001	21	0.2
Human parainfluenza virus	657	1,127	11,605	8,786	4	0.6	<0.001	42	3.7	0.68	226	1.9	<0.001	350	4.0
HPIV1	387	1,127	6,001	8,786	0	0.0	1	0	0.0	1	2	0.0	0.16	0	0.0
HPIV2	387	1,127	6,001	8,786	0	0.0	1	0	0.0	1	4	0.1	0.17	1	0.0
HPIV3	387	1,127	6,001	8,786	0	0.0	<0.001	0	0.0	<0.001	4	0.1	<0.001	324	3.7
HPIV4	387	1,127	6,001	8,786	0	0.0	0.62	0	0.0	0.1	14	0.2	0.73	23	0.3
Rhinovirus	5,15	10,833	11,683	8,786	511	9.9	<0.001	1,194	11.0	<0.001	1,42	12.2	0.11	1,134	12.9
Respiratory syncytial virus	9,912	10,973	11,707	8,786	988	10.0	<0.001	1,196	10.9	<0.001	1,104	9.4	<0.001	56	0.6
SARS-CoV-2	0	**0**	**5,628**	**244,310**	**0**	**0.0**	-	**0**	**0.0**	-	**0**	**0.0**	-	**20748**	**8.5**

HCoV, human coronavirus; HPIV, human parainfluenza virus; SARS-CoV-2, severe acute respiratory syndrome coronavirus 2

Table 4
- Tests performed and positive for PCR detection of gastrointestinal viruses in 2017, 2018, 2019, 2020 and 2021 at Hospital University Institute Méditerranée Infection.

Viruses	Tests						Positive											
	2017	2018	2019	2020	2021	Total 2017-2021	2017		2018		2019		2020		2021		Total 2017-2021	
	N	N	N	N	N	N	N	%	N	%	N	%	N	%	N	%	N	%
Adenovirus	1,674	1,662	1,47	1,166	255	**6,227**	83	5.0	127	7.6	56	3.8	48	4.1	10	3.9	**324**	**5.2**
Astrovirus	0	886	1,458	1,146	255	**3,745**	0	0.0	18	2.0	16	1.1	6	0.5	0	0.0	**40**	**1.1**
Norovirus	1,386	1,666	2,368	2,116	255	**7,791**	60	4.3	91	5.5	101	4.3	63	3.0	10	3.9	**325**	**4.2**
Rotavirus	1,64	1,662	1,471	1,184	255	**6,212**	78	4.8	88	5.3	127	8.6	45	3.8	12	4.7	**350**	**5.6**
Sapovirus	0	886	1,46	1,143	255	**3,744**	0	0.0	12	1.4	38	2.6	8	0.7	1	0.4	**59**	**1.6**

outbreaks in animals which are distinct from outbreaks in humans. It has been demonstrated that the emergence of new variants could be promoted by the intensive captive breeding of certain animals such as mink, which are likely to contaminate humans by being potentially more contagious or more pathogenic for humans [34–36]. Furthermore, it seems likely that a certain number of treatments, including serotherapy with hyper-human sera and antivirals such as remdesivir, can promote the appearance of mutations [37].

Strong reductions in the incidence of some but not all respiratory viruses, and of viral agents of gastrointestinal infections have been also reported in several countries worldwide. This has been particularly noticed for influenza virus infections [28]. Tanislav and Kostev reported fewer non-SARS-CoV-2 respiratory tract infections and gastrointestinal infections during the SARS-CoV-2 pandemic [38]. They collected data from 994 general practitioners and 192 pediatricians in Germany and compared the prevalence of these infections between April 2020-March 2021 and April 2019-March 2020. Substantial falls (-71% for general practices and -90% for paediatrician practices) were observed for influenza virus infections, which was accompanied by a 40% fall of intestinal infections for general practices. Agca et al. reported, in a study on 319 nasopharyngeal samples in Turkey, a 7.5-fold reduction of the proportion of positive testing for influenza virus during March 2020-February 2021 compared to the previous year [2.3% (n= 9 cases) versus 17.3% (133), respectively] [39]. A significant reduction was also observed for other respiratory viruses including RSV but not for rhinoviruses/enteroviruses and metapneumovirus. Ippolito et al. also reported in Italy a strong decrease during the SARS-CoV-2 pandemic of the diagnoses of several seasonal respiratory viruses among hospitalized children younger than two years [40]. Indeed, the number of positive tests was 80% lower during the September 2020-February 2021 period compared to between the same periods of years 2019-2020 and 2018-2019, with a disappearance of influenza viruses and RSV as well as disappearance or strong decreases of other respiratory viruses except rhinoviruses and endemic coronaviruses. In the Southern hemisphere, Yeoh et al., reported decreases by 99% and 98% of diagnoses of influenza viruses and RSV, respectively, in children in Western Australian through winter 2020 [41]. In Singapore, influenza virus positivity rate decreased by 64% during weeks 5–9 of 2020 compared with the preceding years [42].

The present study has several limitations. We focused on the epidemiology of respiratory and gastrointestinal viruses only over the last four cold seasons in our institute. We analyzed here a limited number of seasons but notwithstanding these data further support the unpredictability of the epidemiology of these viruses. We observed considerable variations from season to season throughout the respiratory virus epidemic period regarding the predominant viruses, the time of emergence and duration of winter epidemics, the level of incidence reached at the epidemic peaks, and the time at which this peak occurred [15]. Also, we acknowledge that estimating hospitalization pattern in the typical pre-SARS-CoV-2 season may be the subject of large random variation. Here, we have not analyzed the impact of respiratory viruses on hospital admissions, but their impact on hospital mortality has been

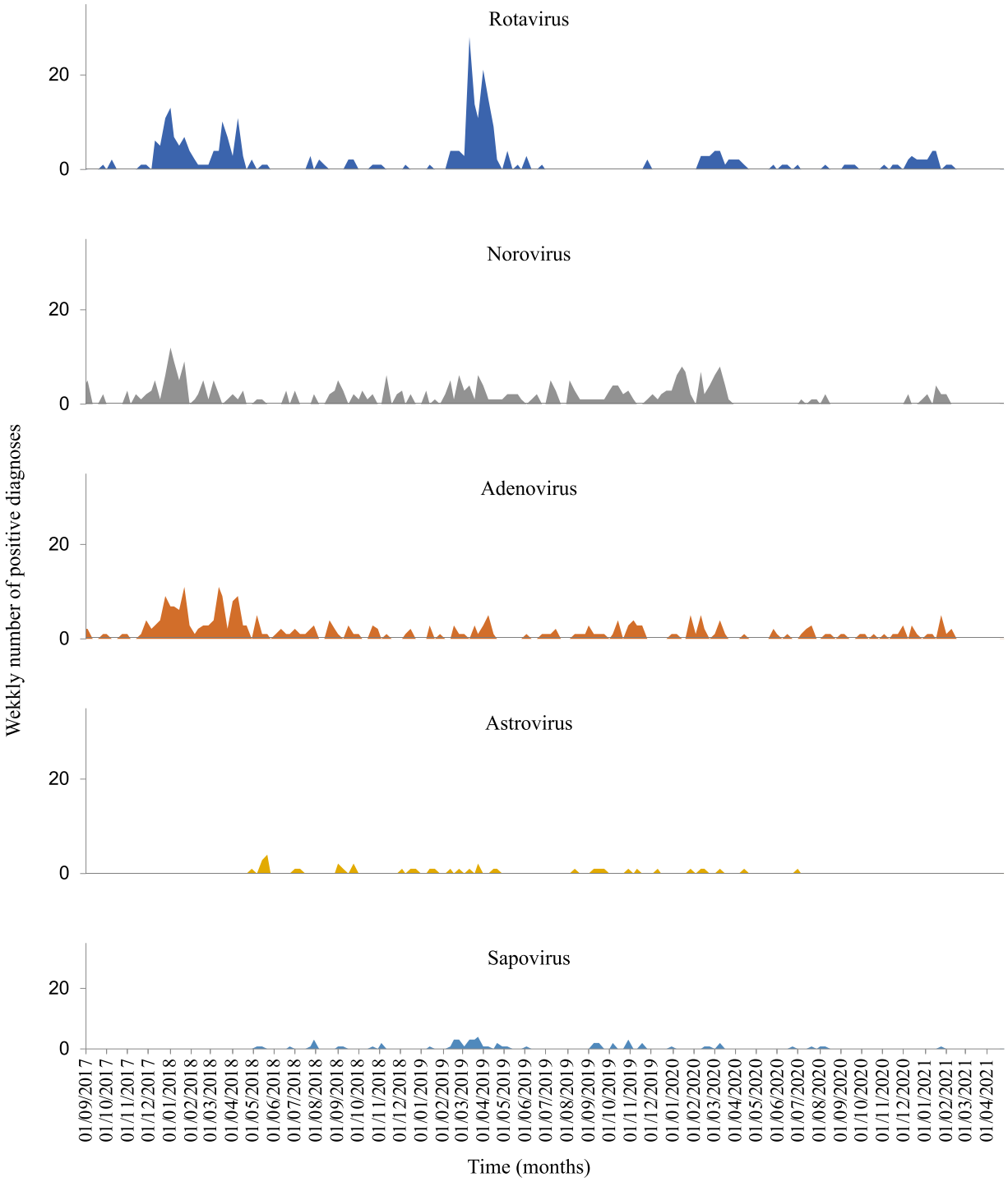


Fig. 2. – Gastrointestinal virus diagnosis between October 2017 and February 2021at Hospital University Institute Méditerranée Infection.*Comparison of winter seasons for gastrointestinal viral infections*

Table 5
- Tests performed and positive for PCR detection of gastrointestinal viruses, during the same cold months in 2017–2018, 2018–2019, 2019–2020 and 2020–2021.

Viruses	Tests				Positive										
	2017-2018	2018-2019	2019-2020	2020-2021	2017-2018			2018-2019			2019-2020			2020-2021	
	N	N	N	N	N	%	p value	N	%	p value	N	%	p value	N	%
Adenovirus	661	369	380	642	61	9.2	<0.001	14	3.8	0.8	31	8.2	0.001	22	3.4
Astrovirus	0	369	368	619	0	0.0	-	6	1.6	0.003	5	1.4	0.01	0	0.0
Norovirus	636	369	1026	664	59	9.3	<0.001	31	8.4	<0.001	56	5.5	<0.001	13	2.0
Rotavirus	661	369	380	661	64	9.7	<0.001	9	2.4	0.3	5	1.3	0.03	24	3.6
Sapovirus	0	369	369	619	0	0.0	-	5	1.4	0.03	8	2.2	0.002	1	0.2

the subject of previous studies [15,33]. Finally, the present work has been conducted in a single institution, and the results could therefore display local specificities.

In conclusion, this study confirms that it is futile to try to make predictions about a disease for which the level of knowledge is limited [43]. The course of the epidemic over the past year was unpredictable and could not be integrated into any predictive models. Caution should be taken when using such models. Furthermore, this leads to the search for different modes of transmission of most respiratory diseases, as had already been mentioned in relation to SARS-CoV, where infections were retrospectively detected at a significant distance from the heart of the SARS-CoV outbreak, with no reasonable explanation [44]. Broad epidemiological surveillance of respiratory and gastrointestinal infections should be pursued in the future, as many changes occur during this pandemic among which public health policies and population behaviours including mask wearing or social distancing [31]. Also, in France, the issue of carriage and transmission by domestic pets has not been resolved and should be the subject of intense research to really understand the reservoirs, transmission and epidemiology of this very atypical virus. A new study on the 2021-2022 winter season should be carried out to better understand the epidemiology of these respiratory and gastrointestinal viruses and the impact of barrier measures on the spread of new SARS-CoV-2 variants with different transmissibiliy [45].

Author contributions

Conceived and designed the study: DR.; Collected data or/and performed experiments: AGG, LK, CB, JPC.; Analysed and interpreted data: AGG, CD, PC, PG, HC and DR.; Wrote the manuscript: AGG, PC, PG and DR.; All authors read and approved the final manuscript.

Ethics

All data have been generated as part of the routine work at Assistance Publique-Hôpitaux de Marseille (Marseille university hospitals), and this study results from routine standard clinical management. The study was approved by the ethical committee of the University Hospital Institute Méditerranée Infection (N°: 2020-029 and 2022-015). Access to the patients' biological and registry data issued from the hospital information system was approved by the data protection committee of Assistance Publique-Hôpitaux de Marseille (APHM) and was recorded in the European General Data Protection Regulation registry under number RGPD/APHM 2019-73.

Funding

This work was supported by the French Government under the "Investments for the Future" programme managed by the National Agency for Research (ANR), Méditerranée-Infection 10-IAHU-03, and was also supported by Région Provence Alpes Côte d'Azur and European funding FEDER PRIMMI (Fonds Européen de Développement Régional - Plateformes de Recherche et d'Innovation Mutualisées Méditerranée Infection), FEDER PA 0000320 PRIMMI.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this article.

References

[1] B. Hu, H. Guo, P. Zhou, Z.L. Shi, Characteristics of SARS-CoV-2 and COVID-19, Nat. Rev. Microbiol. 19 (2021) 141–154, https://doi.org/10.1038/s41579-020-00459-7.

[2] E. Dong, H. Du, L. Gardner, An interactive web-based dashboard to track COVID-19 in real time, Lancet Infect. Dis. 20 (2020) 533–534, https://doi.org/10.1016/S1473-3099(20)30120-1.

[3] P. Lemey, N. Ruktanonchai, S.L. Hong, et al., Untangling introductions and persistence in COVID-19 resurgence in Europe, Nature 595 (2021) 713–717, https://doi.org/10.1038/s41586-021-03754-2.

[4] P Colson, PE Fournier, H Chaudet, et al., Analysis of SARS-CoV-2 variants from 24,181 patients exemplifies the role of globalization and zoonosis in pandemics, Front. Microbiol. 7 (2022), 786233, https://doi.org/10.3389/fmicb.2021.786233.

[5] WT Harvey, AM Carabelli, B Jackson, et al., SARS-CoV-2 variants, spike mutations and immune escape, Nat. Rev. Microbiol. 19 (7) (2021 Jul) 409–424, https://doi.org/10.1038/s41579-021-00573-0.

[6] J Li, S Lai, GF Gao, W Shi, The emergence, genomic diversity and global spread of SARS-CoV-2, Nature 600 (2021) 408–418, https://doi.org/10.1038/s41586-021-04188-6.

[7] C Davis, N Logan, G Tyson, et al., Reduced neutralisation of the Delta (B.1.617.2) SARS-CoV-2 variant of concern following vaccination, PLoS Pathog. 2 (2021), e1010022, https://doi.org/10.1371/journal.ppat.1010022.

[8] TL Dao, VT Hoang, P Colson, et al., SARS-CoV-2 infectivity and severity of COVID-19 according to SARS-CoV-2 variants: Current evidence, J. Clin. Med. 10 (2021) 2635, https://doi.org/10.3390/jcm10122635.

[9] Information coronavirus, Les bons gestes à adopter, affiche – mesures barrières, February 25, 2021. Retrievedfrom, https://www.gouvernement.fr/sites/default/files/coronavirus-mesures_barrieres.pdf.

[10] L. Canini, L. Andréoletti, P Ferrari, et al., Surgical mask to prevent influenza transmission in households: a cluster randomized trial, PloS ONE 5 (2010) e13998, https://doi.org/10.1371/journal.pone.0013998.

[11] Santé Publique France, Prévenir la grippe saisonnière, Sept 2017. Retrieved February 25, 2021, from, https://www.santepubliquefrance.fr/determinants-de-sante/vaccination/documents/depliant-flyer/prevenir-la-grippe-saisonniere-septembre-2017.

[12] Decree No. 2020-260 of 16 March 2020 introducing regulations on movement in the context of the fight against the spread of COVID-19, https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000041728476.

[13] Decree No. 2020-1310 of 29 October 2020 prescribing general measures required to combat the COVID-19 epidemic in the context of the health emergency situation. https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000042475143?r=9BtQcTAF3G.

[14] L. Kaba, A. Giraud-Gatineau, MT. Jimeno, et al., Consequences of the COVID-19 Outbreak Lockdown on Non-Viral Infectious Agents as Reported by a Laboratory-Based Surveillance System at the IHU Méditerranée Infection, Marseille, France, J. Clin. Med. 10 (2021) 3210, https://doi.org/10.3390/jcm10153210.

[15] P. Colson, A. Giraud-Gatineau, P.E. Fournier, et al., Epidemiological surveillance of respiratory viral infections at IHU Méditerranée Infection and its application to SARS-CoV-2, IHU Preprint, 2021, https://doi.org/10.35088/gpgh-wq98. Preprint.

[16] C. Abat, H. Chaudet, P. Colson, et al., Real-time microbiology laboratory surveillance system to detect abnormal events and emerging infections, Marseille, France, Emerg. Infect. Dis. 21 (2015) 1302–1310, https://doi.org/10.3201/eid2108.141419.

[17] P. Colson, J.M. Rolain, C. Abat, et al., EPIMIC: A Simple homemade computer program for real-time EPIdemiological surveillance and alert based on MICrobiological Data, PloS ONE 10 (2015), e0144178, https://doi.org/10.1371/journal.pone.0144178.

[18] D. Raoult. Rapport de mission. http://ifr48 timone univ-mrs fr/files/Documents-Raoult/bioterrorisme2003 pdf 2003Available from: URL: http://ifr48.timone.univ-mrs.fr/files/Documents-Raoult/bioterrorisme2003.pdf.

[19] M. Salmon, D. Schumacher, M. Höhle, Monitoring Count Time Series in R: Aberration Detection in Public Health Surveillance, Journal of Statistical Software 70 (2016) 1–3, https://doi.org/10.18637/jss.v070.i10.

[20] S. Amrane, H. Tissot-Dupont, B. Doudier, et al., Rapid viral diagnosis and ambulatory management of suspected COVID-19 cases presenting at the infectious diseases referral hospital in Marseille, France, January 31st to March 1st, 2020: A respiratory virus snapshot, Travel Med. Infect. Dis 36 (2020) 10163, https://doi.org/10.1016/j.tmaid.2020.101632.

[21] C. Boschi, V.T. Hoang, A. Giraud-Gatineau, et al., Coinfections with SARS-CoV-2 and other respiratory viruses in Southeastern France: A matter of sampling time, J. Med. Virol. 93 (2021) 1878–1881, https://doi.org/10.1002/jmv.26692.

[22] M. Huart, G. Bedubourg, C. Abat, P. Colson, et al., Implementation and initial analysis of a laboratory-based weekly biosurveillance system, Provence-Alpes-Côte d'Azur, France, Emerg. Infect. Dis. 23 (2017) 582–589, https://doi.org/10.3201/eid2304.161399.

[23] R Core Team. A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/2019.

[24] K.H. Chu, N.J. Brown, A. Pelecanos, A.F. Brown, Simpson's paradox: A statistician's case study, Emerg. Med. Australas. 30 (2018) 431–433, https://doi.org/10.1111/1742-6723.12943.

[25] Bulletin épidémiologique grippe, semaine 7. Saison 2020-2021, February 25, 2021. Retrievedfrom, https://www.santepubliquefrance.fr/maladies-et-traumatismes/maladies-et-infections-respiratoires/grippe/documents/bulletin-national/bulletin-epidemiologique-grippe-semaine-7.-saison-2020-2021.

[26] N. Jones, How COVID-19 is changing the cold and flu season, Nature 588 (2020) 388–390, https://doi.org/10.1038/d41586-020-03519-3.

[27] S.J. Olsen, E. Azziz-Baumgartner, A.P. Budd, et al., Decreased influenza activity during the COVID-19 pandemic - United States, Australia, Chile, and South Africa, 2020, Morb. Mortal. Wkly Rep. 69 (2020) 1305–1309, https://doi.org/10.15585/mmwr.mm6937a6.

[28] K. Servick, COVID-19 measures also suppress flu-for now, Science 371 (2020) 224, https://doi.org/10.1126/science.371.6526.224.
[29] Decree No. 2020-860 of 10 July 2020 prescribing general measures required to fight against the COVID-19 epidemic in regions where the health emergency has ended and those in which it has been prolonged. https://www.legifrance.gouv.fr/loda/id/JORFTEXT000042105897/2020-07-11#JORFTEXT000042105897.
[30] M. Zuin, G. Rigatelli, G. Zuliani, L. Roncon, COVID-19 restrictive measures are changing the flu season in Italy, Minerva Med. 10 (2021), https://doi.org/10.23736/S0026-4806.21.07359-6, 2021 Feb 8 Online ahead of print.
[31] E. Bendavid, C. Oh, J. Bhattacharya, J. Ioannidis, Assessing mandatory stay-at-home and business closure effects on the spread of COVID-19, Eur. J. Clin. Invest. 51 (4) (2021) e13484, https://doi.org/10.1111/eci.13484.
[32] Q. De Larochelambert, A. Marc, J. Antero, et al., Covid-19 mortality: A matter of vulnerability among nations facing lmited margins of adaptation, Front. Public Health 8 (2020), 604339, https://doi.org/10.3389/fpubh.2020.604339.
[33] A. Giraud-Gatineau, P. Colson, M.T. Jimeno, et al., Comparison of mortality associated with respiratory viral infections between December 2019 and March 2020 with that of the previous year in Southeastern France, Int. J. Infect. Dis. 96 (2020) 154–156, https://doi.org/10.1016/j.ijid.2020.05.001.
[34] A. Hassanin, P. Grandcolas, G. Veron, Covid-19: natural or anthropic origin ? Mammalia 85 (2021) 1–7, https://doi.org/10.1515/mammalia-2020-0044.
[35] E.I. Patterson, G. Elia, A. Grassi, et al., Evidence of exposure to SARS-CoV-2 in cats and dogs from households in Italy, Nat. Comm. 11 (2020) 6231, https://doi.org/10.1038/s41467-020-20097-0.
[36] F. Fenollar, O. Mediannikov, M. Maurin, et al., Mink, SARS-CoV-2, and the human-animal interface, Front. Microbiol. 12 (2021), 663815, https://doi.org/10.3389/fmicb.2021.663815.
[37] P. Colson, C.A. Devaux, J.C. Lagier, P. Gautret, D. Raoult, A possible role of remdesivir and plasma therapy in the selective sweep and emergence of new SARS-CoV-2 variants, J. Clin. Med. 10 (2021) 3276, https://doi.org/10.3390/jcm10153276.
[38] C. Tanislav, K. Kostev, Fewer non-COVID-19 respiratory tract infections and gastrointestinal infections during the COVID-19 pandemic, J. Med.Virol. 94 (2022) 298–302, https://doi.org/10.1002/jmv.27321.
[39] H. Agca, H. Akalin, I. Saglik, M. Hacimustafaoglu, S. Celebi, B. Ener, Changing epidemiology of influenza and other respiratory viruses in the first year of COVID-19 pandemic, J. Infect. Public Health. 14 (2021) 1186–1190, https://doi.org/10.1016/j.jiph.2021.08.004.
[40] G. Ippolito, A. La Vecchia, G. Umbrello, et al., Disappearance of seasonal respiratory viruses in children under two years old during COVID-19 pandemic: A monocentric retrospective study in Milan, Italy, Front. Pediatr. 9 (2021), 721005, https://doi.org/10.3389/fped.2021.721005.
[41] D.K. Yeoh, D.A. Foley, C.A. Minney-Smith, et al., Impact of coronavirus disease 2019 public health measures on detections of influenza and respiratory syncytial virus in children during the 2020 Australian winter, Clin. Infect. Dis. 72 (2021) 2199–2202, https://doi.org/10.1093/cid/ciaa1475.
[42] R.J.J. Soo, C.J. Chiew, S. Ma, R. Pung, V. Lee, Decreased influenza incidence under COVID-19 control measures, Singapore, Emerg. Infect. Dis. 26 (2020) 1933–1935, https://doi.org/10.1016/j.tmaid.2021.102057.
[43] N.P. Jewell, J.A. Lewnard, B.L. Jewell, Predictive mathematical models of the COVID-19 pandemic: Underlying principles and value of projections, JAMA 323 (2020) 1893–1894, https://doi.org/10.1001/jama.2020.6585.
[44] M.D. Christian, S.M. Poutanen, M.R. Loutfy, et al., Severe acute respiratory syndrome, Clin. Infect. Dis. 38 (2004) 1420–1427, https://doi.org/10.1086/420743.
[45] C. Del Rio, S.B. Omer, P.N. Malani, Winter of Omicron - The evolving COVID-19 pandemic, JAMA 327 (2021) 319–320, https://doi.org/10.1001/jama.2021.24315.

CHAPITRE IV

Etude de périodicité de quelques agents infectieux d'origine bactérienne dans la région PACA

Préambule

La transmission de plusieurs maladies infectieuses s'effectue de façon périodique ou saisonnière ou cyclique. Ces infections présentent une dynamique saisonnière entretenue par des fluctuations périodiques de l'environnement. Certaines études antérieures ont démontré la saisonnalité de plusieurs maladies infectieux mais surtout celles d'origine virale (39). Il a été montré que dans les régions tempérée nord, la recrudescences d'incidence sont fréquemment observées durant l'automne et l'hiver (40,41). Au même moment l'évolution des conditions météorologiques peut être associée à la saisonnalité et contribue à favoriser la survie des agents dans l'environnement et leur transmission, à diminuer la résistance des hôtes susceptibles et à favoriser le changement de comportement des individus par des interactions sociales facilitant la contamination ou la transmission (42–44). Dans la surveillance épidémiologique, la détection précoce de l'émergence des maladies infectieuses constitue une des priorités fondamentales afin de développer des stratégies visant à contrôler la transmission et la propagation de l'agent infectieux. Pour cela, la prédiction du risque d'émergence ou de réémergence d'agents pathogènes à différents moments est essentielle pour le développement des stratégies de santé publique efficaces.

Les saisonnalités étudiées concernent les maladies infectieuses avec des tableaux cliniques évidents et non les agents infectieux. Cependant, la prolifération des agents infectieux et leur transmission peuvent être associée à des évènements périodiques social, religieux ou culturel, mais aussi à des facteurs météorologiques sans pour autant entraîner la manifestation de la forme clinique des maladies (45). Dans ce chapitre, nous visons à mettre en évidence la corrélation de quelques agents bactériens en lien avec la nature et l'origine des prélèvements d'une part et d'autre part avec des facteurs météorologiques.

