



**HAL**  
open science

# Approche métabolomique pour l'identification de marqueurs diagnostiques et pronostiques dans la COVID-19

Céline Occelli

► **To cite this version:**

Céline Occelli. Approche métabolomique pour l'identification de marqueurs diagnostiques et pronostiques dans la COVID-19. Médecine humaine et pathologie. Université Côte d'Azur, 2024. Français. NNT : 2024COAZ6009 . tel-04771354

**HAL Id: tel-04771354**

**<https://theses.hal.science/tel-04771354v1>**

Submitted on 7 Nov 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# THÈSE DE DOCTORAT

## Approche métabolomique pour l'identification de marqueurs diagnostiques et pronostiques dans la COVID-19

**Céline OCCELLI**

Laboratoire TIRO-MATOs (UMR E4320)

**Présentée en vue de l'obtention  
du grade de docteur en Science de la Vie  
et de la Santé  
d'Université Côte d'Azur**

**Dirigée par :** Thierry POURCHER  
**Co-dirigée par :** Jacques LEVRAUT  
**Soutenue le :** 09 juillet 2024

**Devant le jury, composé de :**

Sonia DAGNINO, PU, UniCA, Présidente  
du Jury  
Sandrine CHARPENTIER, PUPH,  
Université Toulouse III, Rapportrice  
Emmanuel MONTASSIER, PUPH, Nantes  
Université, Rapporteur  
Thierry Pourcher, DR, UniCA, Directeur de  
Thèse  
Jacques LEVRAUT, PUPH, UniCA, Co-  
Directeur de thèse, Membre invité

Université Côte d'Azur  
École Doctorale des Sciences de la Vie et de la Santé (ED85)

THESE DE DOCTORAT

Présentée par Céline OCCELLI

**Approche métabolomique  
pour l'identification de marqueurs diagnostiques et pronostiques  
dans la COVID-19**

**Présentée en vue de l'obtention du grade de docteur** en Sciences de la Vie et de la Santé  
d'Université Côte d'Azur

**Dirigée par :** Thierry Pourcher & Pr Jacques Levraut

**Soutenue le :** 9 juillet 2024

Pr Sonia DAGNINO	PU, Université Côte d'Azur, Polytech Nice Sophia	Présidente du Jury
Pr Sandrine CHARPENTIER	PU-PH, Université Toulouse III, CHU Toulouse Rangueil	Rapporteuse
Pr Emmanuel MONTASSIER	PU-PH, Nantes Université, CHU de Nantes	Rapporteur
Dr Thierry POURCHER	Directeur de recherche CEA, Université Côte d'Azur, TIRO, UMR E4320	Directeur de thèse
Pr Jacques LEVRAUT	PU-PH, Université Côte d'Azur, CHU de NICE	Co-Directeur de thèse

Laboratoire TIRO (Transporteurs en Imagerie et Radiothérapie en Oncologie)  
Unité mixte de recherche E4320  
Commissariat à l'Energie Atomique et aux énergies alternatives (CEA) / Université Côte  
d'Azur (UCA)

## RESUME

### **Approche métabolomique pour l'identification de marqueurs diagnostiques et pronostiques dans la COVID-19**

---

Le coronavirus 2 du syndrome respiratoire aigu sévère (SARS-CoV-2) a été responsable d'une pandémie mondiale entachée d'une morbi-mortalité élevée et a eu un impact socio-économique, ainsi que sur le système de santé, majeur. La maladie à coronavirus 2019 (COVID-19) possède un spectre clinique très large, allant de la forme asymptomatique au Syndrome de Détresse Respiratoire Aigu et à la défaillance multiviscérale pouvant tous deux conduire au décès, reflet d'une physiopathologie complexe, dont la connaissance progresse encore actuellement. Dans ce contexte, l'enjeu pour le clinicien est de pouvoir identifier rapidement les patients atteints de COVID-19 et de stratifier le risque de mauvaise évolution clinique. La métabolomique est une approche omique permettant une analyse dynamique des changements moléculaires en réponse à une maladie à l'échelle de l'individu. Elle permet, par le biais d'un profilage métabolique, un phénotypage, une identification de biomarqueurs et des voies métaboliques impactées. L'objectif de ce travail était d'identifier des signatures métabolomiques plasmatiques et urinaires diagnostiques et pronostiques de la COVID-19 et d'analyser les voies métaboliques impactées.

Dans notre première étude, basée sur une approche d'analyse en métabolomique non ciblée sur échantillons plasmatiques de sujets atteints de la COVID-19 et de volontaires sains, nous avons identifié des signatures métabolomiques plasmatiques diagnostiques et pronostiques robustes, impliquant principalement des modifications des taux plasmatiques de lipides et d'acides aminés. Nous avons retrouvé notamment des niveaux élevés de lipides insaturés et des niveaux faibles de lipides saturés dans le plasma des sujets infectés. Nous avons également identifié de nouveaux biomarqueurs plasmatiques putatifs pour l'infection par le SARSCoV-2 et la sévérité de la maladie comme des désoxy-fructosyl-aminoacides et confirmé également l'impact de la voie de la kynurénine dans ces signatures. Dans notre seconde étude, s'intéressant à l'analyse métabolomique d'échantillons urinaires, nous retrouvons également des signatures métabolomiques urinaires diagnostiques et pronostiques performantes, objectivant l'impact du métabolisme urinaire des acides aminés, des purines, du métabolisme énergétique et de la voie de la kynurenine.

Nos résultats confirment l'apport de la métabolomique dans la compréhension de la physiopathologie de la COVID-19 et son intérêt dans la découverte de nouveaux biomarqueurs diagnostiques et pronostiques de la maladie. L'identification et surtout la validation des différents « métabotypes » spécifiques de la maladie, de son évolution et de sa gravité pourrait permettre, en plus de développer des tests diagnostiques robustes et moins invasifs, de prédire le risque de mauvaise évolution clinique d'une infection comme le COVID-19.

*Mots clefs : Métabolomique, COVID-19, Biomarqueurs, Diagnostic, Pronostic*

## **Metabolomics approach for identification of diagnostic and prognostic biomarkers in COVID-19 disease**

---

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has been the source of a global pandemic with high morbidity and mortality, and has had a major socio-economic impact, as well as on the healthcare system. Coronavirus 2019 disease (COVID-19) has a very broad clinical spectrum, from the asymptomatic form to Acute Respiratory Distress Syndrome and multivisceral failure, both of which can lead to death, reflecting a complex pathophysiology, the understanding of which is still improving. In this context, the challenge for clinicians is to be able to quickly identify patients with COVID-19 and stratify the risk of poor clinical outcome. Metabolomics is a omic approach that enables dynamic analysis of molecular changes in response to disease at the individual level. Through metabolic profiling and phenotyping, it enables the identification of biomarkers and the metabolic pathways impacted. The aim of this work was to identify diagnostic and prognostic plasma and urine metabolomic signatures of COVID-19 and to analyze the metabolic pathways affected.

In our first study, based on a non-targeted metabolomics analysis approach on plasma samples from COVID-19-infected subjects and healthy volunteers, we identified robust diagnostic and prognostic plasma metabolomic signatures, mainly involving changes in plasma lipid and amino acid levels. Especially, we found high levels of unsaturated lipids and low levels of saturated lipids in the plasma of infected subjects. We also identified novel plasma biomarkers for SARSCoV-2 infection and disease severity as deoxy-fructosyl-amino acids, and confirmed the impact of the kynurenine pathway in these signatures.

In our second study, based on non-targeted metabolomic analysis of urinary samples, we also found robust diagnostic and prognostic urinary metabolomic signatures, highlighting the impact of urinary amino acid metabolism, purine metabolism, energy metabolism and the kynurenine pathway.

Our results confirm the contribution of metabolomics to understanding the pathophysiology of COVID-19, and its interest in finding new diagnostic and prognostic biomarkers for the condition. The identification and, above all, validation of different “metabotypes” specific to the disease, its course and severity could enable us, in addition to developing robust and less invasive diagnostic tests, to predict the risk of poor clinical outcome of an infection such as COVID-19.

*Key words* : Metabolomics, COVID-19, Biomarkers, Diagnosis, Pronostic

# TABLE DES MATIERES

RESUME .....	2
ABREVIATIONS .....	6
LISTE DES FIGURES .....	8
<b>Partie I : INTRODUCTION .....</b>	<b>10</b>
<i>Chapitre 1 : La COVID-19.....</i>	<i>11</i>
I. L'enjeu de la découverte d'une nouvelle pathologie.....	11
II. Vers une meilleure compréhension de la physiopathologie.....	14
III. La nécessité de développer des outils diagnostiques et pronostiques .....	16
<i>Chapitre 2 : La métabolomique.....</i>	<i>19</i>
I. Introduction .....	19
II. La métabolomique .....	20
1) Les étapes d'une analyse métabolomique.....	21
a. Collecte et préparation des échantillons.....	21
b. Choix de l'approche métabolomique .....	24
c. Plateforme analytique en métabolomique non ciblée.....	25
d. Traitement des données .....	29
e. Analyse statistique.....	33
f. Validation des métabolites .....	38
g. Construction de Heat Map .....	40
h. Analyse des voies métaboliques impliquées .....	40
2) Intérêt de la métabolomique dans le domaine médical .....	43
<b>Partie II : Approche métabolomique à la recherche d'outils diagnostiques et pronostiques dans la COVID-19 .....</b>	<b>45</b>
<i>Chapitre 1 : Objectifs de la thèse .....</i>	<i>46</i>
<i>Chapitre 2 : Identification d'une signature métabolomique plasmatique diagnostique et pronostique de la COVID-19.....</i>	<i>47</i>
1) Préparation des échantillons et analyse LC-MS.....	48
2) Présentation des principaux résultats.....	49
<i>Chapitre 3 : Identification d'une signature métabolomique urinaire diagnostic et pronostique de la COVID-19.....</i>	<i>95</i>
1) Préparation des échantillons et analyse LC-MS.....	96
2) Présentation des principaux résultats.....	97
<i>Chapitre 4 : Discussion - Place de la métabolomique dans la prise en charge et la compréhension de la COVID-19.....</i>	<i>101</i>
<i>Le métabolome sanguin.....</i>	<i>102</i>
<i>Le métabolome urinaire.....</i>	<i>109</i>
<i>Le métabolome de l'air expiré .....</i>	<i>112</i>
<i>Le métabolome salivaire .....</i>	<i>114</i>
<b>Partie IV Perspectives en médecine d'urgence .....</b>	<b>118</b>
ANNEXES .....	128
REFERENCES BIBLIOGRAPHIQUES .....	160

## ABREVIATIONS

---

ACP (PCA en anglais) : Analyse en Composantes Principales  
ACE2 : Enzyme de Conversion de l'Angiotensine 2  
ARN : Acide Ribonucléique  
ARNm : Acide Ribonucléique messenger  
AUROC : Area Under the Receiver Operating Characteristic (aire sous la courbe ROC)  
CHU : Centre Hospitalier Universitaire  
COV : Composés Organiques Volatils  
C19DMap : Cartographie de la COVID-19  
COVID-19 : maladie à coronavirus 2019 (coronavirus disease 2019)  
GC : Chromatographie en phase gazeuse  
HPLC : High Performance Liquid Chromatography (chromatographie liquide à haute performance)  
IDO 1 : Indoléamine 2,3-dioxygénase 1  
IgG : Immunoglobuline de Type G  
IgM : Immunoglobuline de Type M  
IRM : Imagerie par Résonance Magnétique  
QC : Contrôle Qualité  
LC : Chromatographie en phase liquide  
LC-MS : Spectrométrie de Masse avec Chromatographie en phase liquide  
MS : Spectrométrie de Masse  
MERS : Syndrome Respiratoire du Moyen-Orient  
NAD : Nicotinamide Adénine Dinucléotide  
NO : Oxyde Nitrique  
OMS : Organisation Mondiale de la Santé  
PLS : Partial Least Squares (régression des moindres carrés partiels)  
PLS-DA : Partial Least Squares Discriminant Analysis  
POCT : Point Of Care Test (test moléculaire sur le lieu de soins)  
PPRs : Pattern Recognition Receptors  
RMN : Résonance Magnétique Nucléaire  
RT : Temps de Rétention

RT-PCR : Réaction en chaîne par polymérase après Transcription inverse (Reverse Transcription Polymerase Chain Reaction)

SARS-CoV-2 : Coronavirus 2 du Syndrome Respiratoire Aigu Sévère

SDRA : Syndrome de Détresse Respiratoire Aigu

SIRT : Sirtuine

SPM : Specialized Proresolving Mediators (lipides pro-résolutifs de l'inflammation)

SRAS : Syndrome Respiratoire Aigu Sévère

TCA : Cycle des Acides Tricarboxyliques ou Cycle de Krebs

TLRs : Toll-like receptors

VIP : Variable Importance in the Projection

## LISTE DES FIGURES

---

Figure 1. Nombre de cas confirmés de COVID-19 par date et par région selon l’OMS de décembre 2019 à Juin 2020

Figure 2. Physiopathologie de la COVID-19.

Figure 3. Symptômes de la COVID-19

Figure 4. Approches multi-omiques et médecine des systèmes

Figure 5. Principales stratégies omiques dans les études de biologie systémique

Figure 6. Type de métabolites extraits selon la nature du solvant utilisé

Figure 7. Schéma des différentes étapes de l’analyse métabolomique non ciblée

Figure 8. Schéma d’une plateforme analytique en LC-MS

Figure 9. Schéma descriptif des différents temps de la technique LC-MS et des données spectrales générées

Figure 10. Analyse trois dimensions du spectre LC-MS

Figure 11. Schéma de la procédure pré-analytique du logiciel MZmine™

Figure 12. Représentations graphiques des résultats d’une ACP

Figure 13. Exemple de PCA et de PLS-DA avec cross validation et représentation graphique par histogrammes des données de la cross-validation en PLS-DA

Figure 14. Exemple d’une top liste de métabolites issue de l’analyse en PLS-DA

Figure 15. Exemple de courbes ROC

Figure 16. Exemple de Heat Map

Figure 17. Exemple de graphiques d’analyse des voies métaboliques par MetaboAnalyst

Figure 18. Description des groupes de patients et volontaires sains dans le travail de recherche intéressant les prélèvements plasmatiques

Figure 19. Identification d’une signature métabolomique plasmatique diagnostique de la COVID-19

Figure 20. Box plots des 15 métabolites composant notre signature métabolomique diagnostique de l’infection à SARS-CoV-2

Figure 21. Comparaison du métabolome plasmatique des sujets sains avec celui du groupe légèrement atteint par la COVID-19

Figure 22. Identification d’une signature métabolomique plasmatique pronostique du risque d’hospitalisation et du risque d’hospitalisation en service de soins critiques

Figure 23. Diagramme de Venn des voies métaboliques plasmatiques significativement modifiées par l’infection à SARS-CoV-2 et selon le niveau de gravité de l’atteinte de la COVID-19

Figure 24. Description des groupes de patients et volontaires sains dans la cohorte SignCov

Figure 25. Identification d'une signature métabolomique urinaire diagnostique de la COVID-19

Figure 26. Comparaison du métabolome urinaire des sujets symptomatiques non-COVID avec celui des patients atteints de la COVID-19

Figure 27. Identification d'une signature métabolomique urinaire pronostique du risque d'hospitalisation

Figure 28. Les lipides pro-résolutifs de l'inflammation

Figure 29 : Schéma des potentiels biomarqueurs d'intérêt diagnostiques et pronostiques de la COVID-19 selon Bourgin et al.

Figure 30. Analyse multidimensionnelle de l'air expiré basée sur la spectrométrie de masse

Figure 31 : Consensus sur la signature métabolomique des patients COVID-19, selon Bruzzone et al.

Figure 32 : Métabolites associés à la gravité du pronostic de COVID-19

## **Partie I : INTRODUCTION**

# Chapitre 1 : La COVID-19

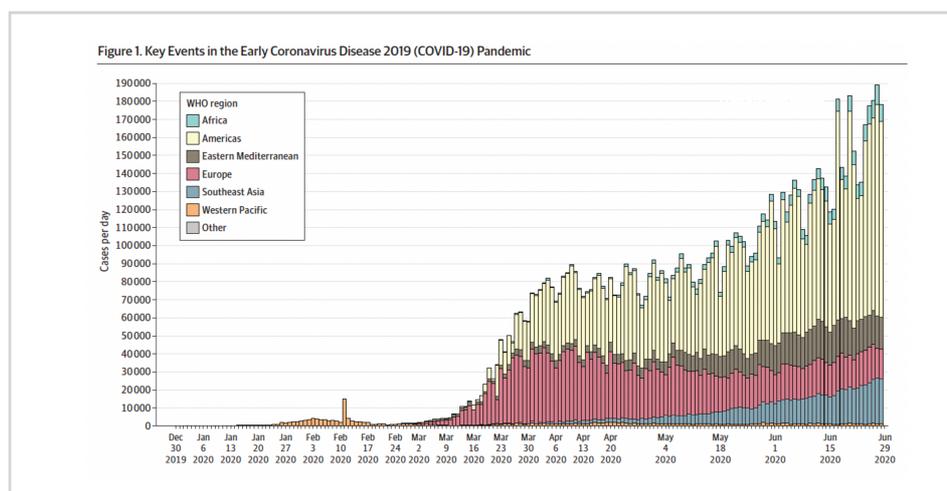
---

## I. L'enjeu de la découverte d'une nouvelle pathologie

En décembre 2019, est identifié un nouveau coronavirus, le coronavirus 2 du syndrome respiratoire aigu sévère (SARS-CoV-2), virus à l'origine d'un groupe de pneumonies, initialement détectées dans la ville de Wuhan, en Chine et entachées d'une morbi-mortalité élevée. Le 11 février 2020, la maladie secondaire à l'infection à SARS-CoV-2 est officiellement nommée *maladie à coronavirus 2019* (COVID-19) par l'Organisation Mondiale de la Santé (OMS) et sur la base de « niveaux alarmants de propagation et de gravité », l'OMS la qualifie de pandémie le 11 mars 2020 (World Health Organization 2020).

