

ÉCOLE NORMALE SUPÉRIEURE DE LYON

Concours d'admission session 2024

Filière universitaire : Second concours

## COMPOSITION DE BIOLOGIE-BIOCHIMIE

Durée : 3 heures

L'utilisation des calculatrices n'est pas autorisée pour cette épreuve.

\*\*\*

# **Le SARS-CoV2, le virus à l'origine de la pandémie de COVID-19**

L'épreuve comprend un sujet de synthèse et deux sous-parties indépendantes d'analyse de documents, le temps conseillé pour chaque partie est indiqué.

Lors de l'évaluation, une importance particulière sera attachée à :

- la rigueur scientifique
- la justification des raisonnements
- la clarté et la concision des réponses
- la qualité et la précision des illustrations
- l'orthographe, la grammaire et la présentation.

## **Partie I – Sujet de synthèse (temps conseillé 45 min)**

### ***La diversité des virus***

Vous rédigerez un exposé structuré et illustré rappelant la définition des virus et traitant de leur diversité. Votre synthèse comprendra une introduction et une conclusion, des exemples et des illustrations, et fera entre 3 et 5 pages.

## **Partie II - Analyse de documents**

Cette analyse comporte deux sous-parties totalement indépendantes à travers l'étude de deux articles publiés en 2020 s'intéressant respectivement à l'entrée du virus dans les cellules et à la réponse immunitaire au cours de l'infection par le SARS-CoV2. Une courte question de conclusion reposant sur les deux parties vous est proposée à la fin de l'exercice.

### ***1. Entrée du virus (temps conseillé 1h10)***

Le SARS-CoV2 est un virus enveloppé exprimant la glycoprotéine Spike à sa surface. Dès le début de la pandémie, la protéine ACE-2 (Angiotensin Converting Enzyme 2, exprimée par de nombreuses cellules à leur surface) a été identifiée comme étant le principal « récepteur » du virus. Le récepteur est une protéine avec laquelle la protéine Spike interagit, ce qui entraîne la fusion de la membrane virale avec la membrane cellulaire grâce aux propriétés d'un domaine particulier de la protéine Spike, appelé peptide fusion. Cela permet l'entrée de protéines et du génome viral dans les cellules.

Les capacités d'entraîner la fusion membranaire de cette protéine Spike sont telles que lorsque l'on infecte des cellules par le virus, on constate la fusion de cellules adjacentes, formant des *syncytia* qui ont pu être observés *in vitro* comme *in vivo* dans certains tissus.

Les auteurs de cette première étude ont mis au point un système permettant de détecter facilement la formation de *syncytia in vitro*: ils mélangent des quantités équivalentes de deux groupes de cellules en culture exprimant toutes le récepteur ACE-2 sur leur membrane plasmique et chacune une moitié de la protéine fluorescente verte GFP dans leur cytosol (GFP-1, correspondant à la première moitié de la GFP, est exprimée par le premier groupe de cellules, et GFP-2, correspondant à la deuxième moitié de la GFP, est exprimée par le second groupe de cellules). Ils infectent ensuite ce mélange de cellules par le virus SARS-CoV2. La fluorescence verte de la GFP ne sera visible que si les deux moitiés de la protéine se retrouvent au sein d'un *syncytium* formé par la fusion de cellules. Les auteurs mesurent la formation des *syncytia* en

mesurant l'étendue (la surface) de cellules présentant une fluorescence verte en microscopie à fluorescence.

Pour mesurer la viabilité cellulaire, ils ajoutent de l'Iodure de Propidium (IP, noté PI sur la figure 1). L'IP est un agent qui fluoresce en rouge lorsqu'il s'intercale dans l'ADN, mais il ne peut pas traverser la membrane plasmique. Il ne marque donc que le noyau des cellules ayant perdu leur intégrité membranaire, c'est-à-dire de cellules endommagées en train de mourir et permet donc d'étudier la viabilité cellulaire.

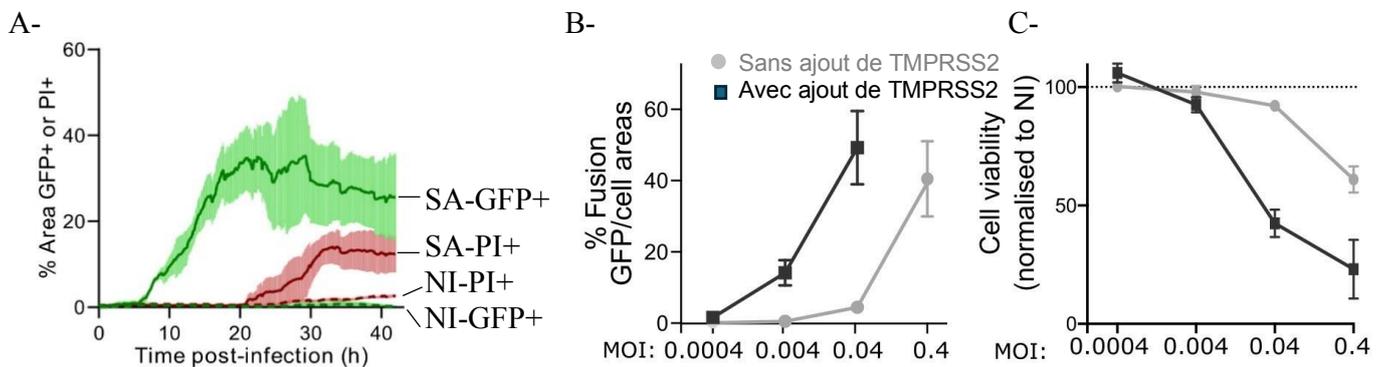
**Question 1** : Expliquez le principe de la fluorescence et les paramètres à vérifier pour pouvoir travailler avec l'IP et la GFP en même temps.