Article 5

Article 5: Influence of infection origin, type of sampling and weather factors on the periodicity of some infectious pathogens in Marseille university hospitals, France

Lanceï Kaba, Audrey Giraud-Gatineau, Philippe Colson, Pierre-Edouard Fournier, Didier Raoult and Hervé Chaudet

En cours de préparation sera soumis dans Frontriers

Influence or infection origin, type of sampling and weather drivers on the periodicity of some infectious pathogens in Marseille University Hospitals

Lancéï Kaba[1,2,3], Audrey Giraud-Gatineau[1,2], Philippe Colson[1,4,5], Pierre-Edouard Fournier[1,2,4], Didier Raoult[1,2,5], Hervé Chaudet[1,2,4*]

[1]IHU Méditerranée Infection, Marseille, France

[2]Aix Marseille Univ., IRD, AP-HM, SSA, VITROME, IHU Méditerranée Infection, Marseille, France

[3]Institut Supérieur des Sciences et de Médecine Vétérinaire (ISSMV) de Dalaba, BP 09, Guinée

[4]Assistance Publique Hôpitaux de Marseille, Marseille, France

[5]Aix Marseille Univ., IRD, AP-HM, MEPHI, Marseille, France

*** Correspondence:**
Hervé Chaudet. Institut Hospitalo-Universitaire Méditerranée-Infection, 19-21 Boulevard Jean Moulin, 13005 Marseille, France (herve.chaudet@gmail.fr)

Keywords: Seasonality, Surveillance system, Meteorological drivers, Community-acquired infections, Hospital-associated infections

Abstract

This big data study aimed at systematically exploring seasonalities of a 16 years series of bacterial identifications in hospitalized patients, considering the infectious site and the community-acquired or hospital-associated origin. Possible meteorological drivers were explored for series with seasonal variations.

Deduplicated bacterial identifications from February 2004 to February 2020 were extracted from the data warehouse of the Institut Hospitalo-Universitaire Mediterranée Infection surveillance system, along with their epidemiological characteristics, and weekly aggregated. Each species' series was processed using a scientific workflow based on TBATS time series model and allowing the systematic analysis of the epidemiological characteristic combinations of the series. Possible co-seasonalities were researched using seasonal peak clustering and series cross-correlations. For all seasonal series, this study explored the possibility of meteorological drivers using MeteoFrance's SYNOP database.

From 575 bacterial species, only 86 had a sufficient number of cases for the study purpose and only the 15 most frequent species were described in detail. About 60% of species exhibited a seasonality for at least one series, and winter is the season with the most of seasonal peaks. The results showed that, as for S. aureus, an overall apparent seasonality or absence of seasonality may hide several different sample-dependant seasonalities. Co-occurrence studies confirms the presence of seasonal complexes associating several species in a same infection location. Possible weather drivers were found for about a third of seasonal series.

Seasonality is a frequent characteristic of bacterial infections. Our results showed significant associations of periodicity between pathogens, origin of infection and type of sampling, as well as significant associations of some species with one or more weather drivers.

1 Introduction

Seasonality is a constantly encountered notion in medicine, along history and space, stated in European medicine since the Hippocratic time or in traditional Taoist medicine since the *Huang ti nei ching* (*Yellow Lord's Inner Canon*), both more than 2,200 years ago. Since then, many infectious diseases are known to be periodical or seasonal, essentially viral infections but also common bacterial infections (Paul, 2012) (*e.g.*, 'influenza' disease takes its name origin from the fact that it is influenced by the cold). However, the introduction of statistical methods adapted to periodicity analyses is more recent, with the systematic use in 1918 of periodograms for investigating measles outbreaks reported in the Bills of Mortality (Brownlee & Fletcher, 1918). As Fisman, we consider seasonal a process having "an incidence associated with a particular calendar period, and which have periodicity, although this is not limited to annual periodicity" (Fisman, 2012). Regarding viral infections, seasonality may be driven by vector seasonality, climatic conditions, abiotic or biotic environment, co-infections, viral antigenic drifts, seasonal human immune variation, human behaviours and seasonality in domestic or wildlife animal hosts (Dowell, 2001; Martinez, 2018). Seasonality studies of infectious diseases have long been based on meteorological and climatic factors that were the most evident associated determinants. For example, Greer et al. reported that norovirus winter outbreaks in Toronto are associated with the seasonal fluctuation of the Lake Ontario temperature favouring the virus survival in winter (Greer *et al.*, 2009). Seasonal viral outbreaks due to seasonal climatic-dependent proliferation of vectors belong to the most well-known mechanisms, as for the Rift Valley fever (Linthicum *et al.*, 2016) or for arboviruses in general (Schuster *et al.*, 2011). Seasonality can be influenced by humans' mobility behaviour (Coletti *et al.*, 2018; Viboud *et al.*, 2006) or by periodical socio-cultural or religious gathering, as during the annual pilgrimage organized in Mecca for the Hajj (Salmon-Rousseau *et al.*, 2016). However, the possibility of confounding factors and then of fallacious associations cannot be eliminated as the subjacent mechanism is rarely understood (Fisman, 2012).

It is only recently that seasonal periodicities of bacterial infections have been systematically studied (Park *et al.*, 2018). Moreover, these systematic studies rarely take into consideration the sites of infection, differentiating, among others, bloodstream, urinary tract, skin and soft tissues, respiratory tract infections, and rarely differentiate infections from colonization. The reason is that large scale systematic studies are mainly done on surveillance data which do not integrate these kinds of information.

The association of various climatic factors such as temperature, precipitation, sunshine, atmospheric pressure, frost and snow with a number of bacterial infectious diseases, has been reported in numerous studies (Lal *et al.*, 2013). As examples, a raised incidence of salmonellosis has been associated with increased temperature and precipitation (Sari Kovats *et al.*, 2005; Stashevsky *et al.*, 2019). *Campylobacter spp.* infections show a peak in spring (Nichols *et al.*, 2012). Studies in the UK and Wales (Djennad *et al.*, 2018; Naumova *et al.*, 1999) reported that *Campylobacter spp.* and *Cryptosporidium spp.* cases were significantly associated with temperature and rainfall. Gram-negative bloodstream infections, including *Escherichia coli*, *Acinetobacter spp.*, and *Klebsiella spp.*, also exhibit seasonal variations, being frequently associated with temperature and rainfall (Al-Hasan *et al.*, 2009; Eber *et al.*, 2011; Freeman *et al.*, 2009). It has also been suggested that seasonal viral infections may drive seasonal bacterial infection variations, as observed for influenza and invasive pneumococcal (Talbot *et al.*, 2005) and meningococcal diseases (Harrison *et al.*, 1991). However, these interactions

are especially difficult to decipher as they may result from a spurious correlation or an unmeasured seasonal factor (Fisman, 2007).

To date, although systematic studies to the search of periodicities were performed for a broad number of pathogens (Cherrie *et al.*, 2018), none took into account both the community-acquired or hospital-associated origin and the type of sampling. In contrast, the present study aimed at exploring systematically the bacterial identifications in hospitalized patients in search of, or in confirmation of, seasonal variations of bacterial infections, taking into account the infectious site and community-acquired or hospital-associated origin, and trying to identify climatic drivers.

2 Material and methods

2.1 Material

2.1.1 Bacterial identifications and related data

The Institut Hospitalo-Universitaire Méditerranée Infection (IHUMI) is the infectious disease-dedicated hospital of Marseille Public hospitals (Assistance Publique – Hôpitaux de Marseille, APHM). It performs all microbiological analyses for all APHM hospitals, representing about 190,000 bacterial cultures per year. Since February 1st, 2014, our microbiological surveillance system (Méditerranée Infection Datawarehouse and Surveillance - MIDaS) allows the weekly monitoring of the routine clinical microbiology activity and further *in-silico* statistical investigations (Abat *et al.*, 2016). We included in this study all routine bacterial identifications at species level present in the MIDaS Datawarehouse, from February 1st, 2014 until January 31st, 2020 (6 years, ending just before the COVID-19 oubreak in France), deduplicated on the basis of the hospital stay and the sample type. Taking in account studies about the effect of data aggregation granularity on time series seasonal analysis (Alarcon Falconi *et al.*, 2020) and the influence of this granularity on case numbering, we aggregated the data to a weekly level.

All routine bacterial identifications were obtained using a Microflex MALDI-TOF mass spectrometer and the Biotyper software (Bruker Daltonics, Bremen, Germany). As recommended by the manufacturer, bacterial identifications at species level were validated when the Biotyper matching log score was ≥ 2. Fastidious bacteria were excluded from this study due to the choice of culture-based routine identifications.

We associated to each identification the epidemiological characteristics available in the data warehouse and required for our study, including the sampling date and origin, a community-acquired (comm) or hospital-associated (hosp) flag (community-acquired if the identification was done within the first 48 hours following the hospital admission), and the patient's age.

2.1.2 Weather data

Weather data for the study period were downloaded from the Météo France SYNOP (surface synoptic observations) opendata service (https://donneespubliques.meteofrance.fr/?fond=produit&id_produit=90&id_rubrique=32). These data gather 57 weather description variables (Supplementary Table 1) collected from 62 stations in France with a timestep of 3 hours. For this study, we selected the following variables: temperature, rain, humidity, wind, and pressure differentials.

In order to detect possible associations between seasonal variations and weather conditions, we selected from the database the dataset collected by the Marignane station (latitude: 43.44° North, longitude:

5.22° East), which is 20 km away from Marseille. We aggregated the data using the same weekly time step than the epidemiological time series, calculating for each week the minimal and maximal values, the mean and the sum.

2.1.3 Legal statement

The surveillance system is in accordance with the Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and the repealing Directive 95/46/EC (General Data Protection Regulation). It is registered by the Data Protection Officer under id 2019-73

2.2 Methods

For each bacterial species, we made the distinction between hospital-infected infections, with further analyses on urines and blood samples, and community-acquired infections, with further analyses on urines, blood, respiratory and skin samples. For both infections, specific analyses were made on ages (0-20, 21-40, 41-60, 61-80, over 80 years-old). A total of 19 times series were then selected for each species (Supplementary Figure 1). Only time series with at least 100 deduplicated observations were considered for statistical analysis, and only the 15 most frequent species will be described in this paper.

For a systematic big-data analysis of our database, we created an analysis workflow of each time series seasonality, as follows:

- A ‘Kwiatkowski–Phillips–Schmidt–Shin’ (KPSS) test for testing a trend-stationarity, with stationarity rejection if $p>0.05$ (Kwiatkowski *et al.*, 1992).
- A ‘Seasonal and Trend decomposition using Loess’ (STL), which is a versatile and robust method for decomposing time series. The loess method used for this decomposition is a method for estimating nonlinear relationships (Cleveland *et al.*, 1990). The advantage of this method is that the seasonal component is allowed to change over time, and it is robust to outliers. For the purpose of this study, we have forced the seasonal component to be identical across years.
- A ‘Trigonometric seasonality, Box-Cox transformation, ARMA errors, Trend and Seasonal components’ (TBATS) analysis for allowing a search for multiple seasonality (De Livera *et al.*, 2011).
- An extraction of the detrended time series using the STL results.

This workflow was applied to each of the 19 species time series with at least 100 observations.

A co-seasonality analysis was performed using an ascending hierarchical classification based on the Euclidean distance between seasonality peaks, with the UPGMA aggregation method. For further in-depth co-seasonality analyses of selected time series, we performed cross-correlation analyses in order to confirm the statistical significance and the time lag (Probst *et al.*, 2012).

Analyses of species periodicities in relation with possible meteorological drivers were done using Poisson mixed-effect time-series regressions on detrended series. Meteorological characteristics with a participation in the regression model significative with $p\leq0.05$ were considered as possible drivers.

The overall statistical analysis workflow is reported in Figure 1. All statistical analyses were done using R (R Core Team, 2021) version 4 with packages 'forecast', 'tseries', 'season', 'urca'. When it applies, we used 0.05 as threshold of statistical significance.

3 Results

The surveillance dataset used for this study included 314,884 bacterial identifications, deduplicated in 228,365 new sample related identifications, which were counted on a weekly basis.

A total of 575 different bacterial species were retrieved in this dataset. Of these, only 86 species (15%) met our inclusion criteria (at least 100 new cases over the study period), and only the first 15 most frequent species are described in this paper. The full results are reported in the supplementary materials.

3.1 Seasonality study

The three most frequent species were *Escherichia coli*, *Staphylococcus aureus*, and *Staphylococcus epidermidis* with median weekly incidences of 145, 74, and 39 cases respectively. The full set of statistics for the weekly incidence series is summarized in Table 1. The systematic periodicity analysis taking in account the sample origin showed that 60% of species (9/15) exhibited at least one seasonality. The full result set for the 15 most frequent species is summarized by the heat map of Figure 2, which reports in green the series with a statistically significant seasonality, and the full set for all 86 bacterial species is reported in the Supplementary Figure 2. Samples coming from hospital-associated infection and from community-acquired infections presented the sample frequency of seasonality (10/15, 66.7%). Seasonality of hospital-acquired infections were specific of 3 species: *K. oxytoca*, *S. haemolyticus* and *G. vaginalis*, while seasonality of community-acquired infections was specific of 4 species: *S. epidermidis*, *S. agalactiae*, *S. hominis* and *P. acnes*. *Pseudomonas aeruginosa* and *Staphylococcus aureus* exhibited a seasonality for most of the samples, with the exception of urine samples coming from hospital-associated infections of patients younger than 20-years old, and urine samples of the 61–80-year-old population with community-acquired infections.

Some bacterial species, apparently without seasonality at the whole series, the global hospital-associated or the global community-acquired levels, may exhibit a seasonality at a narrower level. It is the case for *K. oxytoca*, which presents a seasonality for its blood and respiratory tract levels for its community-acquired infections, or for *P. mirabilis*, which presents a seasonality only at the respiratory tract level for its community-acquired infections.

A further analysis of seasonal peaks was specifically performed for samples from community-acquired infections (Figure 3 and Supplementary Figure 3). More seasonal peaks were observed during the winter season, particularly during week 2, with 15 peaks (Figure 4) for 86 species, and during weeks 4 and 6, both with 9 peaks. Among the 15 peaks of this seasonal maximum, (1/15, 6.7%) were coming from blood (*K. oxytoca*), and (2/15, 13.3%) from respiratory specimens (*E. coli*, *S. aureus*). Most of seasonal peaks belonging to respiratory samples are observed during Winter (10/20, 50%) and Spring (7/20, 35%). Blood samples exhibit a bi-modal seasonality with maxima during Winter (6/15, 40%) and Summer (5/15, 33%). Skin samples show only a seasonal minimum during Summer (2/16, 13%), with *S. aureus* belonging of this group. The apparent seasonality of urinary tract infections is Spring, with peaks (8/17, 47%).

Co-seasonality and cross-correlation analysis

The result of the hierarchical clustering of the seasonal peak distances for community-acquired infections is presented in Figure 5. Kind of sampling with seasonality are grouped according to the proximity of their seasonal peak. The dendrogram shows that samplings of *S. aureus* and *E. coli* follow the same seasonal dynamics, and that seasonality of community-acquired infections of *S. aureus* is the same than its respiratory samples, but opposite to its skin samples. Infections of the respiratory tract of

S. aureus and *E. coli* and synchronous, with a lag between them and *H. influenzae*. Several seasonal complexes may be found: a complex associating *S. hominis* and *epidermidis* with most of their locations; concerning *S. aureus*, *E. coli* and *P. aeruginosa* a first complex associating their respiratory locations and a second associating their blood infections.

The analysis showed that, for *S. aureus*, sample coming for all community-acquired infections and respiratory samples of these infections were synchronous with a strong cross-correlation (r = 0.56), both with their peaks during week 2 (Figure 6). This shows that the apparent overall seasonality of S. aureus community-acquired infections is mostly driven by respiratory infections. In contrast, cutaneous and respiratory *S. aureus* community-acquired infections exhibited a shifted cross-correlation (r=0.24) (Figure 7) with peaks respectively during weeks 38 and 2. These results show that the apparent winter seasonality of *S. aureus* community-acquired infections masks the cutaneous infection seasonality of this species.

3.2 Species periodicities and meteorological drivers

Temperature (°C), rain (mm), humidity (%), wind (m/s) and pressure change (Pa) were tested for significant positive or negative associations with seasonal infection incidences. Significant associations of incidences with possible meteorological drivers for respiratory, urines, blood and cutaneous samples coming from community-acquired and hospital-associated infections, are presented in Table 2 for the 15 most frequent bacterial species. Concerning *S. aureus*, hospital-associated infections exhibited a significant association with temperature, humidity, and pressure change, whereas community-acquired infections were only associated with precipitations.

Community-acquired *H. influenzae* infections were globally associated with humidity and pressure changes (precipitations and pressure change for blood samples, and wind for urine samples).

E. cloacae exhibited significant associations with pressure changes and wind for hospital-associated infections, and with humidity for community-acquired infections.

K. pneumoniae infections were associated to humidity for urine samples in community-acquired infections.

Among a total of 156 combinations of bacterial species, origin, and sample types with a seasonality (Supplementary Table 2), (59/156, 37.8%) were significantly associated with precipitations, (43/156, 27.6%) with pressure change, (38/156, 24.4%) with humidity, (36/156, 23.1%) with wind, and (23/156, 14.4%) with temperature. In contrast, (19/156, 12.2%) combinations had no significant association with any meteorological parameter.

4 Discussion

This seasonality study was based on the university hospital microbiological laboratory activities from an area of the East Mediterranean coast of France, characterized by abundant sunshine, warm and dry summers, windy episodes, associated with rainfalls from October to April. This big-data systematic cross analysis taking in account the infection origin, types of samples and age category, and not only bacterial species identifications (Lal *et al.*, 2013; Djennad *et al.*, 2018; Huang *et al.*, 2011; Warren-Gash *et al.*, 2011) was processed using a statistical workflow in search for seasonal periodicities of the incidence time series. These analyses were supplemented by the search for co-seasonalities and association with meteorological drivers. From a methodological point of view, the originality of this study resides in the distinction made on the sample origin in the systematic analysis, which allows the

decomposition of the bacterial species overall time series into infection-related time series that may have different seasonalities. A workflow was built for controlling series stationarity and for handling varying periodicities. A hierarchical clustering of the time series seasonal peaks allowed a global exploration of co-seasonalities, with possible deeper analyses using cross-correlations. The weekly granularity of the time series brought us more precision in the location of the seasonal peak, and allowed also a more precise search for meteorological drivers. However, this last study must be taken with caution as the linear model used does not take into specific consideration the possible lag between the periodicities of the meteorological drivers and the infections. In the same way, it must be considered that an apparent driver may be in fact the result of a confusion with other determinants, as seasonalities of patients' metabolic conditions (e.g., the seasonality in the skin production of vitamin D), or even more complex factors in the case of the seasonality of hospital-acquired infections. We believe that this study is the first that has systematically examined the association between climatic drivers and bacterial species while taking in account sample kinds.

From 228,365 unique bacterial identifications during a 6-years data collection, 86 bacterial species had at least one time series with at least 100 observations, and able to be analysed by the workflow. Among them, this paper focused specifically on the 15 most frequent bacterial species. A same identification method (MALDI-TOF) with a same protocol has been used along the collection, enforcing the homogeneity of the identification capability.

A first result of this work is that seasonality is more frequent than non-seasonality. An incidence seasonality was retrieved for 55.9% of the time series belonging to the 15 most frequent species. Hospital-associated infection without sample distinction showed a seasonality for 2/3 of these species, and the same frequency of seasonalities in community-acquired infections without sample distinction. It is important to underline here that the seasonality of a specific infection may be hidden in the non-seasonality of the global species incidence, as for *K. oxytoca*, and that an overall apparent seasonality of a species may be the result of the seasonalities of its various forms of infections with different phases. For example, *S. aureus* exhibited a community-acquired respiratory infection peak during winter, a community-acquired skin infection peak during summer, and an overall seasonal peak of community-acquired infections during winter, corresponding to the peak of its most abundant sub-population. Epidemiologically speaking, the microbiological point of view at species level may be different than the clinical infection point of view.

In our study, winter is the season with the most seasonal peaks, and especially the second week.

The association analysis of species with weather variables identified several possible drivers. *H. influenzae* incidence was associated with humidity and pressure change, and *S. capitis* incidence with temperature and precipitations. Our study showed that hospital-associated *S. aureus* blood infections were significantly associated with temperature as shown by Eber et al (Eber *et al.*, 2011), who estimated that an increase in temperature of 5.6°C was associated with a 2.2% increase (95% CI 1.3-3.2) in *S. aureus* frequency. In addition, other studies have shown the presence of a *S. aureus* seasonal peak during summer (Dailiana *et al.*, 2008), or during autumn (Tveten *et al.*, 2002) or both (Leekha Leekha, 2012), depending the kind of clinical picture. In fact, with only a statistical co-occurrence analysis, it is difficult to determine if a species seasonality is driven by a weather condition or if the weather condition is only a characteristic of the season corresponding to the species seasonal peak.

Similar studies of *Campylobacter* and *Salmonella* seasonalities in the UK (Nichols *et al.*, 2012; Cherrie *et al.*, 2018) demonstrated that the prevalence of *Campylobacter spp.* was associated with temperature, while our study showed an association of *Campylobacter* with precipitation and humidity

(supplementary material), and other studies found limited association with temperature (Sari Kovats *et al.*, 2005). Hospital-associated *H. influenzae* infections were associated with pressure change for all age groups, and humidity for the 0-20 year-old age group (Table 2). It was reported by Shaman and Kohn (Shaman & Kohn, 2009) that atmospheric pressure changes could have strong associations with several infectious agents such as the influenza virus, which creates optimal conditions for the manifestation of *H. influenzae*. In our study week 2 gathered the most important number of seasonal peaks, which is may be due to the low temperature and high humidity during this winter period. Most Gram-negative bacteria such as *E. cloacae, E. coli, K. pneumoniae* and *P. aeruginosa* did not exhibit any significant association with temperature except *A. baumannii*, in contrast with other studies (Paul, 2012; Freeman *et al.*, 2009; Richet, 2012) reporting a higher incidence of positive blood cultures for these species during summer months. However, Anderson and al (Anderson *et al.*, 2008) showed in a previous study that *K. pneumoniae* was significantly associated with temperature ($p<0.0001$) and relative humidity ($p<0.0001$) and both were linear predictors of *K. pneumoniae* growth rates.

This systematic big-data study on bacterial identifications of hospitalized patient shows that seasonality is a frequent, but not systematic, characteristic, and that seasonalities of bacterial species is different from the seasonalities of the different infections of the species. Our study of meteorological drivers of seasonality must be expanded with further works taking in account the possible variable lags between disease and meteorological time series.

5 Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

6 Author Contributions

HC: project design. LK, AGG, PC and HC: research, analysis and paper drafting. PEF,DR critical revision for important intellectual content.

7 Funding

Details of all funding sources should be provided, including grant numbers if applicable. Please ensure to add all necessary funding information, as after publication this is no longer possible.

8 References

Abat C., Chaudet H., Rolain J-M., Colson P. & Raoult D. (2016). Traditional and syndromic surveillance of infectious diseases and pathogens. *International Journal of Infectious Diseases*, 48, 22–28. https://doi.org/10.1016/j.ijid.2016.04.021

Al-Hasan M.N., Lahr B.D., Eckel-Passow J.E., Baddour L.M. (2009). Seasonal variation in Escherichia coli bloodstream infection: a population-based study. *Clinical Microbiology and Infection*, 15(10), 947–950. https://doi.org/10.1111/j.1469-0691.2009.02877.x

Alarcon Falconi T.M., Estrella B., Sempértegui F. & Naumova E.N. (2020). Effects of data aggregation on time series analysis of seasonal infections. *International Journal of Environmental Research and Public Health*,17(16), 5887. https://doi.org/10.3390/ijerph17165887

Anderson D.J., Richet H., Chen L.F., Spelman D.W., Hung Y., Huang A.T., Sexton D.J. & Raoult D. (2008). Seasonal variation in Klebsiella pneumoniae bloodstream infection on 4 continents. *The Journal of Infectious Diseases*, 197(5), 752–756. https://doi.org/10.1086/527486

Brownlee J. & Fletcher W.M (1918). VI. An investigation into the periodicity of measles epidemics in London from 1703 to the present day by the method of the periodogram. *Philosophical Transactions of the Royal Society B*, 208(348-359), 225–250. https://doi.org/10.1098/rstb.1918.0006

Cherrie M.P.C., Nichols G., Iacono G.L., Sarran C., Hajat S. & Fleming L.E. (2018). Pathogen seasonality and links with weather in England and Wales: a big data time series analysis. *BMC Public Health*, 18, 1067. https://doi.org/10.1186/s12889-018-5931-6

Cleveland R.B., Cleveland W.S., McRae J.E. & Terpenning I.J. (1990). STL: a seasonal-trend decomposition procedure based on loess. *Journal of Official Statistics*, 6 (1), 3–33.

Coletti P., Poletto C., Turbelin C., Blanchon T. & Colizza V. (2018). Shifting patterns of seasonal influenza epidemics. *Scientific Reports*, 8, 12786. https://doi.org/10.1038/s41598-018-30949-x

Dailiana Z.H., Rigopoulos N., Varitimidis S.E., Poultsides L., Petinaki E. & Malizos K.N. (2008). Clinical and epidemiological features of upper-extremity infections caused by Staphylococcus aureus carrying the PVL gene: a four-year study in Greece. *Medical science monitor: international medical journal of experimental and clinical research*, 14(10), CR511-514.

De Livera A.M., Hyndman R.J. & Snyder R.D. (2011). Forecasting time series with complex seasonal patterns using exponential smoothing. *Journal of the American Statistical Association*, 106(496), 1513–1527. https://doi.org/10.1198/jasa.2011.tm09771

Djennad A., Lo Iacono G., Sarran C., Fleming L.E., Kessel A., Haines A., Haines A. & Nichols G.L. (2018). A comparison of weather variables linked to infectious disease patterns using laboratory addresses and patient residence addresses. *BMC Infectious Diseases*,18, 198. https://doi.org/10.1186/s12879-018-3106-9

Dowell S.F. (2001). Seasonal variation in host susceptibility and cycles of certain infectious diseases. *Emerging Infectious Diseases*, 7(3), 369-374. https://doi.org/10.3201/eid0703.017301.