Les coronavirus sont de grands virus à Acide Ribonucléique (ARN) monocaténaire enveloppés, présents chez l'homme et d'autres mammifères. Le SARS-CoV-2 représente le neuvième coronavirus documenté touchant l'homme et le troisième entraînant une atteinte grave après le Syndrome Respiratoire Aigu Sévère (SRAS) et le Syndrome Respiratoire du Moyen-Orient (MERS) (Wiersinga et al. 2020). L'origine de la pandémie de COVID-19 fait encore aujourd'hui débat. Deux principaux scénarios ressortent : un scénario de "fuite en laboratoire" et l'émergence zoonotique, comme les coronavirus humains précédents (Holmes et al. 2021).

Les premières données suggèrent un taux de transmission élevée, se faisant principalement par projection voire aérosolisation de gouttelettes, une propagation par contact de surface semblant également possible. Très rapidement, et malgré des mesures de confinement et de port de protection personnel le virus se propage pour atteindre au 1<sup>er</sup> avril 2020 le cap de 1 million de cas confirmé de COVID-19 et de 4 million le 9 mai 2020 (Figure 1) (Wiersinga et al. 2020).



**Figure 1 : Nombre de cas confirmés de COVID-19 par date et par région selon l’OMS de décembre 2019 à Juin 2020**

Extrait de Wiersinga, W. Joost et al. “Pathophysiology, Transmission, Diagnosis, and Treatment of Coronavirus Disease 2019 (COVID-19) A Review.” JAMA - Journal of the American Medical Association 8(324):782–793.

Ces mesures de distanciations sociales et de confinement, mises en place au niveau mondial, n’étaient pas dénuées d’impact socio-économique, d’impact sur la santé physique et mentale de la population et surtout ne semblaient pas complètement efficaces pour endiguer la pandémie.

Dans une étude portant sur 44 672 patients atteints de COVID-19 en Chine, publiée en février 2020, 14 % des patients présentaient des manifestations sévères et 5 % des manifestations critiques (Z. Wu and McGoogan 2020). Bien que les premières données de la littérature varient selon les sources, il semblerait que la mortalité hospitalière initiale de la COVID-19 avoisinait les 15-20% et approchait les 40% pour les patients hospitalisés en service de soins critiques. Cette disparité de taux de mortalité dans la littérature pouvant s’expliquer par le fait qu’un pourcentage disproportionné d’hospitalisations et de décès survenait dans les populations à faibles revenus et les minorités et qu’un certain nombre de décès n’était pas comptabilisé du fait de l’absence de test diagnostique réalisé (Wiersinga et al. 2020).

De plus, l’impact sur le système de santé, que ce soit sur les structures hospitalières ou en soins primaires, était non négligeable, avec notamment en France le déclenchement du « plan blanc maximal » dès le 16 mars 2020, entraînant des déprogrammations massives dans le but de faire face à l’arrivée de nombreuses victimes de la COVID-19 (Bourgueil et al. 2020; Carli and Telion 2023).

La lutte contre la COVID-19 devenait alors la cible principale et les avancées scientifiques devenaient urgentes afin de mieux comprendre la physiopathologie de l'infection à SARS-CoV-2 et de développer des outils thérapeutiques et vaccinaux (Chu et al. 2020). En effet, les recommandations de lutte contre la COVID-19 se font au fil de l'eau et principalement selon des données descriptives basées sur les expériences mutuelles. Ainsi, dès février 2020, l'OMS en collaboration avec le *Global Research Collaboration for Infectious Disease Preparedness* organisent un forum de la recherche et de l'innovation sur la thématique du nouveau coronavirus, qui évalue le niveau de connaissances, définit les lacunes et cherche à collaborer pour financer et accélérer les travaux de recherche prioritaires. Ainsi, la course scientifique contre le SARS-CoV-2 a été lancée.

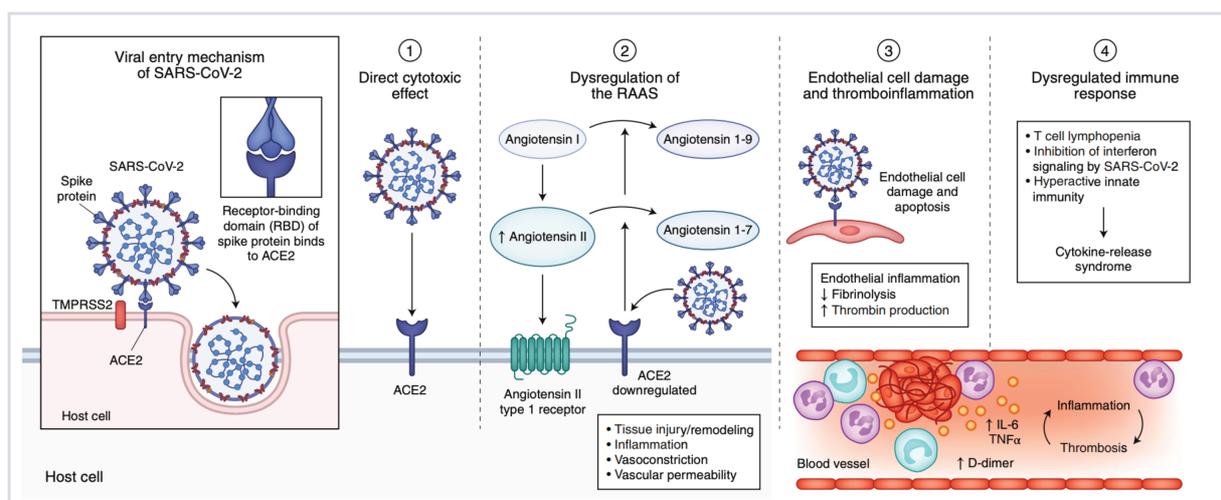
Les sciences omiques, outils très prometteurs dans le domaine de la recherche scientifique en santé, n'échappent pas à la règle. Et les équipes scientifiques spécialisées dans le domaine commencent très rapidement à s'intéresser à la compréhension physiopathologique et à la recherche d'outils diagnostiques, pronostiques, thérapeutiques et vaccinaux de la COVID-19 (Lin et al. 2020).

## II. Vers une meilleure compréhension de la physiopathologie

La COVID-19 possède un spectre clinique très large, allant de la forme asymptomatique au Syndrome de Détresse Respiratoire Aigu (SDRA) et à la défaillance multiviscérale pouvant tous deux conduire au décès (Ramos-Casals, Brito-Zerón, and Mariette 2021).

Le SARS-CoV-2, tout comme le SARS-CoV1, utilise les protéines spikes présentes à sa surface pour entrer dans les cellules hôtes, via le récepteur cellulaire de l'Enzyme de Conversion de l'Angiotensine 2 (ACE2) et le corécepteur TMPRSS2 (Transmembrane Protease Serine 2) permettant ainsi l'endocytose puis la fusion membranaire (Zhou et al. 2020).

La forte densité de récepteur ACE2 au niveau des cellules épithéliales pulmonaire explique le tropisme pulmonaire préférentiel du SARS-CoV-2 (Perrotta et al. 2020). Cependant, le récepteur ACE2 étant également exprimé dans de nombreux tissus et cellules, tels que les cellules à mucus nasales, les cellules bêta-pancréatiques, l'endothélium veineux et artériel, les cellules épithéliales intestinales (où il agit comme un corécepteur dans l'absorption des nutriments, notamment des acides aminés), le myocarde, les tubules proximaux et les podocytes rénaux, les cellules hépatiques ..., cela pourrait participer aux atteintes extra pulmonaires dues à l'infection à SARS-CoV-2 (H. Zhang et al. 2020). La physiopathologie des atteintes multi-organes, secondaires cette infection, semble être multifactorielle. Elle comprend la toxicité virale directe, les lésions des cellules endothéliales, la thrombo-inflammation, le dérèglement de la réponse immunitaire et le dérèglement du système rénine-angiotensine-aldostérone (SRAA) (Gupta et al. 2020) (Figure 2).



**Figure 2. Physiopathologie de la COVID-19.**

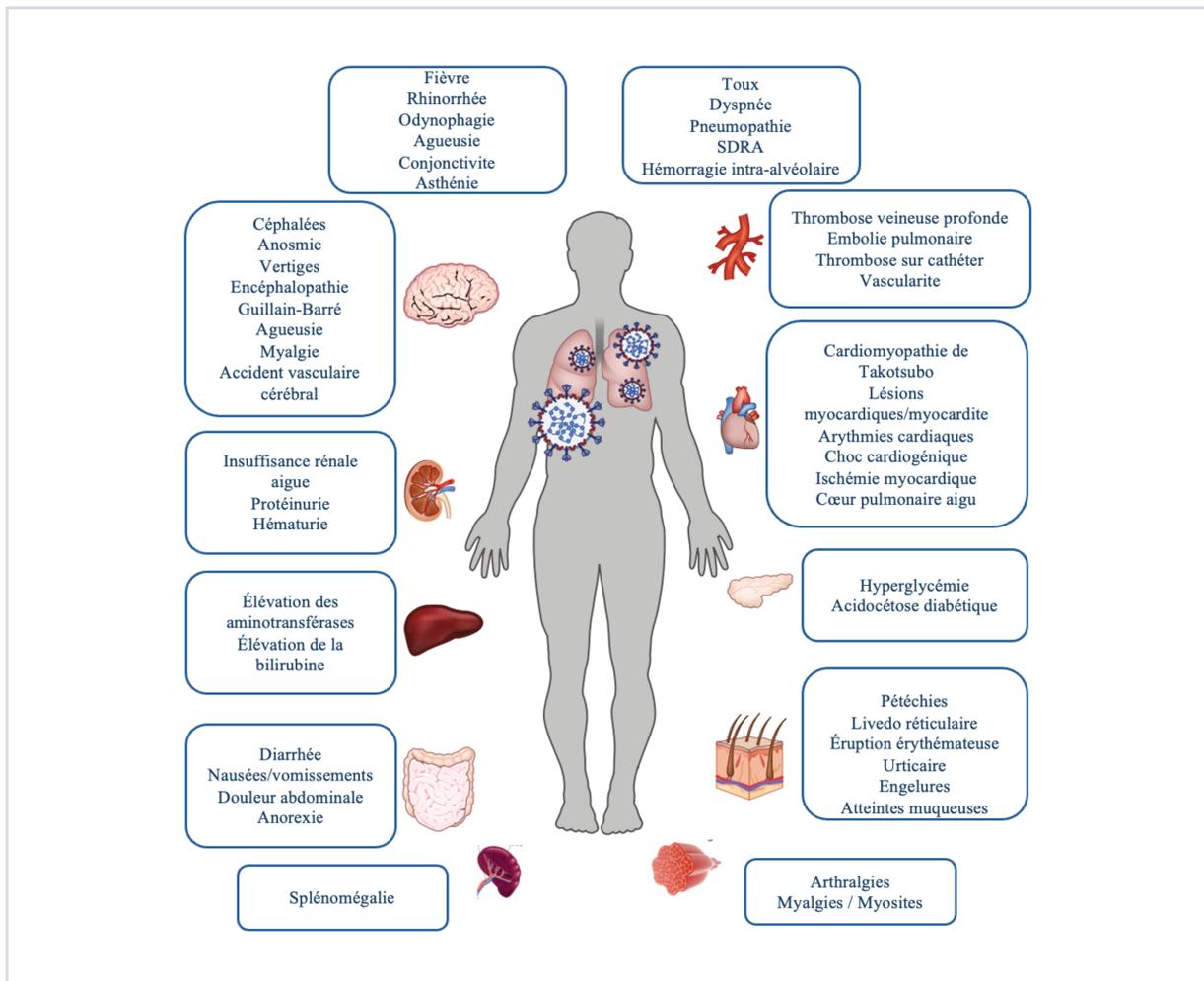
Extrait de Gupta et al. "Extrapulmonary Manifestations of COVID-19." *Nature Medicine* 26(7):1017–32.

Lors de l'invasion virale, les Pattern Recognition Receptors, (PRRs) et les Toll-like receptors (TLRs) reconnaissent les antigènes et vont entraîner la sécrétion de cytokines (TNF- $\alpha$ , IL-1, IL-6, interférons) responsables de la stimulation et du recrutement des cellules de l'immunité innée mais également d'une augmentation de la perméabilité vasculaire favorisant ainsi la dissémination virale. L'IL-6, entre autres, stimule également l'expression de substances pro-coagulantes à la surface de l'endothélium entraînant alors une thrombo-inflammation et un état pro-thrombotique (Bonny et al. 2020). D'autres mécanismes concourent également à la présence de lésions endothéliales et favorisent également cet état pro-thrombotiques (Gupta et al. 2020). Par la suite, les cellules présentatrices d'antigène activent la voie de l'immunité acquise via les lymphocytes. Chez certains patients graves, une dysrégulation de la réponse immunitaire ainsi qu'un syndrome de libération cytokinique dit « orage cytokinique » sont observés, entraînant un état pro inflammatoire majeur, donc entretenant l'état pro-thrombotique, et pouvant donc concourir aux lésions d'organes secondaires.

Les mécanismes immunologiques entraînés par l'infection à SARS-CoV-2 ont fait l'objet d'un grand nombre de travaux scientifiques initiaux puisqu'ils représentaient la clef pour le développement d'outils thérapeutiques et vaccinaux (Guo et al. 2023). Cependant, devant une grande variabilité inter-individuelle de réponse de l'hôte à l'infection à SARS-CoV-2, l'enjeu était de mieux identifier les mécanismes sous-tendant cette variabilité. Les sciences omiques, notamment la métabolomique qui permet d'identifier les dysrégulations métaboliques secondaires à une pathologie et donc d'identifier des phénotypes de celle-ci, représentaient donc des moyens incontournables dans la lutte contre la COVID-19.

### III. La nécessité de développer des outils diagnostiques et pronostiques

Devant la multitude de symptômes peu spécifiques, reflet de la complexité physiopathologique (figure 3), le diagnostic clinique de la COVID-19 représente un réel défi pour les cliniciens. De plus, à la phase initiale de la pandémie, le dépistage des formes asymptomatiques et paucisymptomatiques constituait un élément essentiel du contrôle de cette pandémie.



**Figure 3. Symptômes de la COVID-19**

Adapté de Ramos-Casals et al. 2021. "Systemic and Organ-Specific Immune-Related Manifestations of COVID-19." *Nature Reviews Rheumatology* 17(6):315–32.

Dans l'objectif de répondre aux mieux aux besoins diagnostiques et surtout à la nécessité de dépistage de masse, de nombreux tests ont été développés. Le test de référence pour le dépistage et le diagnostic de l'infection à SARS-CoV-2 est la détection de l'ARN viral par RT-PCR (Réaction en chaîne par polymérase après Transcription inverse ou Reverse Transcription Polymerase Chain Reaction) sur prélèvements rhino-pharyngés par écouvillonnage profond ou

sur ou des prélèvements des voies respiratoires basses (Haute Autorité de Santé 2020). Présentant une spécificité élevée, la sensibilité de la RT-PCR dépend du type d'échantillon (Haute Autorité de Santé 2021), du moment du prélèvement selon le stade de l'infection et de la technique d'échantillonnage (Gala et al. 2020; Marty, Chen, and Verrill 2020).

Peu adaptée aux politiques de dépistage de masse, notamment compte tenu de sa faible disponibilité initiale en médecine libérale et du délai de rendu de résultats au début de la pandémie, d'autres techniques diagnostiques ont rapidement été développées. Parmi elles :

- les tests rapides de détection d'antigènes (reposant sur l'immunochromatographie) qui permettent, sur prélèvement nasal ou nasopharyngé voire oropharyngé, de détecter la présence d'un antigène viral dans un délai de réponse relativement court (moins de 30 minutes) mais qui sont entachés d'une plus faible sensibilité que la RT-PCR (European Commission safety directorate-general for Health and food 2023)
- les tests sérologiques permettant d'objectiver la présence d'anticorps spécifiques (immunoglobulines de type M (IgM) et de type G (IgG)) sur prélèvement sanguin mais également sur prélèvement capillaire par test rapide immunochromatographique sur bandelette de nitrocellulose. Le délai d'apparition des IgM étant de l'ordre de 5 à 7 jours suivant l'apparition des premiers symptômes avec un temps moyen de séroconversion de 10 à 11 jours et de 15 jours pour les IgG, et la positivité de ces tests étant dépendante de la réponse immunitaire de l'hôte (Hantz 2020), ces tests sérologiques semblent peu performants pour répondre aux besoins diagnostiques et de dépistage de masse en vue d'un isolement des cas.

Pour répondre aux besoins diagnostiques, notamment dans un délai acceptable pour le clinicien, les tests moléculaires sur le lieu de soins (Point of Care Test : POCT) pour le SARS-CoV-2 ont eux aussi été rapidement développés. Ces tests ont pour avantage d'être facile d'utilisation, accessibles directement par le clinicien permettant donc une détection rapide avec une performance diagnostique proche de celle de la RT-PCR (Fernandes et al. 2022; Iliescu et al. 2021; Song et al. 2021).