**Question 2** : Résumez schématiquement le principe de l'expérience en faisant figurer des cellules, le virus, les protéines GFP-1, GFP-2 et ACE-2.

**Question 3** : Quels sont selon vous l'intérêt et les limites d'étudier l'entrée virale à travers la formation de syncytia? Quelle peut-être la conséquence de la formation de syncytia *in vivo* ?

Les auteurs ont tout d'abord réalisé des observations à différents temps et différentes multiplicités d'infection (MOI = Multiplicity Of Infection : elle représente la quantité de virus utilisée pour l'infection, comptabilisée en nombre de particules virales par cellule).

Ils ont réalisé une première expérience avec une MOI de 0,4 en mesurant l'évolution des surfaces fluorescentes à la GFP ou à l'IP au cours du temps (figure 1A) ou en utilisant différentes MOI (figure 1B).



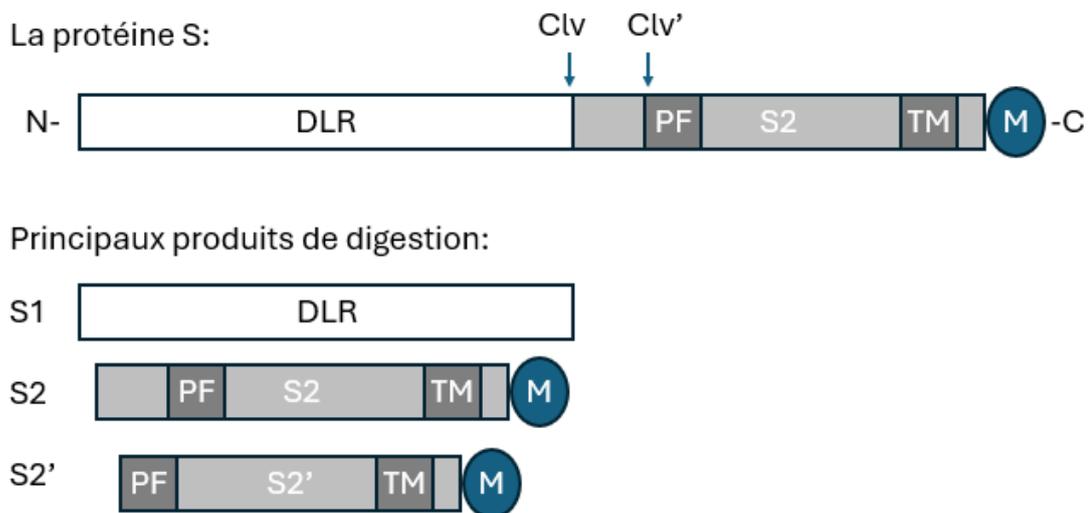
**Figure 1 : Analyse de la formation de syncytia et de la viabilité de cellules exprimant chacune une moitié de la GFP, suite à l'infection par le SARS-CoV2.** La formation de syncytia est quantifiée en % de la surface de culture cellulaire correspondant à des cellules fluorescentes GFP+ (% Area GFP+ ou % Fusion GFP/cell areas qui correspondent au pourcentage de la surface de culture composée de cellules fusionnées contenant les deux moitiés de la GFP). La viabilité cellulaire (% Area PI+ ou Cell viability) est quantifiée en % de la surface de culture cellulaire occupée par des cellules fluorescentes au noyau marqué par l'Iodure de Propidium PI+, normalisée par rapport au contrôle non infecté (normalised to NI).

A- Evolution sur 44h de la formation de syncytia et de la viabilité cellulaire suite à une infection par le SARS-Cov2 à une MOI de 0,4. SA=cellules infectées par le SARS-CoV2 ; NI= cellules non infectées. La courbe représente la moyenne de 3 à 9 mesures.

B- Comparaison de la formation de syncytia et C- de la viabilité cellulaire 20h après infection à différentes MOI, avec ou sans ajout de la protéine TMPRSS2 dans les cellules (voir question 8). La courbe représente la moyenne de 3 à 9 expériences et les barres d'erreurs les écarts-types, en absence de surexpression de TMPRSS2 en gris, avec ajout de TMPRSS2 en noir.

**Question 4 :** Analysez ces résultats. Pour les figures 1B et 1C, vous n'analysez pour l'instant que le résultat obtenu en absence de surexpression de TMRSS2 (donc uniquement les courbes grises). Justifiez pourquoi pour la plupart des expériences suivantes, les auteurs ont réalisé les expériences à une MOI de 0,2 et une observation après 20h.