Eber M.R., Shardell M., Schweizer M.L., Laxminarayan R. & Perencevich E.N. (2011). Seasonal and temperature-associated increases in Gram-negative bacterial bloodstream infections among hospitalized patients. *PLoS ONE*, 6, e25298. https://doi.org/10.1371/journal.pone.0025298

Fisman D. (2012). Seasonality of viral infections: mechanisms and unknowns. *Clinical Microbiology and Infection*, 18(10), 946–954. https://doi.org/10.1111/j.1469-0691.2012.03968.x

Fisman D.N. (2007). Seasonality of infectious diseases. *Annual Review of Public Health*, 28, 127–43. https://doi.org/10.1146/annurev.publhealth.28.021406.144128

Freeman J.T., Anderson D.J. & Sexton D.J. (2009). Seasonal peaks in Escherichia coli infections: possible explanations and implications. *Clinical Microbiology and Infection*, 15(10), 951–953. https://doi.org/10.1111/j.1469-0691.2009.02866.x

359 Greer A.L., Drews S.J. & Fisman D.N. (2009). Why "winter" vomiting disease? Seasonality,
360 hydrology, and Norovirus epidemiology in Toronto, Canada. *EcoHealth*, 6: 192–199.
361 https://doi.org/10.1007/s10393-009-0247-8

362 Harrison L.H., Armstrong C.W., Jenkins S.R., Harmon M.W., Ajello G.W., Miller G.B. Jr & Broome
363 C.V. (1991). A cluster of meningococcal disease on a school bus following epidemic influenza.
364 *Archives of Internal Medicine*, 151(5), 1005–1009.
365 https://doi.org/0.1001/archinte.1991.00400050141028

366 Huang F., Zhou S., Zhang S., Wang H., Tang L. (2011). Temporal correlation analysis between malaria
367 and meteorological factors in Motuo County, Tibet. *Malaria Journal*, 10, 54.
368 https://doi.org/10.1186/1475-2875-10-54

369 Kwiatkowski D., Phillips P.C.B., Schmidt P. & Shin Y. (1992). Testing the null hypothesis of
370 stationarity against the alternative of a unit root. *Journal of Econometrics*, 54(1-3), 159–178.
371 https://doi.org/10.1016/0304-4076(92)90104-Y

372 Lal A., Ikeda T., French N., Baker M.G. & Hales S. (2013). Climate variability, weather and enteric
373 disease incidence in New Zealand: Time Series Analysis. *PLoS ONE*, 8, e83484.
374 https://doi.org/10.1371/journal.pone.0083484

375 Leekha S., Diekema D.J. & Perencevich E.N. (2012). Seasonality of staphylococcal infections. *Clinical*
376 *Microbiology and Infection*, 18(10), 927–933. https://doi.org/10.1111/j.1469-0691.2012.03955.x

377 Linthicum K.J., Britch S.C. & Anyamba A. (2016). Rift Valley Fever: an emerging mosquito-borne
378 disease. *Annual Review of Entomology*, 61, 395–415. https://doi.org/10.1146/annurev-ento-010715-
379 023819

380 Martinez M.E. (2018). The calendar of epidemics: Seasonal cycles of infectious diseases. *PLOS*
381 *Pathogens*, 14(11), e1007327. https://doi.org/10.1371/journal.ppat.1007327

382 Naumova E.N., Christodouleas J., Hunter P.R. & Syed Q. (1999). Effect of precipitation on seasonal
383 variability in cryptosporidiosis recorded by the North West England surveillance system in 1990-1999.
384 *Journal of Water & Health*, 3(2), 185–196. https://doi.org/10.2166/wh.2005.0017

385 Nichols G.L., Richardson J.F., Sheppard S.K., Lane C. & Sarran C. (2012). Campylobacter
386 epidemiology: a descriptive study reviewing 1 million cases in England and Wales between 1989 and
387 2011. *BMJ Open*, 2, e001179. http://dx.doi.org/10.1136/bmjopen-2012-001179

388 Park M.S., Park K.H. & Bahk G.J. (2018). Combined influence of multiple climatic factors on the
389 incidence of bacterial foodborne diseases. *Science of The Total Environment*, 610–611, 10–16.
390 https://doi.org/10.1016/j.scitotenv.2017.08.045

391 Paul M. (2012). Seasonality in infectious diseases: does it exist for all pathogens? *Clinical*
392 *Microbiology and Infection*, 18(10), 925–926. https://doi.org/10.1111/j.1469-0691.2012.03972.x

393 Probst W.N., Stelzenmüller V. & Fock H.O. (2012). Using cross-correlations to assess the relationship
394 between time-lagged pressure and state indicators: an exemplary analysis of North Sea fish population
395 indicators. *ICES Journal of Marine Science*, 69(4), 670–681. https://doi.org/10.1093/icesjms/fss015

R Core Team (2021). A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/

Richet H. (2012). Seasonality in Gram-negative and healthcare-associated infections. *Clinical Microbiology and Infection*, 18(10), 934–940. https://doi.org/10.1111/j.1469-0691.2012.03954.x

Salmon-Rousseau A., Piednoir E., Cattoir V. & de La Blanchardière A. (2016). Hajj-associated infections. *Médecine et Maladies Infectieuses*, 46(7), 346–354. https://doi.org/10.1016/j.medmal.2016.04.002

Sari Kovats R., Edwards S.J., Charron D., Cowden J., D'Souza R.M., Ebi K.L., Gauci C., Gerner-Smidt P., Hajat S., Hales S., Hernández Pezzi G., Kriz B., Kutsar K., McKeown P., Mellou K., Menne B., O'Brien S., van Pelt W. & Schmid H. (2005). Climate variability and campylobacter infection: an international study. *International Journal of Biometeorology*, 49, 207–214. https://doi.org/10.1007/s00484-004-0241-3

Schuster G., Ebert E.E., Stevenson M.A., Corner R.J. & Johansen C.A. (2011). Application of satellite precipitation data to analyse and model arbovirus activity in the tropics. *International Journal of Health Geographics*, 10, 8. https://doi.org/10.1186/1476-072X-10-8

Shaman J. & Kohn M. (2009). Absolute humidity modulates influenza survival, transmission, and seasonality. *Proceedings of the National Academy of Sciences of the United States of America*, 106(9), 3243–3248. https://doi.org/10.1073/pnas.0806852106

Stashevsky P.S., Yakovina I.N., Alarcon Falconi T.M. & Naumova E.N. (2019). Agglomerative clustering of enteric infections and weather parameters to identify seasonal outbreaks in cold climates. *International Journal of Environmental Research and Public Health*, 16(12), 2083. https://dx.doi.org/10.3390/ijerph16122083

Talbot T.R., Poehling K.A., Hartert T.V., Arbogast P.G., Halasa N.B., Edwards K.M., Schaffner W., Craig A.S. & Griffin M.R. (2005). Seasonality of invasive pneumococcal disease: temporal relation to documented influenza and respiratory syncytial viral circulation. *The American Journal of Medicine*, 118(3), 285–291. https://doi.org/10.1016/j.amjmed.2004.09.016

Tveten Y., Jenkins A. & Kristiansen B-E. (2002). A fusidic acid-resistant clone of Staphylococcus aureus associated with impetigo bullosa is spreading in Norway. *Journal of Antimicrobial Chemotherapy*, 50(6), 873–876. https://doi.org/10.1093/jac/dkf217

Viboud C., Bjørnstad O.N., Smith D.L., Simonsen L., Miller M.A. & Grenfell B.T. (2006). Synchrony, waves, and spatial hierarchies in the spread of influenza. *Science*, 312(5772), 447–51. DOI: 10.1126/science.1125237

Warren-Gash C., Bhaskaran K., Hayward A., Leung G.M., Lo S-V., Wong C-M., Ellis J., Pebody R., Smeeth L. & Cowling B.J. (2011). Circulating influenza virus, climatic factors, and acute myocardial infarction: a time series study in England and Wales and Hong Kong. *The Journal of Infectious Diseases*, 203(12), 1710–1718. https://doi.org/10.1093/infdis/jir171

TABLES

Table 1: Descriptive statistics of the weekly incidences series for the 15 most frequent bacterial species

N°	Pathogens	Min	1st Qu	Median	Mean	3rd Qu	Max
1	*Enterobacter cloacae*	3	13	16	16.60	20	34
2	*Enterococcus faecalis*	9	23	27	27.75	32	48
3	*Escherichia coli*	109	135	145	145.4	155	203
4	*Gardnerella vaginalis*	3	16	22	23.46	30	51
5	*Haemophilus influenzae*	1	7	9	10.03	13	28
6	*Klebsiella oxytoca*	0	4	6	6.02	8	16
7	*Klebsiella pneumoniae*	18	31	35	36.06	41	66
8	*Propionibacterium acnes*	0	4	6	6.35	9	22
9	*Proteus mirabilis*	5	11	14	14.03	16	33
10	*Pseudomonas aeruginosa*	18	29	34	33.96	39	55
11	*Staphylococcus aureus*	45	67	74	74.04	80	111
12	*Staphylococcus epidermidis*	22	35	39	39.38	44	66
13	*Staphylococcus haemolyticus*	1	4	6	6.54	8	19
14	*Staphylococcus hominis*	0	6	8	8.04	10	23
15	*Streptococcus agalactiae*	7	18	22	21.52	26	39

437 **Table 2:** Meteorological drivers associated with seasonal times series for the 15 most frequent species

Pathogens	Sample size	Hospital-associated infections		p-value	Community-acquired infections		p-value
		Samples	Drivers		Samples	Drivers	
Staphylococcus aureus	23173	all samples	temperature	0.009	all samples	rain	0.004
			humidity	0.011	urines	rain	0.037
			pressure change	0.013		pressure change	0.039
		urines	rain	0.021			
Staphylococcus epidermidis	12325				resp	rain	0.018
						humidity	0.030
Klebsiella pneumoniae	11287				urines	humidity	0.016
Pseudomonas aeruginosa	10631	blood	rain	0.020	blood	humidity	0.019
					skin	rain	0.043
						pressure change	0.015
Gardnerella vaginalis	7344	all samples	rain	0.013			
		urines	temperature	0.011			
Streptococcus agalactiae	6735	urines	temperature	0.005	resp	rain	0.010
			rain	0.013		humidity	0.029
Enterobacter cloacae	5195	all samples	wind	0.045	all samples	humidity	0.002
			pressure change	0.019			
Haemophilus influenzae	3139	all samples	humidity	0.013	all samples	humidity	0.022
		blood	humidity	0.008		pressure change	0.004
					blood	rain	0.023
						pressure change	0.011
					urines	wind	0.049
Staphylococcus hominis	2516	urines	pressure change	0.043			
Staphylococcus haemolyticus	2048	all samples	humidity	0.024			
		urines	rain	0.023			
			humidity	0.033			
			wind	0.046			
Propionibacterium acnes	1986				all samples	humidity	0.016
					skin	rain	0.001
Klebsiella oxytoca	1884	all samples	rain	0.041			

FIGURES

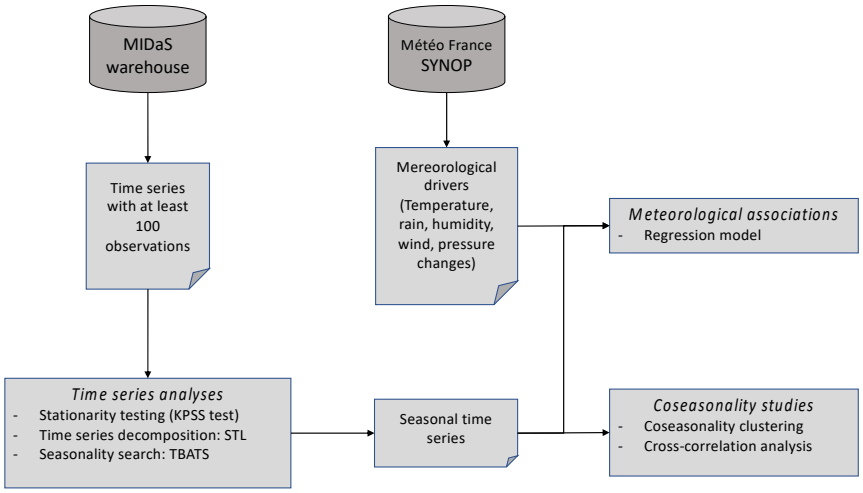


Figure 1: Data analysis workflow

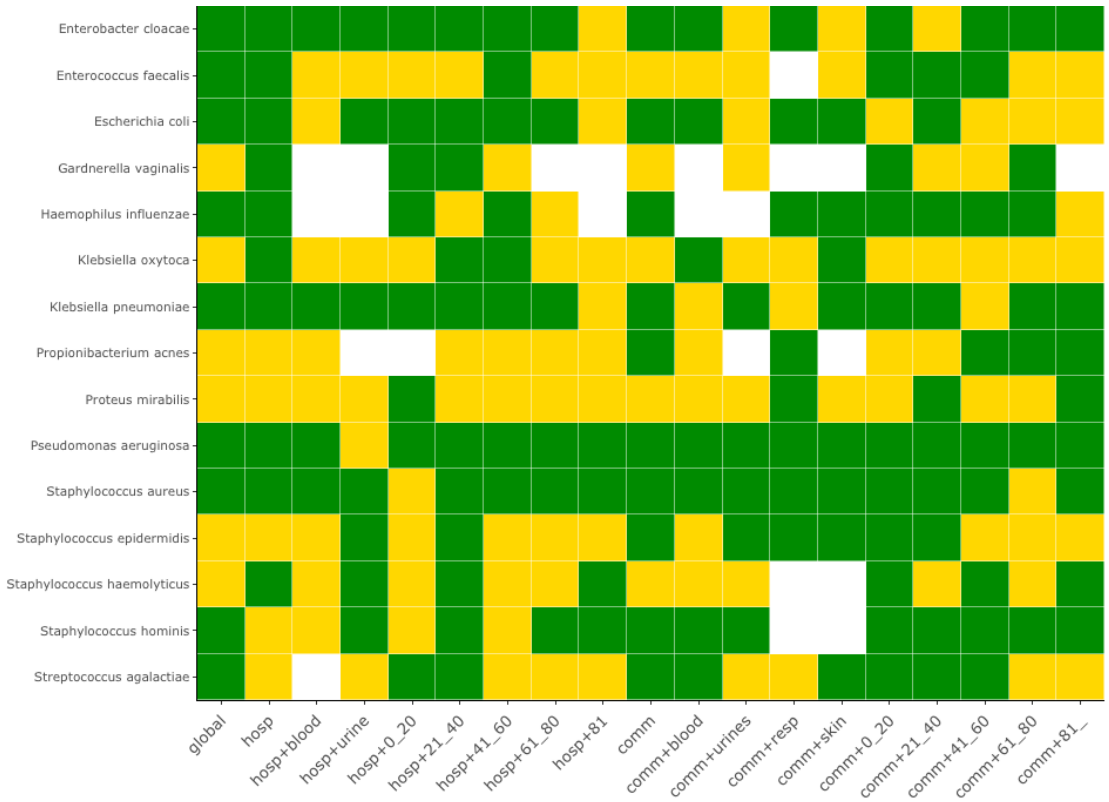


Figure 2: Time series with seasonalities detected by TBATS analysis for the 15 most frequent bacterial species.

Map color: Green: Seasonality detected; Yellow: No seasonality; White: Insufficient number of observations.

Abbreviations: hosp - hospital-associated infections; comm - community-acquired infections; blood - blood samples; resp - respiratory samples; skin - skin samples; urine - urine samples.

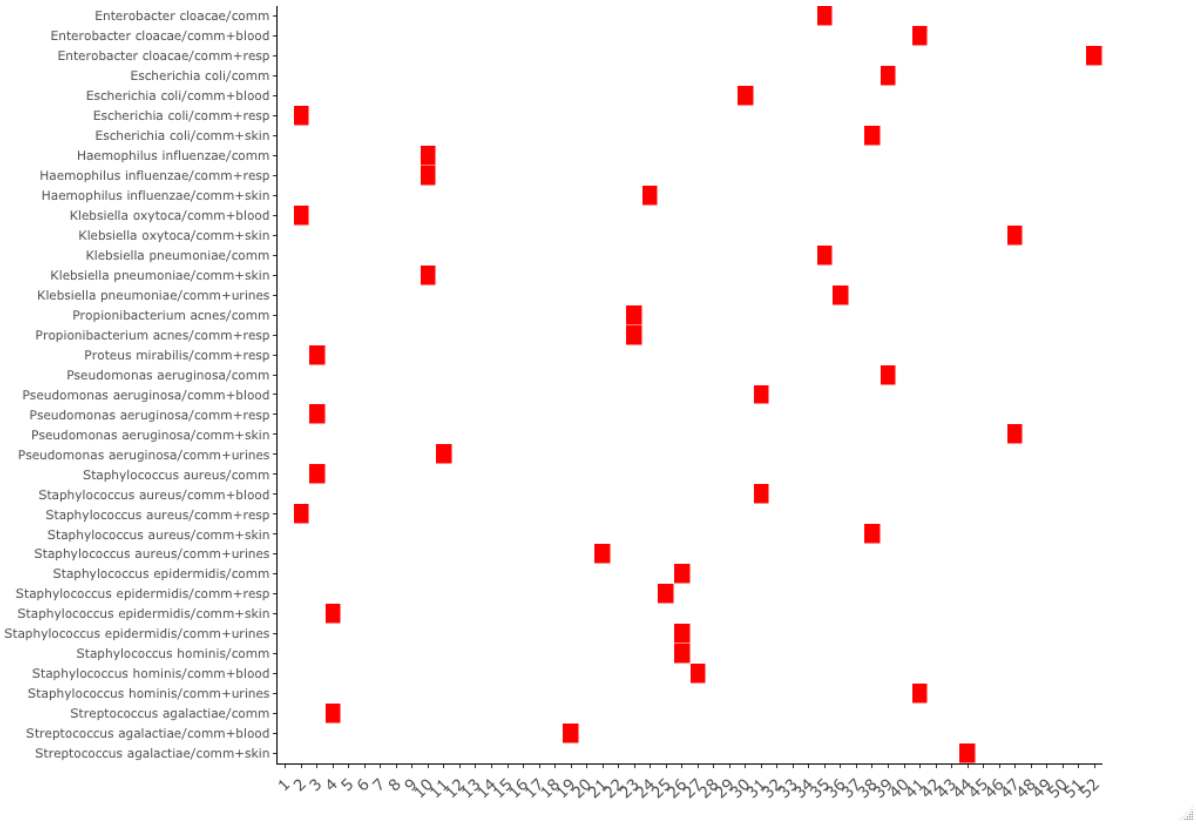


Figure 3: Week of seasonality peak for community-acquired infections of the 15 most frequent bacterial species. Abbreviations: hosp - hospital-associated infections; comm - community-acquired infections; blood - blood samples; resp - respiratory samples; skin - skin samples; urine - urine samples

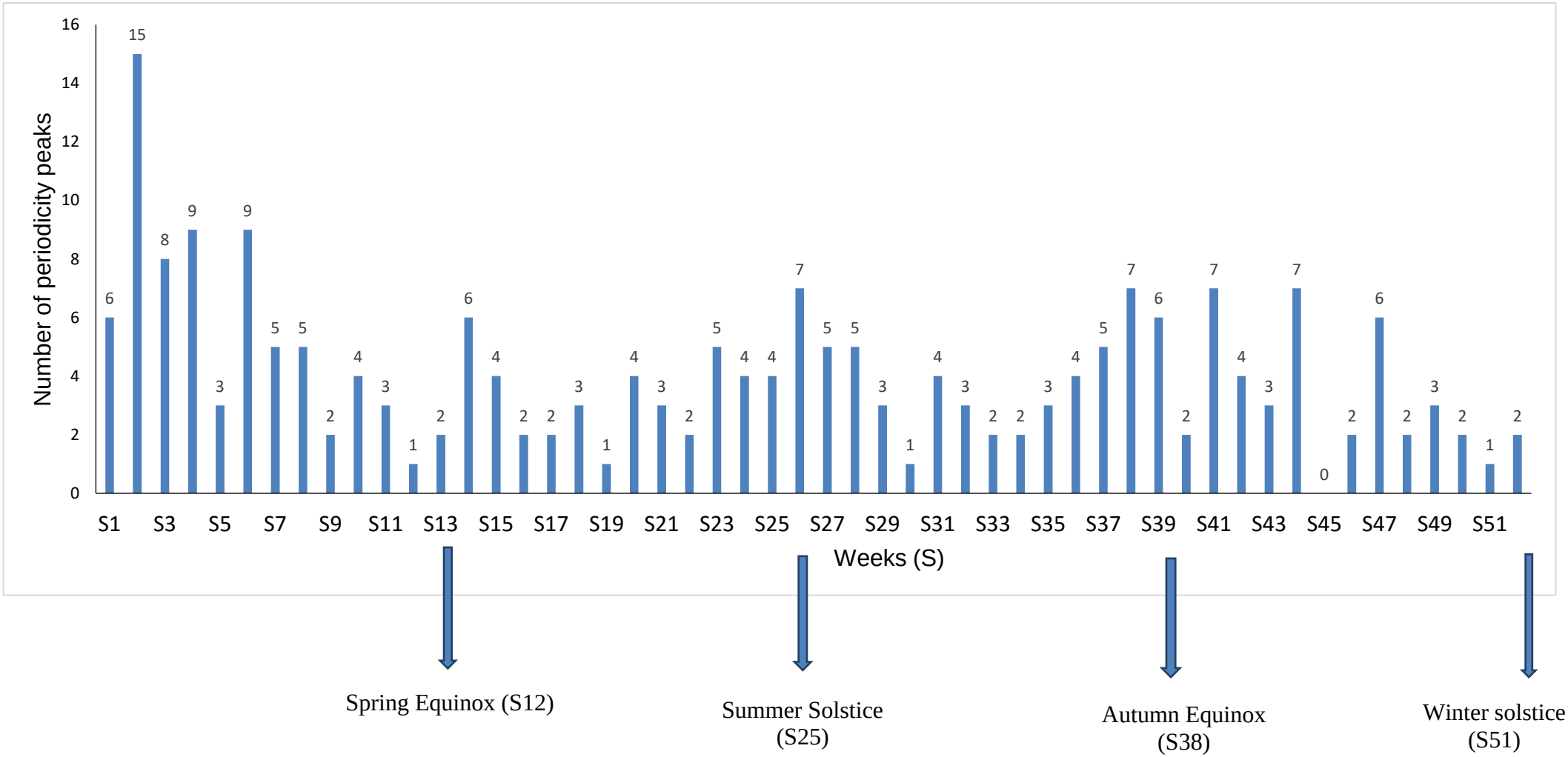


Figure 4: Distribution of the number of s he year for all samples of the 86 species.

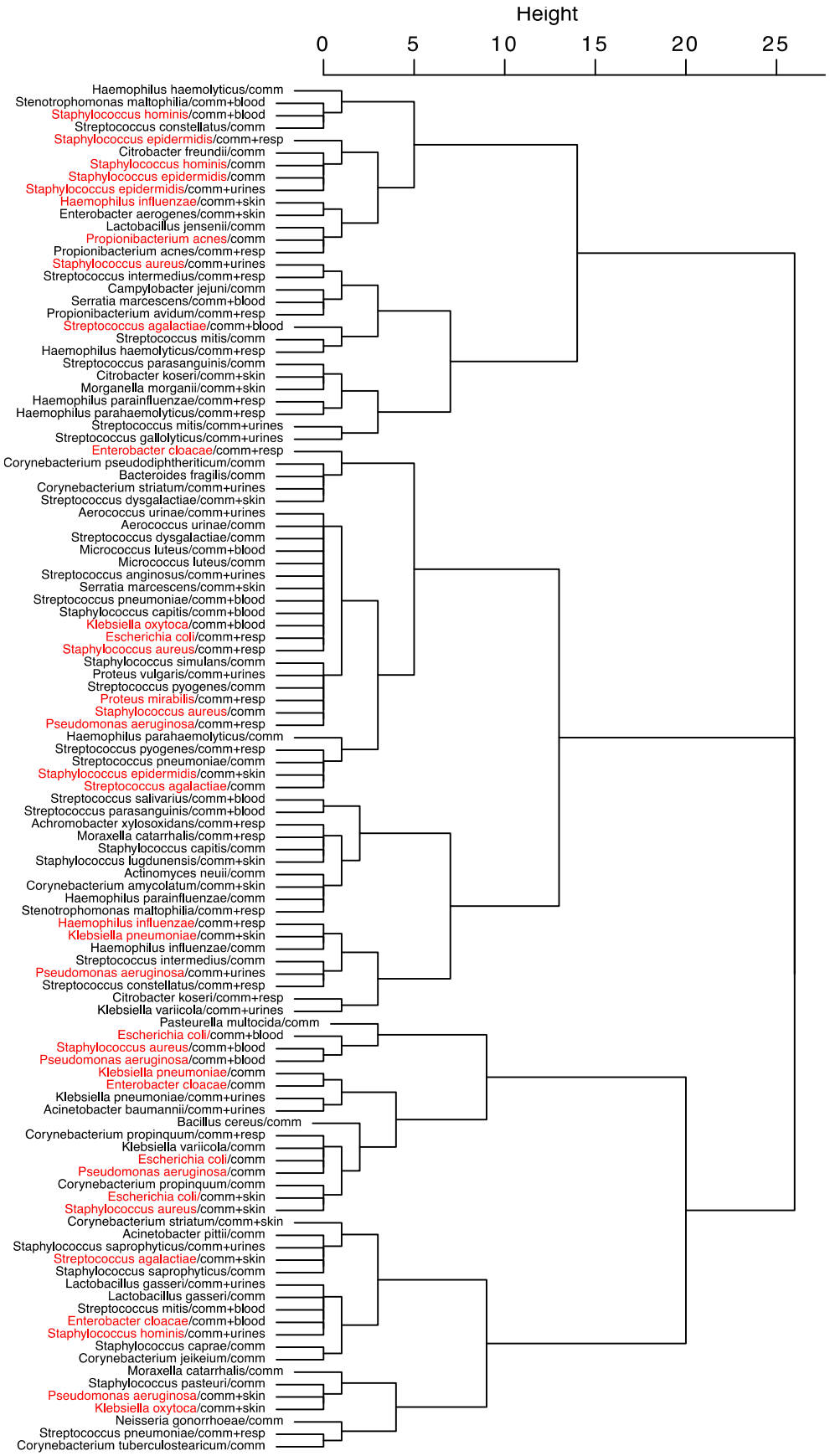


Figure 5: Clustering of seasonal peak distances between community-related infections. Samples from the 15 most frequent species are in red.

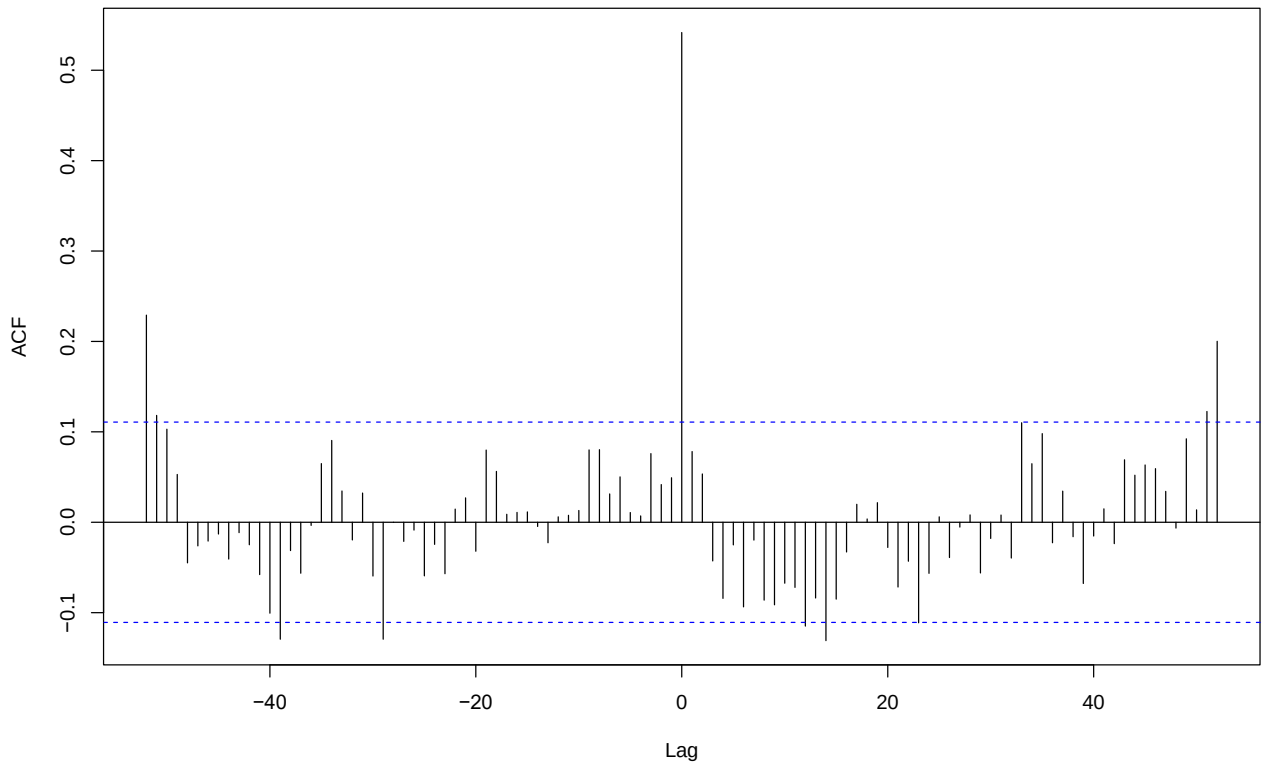


Figure 6: *Staphylococcus aureus* cross-correlation correlogram of community-acquired vs respiratory samples (r= 0.542 for lag=0). The horizontal blue dashed lines indicate the significance levels.