Cependant, la majorité de ces tests se réalisant sur prélèvement nasopharyngé, il semblait pertinent de poursuivre les recherches dans l'objectif de mettre au point des tests diagnostiques moins douloureux et donc plus facilement acceptables pour la population.

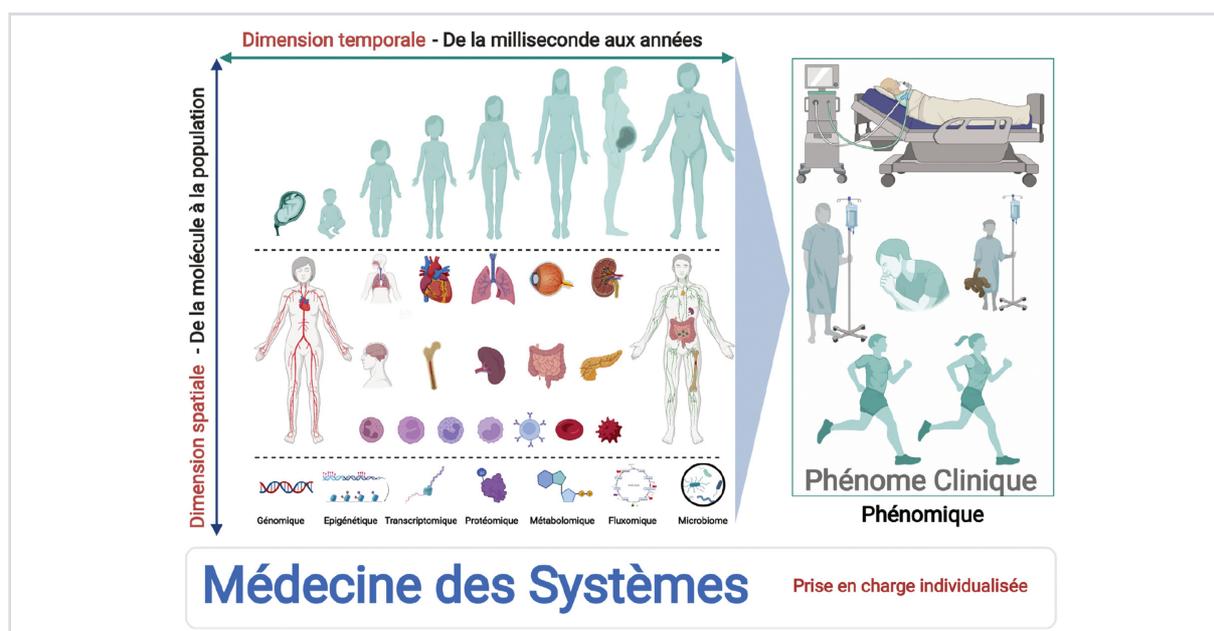
Outre le dépistage, le second grand enjeu de cette pandémie était l'identification des patients à risque de mauvaise évolution. En ce sens, de nombreux scores pronostiques cliniques, scannographiques et biologiques ont été débattus (Izcovich et al. 2020; Wynants et al. 2020) . Cependant, les premiers outils pronostiques décrits dans la littérature étaient principalement issus de scores ou de biomarqueurs usuels, déjà validés dans d'autres pathologies infectieuses ou respiratoires et ne pouvaient donc tenir entièrement compte des spécificités physiopathologiques de la COVID-19. Devant une hétérogénéité non négligeable de réponse de l'hôte à l'infection, certains travaux se sont tournés vers l'identification de phénotypes clinico-biologiques chez les patients infectés par le SARS-CoV-2 et leur corrélation avec le pronostic (Gutiérrez-Gutiérrez et al. 2021). Mais ces phénotypes ne tenaient compte, une nouvelle fois, que de biomarqueurs utilisés en pratique de routine. L'apport des sciences omiques dans la compréhension de ces différents phénotypes, dans la découverte de nouveaux biomarqueurs et dans l'approche multi-marqueurs qu'elle permet, venait donc compléter et améliorer les outils pronostiques préexistants.

## Chapitre 2 : La métabolomique

### I. Introduction

La science des « omiques » est une science en plein essor avec une implication de plus en plus grande dans le domaine médical. Elle permet, entre autres, l'analyse des gènes (génomique) et de leur produits (transcriptomique pour l'ARNm, protéomique pour les protéines) ainsi que l'analyse des métabolites (métabolomiques). Depuis l'avènement du séquençage complet du génome dans les années 2000, des avancées technologiques des méthodes d'analyse à haut débit et le développement de l'intelligence artificielle avec possibilité de traitement des « big datas » ont permis le développement de ces analyses omiques.

A l'heure de la médecine personnalisée et de précision, les méthodes « omiques » permettent d'intégrer la variabilité des individus, leur complexité biologique et biochimique et de prendre en compte l'influence de facteurs exogènes tels que l'environnement ou les pathologies rencontrées (Figure 4) (Tebani and Bekri 2022). Ceci permet d'améliorer nos connaissances sur les changements moléculaires en réponse à une maladie et leur variation dans le temps et l'espace mais également les différences de réponses selon les individus. Ces approches nous permettent ainsi une meilleure compréhension des différents systèmes impliqués dans une maladie et dans la réponse de l'hôte à cette maladie (Benson 2016).



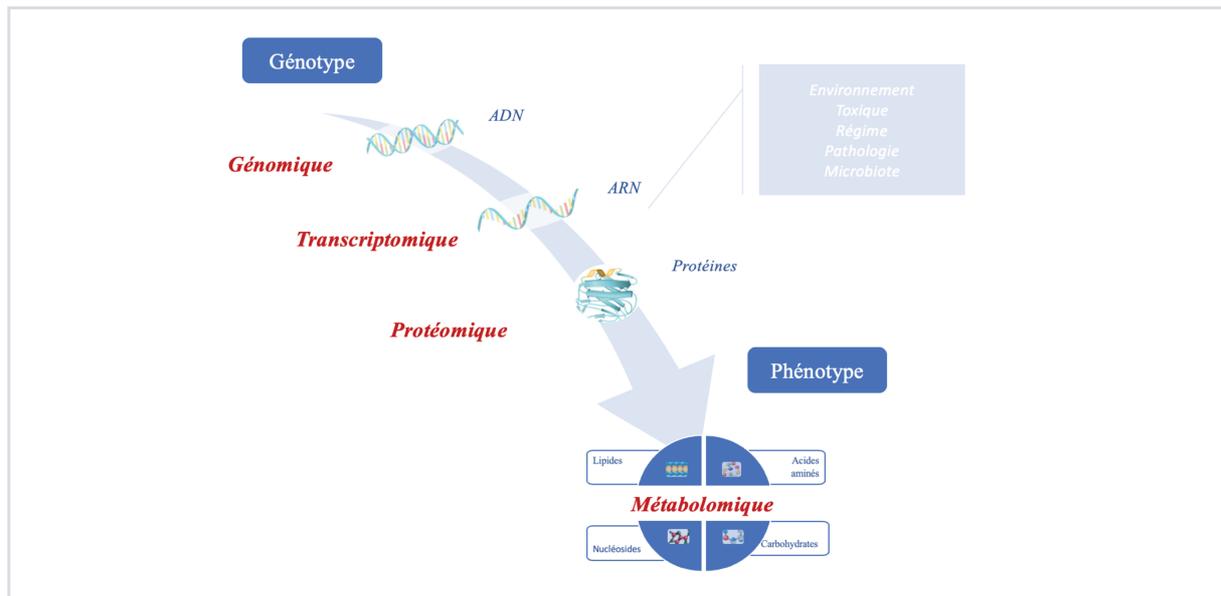
**Figure 4. Approches multi-omiques et médecine des systèmes.**

(extrait de Tebani, A. and Bekri, S. (2022) 'The promise of omics in the precision medicine era', Revue de Medecine Interne, 43(11), pp. 649–660)

## II. La métabolomique

Le terme métabolite désigne des molécules de faible poids moléculaire (inférieures à 1500 Dalton). Ces métabolites peuvent être produits de manière endogène ou présents de manière exogène (aliments, polluants, médicaments, toxiques, etc...). Leur présence ou la variation de leur concentration peut être le reflet d'un état physiologique mais également physiopathologique en réponse à un changement génétique, environnemental, pathologique ou encore toxicologique (Fiehn 2002). Les métabolites endogènes peuvent être classés en métabolites primaires, pour ceux directement impliqués dans les processus essentiels à la vie (par exemple métabolisme cellulaire de base, croissance, division cellulaire, ...) et ayant donc une fonction physiologique (acides aminés, acides organiques, sucres, acides gras, etc...) et métabolites secondaires pour ceux synthétisés lors de fonctions biologiques particulières (Courant et al. 2014).

La métabolomique correspond, quant à elle, à l'étude des métabolites (métabolome) présent dans un échantillon biologique donné (cellules, tissus ou fluides biologiques) (Oliver et al. 1998). Ce profilage métabolique, comprenant une analyse dynamique des variations de taux de métabolites, permet donc d'avoir une vision d'ensemble de processus biochimiques en jeu à un instant donné. La métabolomique se situe au plus proche du phénotype (figure 5), en association ou non aux autres sciences omiques, elle permet comme décrit précédemment, une meilleure compréhension des systèmes biologiques complexes qu'ils soient physiologiques ou physiopathologiques (Klassen et al. 2017).



**Figure 5. Principales stratégies omiques dans les études de biologie systémique**

(adapté de Klassen, A. et al. (2017) 'Metabolomics: Definitions and Significance in Systems Biology.', Advances in experimental medicine and biology. United States, 965, pp. 3–17.) ADN : Acide désoxyribonucléique, ARN : acide ribonucléique.

## 1) Les étapes d'une analyse métabolomique

Une analyse métabolomique est composée de quatre principales étapes :

- Le traitement de l'échantillon : comprenant le prélèvement, la préparation puis l'analyse de l'échantillon (réalisée par spectrométrie de masse dans ce travail)
- Le traitement des données
- L'identification des métabolites d'intérêts et des voies métaboliques impactées
- L'interprétation des résultats

### a. Collecte et préparation des échantillons

La collecte et la préparation des échantillons pouvant être source de variabilité des résultats, les méthodes choisies doivent être les plus reproductibles, les plus simples et rapides possible (pour éviter la perte ou la dégradation des métabolites durant la procédure), afin de garantir la reproductibilité et la fiabilité des résultats (Courant et al. 2014). Ces étapes doivent également être réalisées de manière la plus précautionneuse possible afin de limiter la contamination des échantillons. Les procédures de collecte et de préparation d'échantillons devront être établies selon plusieurs facteurs et doivent notamment prendre en compte le type d'approche métabolomique envisagée (ciblée ou non ciblée).

En tout premier lieu, le métabolisme étant en constante évolution, la concentration des métabolites peut se modifier très rapidement ; une étape « d'arrêt » du métabolisme est donc essentielle afin que les résultats obtenus soit bien le reflet du métabolome du moment du prélèvement de l'échantillon. Après prélèvement de l'échantillon, le métabolisme doit donc être bloqué le plus rapidement possible par dénaturation/extraction (ajout d'un solvant (méthanol) par exemple), par congélation ou adjonction d'azote liquide ou bien par chauffage rapide (Krastanov 2010)

La préparation de l'échantillon vise ensuite un double objectif : elle permet d'extraire les métabolites de l'échantillon analysé mais également d'éliminer des « impuretés », comme les protéines, qui pourraient parasiter l'analyse en spectrométrie de masse (Carneiro et al. 2020).

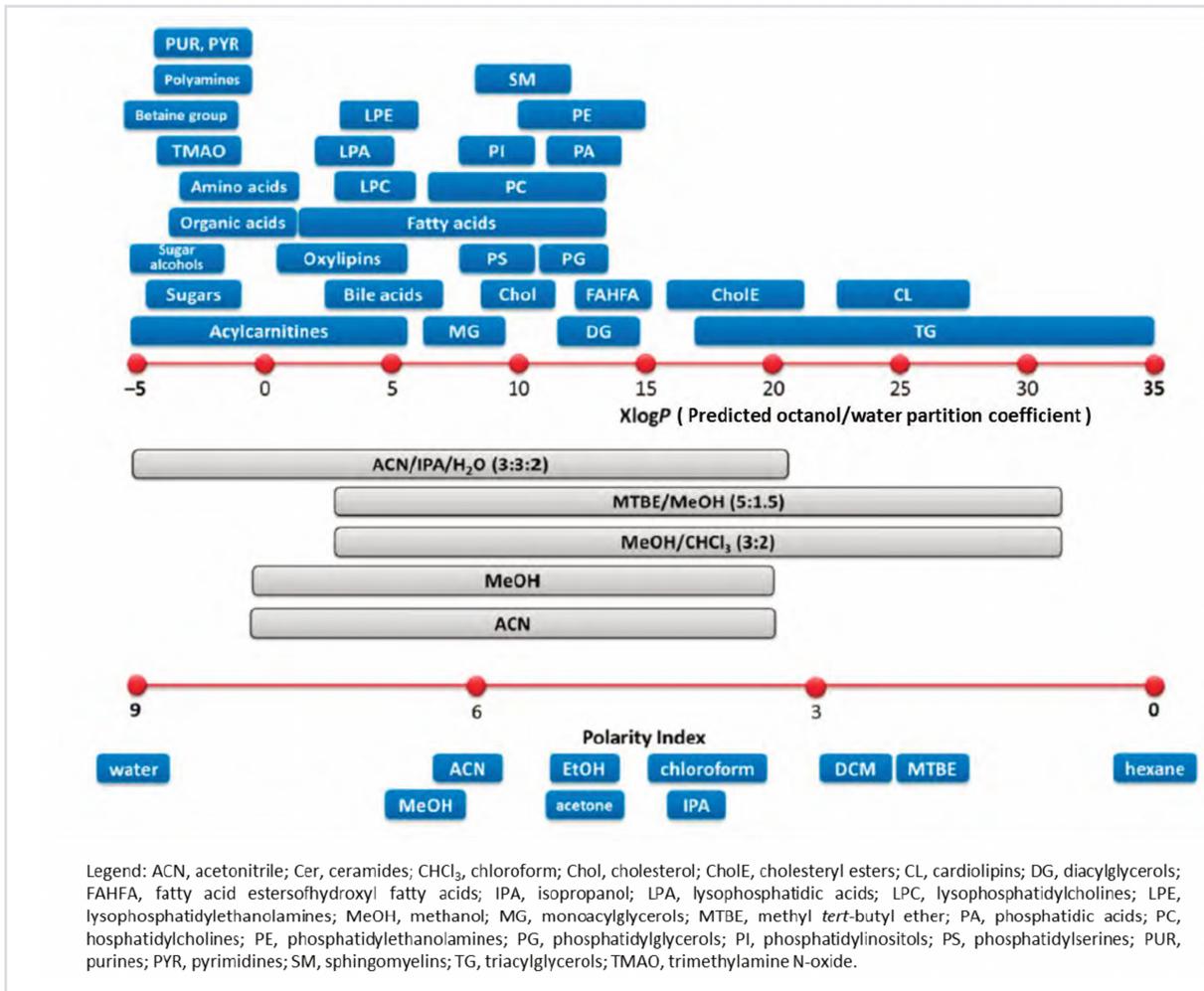
Cette préparation comprend 3 étapes :

- **L'extraction des métabolites** : Il existe plusieurs techniques d'extraction. Le choix de la méthode doit dépendre du type d'échantillon (solide ou liquide, plasma/serum ou urines...) et du type d'analyse envisagée puisque la méthode d'extraction doit être adaptée aux propriétés physico-chimiques et aux classes de métabolites recherchés (Gika et al. 2014).

Pour les études en métabolomique la méthode la plus fréquemment utilisée est l'extraction par adjonction de solvants organiques. Dans cette technique, les types de métabolites extraits sont déterminés par la nature du solvant (ou mélange de solvant) choisi (Figure 6) et par le ratio solvant/échantillon (Cajka and Fiehn 2016). Cette extraction en phase organique est souvent associée à une élimination des protéines par précipitation. A noter que l'analyse des composés apolaires concerne maintenant une technique omique à part entière, non explorée dans ce travail de thèse : la lipidomique.

L'utilisation du type de solvant organique lorsque d'échantillon est constitué de sérum ou de plasma a été largement étudiée et le méthanol est très souvent utilisé (Dunn et al. 2011). Ce solvant utilisé en routine au laboratoire a été choisi dans notre travail.

Il faut également prendre en compte qu'en métabolomique non ciblée (défini ci-après), cette étape doit permettre l'extraction de la plus grande gamme de métabolite possible. L'extraction par solvant organique est donc tout à fait adaptée en raison de sa simplicité et de sa sélectivité.



**Figure 6. Type de métabolites extraits selon la nature du solvant utilisé**

Gamme de coefficients de partage octanol/eau prédits (XlogP) de métabolites courants dans le plasma sanguin et indice de polarité des solvants utilisés pour l'extraction de l'échantillon. Les solvants ou mélanges de solvants typiques (en gris) utilisés indiquent la gamme de polarité des métabolites isolés avec une récupération élevée. Les plages d'extraction augmentent si l'on utilise des mélanges comprenant de l'eau ou si l'on ajoute de l'eau dans des procédures d'extraction.

(Extrait de Cajka, T., and Fiehn, O. (2016). Toward merging untargeted and targeted methods in mass spectrometry-based metabolomics and lipidomics. *Anal. Chem.* 88, 524–545.)