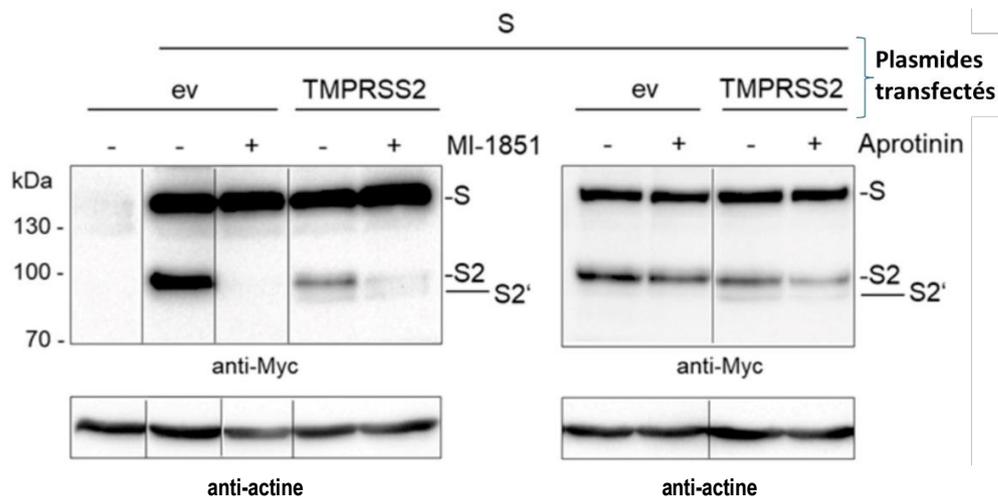
Les auteurs se sont ensuite intéressés à la protéine Spike (S), connue depuis 2020 pour posséder deux sites de clivage par des protéases (notés Clv et Clv' sur la figure 2). La structure de la protéine S utilisée et des principaux produits obtenus après clivage sont résumés dans la figure 2. Pour les besoins de cette étude, une étiquette « Myc » a été ajoutée à l'extrémité C-terminale de la protéine, ce qui permet de détecter facilement la protéine avec un anticorps dirigé contre l'étiquette Myc.



**Figure 2 : Structure de la protéine Spike (S) du SARS-CoV2.** La protéine possède deux sites de clivage Clv et Clv'. Les principales protéines produites après clivage (S1, S2 et S2') sont indiquées. DLR : Domaine de Liaison au Récepteur. PF : Peptide Fusion. TM : domaine TransMembranaire. N- : extrémité N-terminale. -C : extrémité C-terminale. Pour les besoins de l'expérience, une étiquette Myc (M) a été ajoutée au niveau de l'extrémité C-terminale de la protéine.

Deux protéases sont suspectées de cliver la protéine Spike. Il s'agit de la protéase intracellulaire appelée Furine, constitutivement exprimée par de nombreuses cellules et impliquée dans la maturation de nombreuses protéines cellulaires et glycoprotéines virales. L'autre protéase s'appelle TMRSS2 (Type II transmembrane serine protease), c'est une protéase exprimée à la surface de certaines cellules qui est connue pour cliver différentes protéines exprimées à la surface ou en contact avec la membrane plasmique. Afin d'étudier le rôle respectif de ces deux protéases dans la maturation de la protéine S, les auteurs ont analysé par Western-Blot le clivage de la protéine S dans des cellules qui expriment la Furine, mais pas la protéine TMRSS2. Ces cellules ont été transfectées par un plasmide permettant la production de la protéine Spike (S), et par un second plasmide n'exprimant aucune protéine (ev) ou exprimant la protéine TMRSS2 (TMRSS2). L'effet d'un inhibiteur de la Furine (MI-1851) ou de TMRSS2 (Aprotinin) ont aussi été testés.

**Question 5 :** Quelles sont les différentes étapes permettant de réaliser un Western-Blot ?



**Figure 3 : Analyse du clivage de la protéine S en présence de Furine et en présence ou en absence de TMPRSS2 et d'inhibiteurs de la Furine (MI-1851) ou de TMPRSS2 (Aprotinin).** Les plasmides transfectés dans des cellules qui expriment la Furine, sont indiqués en haut de la figure (S : plasmide exprimant la protéine S, ev : plasmide vide, TMPRSS2 : plasmide exprimant TMPRSS2). Par Western-Blot, la protéine Spike (S) et ses produits de digestion (S2 et S2' présentés à la figure 2) ont été détectées avec un anticorps anti-Myc, et l'actine, protéine exprimée de manière constante dans les cellules, a été détectée avec un anticorps anti-actine. Les chiffres indiqués à gauche correspondent à l'échelle de poids moléculaire, et au poids apparent des protéines exprimé en kiloDaltons. Remarque : la différence du niveau d'expression de S entre les expériences avec les deux inhibiteurs ne sera pas prise en compte.

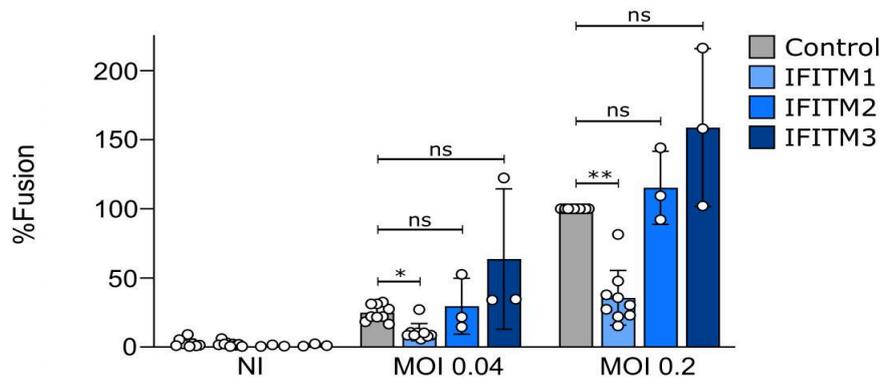
**Question 6 :** Pourquoi a-t-on analysé l'expression de l'actine ? Pourquoi a-t-on réalisé une partie de l'expérience en transfectant un vecteur vide en plus du plasmide S ? Pourquoi ne détecte-t-on pas la protéine S1 ?