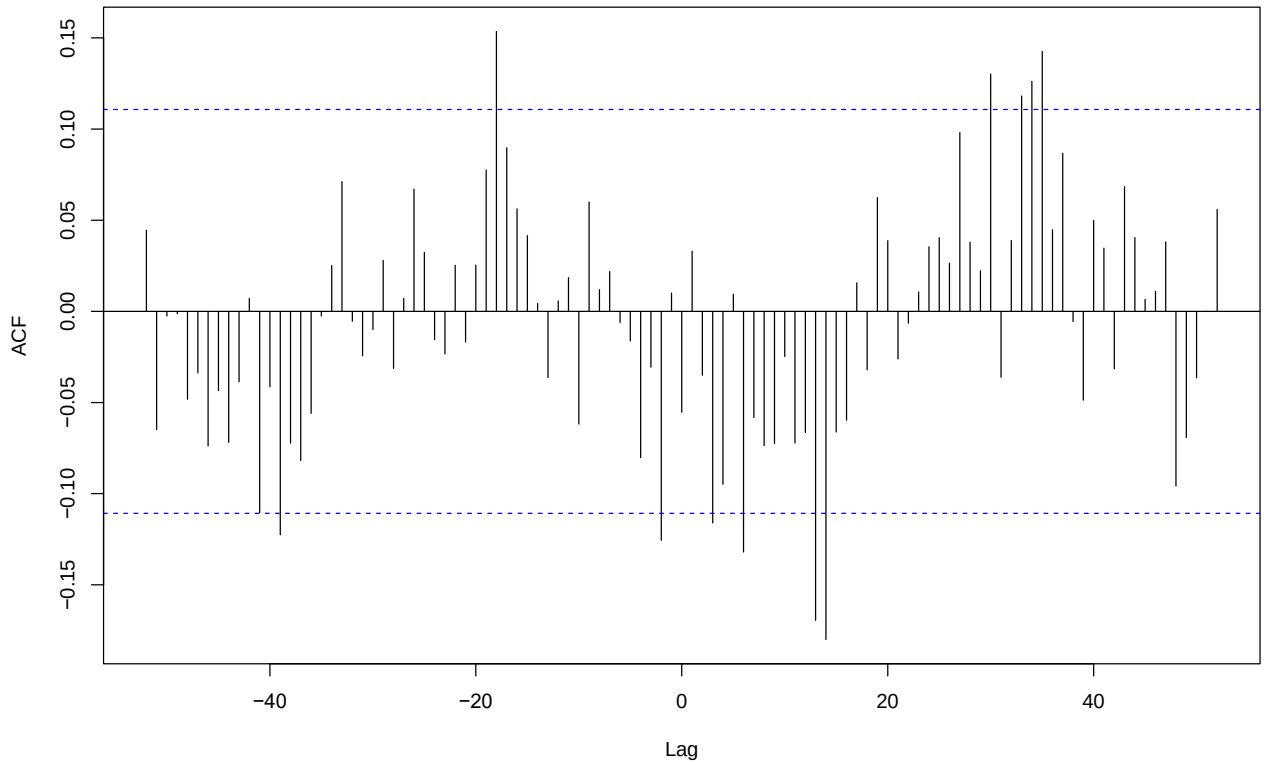


Figure 7: *Staphylococcus aureus* cross-correlation correlogram of community-acquired skin vs respiratory infections (r= 0.154 for lag=-18). The horizontal blue dashed lines indicate the significance levels.

Supplementary Material

1. Supplementary Figures and Tables

1.1 Supplementary Tables

Supplementary Table 1: List of SYNOP weather variables

Description	**type**	**unit**
WMO station identifier	string	
Date (UTC)	string	AAAAMMDDHHMISS
Sea level atmospheric pressure	integer	Pa
Pressure variation in 3 hours	integer	Pa
Barometric trend type	integer	code (0200)
Mean wind direction for 10 min	integer	degree
Mean wind speed for 10 min	float	m/s
Temperature	float	K
Dew point temperature	float	K
Humidity	integer	%
Horizontal visibility	float	m
Current weather	integer	WMO code (4677)
Past weather 1	integer	WMO code (4561)
past weather 2	integer	WMO code (4561)
Total cloudiness	float	%
Low-level cloudiness	integer	octa
Height of the low-level cloud base	integer	m
Low-level cloud kind	integer	WMO code (0513)
Medium-level cloud kind	integer	WMO code (0515)
High-level cloud kind	integer	WMO code (0509)
Station atmospheric pressure	integer	Pa
Barometric level	integer	Pa
Geopotential	integer	m^2/s^2
24 hours pressure variation	integer	Pa
Minimal temperature	float	K
Maximal temperature	float	K
Minimal ground temperature during the last 12 hours	float	K
Wet-bulb temperature method	integer	WMO code (3855)

Wet-bulb temperature	float	K
Gusts during the last 10 min.	float	m/s
Gusts during a period	float	m/s
Period of gust measure	float	min
Ground state	integer	WMO code (0901)
Total snow, ice, other height on ground	float	m
Fresh snow height	float	m
Measure periodicity of the fresh snow	float	1/10 hour
Rain	float	mm
Supplementary information	float	WMO code (3778)

Supplementary Table 2: Meteorological drivers associated with seasonal times series for all species

Bacterial species	Hospital-associated infections		Community-acquired infections	
	Samples	**Drivers**	**Samples**	**Drivers**
Achromobacter xylosoxidans	Urines	Rain		
		Wind		
		Pressure change		
Aerococcus urinae	All samples	Rain		
	Urines	Temperature		
		Rain		
Campylobacter jejuni	Blood	Rain		
		Humidity		
Citrobacter freundii	Blood	Temperature		
		Rain		
Corynebacterium propinquum	All samples	Wind		
	Blood	Rain		
		Wind		
Corynebacterium striatum			Skin	Rain
				Humidity
Corynebacterium urealyticum	All samples	Humidity		
Enterobacter aerogenes	All samples	Wind	Resp	Temperature
		Pressure change		
Enterobacter asburiae	All samples	Rain		
Enterobacter cloacae	All samples	Wind	All samples	Humidity
		Pressure change		
Gardnerella vaginalis	All samples	Temperature		
	Urines	Temperature		
Haemophilus haemolyticus	All samples	Temperature		
		Wind		

	Blood	Temperature		
		Wind		
	Urines	Pressure change		
Haemophilus influenzae	All samples	Humidity	All samples	Humidity
	Blood	Humidity		pressure change
			Blood	Rain
				pressure change
			Urines	Wind
Haemophilus parainfluenzae	All samples	Rain	All samples	Humidity
		Pressure change		pressure change
			Blood	Rain
				Humidity
Klebsiella oxytoca	All samples	Rain		
Klebsiella pneumoniae			Urines	Humidity
Lactobacillus gasseri	Blood	Humidity		
Lactobacillus jensenii	Urines	Temperature		
		Rain		
	All samples	Temperature		
		Rain		
Morganella morganii	Urines	Rain	Blood	Rain
		Humidity		Wind
			Urines	Rain
Neisseria gonorrhoeae	All samples	Rain		
		Humidity		
Propionibacterium acnes			All samples	Humidity
			Skin	Rain
Propionibacterium avidum	All samples	Humidity		
	Urines	Rain		
		Wind		
Pseudomonas aeruginosa	Blood	Rain	Blood	Rain
			Skin	Rain

				pressure change
Pseudomonas putida	All samples	Temperature		
		Pressure change		
Serratia marcescens			Skin	Temperature
				pressure change
Staphylococcus aureus	All samples	Temperature	All samples	Rain
		Humidity	Urines	Rain
		Pressure change		pressure change
	Urines	Rain		
Staphylococcus capitis			All samples	Temperature
				Wind
			Blood	Temperature
				pressure change
Staphylococcus caprae	Blood	Temperature		
Staphylococcus epidermidis			Resp	Rain
				Humidity
Staphylococcus haemolyticus	All samples	Humidity		
	Urines	Rain		
		Humidity		
		Wind		
Staphylococcus hominis	Urines	Pressure change		
Staphylococcus lugdunensis			Resp	Humidity
Staphylococcus pasteuri	All samples	Temperature		
Staphylococcus pettenkoferi	All samples	Pressure change		
Streptococcus agalactiae	Urines	Temperature	Resp	Rain
		Rain		Humidity
Streptococcus anginosus			Resp	Pressure change
Streptococcus constellatus			Urines	Temperature
Streptococcus dysgalactiae	All samples	Temperature		
	Blood	Temperature		
		Wind		

Streptococcus mitis			Blood	Temperature Wind
Streptococcus oralis			Blood	Rain Wind
Streptococcus parasanguinis	Blood	Rain Humidity		
Streptococcus pneumoniae	All samples	Humidity		

Conclusion et perspectives

L'introduction du MALDI-TOF (Matrix Assisted Laser Desorption-Ionization – Time Of Flight) dans l'identification des espèces bactériennes par les laboratoires de microbiologie clinique a élargi la palette des espèces pouvant être directement identifiées par rapport aux méthodes basées sur des kits phénotypiques ou génotypiques. De fait, la diversité et la dynamique apparente des espèces bactériennes s'est largement accrue dans ces laboratoires. Cette diversité dans les laboratoires de microbiologie reste très peu documentée et peu exploitée dans le cadre de la surveillance épidémiologique des maladies infectieuses. En effet, les indicateurs habituels de surveillance (taux d'incidence, de mortalité, …) donnent une vision indépendante des maladies infectieuses sous surveillance, alors que celles-ci sont de plus en plus envisagées dans le cadre d'un « écosystème » microbiologique humain. Une évolution de la surveillance épidémiologique en complétant ces indicateurs est donc envisageable. Cette amélioration passe par une meilleure compréhension de la co-occurrence des agents infectieux et de leur dynamique spatio-temporelle. Les objectifs de cette thèse s'inscrivent dans ce cadre pour (i) faire un état des lieux de l'étude de la diversité bactérienne, (ii) évaluer l'impact des mesures de confinement contre le SARS-CoV-2 sur la population bactérienne, (iii) comprendre la diversité des espèces bactériennes identifiées par les laboratoires de microbiologie, maillon fondamentale de la surveillance épidémiologique, ainsi que sa dynamique et (iv) identifier la corrélation entre des agents bactériens avec certains facteurs météorologiques, en vue de compléter les indicateurs de la surveillance épidémiologique par des indices de diversité.

Nous avons commencé par faire un état des lieux des mesures de diversité appliquées dans les études, ce qui nous a conduit à montrer que les indices couramment utilisés dans l'évaluation de la diversité demeurent principalement la richesse spécifique, l'abondance relative, l'indice de Shannon et l'indice de Simpson, communément appelé diversité alpha. Nous avons ensuite évalué et comparé la diversité alpha et bêta des espèces bactériennes révélées par les laboratoires de microbiologie affiliés au système de surveillance épidémiologique de la région PACA (PACASurvE), qui est l'un des systèmes de surveillance de la base MIDaS de l'Institut Hospitalo-Universitaire Méditerranée Infection (IHU-MI) de Marseille. Ce système assure la surveillance du nombre de patients infectés/colonisés par plus de 600 espèces bactériennes en utilisant les données de microbiologie produites par près de 300 laboratoires privés et publics de la région PACA. Nos résultats ont montré une grande variabilité de diversité entre les laboratoires en termes de volume d'activité c'est-à-dire en nombre d'échantillon analysé d'une part et d'autre part en termes de richesse spécifique et d'abondance relative. La distribution géographique des espèces identifiées variait aussi significativement selon les départements et selon le type du laboratoire. Le mix bactérien augmentait également au fil du temps. Sur la base de la distance de Bray-Curtis, les laboratoires ont été groupés en trois groupes selon leurs similarités. Le groupe 2 composé que de 3 laboratoires APHM, BIOESTEREL et NICE a montré la plus forte

diversité.

L'émergence du Covid 19 en fin 2019 et sa propagation rapide à travers le monde, a amené les Pays à prendre des mesures visant à endiguer et à contrôler la circulation du virus, parmi ces mesures, nous avons les mesures barrières et du confinement. Si ces mesures étaient orientées particulièrement contre le Covid 19, elles auraient également impacté d'autres agents infectieux outre que le SARS-CoV-2. Nous nous sommes également intéressés à évaluer les conséquences de ces mesures sur le top 15 des agents infectieux non-viraux et sur les autres virus respiratoires mais pas le SARS-CoV-2. Nos résultats ont révélé que les mesures appliquées ont entrainé une chute drastique de la fréquence de certains agents comme *Escherichia coli*, *Haemophilus influenzae*, *Klebsiella oxytoca*. En revanche, elles n'ont pas eu d'effet sur d'autres pathogènes telles *Staphylococcus aureus*, *Klebsiella pneumoniae*. De là, ce qu'il faut retenir, ces mesures quoi qu'appliquées contre le Covid 19, ont eu des impacts sur d'autres pathogènes.

Dans la même dynamique de contribuer à améliorer la détection précoce d'éventuelles épidémies par les systèmes de surveillance, nous avons terminé par analysé la périodicité du top 15 des agents infectieux identifiés à l'IHU en lien avec l'origine et le type de prélèvement et l'influence de certains facteurs météorologiques. Il est important de ne pas confondre la périodicité des agents infectieux à celle de la maladie infectieuse sous sa forme clinique qui est plus ou moins largement abordée par de nombreuses études. Dans cette étude, nous avons pu détecter des associations significatives de périodicité entre des pathogènes, l'origine de l'infection (nosocomiale ou communautaire) et le type de prélèvement (respiratoires, urines, hémoculture et cutané), ainsi que des associations significatives avec un ou plusieurs facteurs météorologiques (température, précipitation, humidité, vent et changement de presse). Ces résultats permettent une meilleure compréhension de la périodicité de ces agents infectieux et aideront à planifier des interventions futures pour la prévention d'émergence de ces pathogènes. Néanmoins, nous espérons que nos résultats permettront d'améliorer la surveillance épidémiologique des infections d'origine bactérienne et renforcer les indicateurs des systèmes de surveillance épidémiologique.

Nos travaux ont pu mettre en évidence pour la première fois une diversité bactérienne importante et spécifique dans des laboratoires de microbiologie et une possible introduction des indices de diversité pour compléter les indicateurs des systèmes de surveillance épidémiologique. Dans la même logique, nous avons pu détecter des périodicités associées de certains agents bactériens et non la maladie à un ou plusieurs facteurs météorologiques.

Cependant, tous les aspects n'ont pu être abordés dans cette thèse, il donc serait impérieux pour les recherches futures, d'élargir ces études à des infections d'origine virale et aux facteurs édapho-climatiques des zones de provenance des patients. Dans le cadre de la coopération, renforcer le partenariat avec les chercheurs des pays du sud pour une recherche active afin d'améliorer les systèmes de surveillance dans ces pays où les épidémies font payer de lourde tribu à la population dépourvue de structures de soins adéquates et confrontée à une pauvreté extrême.

Références

1. PEPIN M, BOIREAU P, BOUE F, CASTRIC J, CLIQUET F, DOUZAL Y, et al. Émergence des maladies infectieuses animales et humaines. INRAE Prod Anim [Internet]. 7 juin 2007 [cité 5 mai 2022];20(3):199-206. Disponible sur: https://productions-animales.org/article/view/3455#

2. Paul-Pierre P. Emerging diseases, zoonoses and vaccines to control them. Vaccine. 2009;

3. Astagneau P, Ancelle T, Brucker G. Surveillance épidémiologique principes, méthodes et applications en santé publique. Paris: Lavoisier; 2011.

4. Snowden FM. Emerging and reemerging diseases: a historical perspective. Immunol Rev [Internet]. 2008 [cité 5 mai 2022];225(1):9-26. Disponible sur: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1600-065X.2008.00677.x

5. Brucker G. Le syndrome Respiratoire Aigu Sévère (SRAS). Première épidémie du XXIe siècle. Bull Académie Natl Médecine [Internet]. 1 mai 2003 [cité 5 mai 2022];187(5):977-81. Disponible sur: https://www.sciencedirect.com/science/article/pii/S0001407919339792

6. Reynard O, Volchkov V, Peyrefitte C. Une première épidémie de fièvre à virus Ebola en Afrique de l'Ouest. médecine/sciences [Internet]. 1 juin 2014 [cité 5 mai 2022];30(6-7):671-3. Disponible sur: https://www.medecinesciences.org/articles/medsci/abs/2014/07/medsci2014306-7p671/medsci2014306-7p671.html

7. Gatherer D. The 2014 Ebola virus disease outbreak in West Africa. J Gen Virol. août 2014;95(Pt 8):1619-24.

8. Wiersinga WJ, Rhodes A, Cheng AC, Peacock SJ, Prescott HC. Pathophysiology, Transmission, Diagnosis, and Treatment of Coronavirus Disease 2019 (COVID-19): A Review. JAMA [Internet]. 25 août 2020 [cité 14 févr 2022];324(8):782-93. Disponible sur: https://doi.org/10.1001/jama.2020.12839

9. Ngwa MC, Young A, Liang S, Blackburn J, Mouhaman A, Morris JG. Cultural influences behind cholera transmission in the Far North Region, Republic of Cameroon: a field experience and implications for operational level planning of interventions. Pan Afr Med J [Internet]. 15 déc 2017 [cité 14 févr 2022];28:311. Disponible sur: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5927557/

10. Sack DA, Sack RB, Nair GB, Siddique A. Cholera. The Lancet [Internet]. 17 janv 2004 [cité 14 févr 2022];363(9404):223-33. Disponible sur: https://www.sciencedirect.com/science/article/pii/S0140673603153287

11. Jones KE, Patel NG, Levy MA, Storeygard A, Balk D, Gittleman JL, et al. Global trends in emerging infectious diseases. Nature [Internet]. févr 2008 [cité 6 mai 2022];451(7181):990-3. Disponible sur: https://www.nature.com/articles/nature06536

12. Keesing F, Belden LK, Daszak P, Dobson A, Harvell CD, Holt RD, et al. Impacts of biodiversity on the emergence and transmission of infectious diseases. Nature [Internet]. 2010 [cité 5 mai 2022];468(7324):647-52. Disponible sur: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7094913/

13. Seng P, Rolain JM, Fournier PE, La Scola B, Drancourt M, Raoult D. MALDI-TOF-mass spectrometry applications in clinical microbiology. Future Microbiol. nov 2010;5(11):1733-54.

14. Whittaker RH. EVOLUTION AND MEASUREMENT OF SPECIES DIVERSITY. TAXON [Internet]. mai 1972 [cité 7 janv 2020];21(2-3):213-51. Disponible sur: https://onlinelibrary.wiley.com/doi/abs/10.2307/1218190

15. Fakruddin Md, Mannan KSB, Andrews S. Viable but Nonculturable Bacteria: Food Safety and Public Health Perspective. ISRN Microbiol [Internet]. 2013 [cité 7 janv 2020];2013:1-6. Disponible sur: https://www.hindawi.com/archive/2013/703813/

16. Torsvik V, Goksøyr J, Daae FL. High diversity in DNA of soil bacteria. Appl Environ Microbiol. mars 1990;56(3):782-7.

17. Harpole W. Neutral Theory of Species Diversity. Nat Educ Knowl. 1 janv 2010;1:31.

18. Torsvik V, Daae FL, Sandaa RA, Øvreås L. Novel techniques for analysing microbial diversity in natural and perturbed environments. J Biotechnol [Internet]. sept 1998 [cité 7 janv 2020];64(1):53-62. Disponible sur: https://linkinghub.elsevier.com/retrieve/pii/S0168165698001035

19. Pauls SU, Nowak C, Bálint M, Pfenninger M. The impact of global climate change on genetic diversity within populations and species. Mol Ecol. févr 2013;22(4):925-46.

20. Suggitt AJ, Lister DG, Thomas CD. Widespread Effects of Climate Change on Local Plant Diversity. Curr Biol CB. 9 sept 2019;29(17):2905-2911.e2.

21. Wang J, Chen L, Tang W, Heino J, Jiang X. Effects of dam construction and fish invasion on the species, functional and phylogenetic diversity of fish assemblages in the Yellow River Basin. J Environ Manage. 1 sept 2021;293:112863.

22. Ji F, Yan L, Yan S, Qin T, Shen J, Zha J. Estimating aquatic plant diversity and distribution in rivers from Jingjinji region, China, using environmental DNA metabarcoding and a traditional survey method. Environ Res. août 2021;199:111348.

23. Magurran AE. Measuring biological diversity. 9 [Nachdr.]. Malden, Mass.: Blackwell; 2011. 256 p.

24. Milind G Watve. Questioning a Dogma - Do Bacteria Know When and to Mutate ? Gen Artic [Internet]. 1996;1:34-42. Disponible sur: https://www.ias.ac.in/article/fulltext/reso/001/08/0034-0042

25. Leport C, Guégan J. Les maladies infectieuses émergentes : état de la situation et perspectives [Internet]. undefined. 2011 [cité 8 nov 2020]. Disponible sur: /paper/Les-maladies-infectieuses-%C3%A9mergentes-%3A-%C3%A9tat-de-la-Leport-Gu%C3%A9gan/82ce38dd66d514f181bf28f6706c42dbb8df5f43

26. Global Burden of Disease. Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980-2017 : a systemactic analysis for for the Global Burden of Disease Study 2017. 2017.

27. Yannarell AC, Triplett EW. Geographic and environmental sources of variation in lake bacterial community composition. Appl Environ Microbiol. janv 2005;71(1):227-39.

28. Huart M, Bedubourg G, Abat C, Colson P, Rolain JM, Chaudet H, et al. Implementation and Initial Analysis of a Laboratory-Based Weekly Biosurveillance System, Provence-Alpes-Côte d'Azur, France. Emerg Infect Dis [Internet]. avr 2017 [cité 7 janv 2020];23(4):582-9. Disponible sur: http://wwwnc.cdc.gov/eid/article/23/4/16-1399_article.htm

29. OMS. Coronavirus : l'épidémie de Covid-19 considérée comme une pandémie par l'OMS. Le Monde.fr [Internet]. 11 mars 2020 [cité 8 mars 2022]; Disponible sur: https://www.lemonde.fr/planete/article/2020/03/11/le-point-sur-l-epidemie-due-au-coronavirus-dans-le-monde-l-iran-annonce-63-nouveaux-deces_6032633_3244.html

30. Hu B, Guo H, Zhou P, Shi ZL. Characteristics of SARS-CoV-2 and COVID-19. Nat Rev Microbiol [Internet]. mars 2021 [cité 8 mars 2022];19(3):141-54. Disponible sur: https://www.nature.com/articles/s41579-020-00459-7

31. Sun P, Lu X, Xu C, Sun W, Pan B. Understanding of COVID-19 based on current evidence. J Med Virol [Internet]. juin 2020 [cité 8 mars 2022];92(6):548-51. Disponible sur: https://onlinelibrary.wiley.com/doi/10.1002/jmv.25722

32. 20201012-weekly-epi-update-9.pdf [Internet]. [cité 8 mars 2022]. Disponible sur: https://www.who.int/docs/default-source/coronaviruse/situation-reports/20201012-weekly-epi-update-9.pdf

33. HASSON-FAURÉ PM et N. Coronavirus. Plus de 20 000 décès en France des suites du Covid-19, revivez la journée de lundi [Internet]. Ouest-France.fr. 2020 [cité 8 mars 2022]. Disponible sur: https://www.ouest-france.fr/sante/virus/coronavirus/direct-coronavirus-en-france-le-deconfinement-sera-tres-progressif-6812612

34. Erkoreka A. [Epidemics in nortthern Basque: black death and the Spanish influenza]. Hist Sci Medicales. juin 2008;42(2):113-22.

35. Guénel J. [Spanish influenza in France from 1918-1919]. Hist Sci Medicales. juin 2004;38(2):165-75.

36. Graff I, De Broucker C, Vargas J, Vanoost A, Gondry J, Foulon A. [COVID-19 and lockdown: Impact on pregnancy complications]. Gynecol Obstet Fertil Senol. 14 déc 2021;S2468-7189(21)00342-1.

37. Lamblin G, Golfier F, Peron J, Moret S, Chene G, Nohuz E, et al. [Impact of the COVID-19 Outbreak on the management of patients with gynecological cancers]. Gynecol Obstet Fertil Senol. nov 2020;48(11):777-83.

38. Partinen M, Holzinger B, Morin CM, Espie C, Chung F, Penzel T, et al. Sleep and daytime problems during the COVID-19 pandemic and effects of coronavirus infection, confinement and financial suffering: a multinational survey using a harmonised questionnaire. BMJ Open. 13 déc 2021;11(12):e050672.

39. Fisman D. Seasonality of viral infections: mechanisms and unknowns. Clin Microbiol Infect [Internet]. oct 2012 [cité 12 oct 2021];18(10):946-54. Disponible sur: https://linkinghub.elsevier.com/retrieve/pii/S1198743X14610910

40. Moriyama M, Hugentobler WJ, Iwasaki A. Seasonality of Respiratory Viral Infections. Annu Rev Virol [Internet]. 29 sept 2020 [cité 17 mars 2022];7(1):83-101. Disponible sur: https://www.annualreviews.org/doi/10.1146/annurev-virology-012420-022445

41. Carmona CP, Szava-Kovats R, Partel M. Estimating probabilistic dark diversity based on the hypergeometric distribution [Internet]. Ecology; 2019 mai [cité 29 déc 2019]. Disponible sur: http://biorxiv.org/lookup/doi/10.1101/636753

42. Naumova EN, Christodouleas J, Hunter PR, Syed Q. Effect of precipitation on seasonal variability in cryptosporidiosis recorded by the North West England surveillance system in 1990-1999. J Water Health. juin 2005;3(2):185-96.