- **La séparation des métabolites** : Les métabolites sont ensuite séparés des débris et des agrégats de protéines par centrifugation. Le surnageant est alors récupéré puis transféré dans de nouveaux tubes.

- **La concentration et évaporation** : Ce surnageant est par la suite séché dans un système de concentration par centrifugation sous vide ou « SpeedVac™ ». Il est ensuite conservé à -80°C jusqu'à l'analyse en spectrométrie de masse. Juste avant l'analyse, le culot séché est de nouveau suspendu dans un milieu contenant un mélange d'acétonitrile et d'eau.

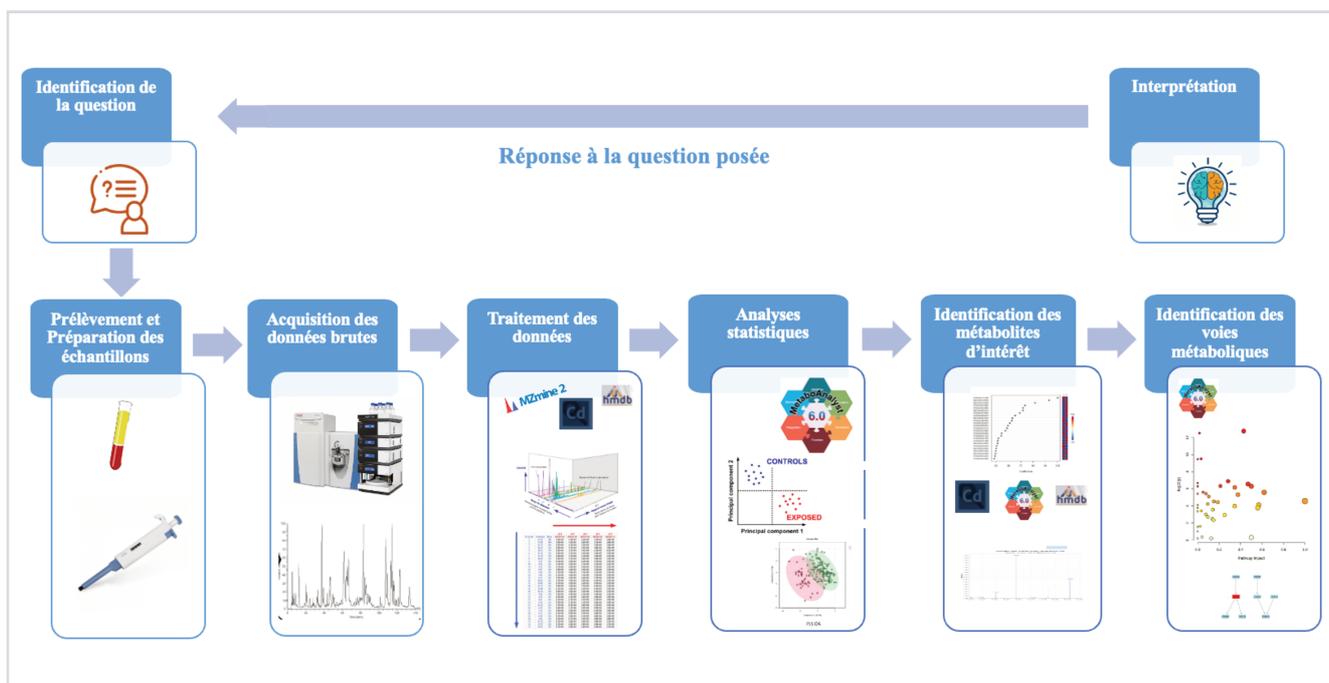
## **b. Choix de l'approche métabolomique**

Deux approches co-existent en métabolomique : l'approche ciblée et l'approche non ciblée (Roberts et al. 2012).

L'approche non ciblée, qui sera l'approche utilisée dans ce travail de thèse, est une approche qui permet d'étudier à une question biologique ou clinique posée, sans à priori c'est-à-dire sans avoir connaissance en amont des potentiels métabolites ou signatures métabolomiques d'intérêt. Dans cette approche, sont analysés tous les métabolites mesurables présents dans l'échantillon, y compris ceux n'étant pas encore identifiés par les banques de données actuelles. Ces données seront ensuite traitées puis statistiquement analysées dans l'objectif d'identifier des biomarqueurs voire des signatures métabolomiques (approche multi-marqueurs) répondant à la question posée. L'approche non ciblée permet une identification qualitative mais ne permet qu'une quantification relative des métabolites. Avec ce type d'approche, il est également possible d'identifier les voies métaboliques impactées par le phénomène d'intérêt étudié et ainsi permettre une meilleure compréhension des phénomènes physiologiques mis en jeu. Il s'agit donc l'approche préférentielle dans la découverte de nouveaux biomarqueurs ou dans la recherche métabolomique visant à comprendre la physiopathologie d'une nouvelle maladie.

L'approche ciblée permet, quant à elle, de valider la présence et de mesurer des variations de niveau, au sein d'un échantillon donné, des métabolites, signatures métaboliques ou voies métaboliques prédéfinies. Cette technique permet aussi une analyse qualitative potentiellement plus précise, mais ne permet cependant pas d'avoir une vision globale du métabolome puisqu'elle ne s'intéresse qu'aux métabolites prédéfinis lors de l'analyse. Elle est particulièrement intéressante pour la validation ou l'utilisation en routine de biomarqueurs/signatures multi-marqueurs préalablement identifiés par une méthode non ciblée. Le choix de l'approche est donc dépendant de la question clinique ou scientifique posée.

Les étapes de l'analyse métabolomique non ciblée sont synthétisées dans la figure 7.



**Figure 7. Schéma des différentes étapes de l'analyse métabolomique non ciblée**

### c. Plateforme analytique en métabolomique non ciblée

Depuis le développement des plateformes instrumentales, il existe de nombreux outils pour l'analyse métabolomique. Pour qu'ils soient performants, ces outils doivent présenter une haute résolution, c'est-à-dire permettre l'identification et l'analyse qualitative et/ou quantitative d'un grand nombre de métabolites mais également permettre une analyse haut-débit, c'est-à-dire permettre l'analyse de nombreux échantillons.

Deux principales méthodes sont décrites dans la littérature : la spectroscopie de Résonance Magnétique Nucléaire (RMN) et la Spectrométrie de Masse (MS pour Mass Spectrometry en anglais).

**La spectroscopie RMN** repose sur les propriétés nucléaires de certains noyaux atomiques (comme l'hydrogène ( $^1\text{H}$ ) et le carbone ( $^{13}\text{C}$ )) entraînant un phénomène de résonance magnétique nucléaire lorsqu'ils sont soumis à un champ magnétique afin d'obtenir les informations structurales sur les métabolites présents. Cette technique a pour avantage d'être non destructrice, hautement reproductible et de permettre une analyse quantitative. Elle est

également capable de mesurer les métabolites extraits par une large gamme de solvants mais est toutefois entachée d'une faible sensibilité ne permettant la détection que des métabolites présents de manière abondante dans l'échantillon (de l'ordre du micromolaire au millimolaire) (Powers and Riekeberg 2017). Ce qui peut être un frein à son utilisation en métabolomique non-ciblée ou ciblé visant des métabolites non majoritaires.

**La Spectrométrie de Masse**, utilise quant-à-elle utilise le procédé suivant :

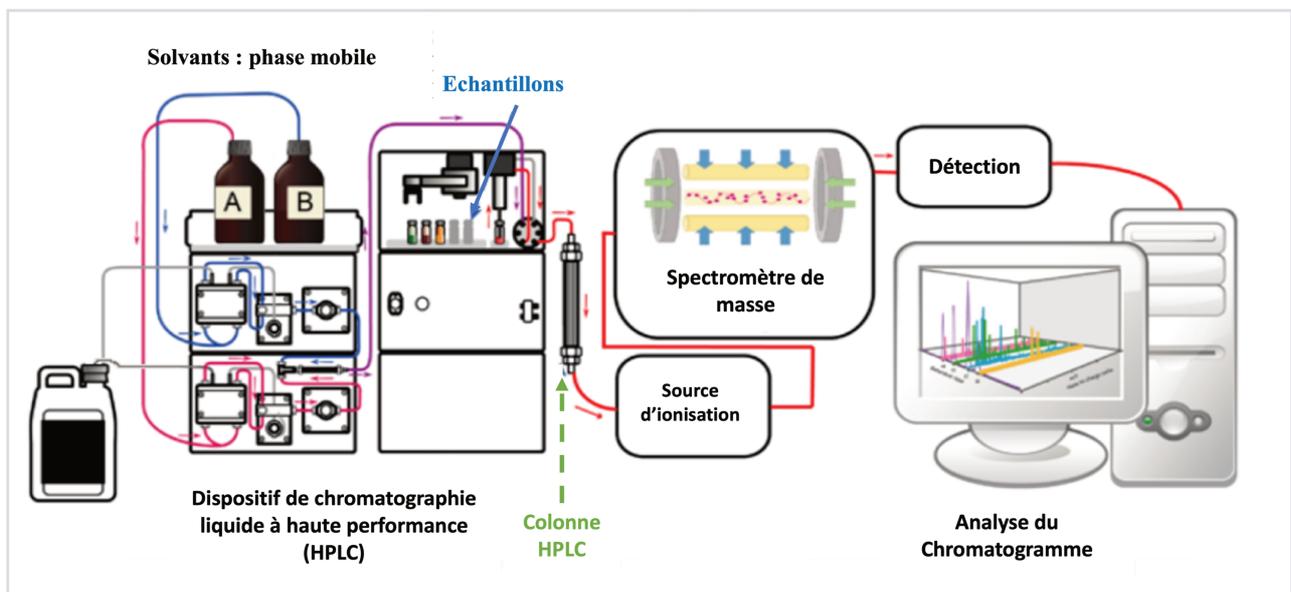
- L'échantillon préalablement préparé peut être conduit directement dans le spectromètre de masse (infusion directe), soit via un système préalable de séparation des métabolites : chromatographie en phase liquide (LC) ou chromatographie en phase gazeuse (GC).
- Ionisation des molécules par une source d'ionisation pour permettre leur entrée dans le spectromètre de masse
- Séparation des molécules chargées en fonction de leur rapport masse sur charge ( $m/z$ ) puis fragmentation des ions majoritaires dans une cellule de collision
- Récupération par le détecteur des composants des ions séparés (avant et après fragmentation) afin de permettre leur caractérisation.

Cette technique est beaucoup plus sensible, donc intéressante en métabolomique non-ciblée pour obtenir un métabolome le plus complet possible ou ciblée sur des métabolites non majoritaires, mais ne permet qu'une quantification relative (variations entre échantillons) puisque l'intensité des pics mesurés va dépendre de l'efficacité d'ionisation, qui peut varier entre les métabolites et est aussi dépendante des conditions expérimentales, ce qui peut entraîner un défaut de reproductibilité (Powers and Riekeberg 2017).

Nous détaillerons ici la technique de **spectrométrie de masse couplée à une chromatographie en phase liquide (LC-MS)** (figure 8, (Kailasam 2021)), technique analytique utilisée dans ce travail. Dans cette technique, une phase liquide (mobile), entraîne les composants de l'échantillon retenus sur une phase stationnaire solide. La séparation des molécules est alors dû à la différence d'affinité des composants pour la phase mobile ou la phase stationnaire et à la différence de solubilité dans la phase liquide. Ainsi, plus la molécule a d'affinité pour la phase stationnaire, moins elle est facilement entraînée par la phase mobile, et donc plus elle est retenue sur la colonne de chromatographie. En faisant varier la composition de phase mobile, les métabolites se séparent puis sortent de la colonne les uns après les autres. La durée entre le temps d'injection et le temps de sortie de colonne d'un métabolite est appelé « **temps de**

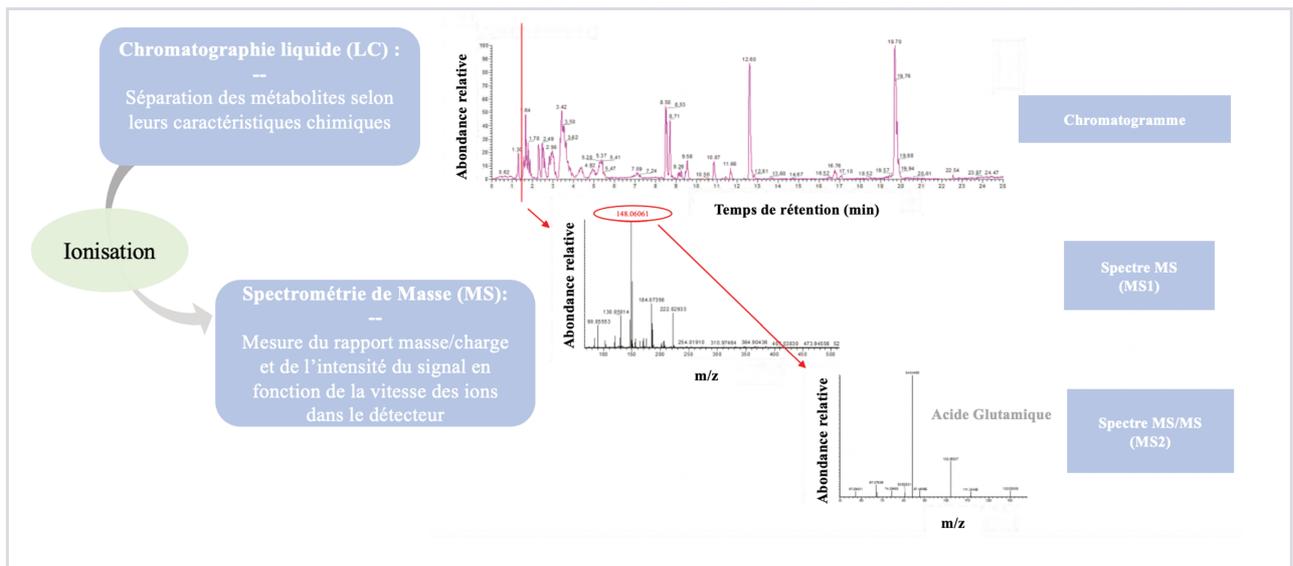
**rétenion (RT) ». Avec les colonnes (C18) et gradients de phase mobiles utilisés dans cette étude, les métabolites sont donc séparés par la chromatographie en phase liquide en fonction de leur polarité avec les plus polaires au début et le apolaires à la fin.**

Par la suite, les molécules sont ionisées par la source d'ionisation puis détectée par le spectromètre de masse permettant d'obtenir des **spectres de masse (Spectre MS1)** par détection des ions selon leur masse sur leur charge ( $m/z$ ). Le courant ionique total correspond à l'intensité de la somme des métabolites détectés à un certain temps de rétenion. L'intensité de chaque métabolite (chaque  $m/z$ ) est aussi identifiée. Des métabolites majoritaires sont ensuite fragmentés dans la chambre de collision et la fragmentation de chaque métabolite donne un **spectre MS/MS (spectre MS2)**. Les spectres MS2 peuvent ensuite être utilisés pour identifier la structure du métabolite. Plusieurs métabolites peuvent avoir le même rapport  $m/z$  mais souvent des temps de rétenions différents, ils sont alors identifiés grâce à leur fragmentation (correspondant aux caractéristiques structurales propres du métabolite) (Figure 9).



**Figure 8. Schéma d'une plateforme analytique en LC-MS**

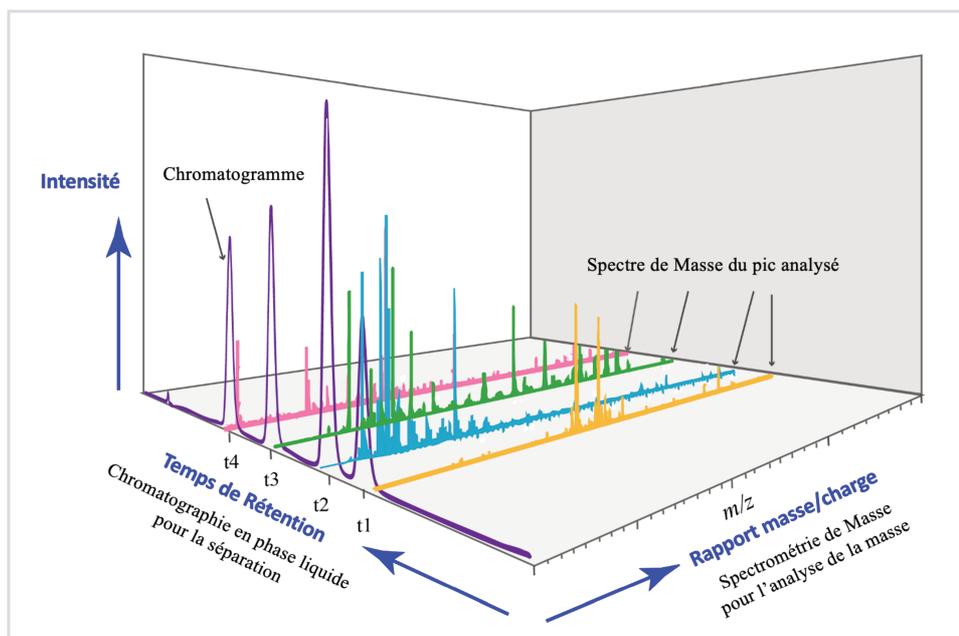
(Extrait de Kailasam, S. (2021) 'Analysis and LC-MS / MS', Technology Networks, (Lc))



**Figure 9. Schéma descriptif des différents temps de la technique LC-MS et des données spectrales générées.**

m/z : rapport masse/charge; MS1 : spectre de masse initial ; MS2 : spectre de masse après collision pour m/z 148,06061

Les variations de métabolites sont généralement dérivées des données MS1 (mesures semi-quantitative des courants ioniques relatifs). Pour exploiter les data bruts du spectromètre, celle-ci doivent être restructurées, débruitées et filtrées en utilisant des logiciels spécifiques (voir ci-après). A la fin du processus, il est donc obtenu des données à trois dimensions (temps de rétention, m/z, et intensité (dépendante de l'abondance du composé) (Figure 10).