**Question 7 :** Analysez les résultats obtenus et déduisez-en les sites de clivage respectifs de la Furine et de TMPRSS2.

Connaissant ces résultats, les auteurs de la première étude ont renouvelé leur expérience en utilisant des cellules exprimant TMPRSS2 en plus d'ACE-2. Il s'agit des résultats présentés en noir dans les figures 1B et 1C.

**Question 8 :** Analysez maintenant l'ensemble des résultats présentés dans les panneaux 1B et 1C. Quel est l'effet de la présence de la protéine TMPRSS2 lors de l'infection ? Quel effet pourrait donc avoir la protéase TMPRSS2 sur l'entrée du virus dans les cellules ?

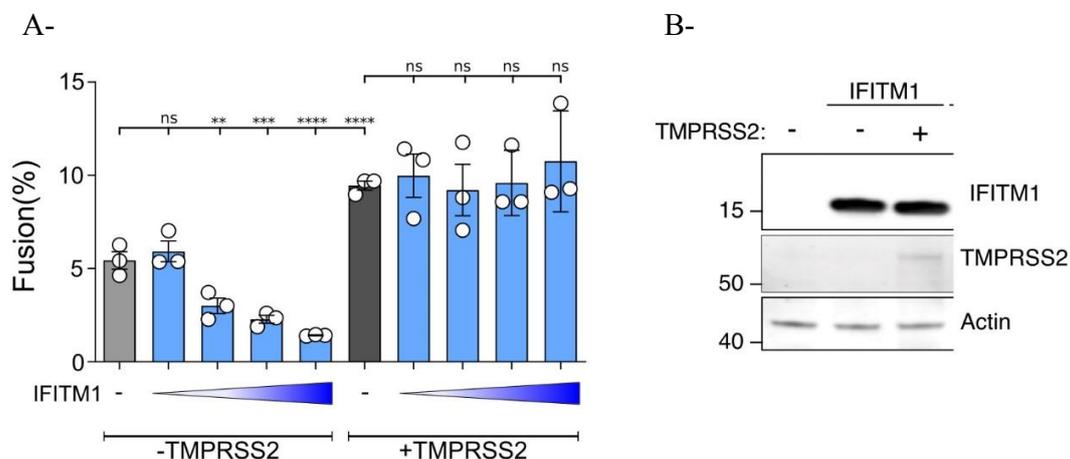
Les auteurs ont ensuite voulu étudier l'influence d'autres protéines connues pour moduler les infections virales : trois protéines de la famille IFITM (protéines transmembranaires induites par l'interféron), IFITM1, IFITM2 et IFITM3. Ils ont utilisé le même système de formation de syncytia fluorescents à la GFP dont l'étendue de la surface est proportionnelle à la fusion cellulaire.



**Figure 4 : Influence de la présence des protéines IFITM1, 2 et 3 sur la formation de syncytia lors de l'infection de cellules par le SARS-CoV2.** Des mélanges de cellules exprimant chacune une moitié de la GFP (voir figure 1) et en plus IFITM1,2 ou 3 ou aucun IFITM (Control) ont été infectés par le SARS-CoV2 à une MOI de 0,04 ou 0,2 (ou non infectées NI). En abscisse est indiqué le % de Fusion calculé en mesurant la surface occupée par des cellules fluorescentes à la GFP, normalisé à 100% pour la surface obtenue après 20h d'infection à une MOI de 0,2. Les histogrammes représentent la moyenne de 3 à 9 expériences (chaque point représentant le résultat d'une expérience), et les barres correspondent à l'écart-type. Les résultats ont été analysés avec un test statistique de type ANOVA, ns signifie non significatif, alors que plus le nombre d'étoiles est élevé plus les différences observées sont significatives.

**Question 9 :** Quel est l'effet de la présence des protéines IFITM1, IFITM2 ou IFITM3 sur la formation des syncytia ?

A la place des lignées cellulaires utilisées à la figure 4, exprimant chacune un niveau donné d'IFITM, les auteurs ont ensuite utilisé des cellules permettant de moduler l'expression des protéines IFITM1 et TMPRSS2. Ils ont alors pu comparer l'effet de quantités croissantes de protéines IFITM1 sur la formation de syncytia en présence ou en absence de TMPRSS2.