43. Warren-Gash C, Bhaskaran K, Hayward A, Leung GM, Lo SV, Wong CM, et al. Circulating Influenza Virus, Climatic Factors, and Acute Myocardial Infarction: A Time Series Study in England and Wales and Hong Kong. J Infect Dis [Internet]. 15 juin 2011 [cité 28 avr 2021];203(12):1710-8. Disponible sur: https://academic.oup.com/jid/article-lookup/doi/10.1093/infdis/jir171

44. Thomson RM, Furuya-Kanamori L, Coffey C, Bell SC, Knibbs LD, Lau CL. Influence of climate variables on the rising incidence of nontuberculous mycobacterial (NTM) infections in Queensland, Australia 2001–2016. Sci Total Environ [Internet]. oct 2020 [cité 28 avr 2021];740:139796. Disponible sur: https://linkinghub.elsevier.com/retrieve/pii/S0048969720333167

45. Carmona P, Gandon S. Winter is coming: Pathogen emergence in seasonal environments. PLOS Comput Biol [Internet]. 6 juill 2020 [cité 23 mars 2022];16(7):e1007954. Disponible sur: https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1007954

ANNEXE

Article 6

Article 6 : Epidemiological surveillance of respiratory viral infections at IHU Méditerranée Infection and its application to SARS-CoV-2. DOI : https://doi.org/10.35088/gpgh-wq98. In the process of submission (2021)

Philippe COLSON, Audrey GIRAUD-GATINEAU, Pierre-Edouard, FOURNIER, Laetitia NINOVE, Christine ZANDOTTI, Marie-Thérèse JIMENO, Céline BOSCHI, Léa LUCIANI, **Lanceï KABA**, Philippe PAROLA, Stéphane RANQUE, Jean-Marc ROLAIN, Philippe GAUTRET, Michel DRANCOURT, Jean-Christophe LAGIER, Bernard LA SCOLA, Hervé CHAUDET, Didier RAOULT

TITLE PAGE

Full-length title:

Epidemiological surveillance of respiratory viral infections at IHU Méditerranée Infection and its application to SARS-CoV-2

Short title (for the running head): Respiratory viral infections at IHU Méditerranée Infection

Author list: Philippe COLSON[1,2], Audrey GIRAUD-GATINEAU[1,3,4], Pierre-Edouard FOURNIER[1,2], Laetitia NINOVE[1], Christine ZANDOTTI[1], Marie-Thérèse JIMENO[1,5], Céline BOSCHI[1,2], Léa LUCIANI[1,2], Lancei KABA[1,2,3], Philippe PAROLA[1,3], Stéphane RANQUE[1,3], Jean-Marc ROLAIN[1,2], Philippe GAUTRET[1,2], Michel DRANCOURT[1,2], Jean-Christophe LAGIER[1,2], Bernard LA SCOLA[1,2], Hervé CHAUDET[1,3,4], Didier RAOULT [1,2]*

Affiliations: [1] IHU Méditerranée Infection, 19-21 boulevard Jean Moulin, 13005 Marseille, France; [2] Aix-Marseille Univ., Institut de Recherche pour le Développement (IRD), Assistance Publique - Hôpitaux de Marseille (AP-HM), Microbes Evolution Phylogeny and Infections (MEPHI), 27 boulevard Jean Moulin, 13005 Marseille, France; [3] Aix Marseille Univ, Institut de Recherche pour le Développement (IRD), Assistance Publique - Hôpitaux de Marseille (AP-HM), Service de Santé des Armées (SSA), Vecteurs - Infections Tropicales et Méditerranéennes (VITROME), Marseille, France; [4] French Armed Forces Center for Epidemiology and Public Health (CESPA), Service de Santé des Armées (SSA), Camp de Sainte Marthe, 408 rue Jean Queillau, 13014 Marseille, France; [5] Service de l'Information Médicale, Assistance Publique - Hôpitaux de Marseille (AP-HM), Hôpital Timone, 234 rue Saint-Pierre, 13385 Marseille, France.

*** Corresponding author:** Didier Raoult, IHU - Méditerranée Infection, 19-21 boulevard Jean Moulin, 13005 Marseille, France. Tel.: +33 413 732 401, Fax: +33 413 732 402; email:

didier.raoult@gmail.com

Keywords (6): Respiratory infections; viral infections; surveillance; epidemiology; clinical laboratory; diagnosis; SARS-CoV-2

Word counts: abstract, 261; text, 3,339

Figures: 4; **Table:** 0; **References:** 40

ABSTRACT

Epidemiological surveillance of infections at IHU Méditerranée Infection is based on in-house systems that use data from our microbiology-virology laboratory and continuously expand and evolve. Until 2020, respiratory samples were the third most frequent clinical samples sent to our laboratory. In 2019 we received ≈18,000 respiratory samples to search for bacteria and fungi and 17,600 to search for viruses. Over the 2015-2019 5-year period, we diagnosed >26,000 infections with respiratory viruses. The onset of the SARS-CoV-2 pandemic has dramatically boosted the number of tests and diagnoses of viral respiratory infections. On December 31st, over 339 days of daily surveillance, 427,787 SARS-CoV-2 tests had been performed for 306,363 patients. The mean number of daily tests was 1.262±930 (range, 8-3,596) and that of new patients tested was 904±688 (7-2,835). A total of 26,327 patients were diagnosed positive, the mean daily number being 78±94 (0-416), corresponding to a rate of new positive patients of 8.6% (mean: 6.1±5.4% (0-25.9%)). We first diagnosed SARS-CoV-2 on February 27th. The number of cases then peaked on March 26th (n= 362), was on average 2.5 between May 9th and July 5th, and increased and peaked again on October 26th (n= 416). Our surveillance strategy allowed observing SARS-CoV-2 temporal and age distributions and coinfections with other respiratory viruses. Data accumulated using and improving our existing tools show that comprehensive real-time surveillance of emerging infections is essential. Indeed it allows observing their epidemiological characteristics that cannot be predicted or extrapolated from other infections as some are new and unexpected and whose timely knowledge is valuable for optimal biological and clinical managements.

TEXT

Principle of the surveillance of infections at IHU Méditerranée Infection

Surveillance of infections has been implemented in our microbiology and virology laboratory since 2003 [1, 2]. It follows the recommendations made in a report on bio-terrorism and infectious diseases by one of us (DR) [3]. This report recommended in particular to implement a surveillance of abnormal events, without *a priori*, including syndromic surveillance, and of mortality. Our laboratory is the only one carrying out microbiology-virology diagnoses for all public and university hospitals (Assistance Publique des Hôpitaux de Marseille (AP-HM)) of Marseille, the second largest city in France with around 860,000 inhabitants (https://www.insee.fr/fr/statistiques/1405599?geo=COM-13055). It performs the diagnoses of all infections including those related to bacterial, fungal, parasitic and viral pathogens. Our syndromic surveillance strategy consists in counting on a weekly basis the number of samples received, classified by their nature, as well as the number of tests carried out, these two elements being situated upstream of the positive diagnosis of infections [4]. This surveillance is supplemented with a “traditional” surveillance corresponding to the follow-up of positive diagnoses for all the microbial and viral pathogens. Since 2003, we have thus followed a “roadmap” leading to monitoring abnormal events related to infections, and this monitoring has adapted from a technical point of view, and to our environment which has been modified over time. In 2012, the creation of the IHU Méditerranée Infection (IHU-MI) made it possible to professionalize surveillance tools with the establishment of a dedicated IT platform (MIDAS), and we were joined on this occasion by a team of epidemiologists of the military health service [2]. In addition, the principle of surveillance based on data from the microbiology-virology laboratory has been extended to the southeastern region of France (Provence-Alpes-Côte d'Azur region, or South region) which includes ≈7% of the population

of metropolitan France. A collaborative network called PACASurvE has been in place since 2013 and the majority of hospitals (n= 17) and around half of private medical biology analysis laboratories (n= 285) participate [5, 6] and up to 386 when considering specialized medical biology analyses [7]. Data from our surveillance systems are examined weekly. Alarms are triggered automatically in the event of an abnormal increase in the number of samples, tests, or positive diagnoses. These events may lead to additional investigations, studies and reports [2]. In addition, since 2014, weekly monitoring of deaths at AP-HM has been integrated into the monitoring [2, 8, 9]. It can detect the infections most frequently associated with death.

Respiratory samples and diagnosis of respiratory infections

Until 2020, respiratory samples were the third most frequent type of sample among those sent to our laboratory, after urine samples and blood cultures. In 2019 we received ≈18,000 respiratory samples to search for bacteria and fungi and 17,600 to search for viruses. Regarding the search for bacteria, during a 63-month period from February 1st 2014 to April 25, 2019, *Staphylococcus aureus*, *Pseudomonas aeruginosa* and *Klebsiella pneumoniae* were the most frequently isolated bacteria from respiratory samples in 6,189, 3,190, and 973 cases, respectively (≈1,180, 610 and 185/year, respectively). There were ≈11,700 searches for mycobateria and 160 positive diagnoses/year (in 2019). Regarding fungi, during a 40-month period from June 1st 2017 to October 31th 2020, 15,976 respiratory samples (≈4,800 /year) from 12,032 patients were analyzed, and ≥1 fungus was isolated from 1,636 (10%) of them (≈490/year). The most frequent species were *Candida albicans* (54%), followed by *C. glabrata* (6%), *C. tropicalis* (6%), and *Aspergillus fumigatus* (6%). Overall, regarding microbiology, 8 of the 15 agents appearing in the top 15 of the most frequently diagnosed microbial agents are strictly or possibly respiratory pathogens: *S. aureus, C. albicans*, *K. pneumoniae*, *P. aeruginosa*, *Haemophilus influenzae*, *Streptococcus pneumoniae*,

Streptococcus pyogenes, and *C. glabrata*.

Viral respiratory infections

Respiratory viruses are an important part of infectious agents in our clinical microbiology-virology laboratory, and the most frequently diagnosed agents of respiratory tract infections. Over the 2015-2019 5-year period, we diagnosed 7,412 influenza A viruses; 2,882 influenza B viruses; 6,754 rhinoviruses; 4,851 respiratory syncytial viruses (RSV); 1,617 metapneumoviruses; 1,239 adenoviruses; 763 infections with human parainfluenza viruses 1 to 4; 480 enteroviruses; and 469 infections with the four seasonal human coronaviruses (HCoV) (229E, NL63, OC43 and HKU1) (**Figure 1**). Since 2010, viral respiratory infection diagnoses have been carried out mainly by real-time PCR (qPCR), based on in-house or commercial simplex or multiplex tests. The numbers of direct diagnoses of respiratory viruses were only exceeded or competed by *Escherichia coli* (5,800 in 2019) and *K. pneumoniae* (1,400) in urines (49,000), and by coagulase negative staphylococci (1,200) in blood cultures (49,000).

An important point in the surveillance of respiratory infections (as for that of other infections) is their unpredictability [10]. This can be observed including for viruses for which we have numerous data such as influenza viruses. Thus, if we consider the PCR diagnoses of influenza infections in our clinical microbiology-virology laboratory during the winters from 2010-2011 to 2019-2020, we observe important variations from year to year of the time of the emergence of the winter epidemic, of its duration, of the level of incidence reached at the epidemic peak, of the period during which this peak is reached, and of the viral types (A and B) and subtypes (H3N2, H1N1) predominant throughout the epidemic period (**Figure 1**). This unpredictability makes surveillance of considerable interest. The unpredictability of respiratory viral infections also applies to the 4 HPIV types circulating in humans as it turns out that these do not have the same seasonality. Thus, HPIV-3 circulates mainly during spring

while HPIV-4 shows peaks of incidence from September to November and between February and March [11, 12]. The epidemic curves of the 4 seasonal HCoV are also not fully superimposed [13-18]. This shows the value of monitoring these viruses separately, which we have performed more comprehensively since 2019. Another point regarding respiratory viral infections is the need for an accurate diagnosis. Attributing cases of respiratory infections to a given virus without documentation by a diagnosis can lead to a very imperfect knowledge of the causes and the epidemiology of these infections [19-21]. In addition, we were able to observe among our diagnoses associations between microbial and viral pathogens in respiratory samples, and their interactions are being investigated in our institute [22, 23]. Finally, the monitoring of the weekly numbers of respiratory samples is a very useful element in addition to that of the diagnoses, since it can lead to earlier alerts triggered by increases in respiratory samples compared to alerts based on positive diagnoses [1].

Another important component of viral respiratory disease surveillance is mortality surveillance (https://www.mediterranee-infection.com/le-global-burden-of-infections-des-hopital-publics-de-marseille-and-the-region-provence-alpes-cote-dazur/). It allows us to observe among the most frequently diagnosed respiratory agents those most frequently associated with death. We do not speculate on the imputability of these infectious agents in the death, but we observe associations between these agents and deaths. A preliminary study carried out between February 1st, 2014 and April 25th, 2019, covering 63 months, measured that among 347,877 patients hospitalized at the AP-HM, 15,235 had died and for 62% of them, i.e. 9,480 patients, ≥1 clinical sample had been sent to our microbiology-virology laboratory. This corresponds to an average of 35 deaths/week for which we analyzed ≥1 clinical sample. We found as agents most frequently associated with death pathogens frequently involved in respiratory infections such as *S. aureus*, *K. pneumoniae*, *P. aeruginosa* as well as *Candida* spp.. However, there were also per year, associated with deaths, around 25

diagnoses of influenza A virus, 32 of influenza B virus, 18 of rhinovirus, and 9 of RSV. In another study that analyzed mortality between weeks 47 and 14 during the winters of 2018, 2019 and 2019-2020, we found that 0.4% of diagnoses of influenza viruses (10 for 2,815 cases), 1.0% of those of rhinoviruses (15 for 1,565 cases), and 1.5% of those of a seasonal HCoV (9 for 615 cases) were in patients who had died [9]. The diagnoses of respiratory infections among travelers are also monitored in our institute [24, 25], in particular those made on return from the Hajj pilgrimage for which for instance the acquisition rates of rhinovirus/enterovirus, HCoV-229E and influenza A virus were determined to be 39%, 20% and 2%, respectively [26]. Our infection surveillance also covers specific populations such as homeless people [27].

Surveillance of SARS-CoV-2 infections

Surveillance of SARS-CoV-2 qPCR tests and positive diagnoses

Our surveillance of respiratory viral infections has been completed and adapted following the emergence of SARS-CoV-2 that has introduced a scale change regarding the number of tests and diagnoses of viral respiratory infections. In fact, while we performed a maximum of between ≈400 and 1,000 tests/week during 2010-2011 and 2018-2019 winters, we received up to ≈20,000 respiratory samples/week during year 2020 (**Figures 2, 3**). Our surveillance went from a weekly rhythm to a daily rhythm. We have set up molecular tests to diagnose infections with this virus as quickly as possible. So three days after the release of the first viral genome (on January 10th, 2020) we had designed and ordered in-house real-time reverse transcription PCR (qPCR) systems. We subsequently used an internationally-validated qPCR system from the virology laboratory of Charité Hospital in Berlin [28]. We tested our SARS-CoV-2 detection tests on January 25th and performed the first test for a patient whose sample was referred to our institute on January 29th. At the beginning of February, we carried out 674

SARS-CoV-2 qPCR tests for 337 people repatriated from China to France [29], and retrospectively tested 137 patients who had died with a respiratory infection between 2018 and 2019, 135 medical students returning in 2018-2019 from Asia, and 144 people in whom a respiratory sample had been collected in Senegal between March 2019 and February 2020; all these tests being negative. The first 280 patients tested between January 29th and March 1st (210 and 60 on return from Italy and Asia, respectively) were negative for SARS-CoV-2, but other respiratory viruses were identified in 49% of the cases (n= 137) [21]. The most frequent viruses detected were influenza A virus (12%); rhinovirus/enterovirus (12%); common HCoV (229E, OC43, NL63 and HKU1 in 1%, 1%, 4% and 7%, respectively); influenza B virus (8%); metapneumovirus (7%); RSV (2%); and adenovirus (1%). In addition, 12 patients (4%) were coinfected with different respiratory viruses, most often with rhinovirus/enterovirus and metapneumovirus.

The first positive SARS-CoV-2 qPCR result was obtained on February 27 th for a patient hospitalized at Nice University Hospital located in the Provence Alpes Côte d'Azur region (PACA; Southern France), since we were at this time the only center in this French region to perform SARS-CoV-2 testing. The first SARS-CoV-2-positive patient hospitalized in Marseille in our institute was detected on March 2nd, after we had routinely performed 4,149 SARS-CoV-2 qPCR tests for 3,417 symptomatic or asymptomatic patients. The surveillance of SARS-CoV-2 infections has been accompanied by daily reports on the IHU Méditerranée Infection website since March 26th in the form of a "Southern France Morning Post" posted every day (https: //www.mediterranee-infection.com/covid-19/). This information available to everyone includes the total number of samples received at the laboratory and of tests performed, the number of positive tests and the percentage of positives. This also includes the numbers of tests and positives for the newly-tested patients, for symptomatic and asymptomatic patients, for patients sampled at IHU MI or at AP-HM, and

still more precisely for patients residing in our department of Bouches-du-Rhône, and in the city of Marseille. We have indeed received a large number of samples from other hospitals and laboratories in the Provence Alpes Côte d'Azur region and also carried out tests for patients domiciled in other French regions (who may have traveled to Marseille to be tested), mainly Auvergne Rhône-Alpes region and Ile-de-France region, particularly Paris (as of July 7th: n= 470, 547 and 235, respectively). As already available for other infectious agents on our intranet platform for the epidemiological surveillance of infections (MIDAS), surveillance charts for SARS-CoV-2 infections were added specifically for SARS-CoV-2, and separately for each French department and each arrondissement of the city of Marseille.

The surveillance of SARS-CoV-2 infections carried out in our institute has been optimized by the testing strategy that has been implemented there. The tests were thus carried out, from the beginning and until now, for all patients regardless of whether they were symptomatic or not, contact-cases or not, and with a medical prescription or not. All these tests were performed by qPCR on nasopharyngeal swabs, the only diagnostic approach that has been used in our institute. In fact, our evaluation of a recommended antigen test on nasopharyngeal swabs from 204 qPCR-positive patients (including 182 (89%) symptomatic) showed a high false-negative rate (21% in symptomatic patients, and 45% in asymptomatic) and positive and negative predictive values of 96% and 72%, respectively [30]. We used various qPCR assays, including in-house techniques in microplates [21, 29] and later during the year a commercial reagent for microplate assays as well as commercial simplex (n= 3) or multiplex (2) tests, including some evaluated in our laboratory [31]. In addition, we tested 7 alternative qPCR systems as backup tests in case of genetic evolution of the viruses that would generate mismatches of primers and/or PCR probes possibly. As of December 31st, 2020, over a period of 339 days, 427,787 tests had been performed for 306,363 patients. The mean number of daily tests was 1.262 (standard deviation, 930; range: 8-3,596; median= 979)

and that of new patients tested was 904±688 (7-2,35; median= 693). A total of 26,327 patients were diagnosed positive, the mean daily number being 78±94 (0-416; median= 35), corresponding to a rate of new positive patients of 8.6% (mean: 6.1±5.4% (0-25.9; median= 5.6%). We observed different phases between February and December. Indeed, the daily number of SARS-CoV-2 diagnoses peaked on March 26th (n= 362), dramatically decreased in May with a mean of 2.5/day for 58 days between May 9th and July 5th (https://www.mediterranee-infection.com/covid-19/). Then, incidence re-increased from early July and peaked again on October 26th (n= 416) before a new drop (**Figure 2**).

Epidemiological features of SARS-CoV-2 infections and associated respiratory viruses

Over the year 2020 and from the first days of the emergence of the SARS-CoV-2 epidemic in our region, we have carried out studies relating to the surveillance of infections by this virus and relying on the observation of our laboratory data. First, we compared the temporal distribution of infections by this emerging coronavirus with that of the 4 other seasonal human coronaviruses [32]. At the end of May, for each of these five coronaviruses, we observed a bell-shaped curve with a lag of a few weeks, SARS-CoV-2 having occurred later than the seasonal HCoVs. These data suggested that the epidemic curve of SARS-CoV-2 may be very similar to that of common HCoV and to that of some other respiratory viruses (**Figure 4**). A second element observed was the age distribution of SARS-CoV-2 infections. A study carried out early until March 14th demonstrated a low proportion of cases in children (0 between 0-1 year, 3 (1%) between 1-5 years and 7 (4%) between 5-10 years), significantly lower than in adults [32]. These results were verified in a larger study carried out on the first 302 pediatric cases (<18 years of age) diagnosed at Marseille university hospital on April 15th, which showed that they corresponded to 5% of the positive patients (n= 5,861) and included 107 (2% of all positive patients) and 70 (1%) children under 10 and 6 years of age,

255 respectively [33]. All these infected children clinically recovered. If we compare these data to
256 those for seasonal HCoVs, we observe that children are spared only by SARS-CoV-2 while
257 they are the age group mainly affected by the four season-endemic coronaviruses [32]. These
258 data showed early in the pandemic that the epidemiology of SARS-CoV-2 could not be
259 predicted based on prior knowledge of other coronavirus infections, nor on that of other
260 respiratory infections such as influenza virus infections. Indeed, other respiratory viral
261 infections affect children extensively, especially the youngest of them [33].
262 Our ability to diagnose all infectious pathogens in the same laboratory, including
263 bacterial, fungal, parasitic and viral pathogens, allows us to analyze possible coinfections.
264 This is another element of respiratory infection surveillance. The multiplex PCR diagnostic
265 approach has been developed since 10 years ago in our laboratory through our POC
266 laboratories [34, 35] but also our core laboratory. It has recently expanded with technical
267 progress and the increasing availability of commercial multiplex tests with rapid results [19,
268 21, 36]. Regarding respiratory infections it is tricky to clinically narrow down a differential
269 diagnosis to a single one due to the significant overlap in the clinical presentations. Multiplex
270 PCR diagnosis allows a more exhaustive coverage of respiratory viruses, as we have shown
271 for example in the context of the first research of SARS-CoV-2 infections in our laboratory
272 [21], and for the diagnosis of respiratory viral coinfections [19]. We thus studied coinfections
273 with SARS-CoV-2 and other respiratory viruses among 4,222 patients during March and
274 April 2020 [19]. A total of 643 patients (15%) were diagnosed with SARS-CoV-2, 1,095
275 (26%) were diagnosed with ≥1 non-SARS-CoV-2 respiratory viruses, and 27 (4% of those
276 SARS-CoV-2-positive) were coinfected with SARS-CoV-2 and another respiratory virus,
277 including a rhinovirus (n= 11), an endemic coronavirus (HCoV-OC43 (2), HCoV-HKU1 (2),
278 HCoV-229E (1)), influenza viruses A (2) or B (2), HPIV 4 (2) and 2 (1), bocavirus (2), and
279 adenovirus (1). The number of coinfections with SARS-CoV-2 and other respiratory viruses

decreased by 3.5 times between March and April while the number of infections with non-SARS-CoV-2 respiratory viruses decreased by 18 times between these 2 months, indicating that the frequency of such coinfections largely depends on the rate of coincidence of these viruses. The surveillance of these coinfections for the more recent period between August and November thus revealed a frequency of coinfections of 0.3% with 46 cases involving overwhelmingly, in 37 cases, rhinoviruses that circulate along the whole year. Interestingly, over the recent period from November to December 2020, rhinoviruses (426 diagnoses (13% of tests)) and adenoviruses (140 (3%)) have been detected, but the incidence rates of infections with other respiratory viruses (apart from SARS-CoV-2) were unexpectedly very low, and lower than those observed during this 2 month-period during the 10 previous years. Thus only two diagnoses of RSV infection, one diagnosis of metapneumovirus infection, and no influenza virus infection were detected in 2020 vs. on average 433±131, 67±44, and 113±134 diagnoses, respectively, during the years 2010 to 2019. The daily surveillance of the number of samples and diagnoses of SARS-CoV-2 infections has been accompanied by other surveillance needs. It included the implementation of the genomic epidemiological surveillance that can reveal some aspects of the SARS-CoV-2 infection that cannot be deduced from the mere observation of the numbers of cases. This surveillance showed that several epidemics have occurred since July that involved different SARS-CoV-2 variants [37-39]. Finally, our daily monitoring of positive diagnoses allowed us to detect SARS-CoV-2 reinfections. We observed that among the 6,799 patients diagnosed positive between February and May, 837 had been retested since June and 15 patients had been found positive again for SARS-CoV-2 >2 months after viral clearance following the first infection. Viral genome sequencing made it possible to demonstrate reinfection with a virus of a different genotype compared to that of the first episode [40].

Conclusion

In total, in the context of the SARS-CoV-2 pandemic, we have completed and adapted our epidemiological surveillance of respiratory viral infections by relying on the strategies and versatile tools that pre-existed in IHU MI. Data accumulated in 2020 show that it is essential to perform a real-time surveillance of emerging infections to be able to observe all their epidemiological characteristics, whose timely knowledge is useful for an optimal biological and clinical management of the cases. These characteristics cannot be predicted or extrapolated from other infections with similar agents since some of them are new and unexpected. Our surveillance strategy, combined with the strategy of massive SARS-CoV-2 screening conducted from February at IHU-MI, allowed us to be the first to observe and communicate on several features of the infections with this emerging virus.

Acknowledgments

This manuscript has been edited by a native English speaker.

Author contributions

Conceived and designed the experiments: PC, HC, DR. Contributed materials/analysis tools: all authors. Analyzed the data: all authors. Wrote the paper: PC and DR.

Funding

This work was supported by the French Government under the “Investments for the Future” program managed by the National Agency for Research (ANR), Méditerranée-Infection 10-IAHU-03 and was also supported by Région Provence Alpes Côte d’Azur and FEDER PRIMMI European funding (Fonds Européen de Développement Régional-Plateformes de

Recherche et d'Innovation Mutualisées Méditerranée Infection), FEDER PA 0000320 PRIMMI.

Conflicts of interest

The authors have no conflicts of interest to declare. Funding sources had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; and preparation, review, or approval of the manuscript.

Ethics

All the data have been generated as part of the routine work at Assistance Publique-Hôpitaux de Marseille (Marseille university hospitals), and this study results from routine standard clinical management. This study has been approved by the ethics committee of our institution (N°2020-029). Access to the patients' biological and registry data issued from the hospital information system was approved by the data protection committee of Assistance Publique-Hôpitaux de Marseille (APHM) and was recorded in the European General Data Protection Regulation registry under number RGPD/APHM 2019-73.

REFERENCES

1. Colson P, Rolain JM, Abat C, Charrel R, Fournier PE, Raoult D. EPIMIC: A Simple Homemade Computer Program for Real-Time EPIdemiological Surveillance and Alert Based on MICrobiological Data. PLoS One **2015**; 10:e0144178.

2. Abat C, Chaudet H, Colson P, Rolain JM, Raoult D. Real-Time Microbiology Laboratory Surveillance System to Detect Abnormal Events and Emerging Infections, Marseille, France. Emerg Infect Dis **2015**; 21:1302-10.

3. Raoult D. Rapport de mission. http://ifr48 timone univ-mrs fr/files/Documents-Raoult/bioterrorisme2003 pdf 2003Available from: URL: http://ifr48.timone.univ-mrs.fr/files/Documents-Raoult/bioterrorisme2003.pdf.

4. Henning KJ. Overview of Syndromic Surveillance What is Syndromic Surveillance? MMWR. Morbidity and mortality weekly report **2004**; 53 Suppl(Suppl):5-11.