**Figure 10. Analyse trois dimensions du spectre LC-MS**

(Extrait de Daniel Norena-Caro, Wikipédia)

La spectroscopie RMN et la LC-MS sont des techniques très utilisées actuellement en métabolomique. Chacune ayant ses avantages et ses faiblesses, la technique devra donc être choisie selon l'objectif de la recherche et la question posée. Elles peuvent même être combinées dans l'objectif d'élargir la couverture des métabolites analysés et de passer d'une quantification relative à une quantification absolue (Bouatra et al. 2013).

Dans ce travail, l'analyse par chromatographie liquide a été réalisée à l'aide d'un système UHPLC (Ultra High Pressure Liquid Chromatography) DIONEX Ultimate 3000 (Thermo Scientific™) et l'analyse par spectrométrie de masse a été réalisée avec un spectromètre de masse haute résolution et haute fréquence Q Exactive Plus Orbitrap (Thermo Scientific™) avec une source d'ionisation par électrospray, HESI II, fonctionnant en mode positif ou négatif.

#### **d. Traitement des données**

À la suite du passage de l'échantillon en spectrométrie de masse dans la plateforme analytique, il est donc obtenu des données instrumentales dites « données brutes ». Il s'agit de données complexes et leur traitement, qui passe par plusieurs étapes détaillées plus loin, ne peut s'effectuer qu'à l'aide de logiciels spécifiques (Sugimoto et al. 2012). Les logiciels utilisés dans ce travail ont été *MZmine 2* (en open source) et *Compound Discoverer™ 3.3* de Thermo Scientific™. Cette étape pré-analytique est indispensable avant de pouvoir procéder à l'analyse statistique et l'interprétation biochimique comprenant l'identification de biomarqueurs ou signatures métabolomiques d'intérêt. L'objectif du traitement des données brutes étant, pour chaque métabolite identifié, d'obtenir la masse, le temps de rétention et l'intensité relative dans chacun des échantillons.

Les principales étapes pré-analytiques sont les suivantes (Figure 11) :

- **Conversion des fichiers** : Les données brutes (*raw data*) n'étant lisibles que par les logiciels constructeurs de l'instrument, une étape de conversion de fichier, en format accessible aux outils de traitement et d'analyse et en format plus compacté, est parfois nécessaire (par exemple format *mzXML*) (Gorrochategui et al. 2016). Cette étape n'a pas été nécessaire avec les 2 logiciels utilisés dans ces études. En effet, ils traitent directement les raw data du spectromètre sous le format produit par le spectromètre.

- **Filtrage des bruits de fond et correction de la ligne de base** : Les instruments analytiques produisent un signal, même lorsqu'une matrice sans composés (blanc) est analysée. Ainsi, il est fréquent que soient détectés des ions isotopes, des induits ou des fragments produits par la source d'ionisation se trouvant au même temps de rétention mais présentant des  $m/z$  différentes. La présence d'un « bruit de fond » dans les données LC-MS peut être aussi causé par l'instrument lui-même ou par les molécules contenues dans le solvant et cela entraîne une gêne dans la détection des pics. Il existe donc une limite de détection de l'instrument pour un métabolite correspondant à la concentration de cette molécule nécessaire pour produire un signal que l'on peut distinguer du bruit (Wells et al. 2011). Une étape de filtrage a pour objectif de réduire voire supprimer ces bruits de fond tout en préservant les pics d'intérêt.

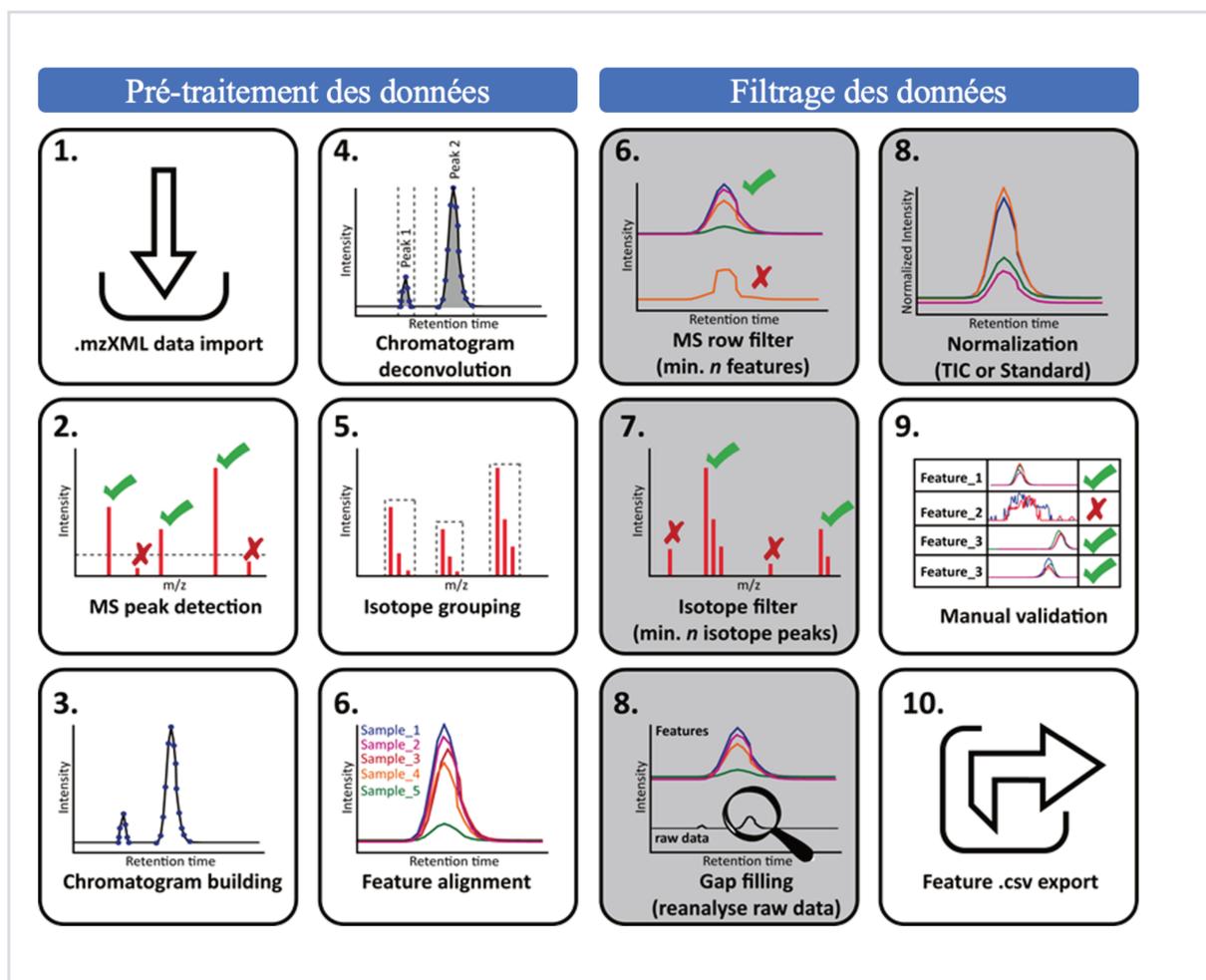
Il existe également un phénomène d'altération de l'intensité de base au fil du temps. L'étape de correction de la ligne de base consiste à soustraire cette dérive du signal brut. Ces modalités correctrices utilisent des algorithmes. Un point de vigilance à avoir est la génération de faux signaux par les algorithmes eux-mêmes.

- **Détection des pics et construction des chromatogramme** : Cette étape permet notamment de sélectionner les signaux sous forme de pics correspondant à de réels ions comparativement à ceux correspondant à de faux signaux, comme par exemple les bruits de fond (Gorrochategui et al. 2016). Ces signaux peuvent être sélectionnés selon plusieurs critères tels qu'une hauteur minimale du pic, un ratio signal sur bruit, la forme du pic ou encore le profil isotopique. Une fois le pic détecté, l'aire sous le pic et la hauteur maximale du pic sont calculées et ces éléments pourront être utilisés dans la quantification relative des métabolites.
- **Alignement des pics** : L'objectif d'une étude en métabolomique étant de rechercher la présence de métabolites mais également de comparer la variation d'un métabolite entre plusieurs échantillons, les pics détectés dans chacun des échantillons doivent être alignés pour chaque ion au niveau de son  $m/z$  et de son temps de rétention. Avec *Compound Discoverer*<sup>TM</sup>, cet alignement se fait sur le chromatogramme complet alors qu'il se fait sur les chromatogrammes d'ion extrait (associés à un seul ion d'intérêt) avec le logiciel *MZmine*<sup>TM</sup>.

Une étape de **filtrage des données** est ensuite nécessaire afin de pouvoir analyser statistiquement des données fiabilisées et pertinentes. Une nouvelle fois, plusieurs techniques co-existent et peuvent être utilisées lors d'une même étude métabolomique.

- **Le MS row filter** : Technique permettant de ne conserver que les pics ayant un minimum d'occurrence c'est à dire retrouvés dans un nombre d'échantillons suffisant, dans l'objectif de ne garder que les métabolites pertinents.
- **Le Filtrage des isotopes** : Cette technique permet de conserver uniquement les pics de molécules accompagnées de leurs isotopes. Ceci permet d'éliminer des pics provenant d'artefact qui ne correspondent donc pas à de réels métabolites. Compound discovered™ regroupe les isotopes mais également les différents adducts selon un seul ion généralement l'adduct majoritaire.
- **Comblement des valeurs manquantes (le « gap filling »)** : Dans les fichiers informatiques princeps, obtenus suite au passage des échantillons dans la plateforme analytique, certains pics peuvent ne pas être visualisés, comme par exemple lorsque le pic est inférieur aux limites de détection fixées par le logiciel traduisant les données brutes en fichier utilisable par nos logiciels de routine. Cette étape de gap filling, permettant d'aller chercher une deuxième fois dans la row data (donnée brute) la présence d'un pic (à temps de rétention et m/z donné) et donc de récupérer certaines valeurs manquantes, est une étape nécessaire.
- **La normalisation** : Une étape de normalisation peut être utilisée afin de limiter des biais pouvant être générés, comme par exemple, des différences entre les différentes mesures dues à la plateforme analytique et non aux réelles variations biologiques. Cette normalisation peut être réalisée par ligne (donc entre les différents échantillons) ou par colonne (donc entre les différents métabolites). Les variations entre les échantillons peuvent toutefois être limitées dès la phase instrumentale, en passant les échantillons dans un ordre aléatoire dans la plateforme analytique ou contrôlées en utilisant des contrôles qualité (QC) par exemple (Dunn et al. 2011). L'échantillon QC contient un volume identique de chaque échantillon et est injecté à intervalle régulier lors de l'analyse instrumentale. Les variations d'intensité dans chaque QC sont alors représentatives de la répétabilité analytique et permettent de calculer des coefficients de variation pour chacun des signaux.
- Pour finir, les logiciels permettent de prédire ou d'identifier le métabolite associé à l'ion détectée. Ainsi, **une prédiction** est réalisée par le logiciel *MZmine*™ à partir de la MS1. Nous utilisons essentiellement la banque de données de métabolites de l'*HMDB* (The Human Metabolome Database). Cette prédiction basée sur le m/z en considérant un induit H- pour les ions négatifs et H+ pour les ions positifs. Bien qu'assez précise, cette étape n'est pas assez satisfaisante et nécessite encore d'être contrôlée et validée par une

**identification manuelle**, sur les métabolites d'intérêt qui sera détaillée par la suite dans le manuscrit. Pour exemple, plusieurs identifications (plusieurs métabolites possibles) peuvent être proposées pour un même pic lors de la pré-identification par le logiciel. Compound Discoverer™ permet une identification via la base ChemSpider par la MS1 et surtout via la base MZ-Cloud par la MS1/MS2. Ainsi, certains métabolites peuvent être directement identifiés par l'utilisation de Compound Discoverer™ et de MZ-Cloud.



**Figure 11. Schéma de la procédure pré-analytique du logiciel MZmine™**

(Extrait adapté de GNPS documentation <https://gnps.ucsd.edu/>)

m/z : rapport masse/charge ; MS : spectre de masse ; n : nombre ; TIC : total ion current/count.

A la fin de ces étapes pré-analytiques, un fichier de sortie est exporté. Cette matrice de données contient l'ensemble des signaux caractérisés par leur intensité, leur m/z et leur temps de rétention dans chaque échantillon.

Dans notre travail, une normalisation linéaire a été effectuée en utilisant l'intensité moyenne de chaque échantillon comme facteur de normalisation. Seuls les métabolites ne présentant aucune valeur nulle, après gap filling, ont été gardés pour l'analyse statistique. Lors du filtrage des données, lorsqu'un métabolite était détecté à la fois dans des modes positifs et négatifs, seul le mode présentant l'intensité moyenne la plus élevée a été pris conservé. De plus, seuls les métabolites présentant des intensités moyennes élevées et ayant été détectés dans un grand nombre d'échantillon ont été conservés (filtrage par l'intensité moyenne et par le raw number). Ces filtrages visent à enrichir en ions qui pourront être identifiés comme des métabolites connus. Devant un grand nombre de molécules détectées (plusieurs milliers par étude), il a été fait le choix de procéder à l'étape de validation manuelle après l'analyse statistique, afin de valider l'identité uniquement des métabolites d'intérêt.

#### **e. Analyse statistique**

L'objectif étant de rechercher de nouveaux biomarqueurs ou de nouvelles signatures multi-marqueurs répondant à une prédiction posée en comparant plusieurs groupes d'échantillons et les matrices de données complexes, l'analyse statistique ne peut se faire qu'en utilisant des méthodes adaptées.

**L'analyse univariée**, telle qu'utilisée en pratique courant en recherche clinique, permet de prouver la significativité des différences entre des échantillons de groupes différents (Wiener et al. 2004). La principale limite réside dans le fait qu'une seule variable peut être évaluée dans chaque test. Cependant, en approche non-ciblée les matrices de données contiennent souvent plusieurs centaines ou milliers de variables pouvant générer de faux positifs. Des mesures correctrices mais plus ou moins conservatrices (correction de Bonferroni, correction du False Discovery Rate...) peuvent être cependant appliquées pour limiter ce biais (Broadhurst and Kell 2007). L'autre frein concernant l'utilisation des méthodes d'analyses univariées vient du fait qu'elles traitent chaque ion analysé comme une variable indépendante, ce qui est très réducteur par rapport à la réalité biochimique puisque chaque métabolite provient d'une cascade biochimique (voie métabolique) entraînant des corrélations entre différents métabolites analysés issus, par exemple, de la même voie métabolique.

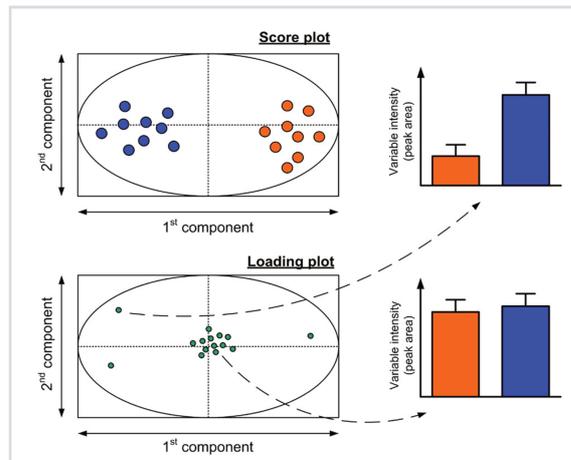
**Les analyses multivariées** apparaissent donc intéressantes en première approche en métabolomique non ciblée (les analyses univariées pouvant parfois être utilisées dans un second temps sur des données plus restreintes). Elles permettent de visualiser dans un espace réduit l'entièreté des données, donc la variabilité des métabolites de chaque échantillon et est donc plus adapté aux données à haute dimension. Dans l'objectif de pouvoir analyser ce grand jeu de données, pouvant ne pas suivre la loi de la normalité et de l'homoscédasticité, une étape préliminaire de transformation des données est parfois nécessaire. Dans notre travail, nous avons utilisé une transformation logarithmique des intensités des sommets des pics.

Il existe deux méthodes d'analyse multivariée : les méthodes non supervisées et supervisées. La méthode non supervisée ne nécessite pas la connaissance de l'appartenance à un groupe et fourni des informations précieuses sur la structure sous-jacente des données. Elle est donc souvent réalisée dans un premier temps, avant la réalisation de l'analyse supervisée (Beisken, Eiden, and Salek 2015) pour contrôler les données.

- **L'analyse multivariée non supervisée** est une analyse descriptive sans à priori, c'est-à-dire sans avoir connaissance des groupes. Elle permet, en analysant les variables (ici des métabolites et leurs variations) de déterminer des groupes d'échantillons et de révéler les relations entre les variables définissant ces groupes. Nous décrivons la méthode non supervisée utilisée dans ce travail, à savoir **l'analyse en composantes principales (ACP ou PCA en anglais)**.