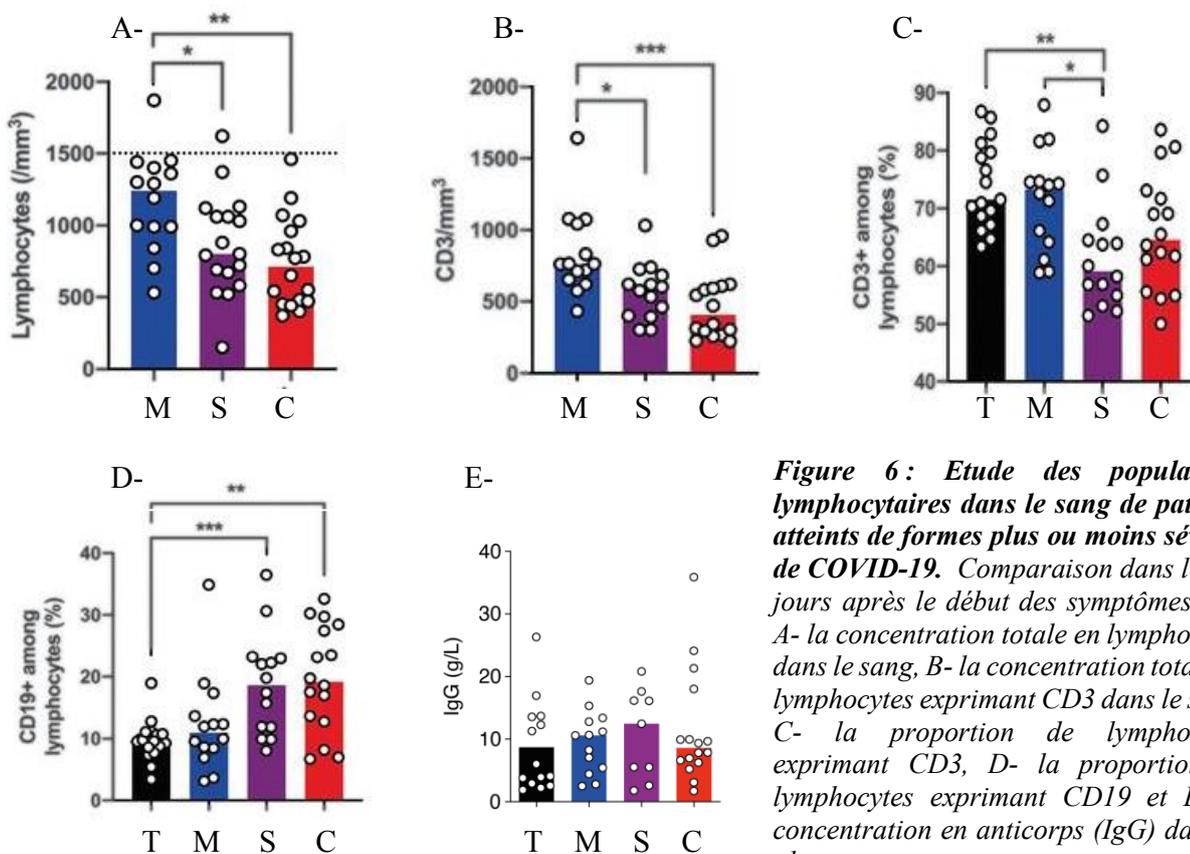
**Figure 5 : Analyse de la formation de syncytia par des cellules infectées par le SARS-CoV2 en présence des protéines IFITM1 et TMPRSS2.** A- Analyse de la fusion cellulaire (en pourcentage de la surface totale de la culture occupée par des cellules fusionnées) en présence de quantités croissantes d'IFITM1 et en présence (à droite) ou en absence (à gauche) de TMPRSS2. B- Western-Blot permettant de détecter les protéines IFITM1, TMPRSS2 et actine. Les chiffres indiqués à gauche correspondent à l'échelle de poids moléculaire exprimés en kiloDaltons.

**Question 10 :** Analysez la figure 5. Quelle hypothèse ont selon vous voulu tester les auteurs en réalisant le Western-Blot présenté à la Figure 5B ? Quels sont les contrôles manquants ? Quelles expériences complémentaires souhaiteriez-vous réaliser ? Proposez à l'aide d'un schéma une conclusion sur les effets respectifs de TMPRSS2 et IFITM1 sur la formation des syncytia et l'entrée du virus dans les cellules.

## 2. Réponse immunitaire au SARS-CoV2 (temps conseillé 1h)

En 2020, la réponse immunitaire d'un groupe de 50 patients COVID-19 présentant des symptômes modérés (15 patients), sévères (17 patients) ou critiques (18 patients) a été analysée en détail, en comparaison de 18 contrôles sains.

Les lymphocytes sont des acteurs centraux de la réponse immunitaire. Ils comprennent deux populations majeures : les lymphocytes T, qui vont notamment reconnaître et détruire les cellules infectées et qui expriment le marqueur CD3 à leur surface, et les lymphocytes B, qui produisent des anticorps et expriment le marqueur CD19 à leur surface. Les auteurs ont commencé par étudier les différentes populations lymphocytaires et la quantité d'anticorps présentes dans le sang des patients.