5. Huart M, Bedubourg G, Abat C, et al. Implementation and Initial Analysis of a Laboratory-Based Weekly Biosurveillance System, Provence-Alpes-Cote d'Azur, France. Emerg Infect Dis **2017**; 23:582-9.

6. Diallo OO, Baron SA, Dubourg G, et al. Major discrepancy between factual antibiotic resistance and consumption in South of France: analysis of 539,037 bacterial strains. Sci Rep **2020**; 10:18262.

7. Colson P, Poveda JD, Trombert Paolantoni S, et al. Weekly surveillance of bacterial, viral and parasitic infections involving private and public medical analysis laboratories through 317833 diagnostic tests in the Provence-Alpes-Côte-d'Azur region, 2014-2019. 30th European Congress of Clinical Microbiology & Infectious Diseases will take place in Paris, France, **2020**; abstr. 8607.

8. Abat C, Rolain JM, Dubourg G, Fournier PE, Chaudet H, Raoult D. Evaluating the Clinical Burden and Mortality Attributable to Antibiotic Resistance: The Disparity of Empirical Data and Simple Model Estimations. Clin Infect Dis **2017**; 65:S58-S63.

9. Giraud-Gatineau A, Colson P, Jimeno MT, et al. Comparison of mortality associated with respiratory viral infections between December 2019 and March 2020 with that of the previous year in Southeastern France. Int J Infect Dis **2020**; 96:154-6.

10. Raoult D. Molecular, epidemiological, and clinical complexities of predicting patterns of infectious diseases. Front Microbiol **2011**; 2:25.

11. Boschi C, Giraud-Gatineau A, Petit P, et al. Infections à virus parainfluenza 3 diagnostiquées dans un centre hospitalo-universitaire au cours des 3 dernières années : à propos de 366 cas. 21ème journées nationales d'infectiologie (JNI), Poitiers, France, **2020**; abstr. RESP-09.

12. Petit P, Boschi C, Zandotti C, Ninove L, La Scola B, Aherfi S. Infections par le virus parainfluenza 4 : une série de 84patients positifs. 21ème journées nationales d'infectiologie (JNI), Poitiers, France, **2020**; abstr. RESP-05.

13. Andreani J, Boschi C, Raoult D, Colson P, La Scola B. Épidémiologie des coronavirus 229E dans les infections respiratoires. 21ème journées nationales d'infectiologie (JNI), Poitiers, France, **2020**; abstr. RESP-03.

14. Boschi C, Dambo M, Aubrey C, Chaudet H, Zandotti C, Parola P. Caractéristiques épidémiologiques et cliniques de 136 infections à coronavirus OC43 diagnostiquées dans des hôpitaux universitaires de 2017 à 2019. 21ème journées nationales d'infectiologie (JNI), Poitiers, France, **2020**; abstr. RESP-08.

15. Dambo M, Aheri S, Ninove L, Zandotti C, La Scola B. Épidémiologie des infections à coronavirus NL63 : à propos d'une série de 64 patients. 21ème journées nationales d'infectiologie (JNI), Poitiers, France, **2020**; abstr. RESP-04.

16. Ninove L, Zandotti C, Dambo M, Colson P, Charrel RN, Nougairede A. Épidémiologie des infections à coronavirus HCoV-HKU1 à Marseille, France. 21ème journées nationales d'infectiologie (JNI), Poitiers, France, **2020**; abstr. RESP-08.

17. Nickbakhsh S, Ho A, Marques DFP, McMenamin J, Gunson RN, Murcia PR. Epidemiology of seasonal coronaviruses: Establishing the context for COVID-19 emergence. J Infect Dis **2020**.

18. Li Y, Wang X, Nair H. Global Seasonality of Human Seasonal Coronaviruses: A Clue for Postpandemic Circulating Season of Severe Acute Respiratory Syndrome Coronavirus 2? J Infect Dis **2020**; 222:1090-7.

19. Boschi C, Hoang VT, Giraud-Gatineau A, et al. Co-infections with SARS-CoV-2 and other respiratory viruses in Southeastern France: a matter of sampling time. J Med Virol **2020** Nov 24;10.1002/jmv.26692. doi: 10.1002/jmv.26692. Online ahead of print.

20. Follin P, Lindqvist A, Nystrom K, Lindh M. A variety of respiratory viruses found in symptomatic travellers returning from countries with ongoing spread of the new influenza A(H1N1)v virus strain. Euro Surveill **2009**; 14:19242.

21. Amrane S, Tissot-Dupont H, Doudier B, et al. Rapid viral diagnosis and ambulatory management of suspected COVID-19 cases presenting at the infectious diseases referral hospital in Marseille, France, - January 31st to March 1st, 2020: a respiratory virus snapshot. **2020**; 36:101632.

22. Parola P, Colson P, Dubourg G, et al. Letter to the editor. Group A streptococcal infections during the seasonal influenza outbreak 2010/11 in South East England. Euro Surveill **2011**; 16:19816.

23. Edouard S, Million M, Bachar D, et al. The nasopharyngeal microbiota in patients with viral respiratory tract infections is enriched in bacterial pathogens. Eur J Clin Microbiol Infect Dis **2018**; 37:1725-33.

24. Hoang VT, Ali-Salem S, Belhouchat K, et al. Respiratory tract infections among French Hajj pilgrims from 2014 to 2017. Sci Rep **2019**; 9:17771.

25. Dao TL, Canard N, Hoang VT, et al. Risk factors for symptoms of infection and microbial carriage among French medical students abroad. Int J Infect Dis **2020**; 100:104-11.

26. Hoang VT, Sow D, Dogue F, et al. Acquisition of respiratory viruses and presence of respiratory symptoms in French pilgrims during the 2016 Hajj: A prospective cohort study. Travel Med Infect Dis **2019**; 30:32-8.

27. Ly TDA, Dao TL, Hoang VT, et al. Pattern of infections in French and migrant homeless hospitalised at Marseille infectious disease units, France: A retrospective study, 2017-2018. Travel Med Infect Dis **2020**; 36:101768.

28. Corman VM, Landt O, Kaiser M, et al. Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. Euro Surveill **2020**; 25:10-7917.

29. Lagier JC, Colson P, Tissot-Dupont H, et al. Testing the repatriated for SARS-Cov2 : should laboratory-based quarantine replace traditional quarantine? **2020**; 34:101624.

30. Fenollar F, Bouam A, Ballouche M, et al. Evaluation of the Panbio Covid-19 rapid antigen detection test device for the screening of patients with Covid-19. J Clin Microbiol **2020** Nov 2;JCM.02589-20. doi: 10.1128/JCM.02589-20. Online ahead of print.

31. Fournier PE, Zandotti C, Ninove L, et al. Contribution of VitaPCR SARS-CoV-2 to the emergency diagnosis of COVID-19. J Clin Virol **2020**; 133:104682.

32. Colson P, Esteves-Vieira V, Giraud-Gatineau A, et al. Temporal and age distributions of SARS-CoV-2 and other coronaviruses, Southeastern France. Int J Infect Dis **2020**; 101:121-5.

33. Morand A, Matteudi T, Fabre A, Minodier P, Bosdure E, Luciani L. Open screening of SARS-CoV-2 infections in the pediatric population in Marseille, Southern France. IHU preprint **2020**; https://www.mediterranee-infection.com/ropen-screening-of-sars-cov-2-infections-in-the-pediatric-population-in-marseille-southern-france/.

34. Cohen-Bacrie S, Ninove L, Nougairede A, et al. Revolutionizing clinical microbiology laboratory organization in hospitals with in situ point-of-care. PLoS One **2011**; 6:e22403.

35. Drancourt M, Michel-Lepage A, Boyer S, Raoult D. The Point-of-Care Laboratory in Clinical Microbiology. Clin Microbiol Rev **2016**; 29:429-47.

36. Ramanan P, Bryson AL, Binnicker MJ, Pritt BS, Patel R. Syndromic Panel-Based Testing in Clinical Microbiology. Clin Microbiol Rev **2017**; 31:e00024-17.

37. Levasseur A, Delerce J, Caputo A, et al. Genomic diversity and evolution of coronavirus (SARS-CoV-2) in France from 309 COVID-19-infected patients. bioRxiv **2020**; doi: https://doi.org/10.1101/2020.09.04.282616.

38. Colson P, Levasseur A, Delerce J, et al. Dramatic increase in the SARS-CoV-2 mutation rate and low mortality rate during the second epidemic in summer in Marseille. IHU preprint **2020**; doi: https://doi.org/10.35088/68c3-ew82.

39. Colson P, Levasseur A, Gautret P, et al. Introduction into the Marseille geographical area of a mild SARS-CoV-2 variant originating from sub-Saharan Africa. Accepted in Travel Med Infect Dis **2020** Nov 15;S0163-4453(20)30706-4. doi: 10.1016/j.jinf.2020.11.011. Online ahead of print.

40. Colson P, Finaud M, Levy N, Lagier JC, Raoult D. Evidence of SARS-CoV-2 re-infection with a different genotype. J Infect **2020**.

FIGURE LEGENDS

Figure 1. Number of diagnoses by qPCR of respiratory viruses during the period from 2010 to 2019

Figure 2. Weekly number of diagnoses by qPCR of respiratory viruses in 2020

Figure 3. Weekly number of respiratory samples sent to our laboratory to test for viruses by qPCR during the period from 2010 to 2020

Figure 4. Weekly numbers of diagnoses by qPCR of respiratory viruses
HCoV, human common coronavirus; HPIV, human parainfluenza virus; RSV, respiratory syncytial virus

Article 7

Article 7: A Non-Invasive Neonatal Signature Predicts Later Development of Atopic Diseases. *J. Clin. Med.* **2022**, *11*, 2749. https://doi.org/10.3390/jcm11102749

Youssouf, S.; Moise, M.; Soraya, M.; Cheick, G.O.; **Lanceï, K**.; Ghiles, G.; Thibault, M.; Jean-Louis, M.; Tu Anh, T.; Pierre, C.; Anne, F.; Joana, V.

Journal of
Clinical Medicine

MDPI

Article

A Non-Invasive Neonatal Signature Predicts Later Development of Atopic Diseases

Youssouf Sereme [1,2], Moïse Michel [1,2,3], Soraya Mezouar [1,2], Cheick Oumar Guindo [1,2], Lanceï Kaba [1,4], Ghiles Grine [1,2,5], Thibault Mura [6,7], Jean-Louis Mège [1,2], Tu Anh Tran [8,9,10], Pierre Corbeau [3,10,11,*,†], Anne Filleron [8,9,10,*,†] and Joana Vitte [1,2,10,12,*,†]

1 IHU Méditerranée Infection, 13005 Marseille, France; seremeyoussouf@yahoo.fr (Y.S.); moise0michel@gmail.com (M.M.); soraya.mezouar@univ-amu.fr (S.M.); cheicko86@gmail.com (C.O.G.); lancekaba@yahoo.fr (L.K.); grineghiles@gmail.com (G.G.); jean-louis.mege@univ-amu.fr (J.-L.M.)
2 IRD, APHM, MEPHI, Aix-Marseille Université, 13284 Marseille, France
3 Immunology Department, University Hospital Nîmes, 30900 Nîmes, France
4 IRD, AP-HM, SSA, VITROME, Aix-Marseille Université, 13284 Marseille, France
5 UFR Odontologie, Aix-Marseille Université, 13284 Marseille, France
6 INSERM, University of Montpellier, U1061, Neuropsychiatry: Epidemiological and Clinical Research, 34093 Montpellier, France; thibault.mura@chu-nimes.fr
7 Laboratoire de Biostatistique, Epidémiologie Clinique, Santé Publique Innovation et Méthodologie (BESPIM), Groupe Hospitalier Caremeau, CHU de Nîmes, Nîmes University Hospital, 30900 Nîmes, France
8 Paediatrics Department, University Hospital Nîmes, 30900 Nîmes, France; tu.anh.tran@chu-nimes.fr
9 INSERM U1183, Institute for Regenerative Medicine & Biotherapy, 34295 Montpellier, France
10 Faculty de Medicine, Montpellier University, 34000 Montpellier, France
11 CNRS UMR 9002, Institute of Human Genetics, 34090 Montpellier, France
12 IDESP, INSERM UMR UA11, Institut Desbrest d'Epidemiologie et de Santé Publique (IDESP) Campus Sante, 34093 Montpellier, France
* Correspondence: pierre.corbeau@igh.cnrs.fr (P.C.); anne.filleron@chu-nimes.fr (A.F.); joana.vitte@inserm.fr (J.V.); Tel.: +33-4-13-73-20-51 (J.V.); Fax: +33-4-13-73-20-52 (J.V.)
† These authors contributed equally to this work.

Citation: Sereme, Y.; Michel, M.; Mezouar, S.; Guindo, C.O.; Kaba, L.; Grine, G.; Mura, T.; Mège, J.-L.; Tran, T.A.; Corbeau, P.; et al. A Non-Invasive Neonatal Signature Predicts Later Development of Atopic Diseases. *J. Clin. Med.* **2022**, *11*, 2749. https://doi.org/10.3390/jcm11102749

Academic Editor: José Joaquín Cerón

Received: 19 March 2022
Accepted: 10 May 2022
Published: 12 May 2022

Abstract: Background: Preterm birth is a major cause of morbidity and mortality in infants and children. Non-invasive methods for screening the neonatal immune status are lacking. Archaea, a prokaryotic life domain, comprise methanogenic species that are part of the neonatal human microbiota and contribute to early immune imprinting. However, they have not yet been characterized in preterm neonates. **Objective:** To characterize the gut immunological and methanogenic Archaeal (MA) signature in preterm neonates, using the presence or absence of atopic conditions at the age of one year as a clinical endpoint. **Methods:** Meconium and stool were collected from preterm neonates and used to develop a standardized stool preparation method for the assessment of mediators and cytokines and characterize the qPCR kinetics of gut MA. Analysis addressed the relationship between immunological biomarkers, Archaea abundance, and atopic disease at age one. **Results:** Immunoglobulin E, tryptase, calprotectin, EDN, cytokines, and MA were detectable in the meconium and later samples. Atopic conditions at age of one year were positively associated with neonatal EDN, IL-1β, IL-10, IL-6, and MA abundance. The latter was negatively associated with neonatal EDN, IL-1β, and IL-6. **Conclusions:** We report a non-invasive method for establishing a gut immunological and Archaeal signature in preterm neonates, predictive of atopic diseases at the age of one year.

Keywords: preterm birth; fecal mediator and cytokine; methanogenic Archaea; allergy; atopy

1. Introduction

Preterm birth, defined as delivery at fewer than 37 completed weeks of gestation, is the leading cause of neonatal mortality and morbidity and has long-term adverse health consequences [1]. The global incidence of preterm births was estimated at 10.6 in 2014, and 9.8 in 2000 [2]. The etiology of preterm birth is multifactorial and not yet fully understood [3]. However, factors related to preterm birth include maternal or fetal medical conditions, genetic and epigenetic [4] influences, environmental exposures, infertility treatments, behavioral and socioeconomic elements [3,5]. Preterm infants experience abnormal immune and metabolic programming, which might exert a lasting influence on the risk of future disease [6,7]. Preterm-born children have been shown to have immune mediator dysregulation [8], impaired innate immunity and adaptive responses characterized by reduced levels of immunoglobulin (Ig) G, opsonization and phagocytosis, and increased activation of Th1 cells compared to that of Th2 cells [9]. Cohort studies show that preterm-born children are at increased risk for preschool wheezing and school-age asthma [10], but not for food allergy [11,12] or atopic dermatitis (AD) [11,13].

During fetal life, maternal microbiota produces compounds that are transferred to the fetus and enhance the generation of innate immune cells [14]. This process is halted prematurely in preterm infants, leaving them vulnerable to disease [9]. Preterm infants have an inflammatory and hypoxic state, which has a negative impact on lung maturation, the risk of respiratory infections, and susceptibility to subsequent exposures [9,14,15]. As early as the neonatal period, the gut microbiota imprints a persistent effect on the immune system through multiple mechanisms, including the modulation of epithelial functions, the production of cytokines, and the recruitment and training of immune cells [16–18].

Archaea, considered a separate domain of life from Eukarya, giant bacteria and viruses, are part of the human microbiota [19–21]. Gut methanogenic Archaea consume hydrogen produced by bacterial fermentation, releasing methane and short chain fatty acids (SCFA) and thus taking part in the energy supply to the host [19]. They interact with the host immune system, triggering innate and adaptive immune responses, generation of specific T and B cells, and hypersensitivity responses in animals and humans [22–24]. We have shown that the neonatal gut is colonized by methanogenic Archaea from the first postnatal hours, possibly starting in utero [21,22]. Gut microbiome establishment is altered in preterm and low- birth-weight infants [25,26].

Clinical investigations and research studies in neonates, including those born before term, are usually performed with peripheral blood. However, the search for non-invasive alternatives has gained momentum in recent years [27,28].

We hypothesize the existence of an association between intestinal methanogenic Archaea and the intestinal immunological signature, understood as a pattern of immune biomarkers including cell-specific products, e.g., mast cell tryptase and eosinophil-derived neurotoxin (EDN) and major pro- and anti-inflammatory cytokines. The validation of this hypothesis would open the prospect of a predictive score for the later occurrence of immune disorders, including atopic diseases.

We addressed this question through the development of a non-invasive standardized method for the assessment of the neonatal gut immune and microbial status, implemented in a cohort of preterm infants. The aims of the present study were: (1) establish a non-invasive method adequate for the investigation of preterm neonates, (2) characterize the gut immune and Archaeal components longitudinally from birth to six weeks in the study cohort, and (3) correlate the results of gut immune and Archaeal investigations at birth and up to six weeks to the later occurrence of allergic or atopic conditions.

2. Methods

2.1. Patients and Sampling

Stool samples from 43 preterm neonates were collected without the use of preservatives at the Nimes and Montpellier University Hospitals and stored at −80 °C. Samples

were collected as meconium (n = 33) and later stool samples at 2 (n = 33), 4 (n = 29), and 6 (n = 24) weeks.

2.2. Ethics Statement

2.2.1. A-Immunological Analysis

Preparation of Fecal Samples

One gram of feces was solubilized in two milliliters of an in-house extraction buffer consisting of phosphate buffered saline supplemented with 4 mM 4-(2-aminoethyl)-benzensulphonyl fluoride, 0.26 mM bestatin, 28 µM E-64, 2 µM leupeptin and 0.6 µM aprotinin, pH 7.4, and a protease inhibitor (Sigma-Aldrich, St. Louis, MN, USA [29]. The stool–buffer mixture was incubated for 20 min at room temperature, prior to centrifugation at 2000 rpm for 15 min at 4 °C. The supernatant liquids were freeze-dried for 24 h, resolubilized in one milliliter of extraction buffer, and used for mediator and cytokine determination (Figure 1).

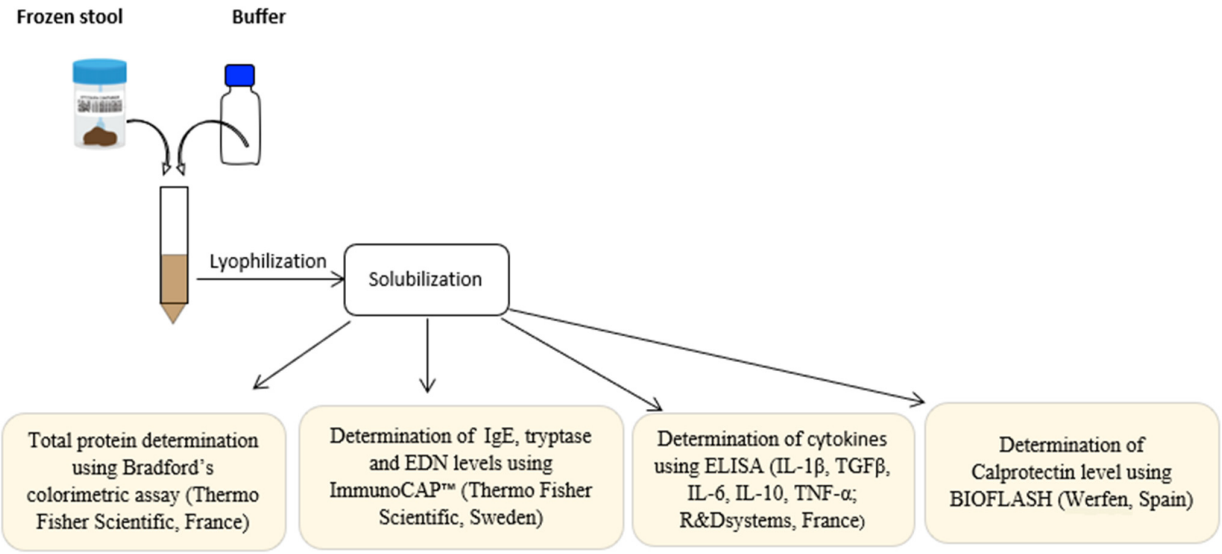


Figure 1. Fecal extraction protocol.

Total IgE, Tryptase, and EDN Determination

The concentration of total IgE, tryptase, and EDN were measured using an automated fluoro-enzymo-immunoassay with the ImmunoCAP™ 250 platform (Thermo Fisher Scientific, Uppsala, Sweden), according to ISO 15,189 standards [30]. The measurement range was 2–5000 kIU/L (4.8–12,000 µg/L) for total IgE, 1–200 µg/L for tryptase, and 2–200 µg/L for EDN.

Calprotectin and Total Protein Determination

Fecal calprotectin was measured using the BIOFLASH (Werfen, Barcelona, Spain) chemo-luminescent analyzer platform according to ISO 15,189 standards. Assay sensitivity was greater than 20 µg/mL. The total protein concentration of the samples was measured by the colorimetric method (BCA Protein assay, Thermo Fisher Scientific).

Immunoassays

Cytokines (IL-6, IL-10, IL-1β, TGF-β, and TNF-α) were measured by ELISA using specific immunoassay kits according to the manufacturer's protocols (R&D systems, Minneapolis, MN, USA). The sensitivity of the assays was 1.0 pg/mL.

2.2.2. A-Microbiological Analysis: Methanogenic Archaea by qPCR

DNA Extraction and PCR Assays

For DNA extraction, 0.2 g of each stool sample were mixed in 1.5 mL tubes with 500 µL of G2 lysis buffer from an EZ1®DNA Tissue Kit (QIAGEN, Hilden, Germany). Then, 0.3 g of acid-washed beads ≤ 106 µm (Sigma-Aldrich, Saint Quentin Fallavier, France) were added in each tube and shaken in a FastPrep BIO 101 device (MP Biomedicals, Illkirch, France) for 45 s for mechanical lysis before 10 min incubation at 100 °C. A 180 µL volume of the mixture was then incubated with 20 µL of proteinase K (QIAGEN, Hilden, Germany) at 56 °C overnight before a second mechanical lysis was performed. Total DNA was finally extracted with an EZ1 Advanced XL extraction kit (QIAGEN) and 50 µL eluted volume. Sterile phosphate buffered saline (PBS) (Fisher Scientific, Illkirch, France) was used as a negative control in each DNA extraction run. Extracted DNA was incorporated into real-time PCR performed using Metha_16S_2_MBF: 5′-CGAACCGGATTAGATACCCG -3′ and Metha_16S_2_MBR: 5′-CCCGCCAATTCCTTTAAGTT-3′ primers and the FAM_Metha_16S_2_MBP 6FAM- CCTGGGAAGTACGGTCGCAAG probe targeting the 16S DNA gene of methanogens, designed in our laboratory (Eurogentec, Angers, France) as previously described [31]. PCR amplification was done in 20 µL volume including 15 µL of mix and 5 µL of extracted DNA. Five microliters of ultra-pure water (Fisher Scientific, Illkirch, France) were used instead of DNA in the negative controls. The amplification reaction was performed in a CFX96 thermocycler (BioRad, Marnes-la-Coquette, France) incorporating a protocol with a cycle of 50 °C for 2 min, followed by 39 cycles of 95 °C for 45 s, 95 °C for 5 s and finally 60 °C for 30 s. Samples with a CT < 40 were considered positive. Gene amplification and PCR sequencing were performed as previously described [25,26,32–34].

Statistical Analysis

The responses for each quantitative parameter were described using median and 25–75 percentile (interquartile range, IQR) unless otherwise stated. Analyses were performed using the Wilcoxon test when two groups were compared, and the Kruskal–Wallis test when more than two groups were compared. The association between the different biomarkers of interest were analyzed, at each sampling time, using Spearman's correlation coefficient. The association profiles between different biomarkers were also analyzed using a principal component analysis method. Statistical analyses were performed at the conventional two-tailed α level of 0.05, using R 2.13.2 statistical software (R Foundation for Statistical Computing, https://www.r-project.org (accessed on 18 March 2022), Vienna, Austria).

3. Results

3.1. Demographic and Clinical Characteristics of Preterm Infants

The 43 preterm neonates included in our study had at birth an average weight of 1160.41 g (range 440–1750 g), an average gestational age of 29 weeks (range 24–32 weeks) and an average height of 37.11 cm (range 32–47 cm). Thirty-five (81%) were born by cesarean section and 8 (19%) by vaginal delivery. Only five mothers (11.6%) had received antibiotic therapy during the peripartum period. As part of the cohort follow-up, clinical evaluation (AF) was conducted at 1 year and assessed the presence or absence of health conditions, including atopic diseases. When necessary, allergy diagnosis was carried out according to current recommendations [35,36]. A total of nine children developed an atopic condition during the first year, manifested as asthma or cow's milk allergy (CMA) in eight and AD in three, with two patients presenting an association of AD, asthma, and CMA (Table 1).

Table 1. Clinical data for preterm infants investigated for the presence of fecal biomarkers. VD: vaginal delivery; CMA, cow's milk allergy; CS: cesarean section.