L'ACP est une transformation orthogonale de données multivariées permettant de découvrir des structures internes cachées en construisant des composantes principales (en général 2 ou 3 composantes) qui décrivent la variance maximale des données (Boccard, Veuthey, and Rudaz 2010). Généralement, la première composante explique la variance la plus important puis les variances plus réduites. Les échantillons sont ensuite représentés, au sein d'espaces plus réduits, dans des graphiques à 2 ou 3 dimensions (Figure 12). Le *graphique de score* ou *score plot* représente la projection en deux dimensions des échantillons où les échantillons appartenant à un même groupe sont proches les uns des autres. Le *graphique de poids* ou *loading plot* représente la contribution de chaque variable à chacune des composantes principales où les variables responsables de la séparation des groupes sont éloignées du centre (Roux et al. 2011). Cette construction graphique est également utile dans l'identification de valeurs aberrantes, permettant de mettre en avant des échantillons pouvant présenter des

données anormales (un problème d'échantillonnage, effet pouvant être dû à des analyses LC-MS différentes d'échantillons d'une même série par exemple).



**Figure 12. Représentations graphiques des résultats d'une ACP**

(extrait de Roux A, et al. (2011). "Applications of Liquid Chromatography Coupled to Mass Spectrometry-Based Metabolomics in Clinical Chemistry and Toxicology : A Review." *Clinical Biochemistry* 44(1):119–35.

Le graphique des scores (score plot) représente la projection en deux dimensions des échantillons sur les composantes principales. Le graphique de poids (loading plot) représente le poids des variables sur les composantes. Les variables responsables de la discrimination entre les groupes sont éloignées du centre.

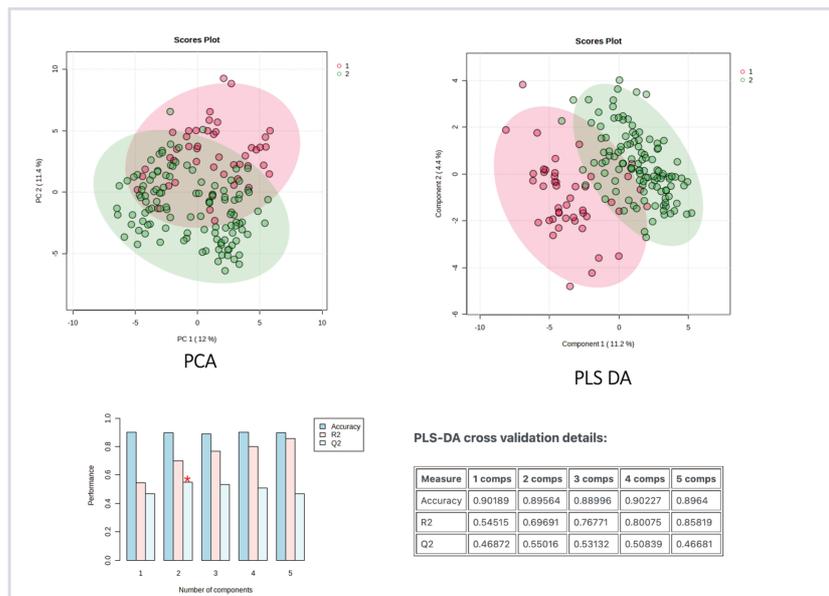
- Lors de **l'analyse multivariée supervisée**, les échantillons sont regroupés en classes. Pour exemple, dans notre travail, les classes étaient représentées par les niveaux de gravité de l'atteinte par la COVID-19. Cette analyse a pour objectif d'expliquer ces groupes par les variables explicatives. Parmi les méthodes d'analyse supervisée, nous avons utilisé la **régression des moindres carrés partiels** encore nommée **PLS (Partial Least Squares)**. Cette technique de classification cherche à maximiser la corrélation entre les variables (X) et les variables à expliquer (Y). Elle combine la réduction de la dimensionnalité et l'analyse discriminante, en regroupant l'ACP et la régression multiple, en un seul algorithme et s'applique particulièrement à la modélisation des données de haute dimension (Lee, Liang, and Jemain 2018). Lorsque les variables à expliquer sont qualitatives (comme nos groupes prédéfinis de patients), la **PLS-DA (Partial Least Squares Discriminant Analysis)** est utilisée.

L'algorithme construit donc un espace de faible dimension basé sur des combinaisons linéaires des variables X, tout en ajustant le modèle pour y intégrer les variations liées à Y. Dans la PLS-DA, l'algorithme cherche à obtenir une séparation maximale entre les différents groupes prédéfinis. Comme c'est le cas pour l'ACP, la discrimination des

différentes classes est visualisée sur le score plot et la contribution de chaque métabolite à la discrimination sur le loading plot.

La validation du modèle construit par la PLS-DA s'évalue par la **validation croisée (cross validation en anglais)**. Une partie des échantillons est alors retirée du modèle pour servir de série de validation. Le modèle de discrimination est construit avec les échantillons restants (série d'apprentissage) et le pouvoir prédictif du modèle est estimé grâce à la série de validation retirée initialement. Ce processus est répété jusqu'à ce que tous les échantillons aient été retirés au moins une fois. Ce processus permet de valider la signification statistique de chaque modèle. Cette dernière est estimée par la **précision de la performance (accuracy)**, le **R<sup>2</sup>** (proportion de variance expliquée par le modèle PLS-DA) et le **Q<sup>2</sup>** (capacité prédictive du modèle) et l'**aire sous la courbe ROC (AUROC)** (Szymanska et al. 2012). Le modèle est considéré comme fiable quand le R<sup>2</sup> et le Q<sup>2</sup> sont supérieurs à 0,5 et les valeurs de ces derniers doivent se rapprocher le plus de 1. Un **test de permutation** peut également être réalisé. Dans ce cas, les échantillons sont répartis dans les groupes au hasard. Ceci devant entraîner une diminution du R<sup>2</sup> et du Q<sup>2</sup>, prouvant alors que les résultats obtenus initialement ne sont pas dus au hasard.

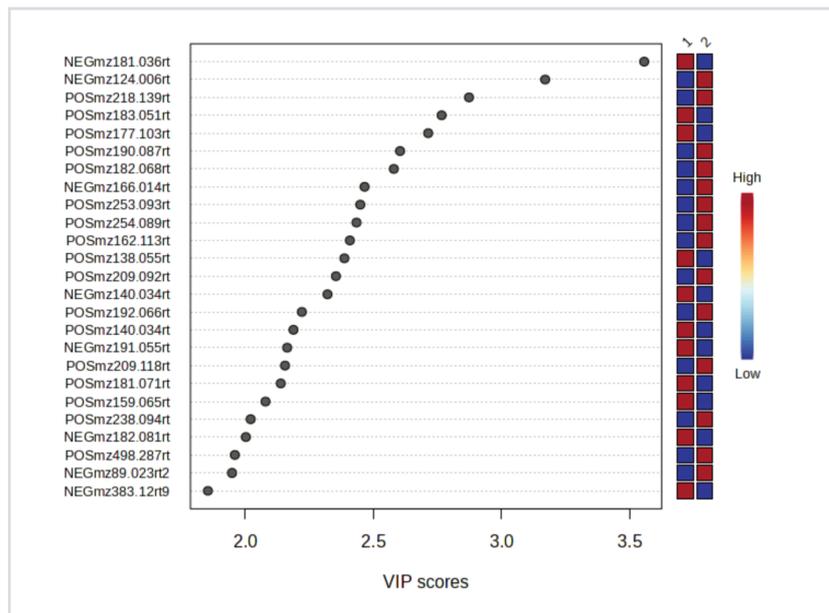
La figure 13 représente un exemple de PCA et de PLS-DA avec cross validation et de représentation graphique par histogrammes des données de la cross-validation en PLS-DA.



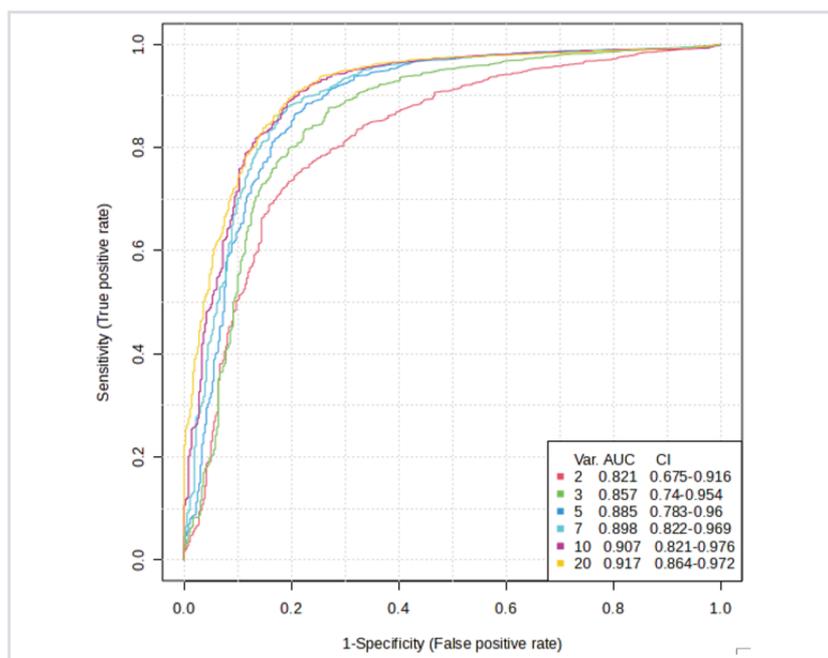
**Figure 13. Exemple de PCA et de PLS-DA avec cross validation et représentation graphique par histogrammes des données de la cross-validation en PLS-DA.**  
PC : principal component

Une fois le modèle obtenu par la PLS-DA validé, le poids de chaque variable (métabolite) sur le modèle peut être estimé par le calcul de la **Variables Importance in Projection (VIP)**. Plus la VIP est haute, plus le métabolite participe à la discrimination des groupes. Habituellement, les variables ayant un score de VIP supérieur à 1 sont sélectionnées mais cette valeur seuil est encore débattue (Gorrochategui et al. 2016)

L'analyse PLS-DA permet d'extraire une « top liste » de métabolites d'intérêt correspondant à  $n$  métabolites les plus intéressants dans le modèle et représentant la **signature métabolomique** de la question posée (figure XX). Le nombre de métabolites à intégrer dans cette signature métabolomique pourra être évalué par les performances données par les calculs d'AUROC (Figures 14 et 15).



**Figure 14. Exemple d'une top liste de métabolites issue de l'analyse en PLS-DA**



**Figure 15. Exemple de courbes ROC**

Comparaison des performances d'une top liste intégrant 20, 10, 7, 5, 3 et 2 métabolites.

Dans notre travail, l'ensemble de ces analyses statistiques a été conduit avec le logiciel *MetaboAnalyst*<sup>TM</sup> version 5.0 puis 6.0 (<https://www.metaboanalyst.ca/>) (Pang, Chong, et al. 2021).

#### f. Validation des métabolites

Comme expliqué précédemment, une étape de d'identification manuelle des métabolites a été nécessaire. La validation des métabolites consiste en la comparaison entre le spectre MS2 obtenu pour un ion et les données spectrales (expérimentales ou théoriques) des bases de données préétablies. Dans notre travail, la base de données *The Human Métabolome DataBase version 5.0 (HMBD)* (<https://hmdb.ca/>) a été principalement utilisée (Wishart et al. 2022). Elle contient actuellement 220 945 métabolites et est reliée par des hyperliens à différentes autres bases de données (KEGG, PubChem, MetaCyc, ChEBI, PDB, UniProt et GenBank).

Il s'agit donc de comparer les valeurs de m/z mesurées (MS1 et MS2) avec celles des molécules référencées dans cette base de données.

Une même valeur de m/z pouvant correspondre à plusieurs molécules, la MS2 est utilisée pour procéder à cette identification (Vinaixa et al. 2016). Cette stratégie à profit le fait que les fragments obtenus par MS2 sont généralement assez spécifiques pour un métabolite donné.

Ainsi, la correspondance de chaque MS2 expérimentale de nos analyses à la MS2 de la base HMDB a été « manuellement » vérifiée et ce pour une correspondance des fragments de chaque métabolite d'intérêt.

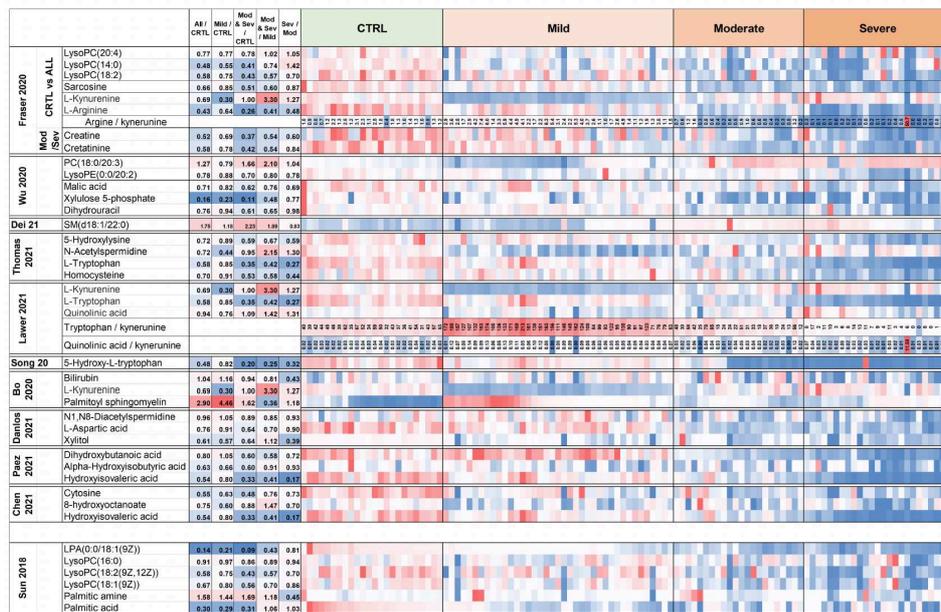
Parallèlement au traitement des données de sortie du spectromètre par MZmine, nous avons aussi utilisé un autre logiciel Compound Discoverer™ pour le traitement ces raw data. Il effectue des opérations similaires à MZmine mais avec des algorithmes différents. Nous avons gardé les ions identifiés par les 2 logiciels afin de ne garder que des ions les plus fiables. A noter que la quantification relative basée sur des calculs d'aires sous les pics réalisée par ce logiciel s'était avérée beaucoup moins fiable pour nos analyses que celle de MZmine basée sur les hauteurs de pics. Pour caractériser les ions d'intérêt, nous avons aussi mis à profit la capacité de Compound Discoverer™ à identifier et regrouper les différents adducts et isotopes d'une molécule mais également l'utilisation de base de données pour identifier l'ion via sa MS2. En effet, selon notre paramétrage par défaut, l'analyse par Compound Discoverer™ permet une identification via la base MZ-Cloud.

Dans notre travail, il a été choisi de ne valider que les métabolites entrant dans les « top listes » des modèles prédictifs. Lorsqu'un des métabolites sélectionnés n'était pas été identifié par sa MS2, il était retiré et une nouvelle analyse PLS-DA était alors réalisée (afin de ne pas sous-estimer l'influence du retrait de ce métabolite dans la construction du modèle).

Une fois toutes ces étapes analytiques réalisées, les courbes ROC et le calcul de l'AUROC ont été établies. Cette stratégie a permis de prédire le nombre de métabolites valisés par MS2, à intégrer à la signature métabolique, le plus performant statistiquement pour répondre à la question posée.

### g. Construction de Heat Map

La création sur le logiciel Excel™ d'une carte thermique (Heat Map) permet une visualisation didactique du tableau de données brutes (Figure 16). Elle permet d'afficher, selon les intensités de couleur, les abondances relatives des métabolites détectés entre chaque échantillon et entre les groupes d'intérêt (Benton et al. 2015). Elle permet également de mettre en avant certaines données aberrantes.