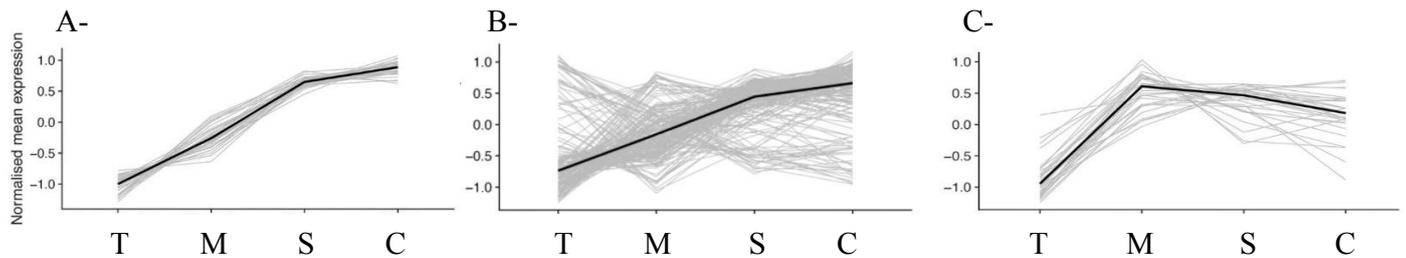


**Figure 6 :** Etude des populations lymphocytaires dans le sang de patients atteints de formes plus ou moins sévères de COVID-19. Comparaison dans les 10 jours après le début des symptômes de : A- la concentration totale en lymphocytes dans le sang, B- la concentration totale de lymphocytes exprimant CD3 dans le sang, C- la proportion de lymphocytes exprimant CD3, D- la proportion de lymphocytes exprimant CD19 et E- la concentration en anticorps (IgG) dans le plasma.

T = Témoin non malade, M = patient Modérément atteint, S = patient Sévèrement atteint, C = patient dans un état Critique. Chaque point représente un patient. La barre d'histogramme représente la moyenne. Le test statistique utilisé est le test de Kruskal-Wallis, plus le nombre d'étoiles est élevé, plus la différence est significative (l'absence d'indication signifie une absence de différence significative). La ligne en pointillés représente le niveau du témoin.

**Question 11 :** Analysez l'ensemble des résultats présentés à la figure 6 et tentez d'expliquer les différences observées en fonction de la gravité de la maladie.

Les auteurs ont analysé l'expression des gènes dans les lymphocytes de ces patients, et se sont en particulier intéressés à l'ensemble des gènes de la réponse inflammatoire et immunitaire (figure 7A), à l'expression des gènes de cytokines (messagers spécifiques de la réponse immunitaire qui agissent de façon très localisée) et des gènes activés par ces cytokines (figure 7B) et enfin des gènes de la réponse interféron, une réponse connue pour jouer un rôle clé dans la réponse immunitaire antivirale (figure 7C).

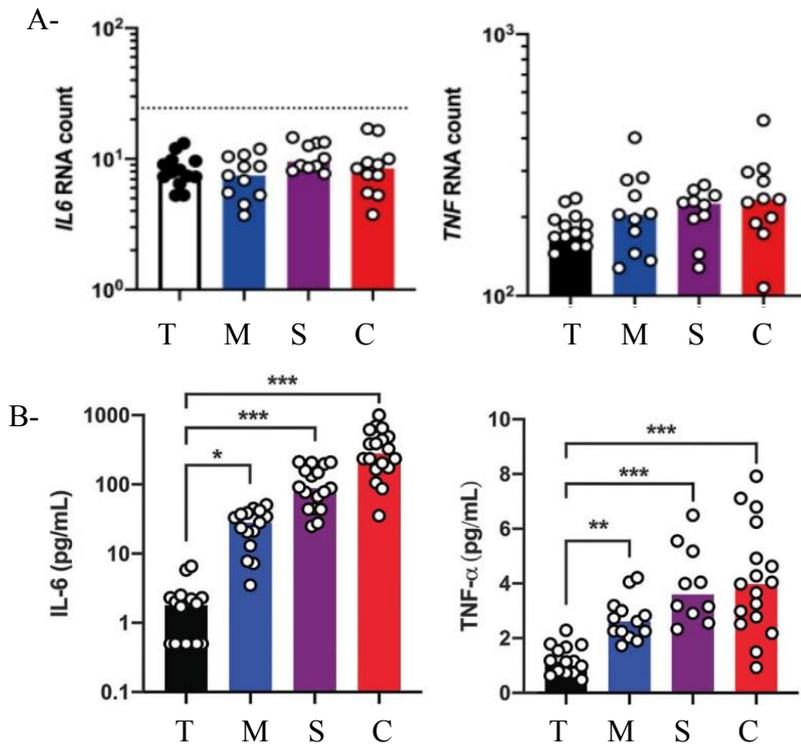


**Figure 7 :** Analyse de l'expression de gènes impliqués dans la réponse immunitaire par les lymphocytes de patients atteints de formes plus ou moins sévères de COVID-19. A- Ensemble des gènes de la réponse inflammatoire et immunitaire. B- Gènes de cytokines et de réponses à l'activation par les cytokines. C- Gènes de la réponse interféron. Le niveau moyen d'expression de ces gènes a été normalisé par rapport à l'expression de gènes dont l'expression est constante (« Normalised mean expression »), et représenté en fonction de la gravité de la maladie. Les courbes en gris relient des points individuels et ne seront pas interprétées, la courbe noire représente la moyenne de tous les patients d'un groupe. T = Témoin non malade, M = patient Modérément atteint, S = patient Sévèrement atteint, C = patient dans un état Critique.

**Question 12 :** A votre avis, quelle(s) technique(s) ont employé les auteurs pour mesurer le niveau d'expression des gènes dans les cellules sanguines ? Quelles en sont les étapes essentielles ?