Code	Meconium (M)	Two-weeks (W2)	Four-Week (W4)	Six-weeks (W6)	Peripartum Maternal Antibiotic Therapy	Mode of Delivery	Gestational Age	Weight	Size	Asthma or CMA	Atopic Dermatitis
1	0	W2	W4	W6	No	VD	30	1275	37	Yes	Yes
2	M	W2	W4	W6	Yes	CS	27	925	34	Yes	No
3	M	W2	W4	W6	No	CS	26	565	31	Yes	No
4	0	0	W4	0	No	VD	25	820	34	Yes	No
5	M	W2	W4	0	No	CS	32	1260	39	Yes	No
6	M	W2	W4	W6	Yes	CS	27	680	31	Yes	No
7	M	0	W4	W6	No	CS	29	1565	42	No	No
8	M	W2	W4	W6	No	CS	28	890	33	No	No
9	0	0	0	W6	Yes	CS	30	1150	38	No	No
10	M	W2	0	0	No	CS	31	1570	43	No	No
11	M	W2	W4	0	No	CS	32	1575	44	No	No
12	M	W2	W4	0	No	CS	30	1360	39	No	Yes
13	M	W2	0	W6	No	CS	25	870	34	Yes	Yes
14	M	W2	W4	0	No	CS	32	1155	39	Yes	No
15	0	W2	W4	W6	No	CS	25	440	28	No	No
16	0	0	0	W6	No	VD	30	1590	41	No	No
17	M	W2	W4	W6	No	CS	24	530	31	No	No
18	M	W2	0	0	No	CS	26	925	35	No	No
19	M	W2	W4	0	No	VD	30	1480	41	No	No
20	M	W2	W4	0	No	VD	30	1460	38	No	No
21	M	W2	W4	W6	No	CS	29	880	35	No	No
22	0	W2	W4	W6	No	CS	28	840	35	No	No
23	M	W2	W4	0	Yes	VD	30	1670	43	No	No
24	M	W2	W4	W6	No	CS	31	1120	38	No	No
25	0	W2	0	0	No	CS	28	915	36	No	No
26	M	0	W4	W6	No	CS	26	925	35	No	No
27	M	W2	W4	W6	Yes	CS	30	1335	39	No	No
28	M	W2	W4	W6	Yes	CS	30	1355	47	No	No
29	M	W2	0	W6	No	CS	30	1480	39	No	No
30	0	W2	W4	W6	Yes	CS	28	1010	35	No	No
31	M	W2	W4	W6	No	CS	29	1050	39	No	No
32	M	W2	W4	0	No	CS	29	1190	38	No	No
33	0	W2	W4	W6	No	CS	30	1175	39	No	No
34	M	0	0	0	No	CS	32	1930	44	No	No
35	0	W2	W4	0	Yes	CS	29	1430	30	No	No
36	M	W2	W4	0	No	CS	30	1750	43	No	No
37	M	0	0	0	No	CS	27	600	29	No	No
38	M	W2	W4	W6	No	VD	25	750	32	No	No
39	M	W2	0	W6	No	CS	31	980	36	No	No
40	M	W2	0	W6	No	CS	31	1410	39	No	No
41	M	0	0	0	No	CS	30	770	33	No	No
42	M	0	0	0	Yes	VD	32	1568	41	No	No
43	M	0	0	0	No	CS	30	1680	39	No	No

3.2. Immune Profiling

3.2.1. Total Protein Determination

First, we measured the total protein content in all samples. The median concentration of fecal proteins was stable from birth to six weeks, ranging from 4.53 to 9.18 g/L ($p = 0.10$; Kruskal–Wallis) (Table 2).

Table 2. Determination of fecal immune biomarkers.

	Meconium		Two Weeks		Four Weeks		Six Weeks			
	n = 33		n = 33		n = 29		n = 24			
	n (%) Detectable	Median IQR	n (%) Detectable	Median IQR	n (%) Detectable	Median IQR	n (%) Detectable	Median IQR	*p*-Value (Frequency)	*p*-Value (Levels)
Total Proteins (g/L)	33 (100)	9.18 (4.51–13.54)	33 (10,055)	5. 4.23–6.05)	29 (100)	4.53 (3.00–5.52)	24 (100)	6.46 (5.39–7.76)	NS	0.10
Total IgE (μg/L)	30 (90.90)	7.3 (6.4–9.9)	32 (97)	8.47 (6.8–9.8)	27 (93.10)	9.74 (3.39–0.26)	24 (100)	115.08 (41.00–193.70)	0.41	<0.0001
Tryptase (μg/L)	3 (9.1)	<1	3 (9.1)	<1	4 (13.79)	<1	14 (58.33)	1.8 (0.0–3.4)	<0.0001	0.61
Calprotectin (μg/L)	33 (100)	310.4 (151.1–771.3)	33 (100)	291.23 (189.41–487.87)	29 (100)	402.44 (300.06–607.3)	24 (100)	422.37 (335.53–823.30)	NC	0.13
EDN (μg/L)	33 (100)	83.2 (19.3–165.0)	33 (100)	70.1 (17.8–152.5)	29 (100)	109.0 (44.2–200.0)	24 (100)	98.1 (57.5–200.0)	NC	0.21
TGF-β (pg/L)	24 (72.7)	121.3 (4.6–258.9)	30 (91)	267.43 (61.71–1000)	26 (89.65)	384.57 (129.60–936)	22 (91.66)	466 (104.36–1430.29)	0.09	0.014
IL-1β (pg/L)	13 (39.4)	0.12 (0.1–2.7)	28 (84.8)	1.53 (0.37–6.53)	25 (86.20)	3.27 (0.31–10.76)	22 (91.66)	6.23 (1.66–20.84)	<0.0001	0.001
IL-10 (pg/L)	4 (12.12)	3.9 (3.9–3.9)	6 (18.18)	3.9 (3.9–3.9)	5 (17.24)	3.9 (3.9–3.9)	5 (20.83)	3.9 (3.9–3.9)	0.85	0.53
IL-6 (pg/L)	25 (75.75)	11.6 (0.5–43.7)	7 (21.21)	0.2 (0.2–0.2)	20 (68.96)	0.2 (0.2–0.2)	20 (83.33)	3.77 (1.66–19.25)	<0.0001	<0.001

Concentrations are expressed as median and interquartile ranges (IQR). n (%): number of samples in which the biomarker was detected (relative frequency of detection). The median and IQR were calculated by restricting the results above the lower LOQ (limit of quantitation) for each analyte. Statistical test: chi-square (frequency), Kruskal–Wallis (concentration). NC, not calculable (calprotectin and EDN were detectable in all samples and at all sampling times).

3.2.2. Immune Cell Markers and Cytokines

Total IgE was detectable in over 90% of the samples at all ages, in increasing amounts between birth (meconium) and six weeks ($p < 0.0001$; Kruskal–Wallis).

Conversely, tryptase detection increased with sampling age, reaching 58% in samples at six weeks, up from less than 15% at earlier times ($p < 0.0001$; Chi-square). As most values were lower than the quantification limit, quantitative comparison was not significant ($p = 0.61$, Kruskal–Wallis) (Figure 2 and Table 2).

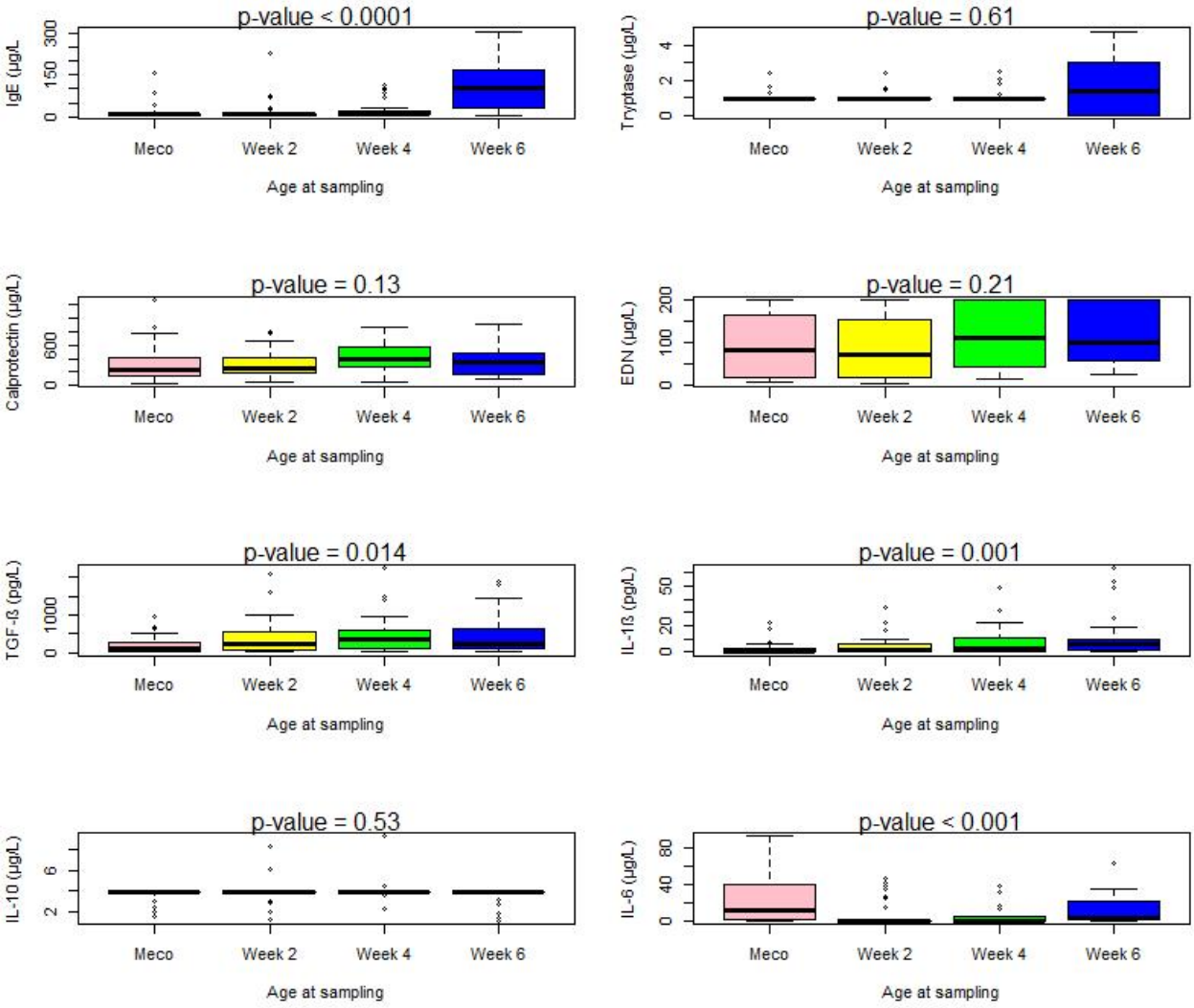


Figure 2. Statistical test: chi-square (frequency), Kruskal–Wallis (concentration). NC, not calculable (calprotectin and EDN were detectable in all samples and at all sampling times).

Calprotectin and EDN were detected in all samples at comparable levels irrespective of age ($p = 0.13$ and 0.21, Kruskal–Wallis) (Figure 2 and Table 2).

All cytokines except TNF-α were detectable in meconium and fecal samples. TGF-β and IL-6 were the most prevalent, detected in up to 90% of samples, while IL-10 was the less prevalent, found in 20% or less of the fecal samples. The frequency of detection of IL-6 and IL-1β increased with age ($p < 0.0001$; chi-square), although there was a sharp drop in IL-6 frequency of detection and measured levels between meconium (75%, median 11.6 pg/L) and samples at two weeks (21%, median 0.2 pg/L).

TGF-β and IL-1β median concentrations increased with age ($p = 0.014$ and 0.001, respectively; Kruskal–Wallis). The median level of IL-6 was the highest in meconium samples and increased again at six weeks ($p < 0.0001$; Kruskal–Wallis). IL-10 median concentrations did not vary with age (Figure 2 and Table 2).

Maternal antibiotic therapy and route of delivery did not significantly affect the meconium levels of cytokines, total IgE, tryptase, calprotectin, and EDN (Supplementary Table S1). However, analysis according to the development of atopic disease during the first year showed that meconium calprotectin levels were lower in neonates who subsequently developed asthma or CMA compared to those who did not ($p = 0.02$; Wilcoxon test) (Figure 2). Levels of other mediators and cytokines were not associated with the occurrence of an atopic disease (Table 3).

Table 3. Comparison of mediators and cytokines and the occurrence or absence of an atopic condition between years 0 and 1. Statistical test used: Wilcoxon test.

Variables	Allergic Condition (APLV and Asthma)	*p*-Value			
		Meconium	2 Weeks	4 Weeks	6 Weeks
IgE	Yes / No	0.27	0.06	0.12	**0.03**
Calprotectin	Yes / No	0.27	0.18	0.91	0.61
EDN	Yes / No	0.59	0.19	0.41	0.87
TGF-β	Yes / No	0.09	0.18	0.76	0.76
IL-1β	Yes / No	0.37	1.00	0.28	0.91
IL-10	Yes / No	0.62	1.00	0.89	0.13
IL-6	Yes	0.61	0.28	0.37	0.75

Comparing the different biomarkers according to the presence or absence of an atopic condition, we observed a significant difference at week six for IgE between the presence and absence of cow's milk allergy and asthma. No significant difference was observed for AD (Supplementary Table S2)

3.2.3. Correlation between Biomarkers

As an expected control, significant correlations were found between weight and height (R = 0.90; $p < 0.0001$), between gestational age and height (R = 0.75; $p < 0.0001$), and between gestational age and weight (R = 0.75; $p < 0.0001$).

Total IgE and tryptase levels were strongly correlated in samples taken at any age. IL-10 and IL-6 were correlated at all ages except at two weeks (Table 4).

Table 4. Result of the detection of methanogenic Archaea.

	Meconium (n = 33)		Two Weeks (n = 33)		Four Weeks (n = 29)		Six Weeks (n = 26)		*p*-Value (Frequency)	*p*-Value (CT)
	n (%)	Median IQR	n (%)	Median IQR	n (%)	Median IQR	n (%)	Median IQR		
CT qPCR	30 (90.9)	36.74 (33.85–38.24)	27 (81.81)	37.20 (36.07–38.33)	23 (79.31)	37.75 (36.13–38.50)	19 (73.03)	38.28 (37.27–39.96)	0.34	0.12

CT methanogenic Archaea are expressed as median and interquartile ranges (IQR). n (%): number of samples in which methanogenic Archaea were detected (relative frequency of detection). The

stool concentration factor and median and RDI were not included in our calculations, and the median and RDI were calculated by restricting the results above the lower LOQ (limit of quantitation) for each analyte. Statistical test: Kruskal–Wallis.

3.2.4. Meconium Samples

Tryptase levels were correlated to levels of IL-10 (R = 0.48, $p = 0.001$) and IL-6 (R = 0.46; $p = 0.001$). Tryptase levels were correlated to levels of IgE (R = 0.91, $p = 0.0001$), and strong correlations were observed between levels of calprotectin and IL-1β (R = 0.90; $p < 0.0001$). A negative correlation between total protein concentration and TGF-β (R = −0.36; $p = 0.01$) was observed (Table 3).

3.2.5. Samples at Two Weeks

Strong correlations were observed between levels of calprotectin and IL-1β (R = 0.90; $p < 0.0001$), tryptase and IL-10 (R = 0.88; $p < 0.0001$), and total IgE and IL-10 (R = 0.85; $p < 0.0001$), while total protein concentration and IL-6 were negatively correlated (R = −0.61; $p = 0.0002$) (Table 3).

3.2.6. Samples at Four Weeks

Again, calprotectin and IL-1β were strongly correlated (R = 0.74; $p < 0.0001$). IgE and total protein were also correlated (R = 0.38; $p < 0.04$) (Figure 3).

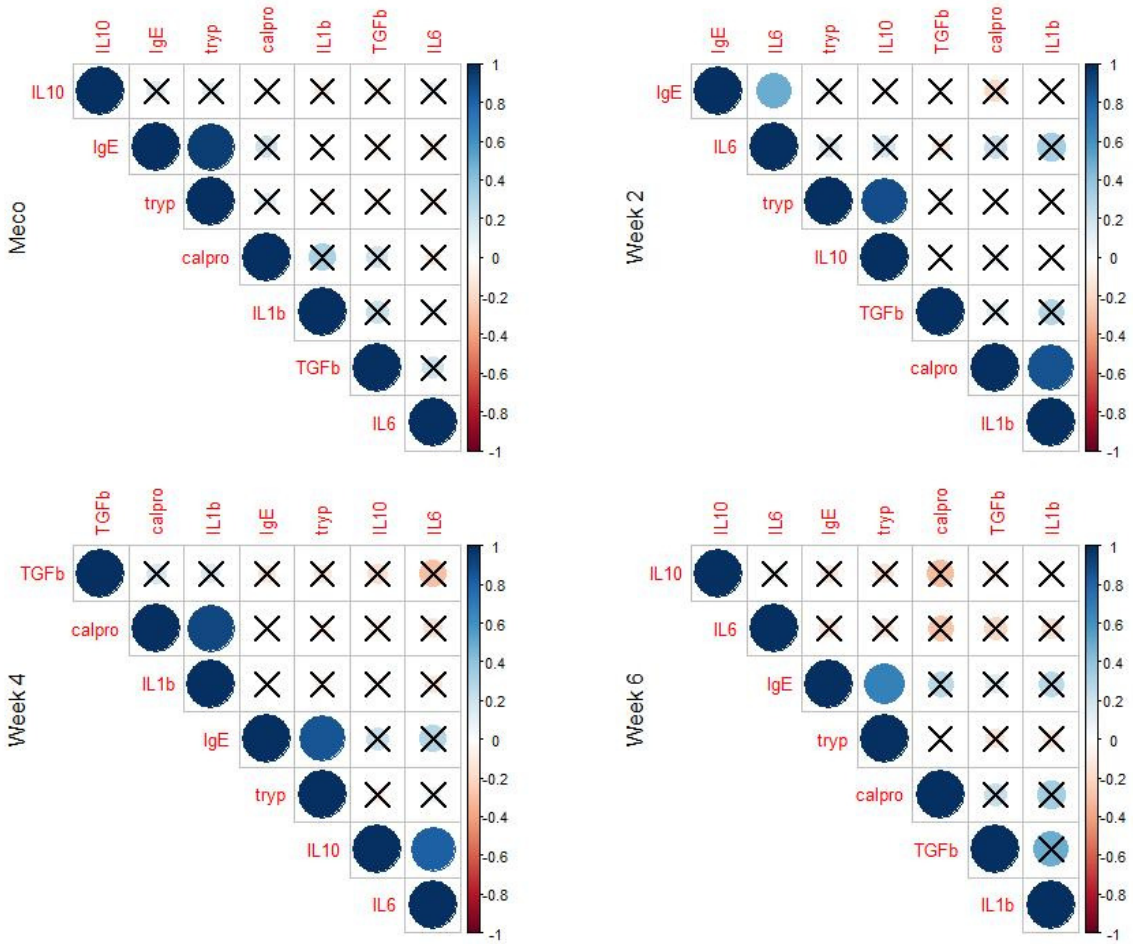


Figure 3. Correlation figure of fecal immune biomarkers.

3.2.7. Samples at Six Weeks

TGF-β was correlated with IL-1β (R = 0.49; $p < 0.01$) and with total proteins (R = 0.43; $p < 0.03$). Total proteins were correlated with IL-1β (R = 0.47; $p < 0.02$) and negatively with IL-6 (R = -0.68; $p = 0.0003$) (Figure 3).

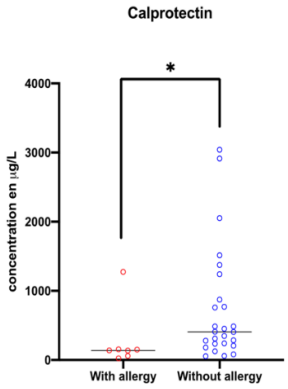
3.3. Frequency of Detection of Methanogenic Archaea and Relationship with the Subsequent Development of Atopic Diseases

Using real-time PCR with 16S rRNA archaeal gene PCR primers, we detected methanogenic Archaea DNA in 30/33 (90%) meconium samples, 27/33 (81%) two-week-old samples, 23/29 (79%) four-week-old samples, and 19/24 (73%) six-week-old samples, respectively. We found no significant difference in the frequency of detection nor in CTs according to age at sampling (Table 4).

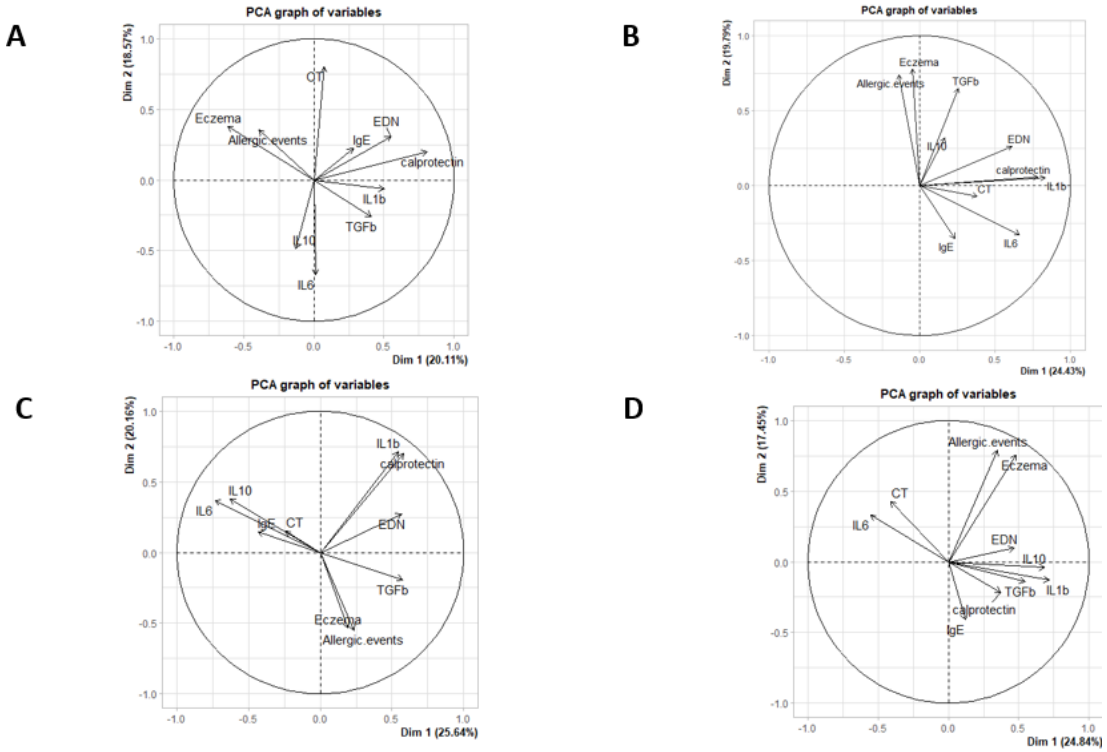
3.4. Unsupervised Analysis of Immunological Markers and Methanogenic Archaea Atthe Neonatal Period, and the Subsequent Occurrence of AD, Asthma and CMA during the First Year

We performed unsupervised analysis of the immunological data, CT of Archaea, and the clinical information of the occurrence of allergic events during the first year of life. Data were analyzed for each of the four sampling times.

For meconium, calprotectin, EDN, and IL-1β levels were negatively and significantly ($p < 0.001$) correlated (r = −0.64) with subsequent development of AD. Calprotectin, EDN and IL-1β had the largest and most significantly ($p < 0.01$) correlated positive correlation coefficients, which were 0.79, 0.53, and 0.51, respectively. No correlation was observed for the Archaea CT with the other parameters (Figure 4b(A)).



(**a**)



(**b**)

Figure 4. (**a**). Comparison of meconial calprotectin concentration according to the later occurrence of asthma or cow's milk allergy. (**b**). Principal component analysis of neonatal immune and archaeal biomarkers as a function of later occurrence of atopic conditions. (**A**) Negative and significant ($p < 0.001$) correlation (r = −0.64) between the occurrence of atopic dermatitis with calprotectin, EDN, and IL-1β. Strong positive and significant correlation ($p < 0.01$) between calprotectin (r = 0.79), EDN (0.53), and IL-1b (0.51). No correlation between Archaea TCs and other parameters. (**B**) Correlation between allergic events (asthma or cow's milk allergy) and atopic dermatitis. Strong positive and significant ($p < 0.001$) correlation between IL-1b (r = 0.88) and calprotectin (r = 0.82), positive correlation between IL-6 (r = 0.62) and EDN (r = 0.59), significant ($p < 0.001$). Low positive (r < 0.5) and significant correlation between Archaea TCs and biomarkers IL-1b, calprotectin, IL-6, and EDN. (**C**) Positive and significant correlation ($p < 0.001$) between atopic dermatitis, allergic events with the markers calprotectin (r = 0.61), IL-1b (r = 0.58), EDN (r = 0.57), and TGF-β (r = 0.57). Negative correlation between IL-6 (r = −0.71) and IL-10 (r = −0.61) with atopic dermatitis and allergic events. Positive and significant correlation between Archaea TCs with calprotectin, IL-1b, IL-6, and EDN. Negative correlation between Archaea Ct and allergic events. (**D**) Negative correlation between IL-6 (r = −0.51) with other biomarkers and allergic events. Strong positive and significant correlation ($p < 0.001$) between IL-1b (r = 0.68), IL-10 (r = 0.67), and atopic dermatitis (r = 0.58). Low correlation (r < 0.5) between Archaea TCs with IL-1β, calprotectin, IL-6, and EDN. Negative correlation between allergic events and TC of Archaea.

At two weeks, IL-1β (r= 0.88) and calprotectin (r = 0.82) had a strong positive correlation with each other, followed by IL-6 (r = 0.62) and EDN (r = 0.59). These biomarkers were significantly ($p < 0.001$) associated. Archaea CTs had a weak (r < 0.5) but positive and significant ($p < 0.01$) association with the biomarkers IL-1β, calprotectin, IL-6, and EDN (Figure 4b(B)).

At four weeks, later occurrence of AD, CMA, and asthma was positively correlated with calprotectin (r = 0.61), IL-1β (r = 0.58), EDN (r = 0.57), and TGF-β (r = 0.57), and negatively correlated with IL-6 (r = −0.71) and IL-10 (r = −0.61). Calprotectin, IL-1β, and EDN were significantly associated with each other ($p < 0.001$). Archaea CTs were positively associated with calprotectin, IL-1β, IL-6, and EDN with a significant correlation ($p < 0.001$),

however, they were inversely correlated with allergic events, although the correlation coefficient was low (Figure 4b(C)).

At six weeks, only IL-6 correlated negatively (r = −0.51) with the other biomarkers and allergic events. IL-1β (r = 0.68), IL-10 (r = 0.67), and AD (r = 0.58) showed the strongest positive correlations, with AD significantly ($p < 0.001$) associated with IL-1β and IL-10. Archaea CTs were weakly correlated (r < 0.5) with IL-1β, calprotectin, IL-6, and EDN. However, allergic events were negatively associated with CT, and AD had almost no correlation (Figure 4b(D)).

4. Discussion

In this study, we describe a non-invasive screening method for profiling neonatal immunity and its validation in a preterm neonate cohort as a predictive tool for subsequent development of atopic diseases. The method was also applied to meconium samples, which reflect intrauterine processes and contain almost 1000 identified proteins with important functions [37]. The clinical endpoints of this study were evaluated at the age of one year, while the total duration of cohort follow-up will be three years.

Total IgE was detected in over 90% of the samples, at increasing concentrations with age. Transplacental delivery of allergens and preterm sensitization have long been recognized, possibly inducing sensitization and detectable meconial IgE [38,39]. Transplacental transport of maternal IgE able to sensitize fetal mast cells has been recently demonstrated [40], but its role in neonatal immune defenses or subsequent immune disorders is only speculative.