**Figure 16. Exemple de Heat Map.**

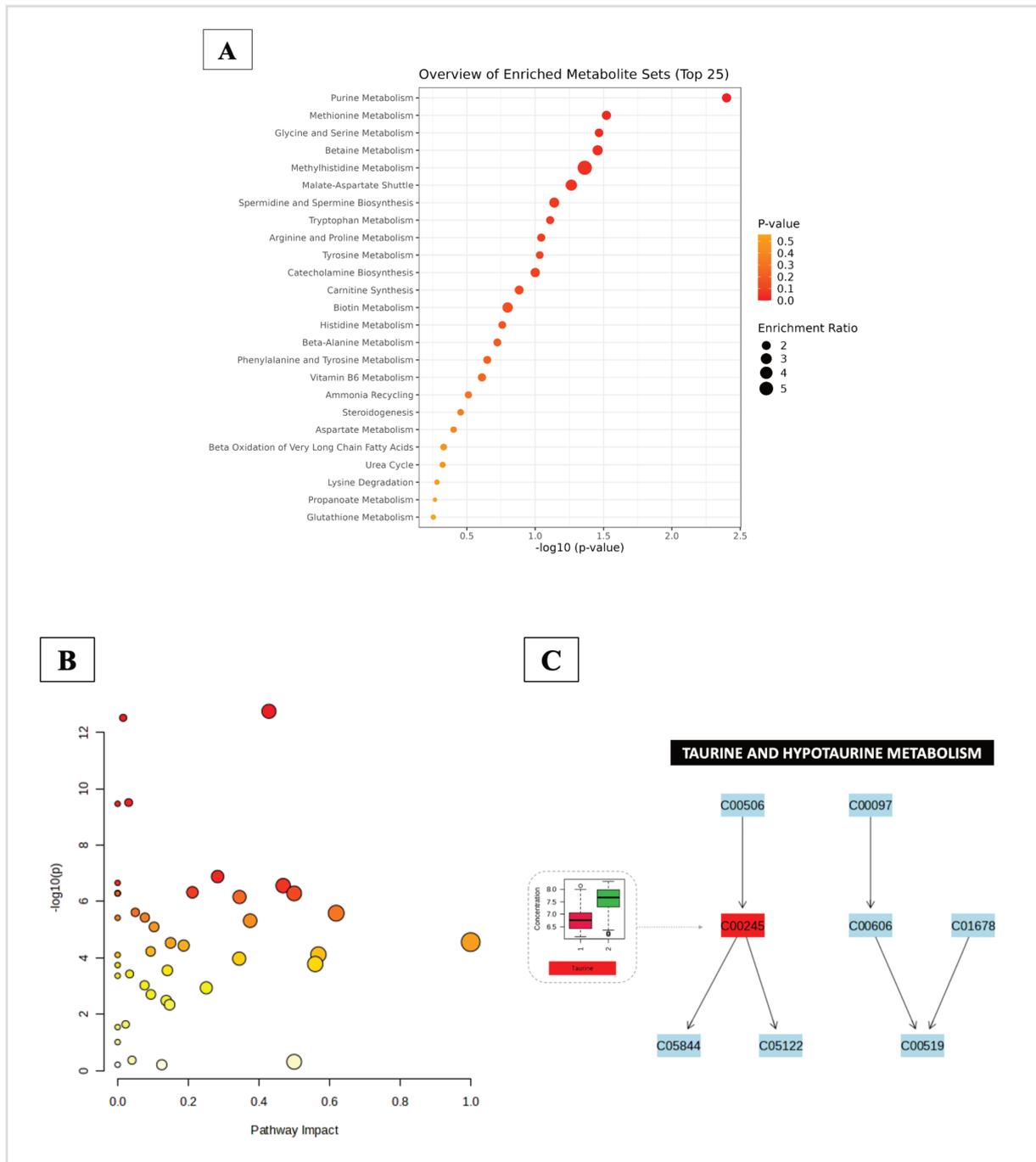
Extrait de Occelli et al., (2022) Untargeted plasma metabolomic fingerprinting highlights several biomarkers for the diagnosis and prognosis of coronavirus disease19. Front. Med. 9:995069.

### h. Analyse des voies métaboliques impliquées

Une fois les signatures métabolomiques et les potentiels nouveaux biomarqueurs identifiés, il reste à comprendre les mécanismes biologiques et biochimiques concernés. Cette étape passe par l'analyse des voies métaboliques impactés par le phénomène étudié. De nombreux logiciels permettent cette analyse (Marco-ramell et al. 2018). Dans la continuité de nos analyses statistiques, il a été décidé d'utiliser la plateforme *MetaboAnalyst* qui propose cette fonctionnalité (Chong, Wishart, and Xia 2019)

L'étude des voies métaboliques entrant en jeu dans la question posée combine plusieurs analyses (Figure 17). Tout d'abord, **l'analyse de surreprésentation** qui vérifie, par un test exact de Fisher ou un test hypergéométrique, si le nombre de métabolites associés à une voie métabolique est dû au hasard ou bien secondaire au phénomène étudié. La principale limite de l'analyse de surreprésentation étant que l'abondance relative des métabolites n'est pas intégrée et donc les métabolites ayant une très faible variation pourrait peser autant que des métabolites à très forte variation. Lorsque les concentrations de métabolites sont connus, il est possible d'étudier le **phénomène d'enrichissement**, qui compare la distribution des métabolites appartenant à chaque voie métabolique à une banque de données de référence. Si les métabolites appartenant à une voie sont observés plus fréquemment ou en plus grande quantité que la référence, la voie métabolique identifiée est alors considérée comme enrichie, c'est-à-dire ayant été impactée par le phénomène étudié (Manzoni et al. 2018). Une **analyse topologique** des voies peut venir compléter ces analyses, en permettant d'évaluer pour chaque métabolite d'intérêt, sa localisation par rapport au réseau de voies métaboliques et son degré de connectivité avec les métabolites voisins.

Pour chaque voie métabolique est rendu un taux d'enrichissement et une valeur-p correspondant à la probabilité d'un tel enrichissement. Ces données peuvent être traduites en graphique comprenant en axe des ordonnées la valeur- p et en axe des abscisses le taux d'enrichissement.



**Figure 17. Exemple de graphiques d'analyse des voies métaboliques par MetaboAnalyst**

A : Top liste de voies métaboliques enrichies issues d'une analyse d'enrichissement des voies métaboliques

B : Graphique représentant l'analyse globale des voies métaboliques impactées. Chaque cercle représente une voie métabolique spécifique

C : Schéma d'une voie métabolique impactée (obtenue en cliquant sur le cercle du graphique A). Les métabolites identifiés dans les échantillons et entrant en jeu dans la voie métabolique sont représentés en couleur. Il est possible, pour chaque métabolite, d'obtenir les différences de concentration dans les échantillons des deux groupes étudiés.

## 2) Intérêt de la métabolomique dans le domaine médical

La métabolomique est un outil puissant permettant de caractériser, à un instant donné et au sein d'une grande gamme de type d'échantillon (plasma, serum, urines, air expiré...), le profil métabolique d'un système biologique (S. Claus and Guillou 2019).

Elle trouve alors sa place dans de nombreux domaines du monde médical :

- **En pharmacologie** : pour mieux comprendre les mécanismes d'action des médicaments, rechercher des biomarqueurs reflétant leur effet sur l'organisme et notamment leur toxicité ou leur efficacité. A l'heure de la médecine personnalisée, la pharmaco-métabolomique permettrait d'évaluer les effets d'une molécule dans des sous-groupes de patient voire à l'échelle d'un patient et donc de prédire la réponse au traitement (Clayton et al. 2006).
- **En toxicologie** : afin de mieux comprendre les effets physiopathologiques du toxique ou d'identifier de nouveaux biomarqueurs ou signatures d'exposition (Lenski and Bruno 2023).
- **Dans l'agro-alimentaire et la nutrition** : où la nutrimétabolomique permet aujourd'hui de mieux comprendre les réponses métaboliques aux changements alimentaires ou les variations inter-individuelles des réponses à des nutriments, mais également l'impact sur le long terme de ces effets (S. P. Claus and Swann 2013). Des analyses des aliments par spectrométrie de masse permettent également d'évaluer la qualité d'un produit ou la présence d'un polluant dans ce dernier.
- **En cancérologie** : Dans ce domaine, la métabolomique est très prometteuse et trouve sa place dans chaque étape, allant du diagnostic en permettant l'identification de nouveaux biomarqueurs, au pronostic (Bailleux et al. 2023; Chardin et al. 2023; D'Andréa et al. 2023), en passant par le développement de nouvelles thérapies ciblées (Armitage and Barbas 2014).
- **Dans la recherche sur le microbiote** : la métabolomique permettrait une meilleure compréhension des effets du microbiote sur son hôte mais également des effets de l'environnement et de l'hôte lui-même sur le microbiote (McDonald et al. 2018).

- **En médecine légale** : où l'analyse de prélèvements réalisés post-mortem pourrait par exemple aider à identifier la présence de pathologies (Nariai et al. 2022).
- **Dans les maladies cardiovasculaires** : où la découverte de signatures métabolomiques laisse espérer une meilleure prédiction du risque d'évènement cardiovasculaire, notamment dans la coronaropathie et permettrait de mieux identifier les cibles thérapeutiques (Doğan 2024) et dans le pronostic et l'étiologie des accidents vasculaires cérébraux (Suissa et al. 2020, 2021).
- Dans l'aide à la **compréhension physiopathologique des nouvelles maladies et dans l'identification de nouveaux biomarqueurs diagnostiques et pronostiques**, comme l'illustre ce travail de thèse portant sur la COVID-19 et comme l'a largement montré la communauté scientifique travaillant sur les sciences omiques dans cette pathologie. Cette discipline est également très prometteuse pour des investigations dans des pathologies déjà bien connues et continuer de faire progresser les connaissances biomédicales et permettre de pratiquer une médecine plus personnalisée au patient.

La place de la métabolomique dans le domaine médical n'est donc plus à prouver et le développement constant des plateformes analytiques et des logiciels de traitement de données lui promet un bel avenir.

**Partie II : Approche métabolomique à la  
recherche d'outils diagnostiques et pronostiques  
dans la COVID-19**

## Chapitre 1 : Objectifs de la thèse

---

L'objectif de ce travail de thèse était d'identifier des signatures métabolomiques plasmatiques et urinaires diagnostiques et pronostiques de la COVID-19

Notre premier travail s'est déroulé sur une bio-banque d'échantillons plasmatiques issue d'une cohorte de patient atteint de la COVID-19 et pris en charge sur le Centre Hospitalier Universitaire (CHU) de NICE lors de la première vague COVID et de volontaires sains représentés par du personnel soignant (Ruetsch et al. 2021) (NCT04355351). Nous avons débuté ce travail entre la première et la deuxième vague COVID. L'utilisation d'outils de biologie moléculaire délocalisés dans les services d'urgences n'était pas encore généralisée et il existait un réel enjeu dans la découverte d'outils pronostiques afin de pouvoir caractériser de la manière la plus fiable et précoce possible les patients à risque de mauvaise évolution.

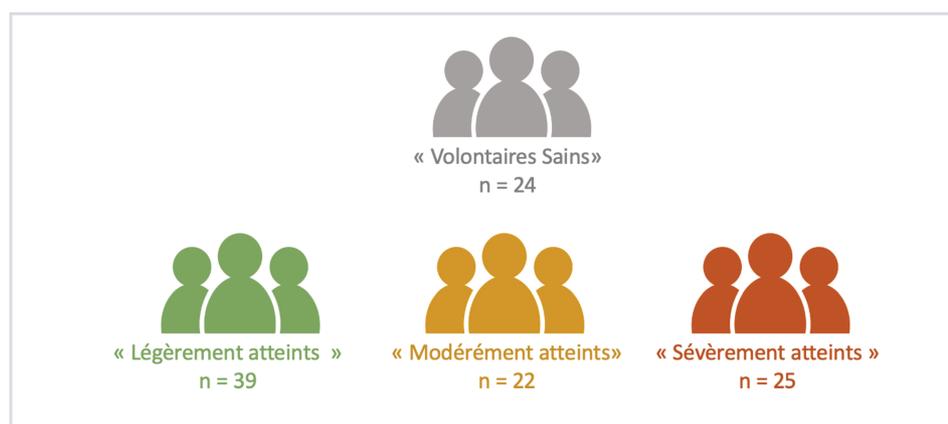
L'objectif principal de ce travail, basé sur une approche d'analyse en métabolomique non ciblée, était d'identifier des signatures métabolomiques plasmatiques permettant une classification robuste des patients COVID-19 asymptomatiques, légèrement, modérément et sévèrement atteints, et donc de prédire l'évolution de l'état de santé du patient. Nous souhaitons également évaluer si la métabolomique nous permettait d'améliorer nos connaissances sur les mécanismes biologiques sous-jacents de la maladie.

La seconde partie de cette thèse, basée sur l'analyse du métabolome urinaire des patients infectés par le SARS-CoV-2, pour objectif de rechercher des signatures diagnostiques et pronostiques pouvant être obtenues sur des prélèvements moins invasifs, moins douloureux et donc plus acceptable par la population.

Pour ce faire, nous avons mené une étude prospective (*Etude SignCov*, NCT04497272) au sein du Département Hospitalo-Universitaire de Médecine d'Urgence et du service de Médecine Intensive et Réanimation du CHU de Nice dont les inclusions se sont déroulées entre les mois d'août 2020 et mars 2021.

## Chapitre 2 : Identification d'une signature métabolomique plasmatique diagnostique et pronostique de la COVID-19

Un total de 110 échantillons plasmatiques issus de la bio-banque (Ruetsch et al. 2021) (NCT04355351) a été analysé, 86 correspondaient à prélèvements réalisés sur des patients atteints de la COVID-19 et 24 à des volontaires sains. A noter que les prélèvements analysés étaient des prélèvements réalisés lors de l'admission du patient en structure hospitalière. Les patients ont été considérés comme des cas de COVID-19 selon la classification de l'OMS : présence d'une atteinte cutanée en faveur (engelures) ou présence d'un scanner présentant des lésions pulmonaires caractéristiques de la COVID-19 ou deux tests RT-PCR positifs consécutifs pour le virus SARS-CoV-2 ou des tests sérologiques positifs. Les volontaires sains étaient des professionnels de santé qui présentaient un statut sérologique négatif pour le SARS-CoV-2. Dans l'objectif d'évaluer l'apport pronostique potentiel d'une signature métabolomique, les prélèvements étaient en 4 groupes : « **volontaires sains** », « **légèrement atteint** », « **modérément atteint** » et « **sévèrement atteint** ». Ainsi, le groupe « légèrement atteint » (n=39) comprenait les échantillons de patients présentant des symptômes bénins (engelures ou symptôme ne nécessitant pas de suivi hospitalier), le groupe « modérément atteint » (n=22) concernait les patients hospitalisés dans des services d'hospitalisation conventionnelle et le groupe « sévèrement atteint » (n=25) concernait les patients nécessitant une hospitalisation dans un service de soins intensifs (Figure 18).



**Figure 18. Description des groupes de patients et volontaires sains dans le travail de recherche intéressant les prélèvements plasmatiques**

## 1) Préparation des échantillons et analyse LC-MS

Un volume de 100 µl de plasma a été mélangé à 100 µl d'H<sub>2</sub>O (qualité HPLC, Merck Millipore, USA), et 600 µl de méthanol (qualité HPLC, Merck Millipore, USA) ont été ajoutés. Les échantillons ont été vortexés et incubés pendant une nuit à - 20 °C pour la précipitation des protéines. Après 15 minutes de centrifugation à 4°C (force 13 000 g), le surnageant a été retiré, puis séché à l'aide d'un évaporateur centrifuge SpeedVAC (SVC100H, SAVANT, Thermo Fisher Scientific, Illkirch, France), pour enfin être remis en suspension dans 80 µl de mélange d'un mélange acétonitrile-H<sub>2</sub>O (ratio 20/80) (qualité HPLC, Merck Millipore) et conservé à - 20 °C jusqu'à son utilisation pour les analyses en spectrométrie de masse.

Les chromatographies ont été effectuées avec le système HPLC Ultimate 3000 de DIONEX couplé à une colonne chromatographique (Phenomenex Synergi 4 u Hydro-RP80A 250 3,0 mm) réglée à 40 °C et à un débit de 0,9 ml/min. Les gradients des phases mobiles (phase mobile A : 0,1 % d'acide formique dans l'eau et phase mobile B : 0,1 % d'acide formique dans l'acétonitrile) ont été effectués sur une durée totale de 25 min. Les analyses MS ont été effectuées sur un spectromètre de masse Thermo Fisher Scientific Exactive Plus Benchtop Orbitrap. La source d'ionisation électrospray chauffée (HESI II) a été utilisée en mode ionique positif et négatif. L'instrument a été utilisé en mode de balayage complet de m/z 60 à m/z.

Le post-traitement des données a été réalisé à l'aide de MZmine version 2.39. Le remplissage des lacunes a été effectué à l'aide de la méthode "Same RT and m/z range gap filler" du logiciel MZmine. La tolérance m/z a été fixée à 0,001 m/z ou 10,0 ppm. Les métabolites ont été putativement annotés [correspondant au niveau 2 de la Metabolomics Standard Initiative (MSI)]<sup>2</sup> à l'aide de la base de données du métabolome humain version 6.03. Seuls les ions identifiés comme des adduits [M + H]<sup>+</sup> en mode positif et [M-H]<sup>-</sup> en mode négatif et les ions trouvés dans tous les échantillons après avoir comblé les lacunes (gap-filling) ont été inclus. Les métabolites identifiés comme des médicaments ou des composés exogènes ont été retirés.

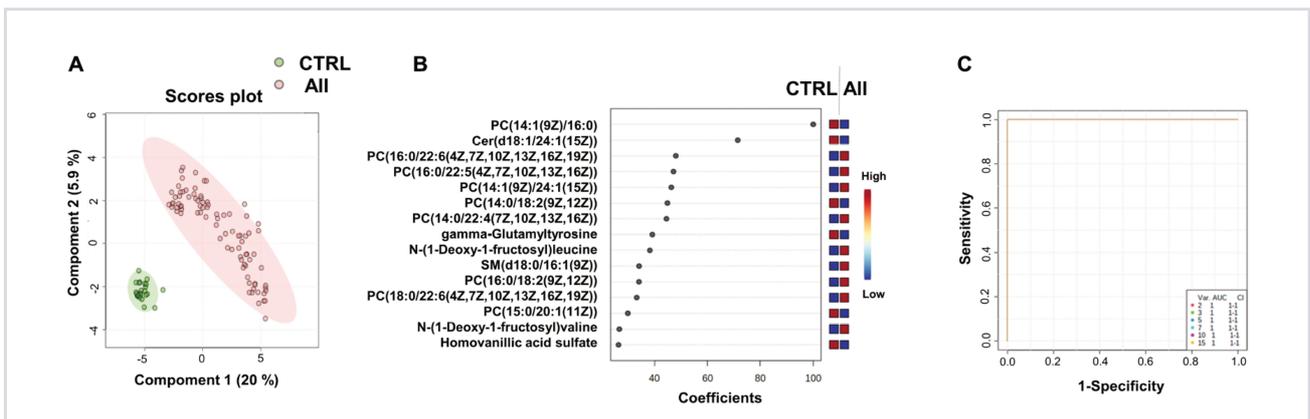
## 2) Présentation des principaux résultats

Dans ce travail, nous avons fait le choix de retenir des top listes de 15 métabolites, compromis qui permettait d'obtenir des performances diagnostiques et pronostiques satisfaisantes dans chacune de nos comparaisons, tout en ayant un nombre de métabolites limité.