**Question 13 :** Analysez les résultats obtenus pour les différents types de gènes analysés, en fonction de la gravité de la maladie.

Une inflammation excessive est connue pour être observée dans les cas les plus graves de la maladie et notamment une production excessive des cytokines inflammatoires comme l'InterLeukine 6 (IL-6) et le Tumor Necrosis Factor-alpha (noté TNF- $\alpha$ , TNF-  $\alpha$ 2 ou TNF). Les auteurs ont analysé d'une part l'expression des gènes codant ces cytokines par les cellules sanguines (figure 8A) et d'autre part la quantité de cytokines inflammatoires présentes dans le sang (figure 8B).



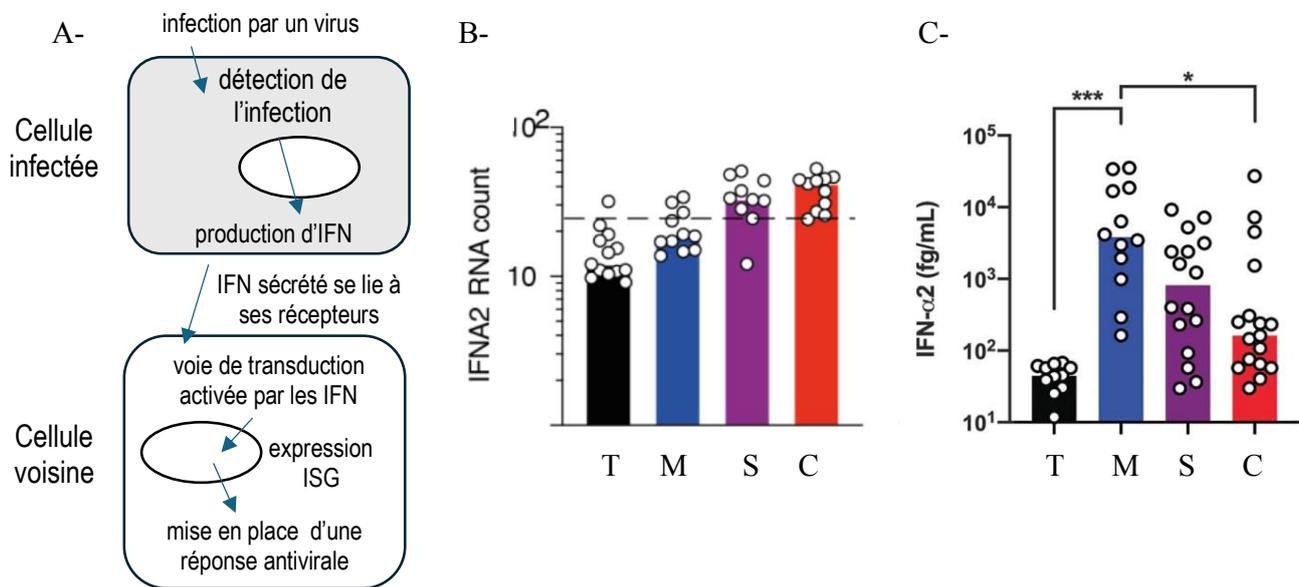
**Figure 8 : Analyse du niveau d'ARN messagers d'IL6 et de TNF exprimés dans les lymphocytes (A-) et du taux de protéines IL6 et TNF détectés (B-) dans le sang de patients atteints de formes plus ou moins sévères de COVID-19.** A- La quantité d'ARN est mesurée en unités arbitraires (RNA count). La ligne pointillée représente le seuil de détection. B- la quantité de cytokines est exprimée en pg/mL de plasma. T = Témoin non malade, M = patient Modérément atteint, S = patient Sévèrement atteint, C = patient dans un état Critique. Chaque point représente un patient. La barre d'histogramme représente la moyenne. Le test statistique utilisé est le test de Kruskal-Wallis, plus le nombre d'étoiles est élevé, plus la différence est significative (l'absence d'indication signifie une absence de différence significative).

**Question 14 :** Analysez les résultats obtenus pour ces deux cytokines. Formulez une ou des hypothèse(s) pour les expliquer.

La réponse interféron (IFN), représentée schématiquement dans la figure 9A, est déclenchée par l'entrée d'un virus dans une cellule, qui va se mettre à produire et sécréter des interférons, molécules qui vont, par une voie de transduction spécifique, provoquer une réponse antivirale dans les cellules avoisinantes en activant l'expression de gènes (dont les IFITM vus précédemment), collectivement appelés ISG (Interferon-Stimulated Genes). Les protéines antivirales produites à partir de ces gènes permettent en général de lutter efficacement contre l'infection virale.

Les auteurs ont étudié les gènes de la voie de transduction activée par les interférons (données non présentées) : leur expression augmente avec la gravité de la maladie, les cellules conservent donc, voire augmentent, leur capacité à développer une réponse antivirale suite à une stimulation par l'interféron.