Addressing mast cell tryptase in meconium and later samples, we found that it was detectable only in a minority of meconium samples and during the first month of life, however, it became a common finding at the end of the neonatal period, represented by samples collected at six weeks. Fecal tryptase and IgE levels were strongly associated at each of the studied time points. These results suggest that gastrointestinal mast cells, as opposed to skin mast cells [40], are mostly recruited postnatally, and mature after birth with IgE levels exerting a positive effect. Conversely, maturity of gastrointestinal mast cell populations might be attained during late pregnancy. Tryptase is a serine protease able of autocrine activation of mast cells and induction of proinflammatory effects such as proteolytic cleavage and activation of PAR2 receptors and inactivation of VIP (Vasoactive Intestinal Peptide), associated with smooth muscle relaxation [41]. Through PAR-2 activation, luminal tryptase can contribute to the dysfunction of the gut epithelial barrier [42]. The presence of tryptase in stool samples has been associated with food allergic diseases, dietary exposure and/or mast cell stimulation or increased intestinal mast cell count [43,44]. In addition, fecal tryptase has also been shown to be associated with inflammatory bowel disease and irritable bowel syndrome [45,46].

Focusing on two secreted biomarkers of innate immune cells, neutrophil-derived calprotectin, and eosinophil-derived EDN, we found that fecal samples at all studied ages contained detectable and stable levels of both biomarkers. The levels measured in our preterm cohort were much lower than the reference values [47], but similar to those reported during the first postnatal month in another preterm cohort [48]. These results suggest that the preterm gut contains small numbers of granulocytes, or that such granulocytes are not activated. Indeed, lower neutrophils were reported in preterm infant cord blood [49]. An association between low levels of fecal calprotectin and adverse health conditions, including obesity and sepsis, by age two has been suggested [48].

Proinflammatory and anti-inflammatory cytokines were detected in the meconium and later samples of preterm infants, with the notable exception of TNF-α which was not demonstrated in any sample. Different temporal patterns were demonstrated: IL-6 levels were higher in the meconium than in later samples, while IL-10 was seldom detected and TGF-β and Il-1β displayed a progressive increase between birth and six weeks. Mostly undetectable IL-10 levels were also reported in a pilot study of fecal biomarkers in preterm infants [8]. Although we did not determine the cellular source of fecal cytokines, a shift in

immune cells lining the intestine has been demonstrated for macrophages, with resident fetal macrophages being replaced after birth by bone-marrow-derived macrophages [50]. Macrophage cytokine production, most notably of proinflammatory IL-6 and IL-1β, can be persistently altered by metabolic conditions [51]. The increase in TGF-β levels from birth to six weeks might provide a counter-acting mechanism in a proinflammatory environment. Allergic events (asthma or cow's milk allergy) and atopic dermatitis were also positively correlated with EDN, IL-1β, IL-10, and IL-6 at four and six weeks. The high production of fecal EDN, IL-1β, and IL-10 during the first weeks of life may therefore be an indicator for later risk of allergic diseases.

The neonatal period is paramount for the establishment of the intestinal microbiota. Intrauterine life is associated with low levels of maternal microbial translocation [52]. However, we have recently demonstrated the presence of the viable methanogenic Archaea *Methanobrevibacter smithii* in the meconium, suggesting intrauterine colonization of the fetus by this microorganism [26]. Here, we provided evidence for postnatal persistence of methanogenic Archaea in fecal samples and suggest a possible role in the orientation of intestinal immunity, supported by the negative association between Archaea abundance (inversely proportional to CT values) and the concentrations of EDN, IL-1β, and IL-6. We also found that Archaea abundance at four and six weeks was positively associated with later occurrence of allergic events. Archaea have been shown to produce SCFA which induce regulatory T cell differentiation, downregulate proinflammatory cytokines, and may protect against the occurrence of atopic conditions [19,53–57] However, in a cohort study, the protective effect of methanogenic Archaea was restricted to the species *Methanobrevibacter stadtmanae* [58]. A decrease in the load of beneficial methanogenic Archaea during the first years of life could therefore favor the occurrence of allergic events during the first years of life.

A third line of contribution to protection or increased risk of developing atopic conditions is the genetic background. As an example, a del/del genotype (−2549 −2567 del18) of Vascular Endothelial Growth Factor (VEGF) has been associated with asthma occurrence and irreversible bronchoconstriction [59].

The strengths of our study are methodological and medical:

(1) miniaturization and standardization, using small quantities of stool (1 g) and small volumes of extraction buffer (2 mL). The dilution of the samples was corrected by the freeze-drying process, as the lyophilizates were contained in 1 mL of buffer.
(2) prevention, thanks to the use of protease inhibitors, of the risk of potential contamination of the handler.
(3) suitability for a microarray platform yielding patterns of immune responses rather than individual measurements.
(4) suitability for combined immune and microbiological assessment.
(5) proof of concept of the immune profiling of fecal mediators in meconium and neonatal samples as predictors of later development of atopic disorders.
(6) proof of concept for non-invasive investigation of the immune status of preterm neonates.

The main weakness of this study is the lack of microbiological data outside Archaea. Further studies are warranted for longitudinal immuno-microbiological profiling of meconium and neonatal samples, in preterm and at-term infants. Its validation as a non-invasive diagnostic method will be in line with the currently unmet needs in terms of non-invasive diagnosis of allergy.

5. Conclusions

This study allowed us to highlight the presence of mediators in the meconium and feces of preterm infants. We provide proof of concept of the feasibility and value of a standardized fecal mediator assay for non-invasive profiling of neonatal immunity. Such assays can be used for early characterization of the immune status of a newborn. Technical

optimization for a multiplex assay could facilitate the implementation of fecal immune profiling in clinical and research laboratories. We report evidence of a correlation between the meconial and neonatal load of methanogenic Archaea and selected fecal biomarkers, and later occurrence of atopic conditions in preterm children.

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/jcm11102749/s1, Table S1: Comparison of biomarkers and CT of Archaea according to the presence or absence of maternal antibiotic therapy and according to the mode of delivery, Table S2: Comparison of median CT of Archaea in preterm infants according to the presence or absence of allergic events during the first year of life, Table S3: Comparison of median CTs of Archaea in preterm infants according to the presence or absence of atopic dermatitis during the first year of life.

Author Contributions: Y.S. developed the stool preparation method, performed the immunology experiments, analyzed the data, and wrote the first draft of the manuscript. M.M. and S.M. performed the experiments, analyzed the data, and proofread the manuscript. Y.S., C.O.G., L.K. and G.G. performed the microbiology experiments and analyzed the data. T.M. analyzed the data. G.G. and T.M. proofread the manuscript. J.-L.M., T.A.T., P.C., A.F. and J.V. designed the study and supervised the experiments. T.A.T. and A.F. collected and analyzed the clinical data. A.F. included patients and performed clinical examination. T.A.T., P.C., A.F. and J.V. wrote the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: Y.S. is supported by a "Fondation Méditerranée Infection" doctoral position. S.M. was first supported by a "Fondation pour la Recherche Médicale" postdoctoral fellowship (SPF20151234951) and then by a "Fondation Méditerranée infection" postdoctoral fellowship. This work was supported by the French Government under the "Investissements d'avenir" (Investments for the Future) program managed by the "Agence Nationale de la Recherche" (reference: Méditerranée Infection 10-IAHU-03), by the Région Provence-Alpes-Côte d'Azur and the European funding FEDER IHU PRIMMI, and the PHRC program for clinical research Primibiota.

Institutional Review Board Statement: This study was embedded in the Primibiota "Influence of Intestinal Microbiota Implantation in Preterm Infants on Microbiota and Immune Orientation at 3 Years" (NCT02738411, principal investigator AF), a population-based prospective cohort study from birth to the age of 3 years, enrolling children born preterm in the University Hospitals of Nîmes and Montpellier, France. The ancillary study presented here was approved by a joint committee of the Clinical Research Departments of the University Hospitals of Nîmes, France and the University Hospitals of Marseille, France (Research collaboration agreement 2018.1238).

Informed Consent Statement: Informed consent was obtained from both parents at infants' birth.

Data Availability Statement: Not applicable.

Acknowledgments: We thank Michel Drancourt and Bernard La Scola for conceptual and methodological assistance; Patricia Blanchard, Céline Chartier, Robert Perrier, and Simon Pinchemel for technical assistance with the ImmunoCAP platform.

Conflicts of Interest: J.V. reports speaker and consultancy fees in the past five years from Astra Zeneca, Meda Pharma (Mylan), Novartis, Sanofi, Thermo Fisher Scientific, outside the submitted work. The other authors declare no competing interests in relation to this study.

Abbreviations

AD: atopic dermatitis; CMA, cow's milk allergy; CT, cycle threshold; IQR, interquartile range; SCFA, short chain fatty acid; EDN, eosinophil-derived neurotoxin

References

1. Crump, C.; Sundquist, J.; Winkleby, M.A.; Sundquist, K. Gestational age at birth and mortality from infancy into mid-adulthood: A national cohort study. *Lancet Child Adolesc. Health* **2019**, *3*, 408–417.
2. Chawanpaiboon, S.; Vogel, J.P.; Moller, A.B.; Lumbiganon, P.; Petzold, M.; Hogan, D.; Landoulsi, S.; Jampathong, N.; Kongwattanakul, K.; Laopaiboon, M.; et al. Global, regional, and national estimates of levels of preterm birth in 2014: A systematic review and modelling analysis. *Lancet Glob. Health* **2019**, *7*, e37–e46.
3. Goldenberg, R.L.; Culhane, J.F.; Iams, J.D.; Romero, R. Epidemiology and causes of preterm birth. *Lancet* **2008**, *371*, 75–84.
4. Schuster, J.; Uzun, A.; Stablia, J.; Schorl, C.; Mori, M.; Padbury, J.F. Effect of prematurity on genome wide methylation in the placenta. *BMC Med. Genet.* **2019**, *20*, 116.

5. Beck, S.; Wojdyla, D.; Say, L.; Betran, A.P.; Merialdi, M.; Requejo, J.H.; Rubens, C.; Menonf, M.; Look, V.F.P. The worldwide incidence of preterm birth: A systematic review of maternal mortality and morbidity. *Bull. World Health Organ.* **2010**, *88*, 31–38.
6. Gagneur, A.; Pinquier, D.; Quach, C. Immunization of preterm infants. *Hum. Vaccines Immunother.* **2015**, *11*, 2556–2563.
7. Marshall, H.; Clarke, M.; Rasiah, K.; Richmond, P.; Buttery, J.; Reynolds, G.; Andrews, R.; Nissen, M.; Wood, N.; McIntyre, P . Predictors of Disease Severity in Children Hospitalized for Pertussis During an Epidemic. *Pediatr. Infect. Dis. J.* **2015**, *34*, 339–345.
8. Gómez, M.; Moles, L.; Espinosa-Martos, I.; Bustos, G.; de Vos, W.; Fernández, L.; Rodríguez, M.J.; Fuentes, S.; Jiménez, E. Bacteriological and Immunological Profiling of Meconium and Fecal Samples from Preterm Infants: A Two-Year Follow-Up Study. *Nutrients* **2017**, *9*, 1293.
9. Helmo, F.R.; Alves, E.A.R.; Moreira, R.A.D.A.; Severino, V.O.; Rocha, L.P.; Monteiro, M.L.G.D.R.; Marlene Antônia dos Reis, D.A.M.; Etchebehere, M.R.; Machado, R.J.; Corrêa, M.R.R. Intrauterine infection, immune system and premature birth. *J. Matern.-Fetal Neonatal Med.* **2018**, *31*, 1227–1233.
10. Sonnenschein-van der Voort, A.M.; Arends, L.R.; de Jongste, J.C.; Annesi-Maesano, I.; Arshad, S.H.; Barros, H.; Basterrechea, M.; Bisgard, H.; Chatzi, L.; Corpeleijn, E.; et al. Preterm birth, infant weight gain, and childhood asthma risk: A meta-analysis of 147,000 European children. *J. Allergy Clin. Immunol.* **2014**, *133*, 1317–1329.
11. Wooldridge, A.L.; McMillan, M.; Kaur, M.; Giles, L.C.; Marshall, H.S.; Gatford, K.L. Relationship between birth weight or fetal growth rate and postnatal allergy: A systematic review. *J. Allergy Clin. Immunol.* **2019**, *144*, 1703–1713.
12. Mitselou, N.; Hallberg, J.; Stephansson, O.; Almqvist, C.; Melén, E.; Ludvigsson, J.F. Cesarean delivery, preterm birth, and risk of food allergy: Nationwide Swedish cohort study of more than 1 million children. *J. Allergy Clin. Immunol.* **2018**, *142*, 1510–1514.e2.
13. Barbarot, S.; Gras-Leguen, C.; Colas, H.; Garrot, E.; Darmaun, D.; Larroque, B.; Roze, J.C.; Ancel. P. Lower risk of atopic dermatitis among infants born extremely preterm compared with higher gestational age. *Br. J. Dermatol.* **2013**, *169*, 1257–1264.
14. Al Nabhani, Z.; Eberl, G. Imprinting of the immune system by the microbiota early in life. *Mucosal Immunol.* **2020**, *13*, 183–189.
15. Stokholm, J.; Blaser, M.J.; Thorsen, J.; Rasmussen, M.A.; Waage, J.; Vinding, R.K.; Schoos, M.A.; Kunøe, A.; Fink, R.N.; Chawes, C.B.; et al. Maturation of the gut microbiome and risk of asthma in childhood. *Nat. Commun.* **2018**, *9*, 141. https://doi.org/10.1038/s41467-017-02573-2.
16. Gomez Perdiguero, E.; Klapproth, K.; Schulz, C.; Busch, K.; Azzoni, E.; Crozet, L.; Garner, H.; Trouillet, C.; Bruijn, D.F.M.; Geissmann, F.; et al. Tissue-resident macrophages originate from yolk-sac-derived erythro-myeloid progenitors. *Nature* **2015**, *518*, 547–551.
17. Schirmer, M.; Smeekens, S.P.; Vlamakis, H.; Jaeger, M.; Oosting, M.; Franzosa, E.A.; Horst, T.R.; Jansen, A.; Jacobs, L.; Bonder, J.M.; et al. . Linking the Human Gut Microbiome to Inflammatory Cytokine Production Capacity. *Cell* **2016**, *167*, 1897.
18. Martens, E.C.; Neumann, M.; Desai, M.S. Interactions of commensal and pathogenic microorganisms with the intestinal mucosal barrier. *Nat. Rev. Microbiol.* **2018**, *16*, 457–470.
19. Sereme, Y.; Mezouar, S.; Grine, G.; Mege, J.L.; Drancourt, M.; Corbeau, P.; Vitte, J.L. Methanogenic Archaea: Emerging Partners in the Field of Allergic Diseases. *Clin. Rev. Allergy Immunol.* **2019**, *57*, 456–466.
20. Dridi, B.; Henry, M.; El Khéchine, A.; Raoult, D.; Drancourt, M. High prevalence of Methanobrevibacter smithii and Methanosphaera stadtmanae detected in the human gut using an improved DNA detection protocol. *PLoS ONE* **2009**, *4*, e7063.
21. Gaci, N.; Borrel, G.; Tottey, W.; O'Toole, P.W.; Brugère, J.F. Archaea and the human gut: New beginning of an old story. *World J. Gastroenterol.* **2014**, *20*, 16062–16078.
22. Bang, C.; Weidenbach, K.; Gutsmann, T.; Heine, H.; Schmitz, R.A. The Intestinal Archaea Methanosphaera stadtmanae and Methanobrevibacter smithii Activate Human Dendritic Cells. *PLoS ONE* **2014**, *9*, e99411.
23. Vierbuchen, T.; Stein, K.; Heine, H. RNA is taking its Toll: Impact of RNA-specific Toll-like receptors on health and disease. *Allergy* **2019**, *74*, 223–235.
24. Bernatchez, E.; Gold, M.J.; Langlois, A.; Blais-Lecours, P.; Boucher, M.; Duchaine, C.; Marsolais, D.; McNagny, M.K.; Blanchet, R.M. Methanosphaera stadtmanae induces a type IV hypersensitivity response in a mouse model of airway inflammation. *Physiol. Rep.* **2017**, *5*, e13163.
25. Grine, G.; Boualam, M.A.; Drancourt, M. Methanobrevibacter smithii, a methanogen consistently colonising the newborn stomach. *Eur. J. Clin. Microbiol. Infect. Dis.* **2017**, *36*, 2449–2455.
26. Sereme, Y.; Guindo, C.O.; Filleron, A.; Corbeau, P.; Tran, T.A.; Drancourt, M.; Vitte, J.; Grine, G. Meconial Methanobrevibacter smithii suggests intrauterine methanogen colonization in preterm neonates. *Curr. Res. Microb. Sci.* **2021**, *2*, 100034. https://doi.org/10.1016/j.crmicr.2021.100034.
27. Baumann, R.; Untersmayr, E.; Zissler, U.M.; Eyerich, S.; Adcock, I.M.; Brockow, K.; Biedermann, T.; Ollert, M.; Chaker, M.A.; Pfaar, O.; et al. Noninvasive and minimally invasive techniques for the diagnosis and management of allergic diseases. *Allergy* **2021**, *76*, 1010–1023.
28. Bao, R.; Hesser, L.A.; He, Z.; Zhou, X.; Nadeau, K.C.; Nagler, C.R. Fecal microbiome and metabolome differ in healthy and food-allergic twins. *J. Clin. Investig.* **2021**, *131*, e141935.
29. Sereme, Y.; Zarza, S.M.; Medkour, H.; Amona, I.; Fenollar, F.; Akiana, J.; Mezouar, S.; Orain, N.; Vitte, J.; Davout, D.; et al. Stool Serology: Development of a Non-Invasive Immunological Method for the Detection of Enterovirus-Specific Antibodies in Congo Gorilla Faeces. *Microorganisms* **2021**, *9*, 810.

30. Lambert, C.; Sarrat, A.; Bienvenu, F.; Brabant, S.; Nicaise-Roland, P.; Alyanakian, M.-A.; Apoil, P.A.; Capron, C.; Couderc, R.; Evrard, B.; et al. The importance of EN ISO 15189 accreditation of allergen-specific IgE determination for reliable in vitro allergy diagnosis. *Allergy* **2015**, *70*, 180–186.
31. Drancourt, M.; Djemai, K.; Gouriet, F.; Grine, G.; Loukil, A.; Bedotto, M.; Levasseur, A.; Lépidi, H.; Khalil, B.J.; Khelaifia, S.; et al. Methanobrevibacter smithii archaemia in febrile patients with bacteremia, including those with endocarditis. *Clin. Infect. Dis.* **2020**, *73*, e2571–e2579. https://doi.org/10.1093/cid/ciaa998. Online ahead of print.
32. Grine, G.; Terrer, E.; Boualam, M.A.; Aboudharam, G.; Chaudet, H.; Ruimy, R.; Drancourt, D. Tobacco-smoking-related prevalence of methanogens in the oral fluid microbiota. *Sci. Rep.* **2018**, *8*, 9197.
33. Nkamga, V.D.; Huynh, H.T.T.; Aboudharam, G.; Ruimy, R.; Drancourt, M. Diversity of human-associated Methanobrevibacter smithii isolates revealed by multispacer sequence typing. *Curr. Microbiol.* **2015**, *70*, 810–815.
34. Guindo, C.O.; Davoust, B.; Drancourt, M.; Grine, G. Diversity of Methanogens in Animals' Gut. *Microorganisms* **2020**, *9*, 13.
35. Muraro, A.; Werfel, T.; Hoffmann-Sommergruber, K.; Roberts, G.; Beyer, K.; Bindslev-Jensen, C.; Bindslev-Jensen, C.; Cardona, V.; Dubois, A.; duToit, G.; et al. EAACI food allergy and anaphylaxis guidelines: Diagnosis and management of food allergy. *Allergy* **2014**, *69*, 1008–1025.
36. Papadopoulos, N.G.; Arakawa, H.; Carlsen, K.H.; Custovic, A.; Gern, J.; Lemanske, R.; Le Souef, P.; Mäkelä, M.; Roberts, G.; Wong, G.; et al. International consensus on (ICON) pediatric asthma. *Allergy* **2012**, *67*, 976–997.
37. Lisowska-Myjak, B.; Skarżyńska, E.; Wojdan, K.; Nasierowska-Guttmejer, A. Protein and peptide profiles in neonatal meconium: Classification of meconium proteins. *J. Obstet. Gynaecol. Res.* **2019**, *45*, 556–564.
38. Kolmannskog, S.; Marhaug, G.; Haneberg, B. Fragments of IgE Antibodies in Human Feces. *Int. Arch. Allergy Immunol.* **1985**, *78*, 358–363.
39. Szépfalusi, Z.; Pichler, J.; Elsässer, S.; van Duren, K.; Ebner, C.; Bernaschek, G.; Urbanek, R. Transplacental priming of the human immune system with environmental allergens can occur early in gestation. *J. Allergy Clin. Immunol.* **2000**, *106*, 530–536.
40. Msallam, R.; Balla, J.; Rathore, A.P.S.; Kared, H.; Malleret, B.; Saron, W.A.A.; Liu, Z.; Hang, W.J.; Dutertre, A.C.; Larbi, A.; et al. Fetal mast cells mediate postnatal allergic responses dependent on maternal IgE. *Science* **2020**, *370*, 941–950.
41. Lyons, J.J.; Yi, T. Mast cell tryptases in allergic inflammation and immediate hypersensitivity. *Curr. Opin. Immunol.* **2021**, *72*, 94–106.
42. Edogawa, S.; Edwinson, A.L.; Peters, S.A.; Chikkamenahalli, L.L.; Sundt, W.; Graves, S.; Gurunathan,V.S.; Breen-Lyles, K.M.; Johnson, S.; Dyer, B.R.; et al. Serine proteases as luminal mediators of intestinal barrier dysfunction symptom severity in IBS. *Gut* **2020**, *69*, 62–73.
43. Raithel, M.; Winterkamp, S.; Pacurar, A.; Ulrich, P.; Hochberger, J.; Hahn, E.G. Release of mast cell tryptase from human colorectal mucosa in inflammatory bowel disease. *Scand. J. Gastroenterol.* **2001**, *36*, 174–179.
44. Peterson, C.G.B.; Hansson, T.; Skott, A.; Bengtsson, U.; Ahlstedt, S.; Magnussons, J. Detection of local mast-cell activity in patients with food hypersensitivity. *J. Investig. Allergol. Clin. Immunol.* **2007**, *17*, 314–320.
45. Carroccio, A.; Brusca, I.; Mansueto, P.; Soresi, M.; D'Alcamo, A.; Ambrosiano, G.; Pepe, I.; Iacono, G.; Lospalluti, L.M.; Chiusa, L.M.S.; et al. Fecal assays detect hypersensitivity to cow's milk protein and gluten in adults with irritable bowel syndrome. *Clin. Gastroenterol. Hepatol.* **2011**, *9*, 965–971.e3.
46. Lettesjö, H.; Hansson, T.; Peterson, C.; Ung, K.-A.; Ringström, G.; Abrahamsson, H.; Simrén, M. Detection of inflammatory markers in stools from patients with irritable bowel syndrome and collagenous colitis. *Scand. J. Gastroenterol.* **2006**, *41*, 54–59.
47. Roca, M.; Rodriguez Varela, A.; Donat, E.; Cano, F.; Hervas, D.; Armisen, A.; Vaya, J.M.; Sjölander, A.; Ribes-Koninckx, C. Fecal Calprotectin and Eosinophil-derived Neurotoxin in Healthy Children Between 0 and 12 Years. *J. Pediatr. Gastroenterol. Nutr.* **2017**, *65*, 394–398.
48. Willers, M.; Ulas, T.; Völlger, L.; Vogl, T.; Heinemann, A.S.; Pirr, S.; Pagel, J.; Fehlhaber, B.; Halle, O.; Schöning, J.; et al. S100A8 and S100A9 Are Important for Postnatal Development of Gut Microbiota and Immune System in Mice and Infants. *Gastroenterology* **2020**, *159*, 2130–2145.e5.
49. Olin, A.; Henckel, E.; Chen, Y.; Lakshmikanth, T.; Pou, C.; Mikes, J.; Gustafsson, A.; Bernhardsson, K.A.; Zhang, C.; Bohlin, K.; et al. Stereotypic Immune System Development in Newborn Children. *Cell* **2018**, *174*, 1277–1292.e14.
50. Torow, N.; Marsland, B.J.; Hornef, M.W.; Gollwitzer, E.S. Neonatal mucosal immunology. *Mucosal Immunol.* **2017**, *10*, 5–17.
51. Goretzki, A.; Lin, Y.J.; Schülke, S. Immune metabolism in allergies, does it matter?—A review of immune metabolic basics and adaptations associated with the activation of innate immune cells in allergy. *Allergy* **2021**, *76*, 3314–3331. https://doi.org/10.1111/all.14843.
52. Macpherson, A.J.; de Agüero, M.G.; Ganal-Vonarburg, S.C. How nutrition and the maternal microbiota shape the neonatal immune system. *Nat. Rev. Immunol.* **2017**, *17*, 508–517.
53. Singh, N.; Gurav, A.; Sivaprakasam, S.; Brady, E.; Padia, R.; Shi, H.; Thangaraju, M.; Prasad, D.P.; Manicassamy, S.; Munn, H.D.; et al. Activation of Gpr109a, receptor for niacin and the commensal metabolite butyrate, suppresses colonic inflammation and carcinogenesis. *Immunity* **2014**, *40*, 128–139.
54. Furusawa, Y.; Obata, Y.; Fukuda, S.; Endo, T.A.; Nakato, G.; Takahashi, D.; Nakanishi, Y.; Uetake, C.; Kato, K.; Kato, T.; et al. Commensal microbe-derived butyrate induces the differentiation of colonic regulatory T cells. *Nature* **2013**, *504*, 446–450.
55. Atarashi, K.; Tanoue, T.; Ando, M.; Kamada, N.; Nagano, Y.; Narushima, S.; Suda, W.; Imaoka, A.; Setoyama, H.; Nagamori, T.; et al. Th17 Cell Induction by Adhesion of Microbes to Intestinal Epithelial Cells. *Cell* **2015**, *163*, 367–380.

56. Vonk, M.M.; Blokhuis, B.R.J.; Diks, M.A.P.; Wagenaar, L.; Smit, J.J.; Pieters, R.H.H.; Garssen, J.; Knippels, J.M.L.; Esch, V.M.A.C.B. Butyrate Enhances Desensitization Induced by Oral Immunotherapy in Cow's Milk Allergic Mice. *Mediat. Inflamm.* **2019**, *2019*, 9062537.
57. Roduit, C.; Frei, R.; Ferstl, R.; Loeliger, S.; Westermann, P.; Rhyner, C.; Schiavi, E.; Barcik, W.; Rodriguez-Perez, N.; Wawrzyniak, N.; et al. High levels of butyrate and propionate in early life are associated with protection against atopy. *Allergy* **2019**, *74*, 799–809.
58. Barnett, D.J.M.; Mommers, M.; Penders, J.; Arts, I.C.W.; Thijs, C. Intestinal archaea inversely associated with childhood asthma. *J. Allergy Clin. Immunol.* **2019**, *143*, 2305–2307.
59. Gomulka, K.; Liebhart, J.; Jaskula, E.; Lange, A.; Medrala, W. The –2549 –2567 del18 Polymorphism in VEGF and Irreversible Bronchoconstriction in Asthmatics. *J. Investig. Allergol. Clin. Immunol.* **2019**, *29*, 431–435.