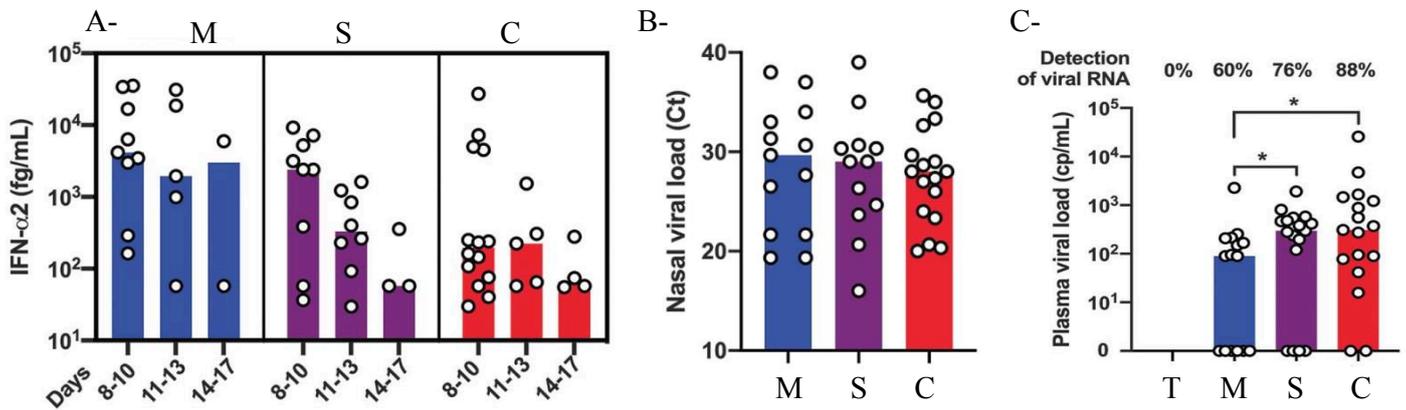
Ils ont aussi étudié la production de l'un des interférons, l'interféron A2 (IFNA2 ou IFN-α2) en mesurant la quantité d'ARN codant l'IFNA2 dans les cellules sanguines (figure 9B) et la quantité de protéines interféron IFN-α2 dans le plasma (figure 9C).



**Figure 9 : Analyse de la réponse interféron dans le sang de patients atteints de formes plus ou moins sévères de COVID-19.** A- Représentation schématique simplifiée de la voie interféron (IFN). B- Analyse de la quantité d'ARNm codant l'IFN2A dans les lymphocytes, exprimée en unités arbitraires, la ligne pointillée représentant la moyenne des mesures. C- Mesure de la quantité de protéine IFN-α2 dans le plasma, en fg/mL. T = Témoin non malade, M = patient Modérément atteint, S = patient Sévèrement atteint, C = patient dans un état Critique. Chaque point représente un patient. La barre d'histogramme représente la moyenne. Le test statistique utilisé est le test de Kruskal-Wallis, plus le nombre d'étoiles est élevé, plus la différence est significative (l'absence d'indication signifie une absence de différence significative). IFN : interféron. ISG : Gènes Stimulés par l'Interféron.

**Question 15 :** Analysez les documents présentés à l'aide des informations données. Quelle est l'étape de la réponse interféron qui semble la plus affectée lorsque la gravité de la maladie augmente ?

Les auteurs ont enfin présenté les données complémentaires suivantes : l'évolution de la concentration plasmatique en IFN-α2 au cours du temps (figure 10A) et la mesure de la charge virale au niveau nasal (figure 10B) et la concentration en ARN viral dans le sang (figure 10C), 10 jours après le début des symptômes.



**Figure 10 :** Analyse de paramètres immunologiques et virologiques dans le sang et le nez de patients atteints de formes plus ou moins sévères de COVID-19 10 jours après le début des symptômes. A- Evolution de la concentration plasmatique en protéine IFN- $\alpha$ 2 au cours du temps (l'analyse statistique n'a pas été réalisée pour cette figure complémentaire). B- Mesure de la charge virale en unités arbitraire (Ct) au niveau nasal. C- Mesure de la charge virale dans le plasma. T = Témoin non malade, M = patient Modérément atteint, S = patient Sévèrement atteint, C = patient dans un état Critique. Chaque point représente un patient. La barre d'histogramme représente la moyenne. Le test statistique utilisé est le test de Kruskal-Wallis, plus le nombre d'étoiles est élevé, plus la différence est significative (l'absence d'indication signifie une absence de différence significative, sauf en A-).

**Question 16 :** Analysez la figure 10. Que vous apprend-elle de nouveau ?

**Question 17 :** D'après ces données, quel(s) paramètre(s) biologiques suggèreriez-vous d'analyser pour estimer la sévérité de la maladie chez un patient se présentant à l'hôpital 10 jours après le début des symptômes ? Quels paramètres biologiques ne seraient au contraire pas de bons indicateurs ?

### 3. Elargissement (5 min)

**Question 18 :** A la lumière des données étudiées dans la partie II, quel(s) traitement(s) suggèreriez-vous de tester contre le SARS-CoV2 ? Vous pouvez aussi comparez ces stratégies avec les autres méthodes de lutte contre le SARS-CoV2 que vous pouvez connaître (maximum 15 lignes).

## Bibliographie

Certains éléments ont été simplifiés ou omis pour les besoins de l'exercice.

### Sous-Partie 1

Buchrieser J *et al.*, EMBO J. 2020; 39 (23)

Bestle D *et al.*, Life Sci Alliance. 2020; 3 (9)

### Sous-Partie 2

Hadjadj *et al.*, Science. 2020; 369 (718–724)