### ▪ Signature métabolomique diagnostique

Dans l'objectif d'identifier une signature métabolomique diagnostique de l'infection à SARS-CoV-2 nous avons souhaité comparer le métabolome plasmatique des volontaires sains à celui de tous les sujets infectés (comprenant donc tous les échantillons des groupes légèrement, modérément et sévèrement atteints de la COVID-19). Après post traitement des données brutes issues de la LC-MS, nous avons sélectionnés 605 métabolites.

Dans cette comparaison, la signature métabolomique à 15 métabolites possédait une performance diagnostique avec une AUC à 1 (IC 95%, 1 – 1). (Figure 19). A noter que le modèle PLS-DA le plus performant était un modèle à 5 composantes permettant d'obtenir un  $Q^2$  à 0,99. Au sein de cette top liste de métabolites, les lipides semblaient fortement contribuer au modèle PLS-DA.



### Figure 19. Identification d'une signature métabolomique plasmatique diagnostique de la COVID-19

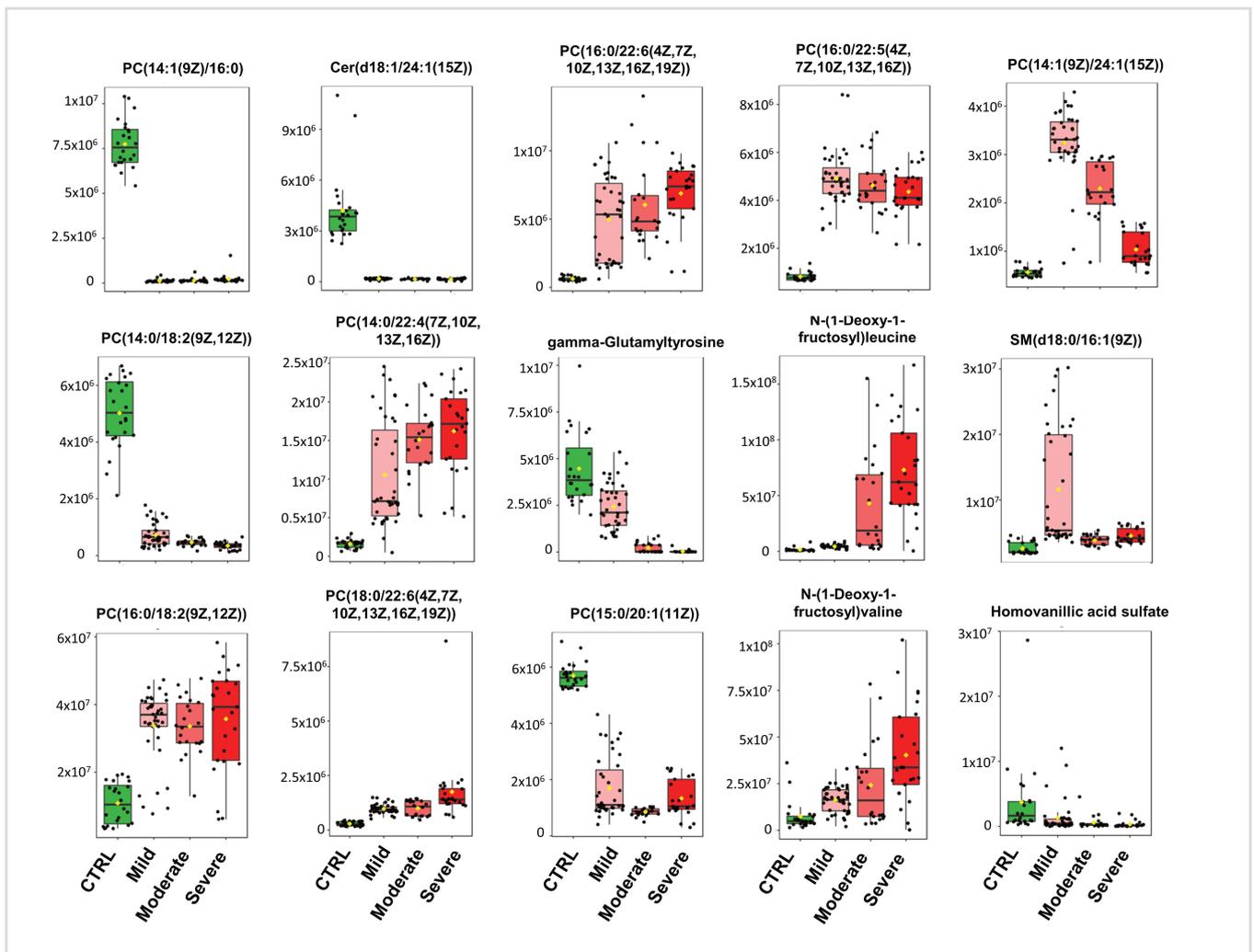
Les points verts représentent les échantillons plasmatiques du groupe de volontaires sains. Les points rouges représentent les échantillons plasmatiques de tous les patients atteints de la COVID-19.

A : Analyse PLS-DA. Le diagramme de score est représenté par une ellipse de confiance de 95 %

B : Diagramme représentant la top liste de 15 métabolites sélectionnées sur la moyenne des cinq premières composantes du modèle PLS-DA. Les métabolites identifiés ont été validés par des analyses MS2 et sont représentés sous forme de noms chimiques.

C : Comparaisons des aires sous les courbes ROC multivariées de chaque top liste de métabolite (comprenant respectivement 15, 10, 7, 5, 3 et 2 métabolites)

Afin de mieux comprendre les modifications de taux de ces métabolites, nous nous sommes intéressés aux variations (Box plot) de ces 15 métabolites d'intérêt. Nous avons ainsi remarqué que les taux plasmatiques de lipides saturés étaient plus faibles chez les patients atteints de COVID-19 comparativement au groupe contrôle (Figure 20) alors que les taux de lipides insaturés étaient quant-à-eux plus élevés. Nous avons également identifiés deux acides aminés (*N*-(1-désoxy-1-fructosyl)leucine et *N*-(1-désoxy-1-fructosyl)valine) dont les taux étaient plus élevés chez les patients infectés.

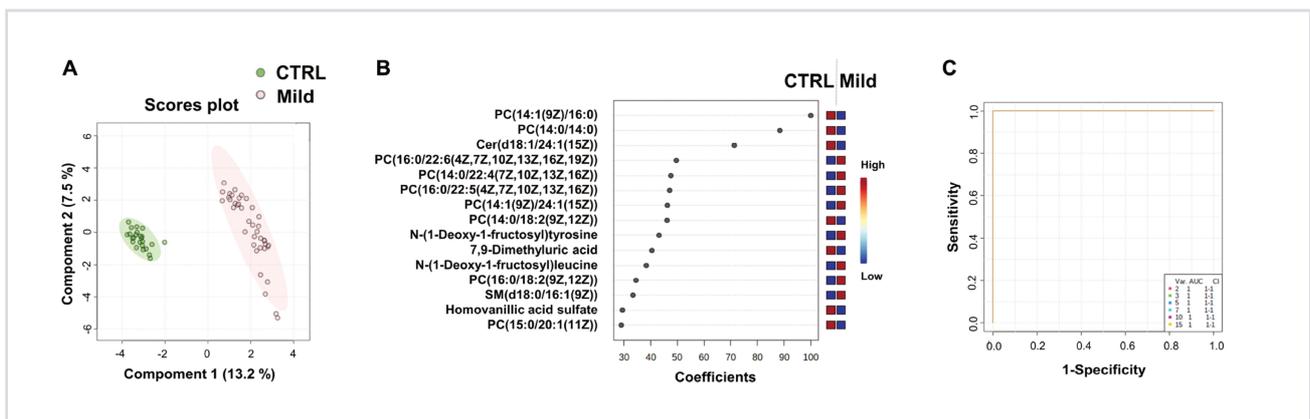


**Figure 20. Box plots des 15 métabolites composant notre signature métabolomique diagnostique de l'infection à SARS-CoV-2.**

Les métabolites ont été regroupés à l'aide de la méthode Ward sur la base des valeurs des tests t et ANOVA.

Dans l'objectif d'identifier une signature métabolomique qui serait cliniquement pertinente, nous avons souhaité poursuivre nos travaux sur la comparaison du métabolome des sujets sains avec celui des patients légèrement atteints de la COVID-19. Lors de cette comparaison, après post traitement des données, nous avons obtenus 574 métabolites.

Comme attendu, la top liste de 15 métabolites issue de cette comparaison comprenait des métabolites d'intérêts similaires. Une nouvelle fois, nous avons identifié des variations de taux de lipides et de deux acides aminés (*N*-(1-désoxy-1-fructosyl)tyrosine et *N*-(1-désoxy-1-fructosyl)leucine) (Figure 21). Concernant les variations de taux de lipides nous observons une nouvelle fois une baisse de taux pour les lipides saturés. Cette signature métabolomique plasmatique possédait, une nouvelle fois, une performance diagnostique forte avec une AUC à 1 (IC 95%, 1-1) que nous considérons 2 ou 15 métabolites.



**Figure 21. Comparaison du métabolome plasmatique des sujets sains avec celui du groupe légèrement atteint par la COVID-19**

Les points verts représentent les échantillons plasmatiques du groupe de volontaires sains. Les points rouges représentent les échantillons plasmatiques du groupe légèrement atteint par la COVID-19.

A : Analyse PLS-DA. Le diagramme de score est représenté par une ellipse de confiance de 95 %

B : Diagramme représentant la top liste de 15 métabolites sélectionnées sur la moyenne des cinq premières composantes du modèle PLS-DA. Les métabolites identifiés ont été validés par des analyses MS2 et sont représentés sous forme de noms chimiques.

C : Comparaisons des aires sous les courbes ROC multivariées de chaque top liste de métabolite (comprenant respectivement 15, 10, 7, 5, 3 et 2 métabolites)

## ▪ Signature métabolomique pronostique

Un des principaux enjeux au cours des premières vagues COVID, était de pouvoir identifier les patients infectés par le SARS-CoV-2 à risque de mauvaise évolution clinique et notamment à risque d'hospitalisation en service de soins critiques. Ces interrogations nous ont conduit à tenter d'identifier une signature métabolomique robuste qui permettrait de répondre à cette question. Dans cet objectif, nous avons par la suite réalisé deux nouvelles comparaisons. Tout d'abord nous avons comparé le groupe légèrement atteint à un groupe comprenant les sujets ayant nécessité une hospitalisation (patients modérément et sévèrement atteints). Puis, nous avons effectué la comparaison du groupe modérément atteint au groupe sévèrement atteint.

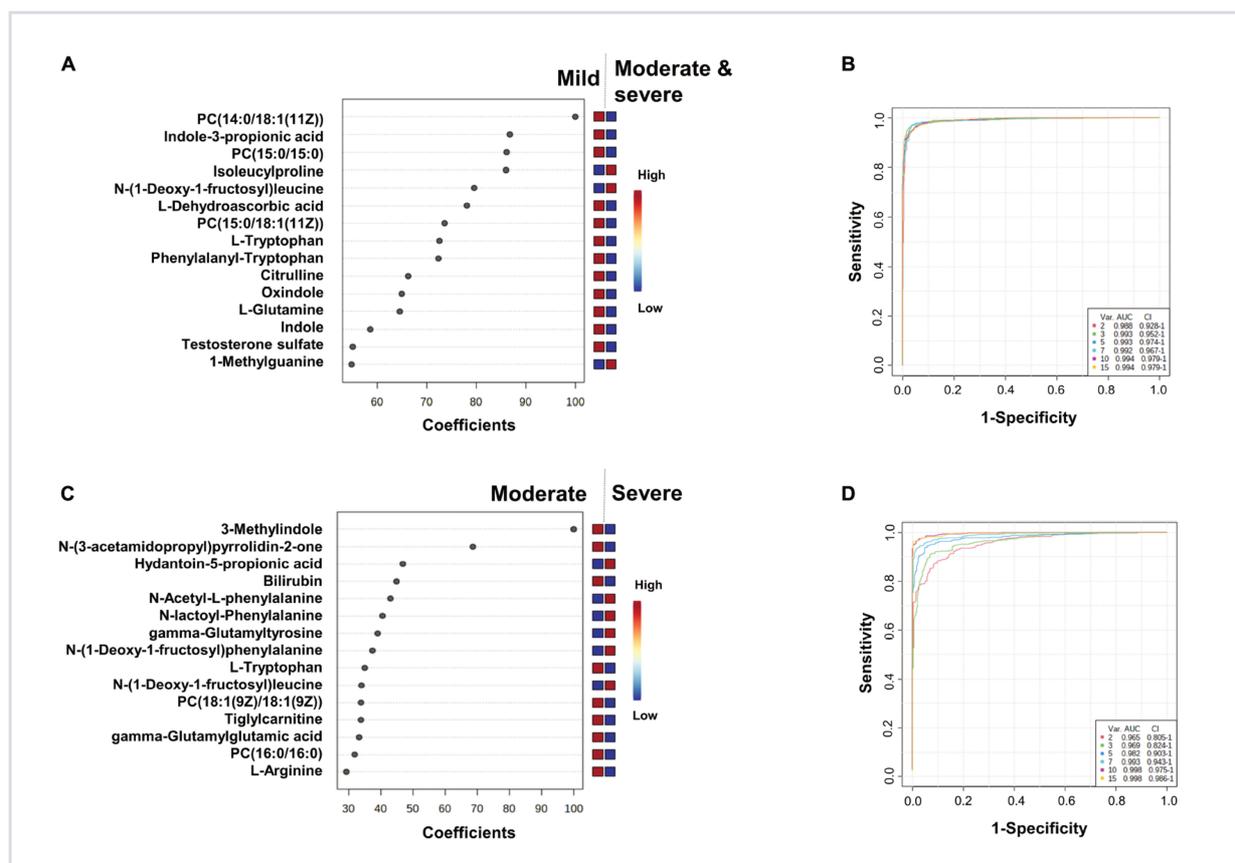
Lors de la première comparaison 547 métabolites ont été sélectionnés. Le modèle PLS-DA du métabolome plasmatique le plus pertinent pour la différenciation des échantillons du groupe léger de ceux des groupes modérément et sévèrement atteints a été obtenu avec cinq composantes (valeur  $Q^2$  à 0,86).

Nous avons obtenu une signature métabolomique plasmatique pronostique du risque d'hospitalisation robuste avec une AUC à 0,994 (IC 95%, 0,979 - 1) pour 15 métabolites. Nous avons observé que les taux plasmatiques de lipides saturés étaient plus faibles dans le groupe de patients hospitalisés. Le niveau de *N-(1-désoxy-1-fructosyl) leucine* était quant à lui plus élevé dans les plasmas du groupe modéré et sévère. Les taux d'*indole*, d'*oxindole*, d'*acide indole-3-propionique* et de *3-méthylindole* étaient plus faibles dans les plasmas du groupe modéré et sévère que dans ceux du groupe léger, de même que les taux d'*acide L-déhydroascorbique*, de *phénylalaninyl-tryptophane* et de *citrulline*. Nous avons également remarqué que le métabolisme du *tryptophane* était impliqué dans cette signature métabolomique pronostique (Figure 22, A et B).

Concernant la seconde comparaison, visant à identifier une signature pronostique du risque d'hospitalisation en service de soins critiques (comparaison des groupes modérément atteints et sévèrement atteints), 527 métabolites ont été sélectionnés. Le modèle PLS-DA du métabolome plasmatique le plus pertinent pour distinguer le groupe modéré du groupe de patients gravement atteints a été obtenu avec quatre composantes (valeur  $Q^2$  de 0,72).

Une nouvelle fois, nous avons observé que les niveaux de deux lipides saturés présents dans la top liste étaient plus faibles dans les plasmas du groupe sévère par rapport à ceux du groupe modéré. Les niveaux des deux acides aminés étaient, là aussi, plus élevés dans les plasmas des

patients les plus graves. En revanche, les taux de *3-méthylindole* étaient plus faibles dans les plasmas du groupe sévère que dans ceux du groupe modéré, de même que les taux de *L-tryptophane*, de *tiglylcarnitine*, d'*acide gamma-glutamylglutamique* et de *L-arginine*. Les taux de *N-acétyl-L-phénylalanine*, de *N-lactoyl-phénylalanine* et de *gamma-glutamyltyrosine* étaient plus élevés dans les plasmas du groupe sévère. Cette signature métabolomique pronostique de 15 métabolites possède une AUC de 0,998 (IC à 95 % (0,986-1)) pour prédire l'appartenance au groupe de patient hospitalisé en service de soins critiques (Figure 22, C et D).



**Figure 22. Identification d'une signature métabolomique plasmatique pronostique du risque d'hospitalisation et du risque d'hospitalisation en service de soins critiques**

A : Diagramme représentant la top liste de 15 métabolites sélectionnées sur la moyenne des cinq premières composantes du modèle PLS-DA pour la comparaison du groupe légèrement atteint à un groupe comprenant les sujets ayant nécessité une hospitalisation (patients modérément et sévèrement atteints). Les métabolites identifiés ont été validés par des analyses MS2 et sont représentés sous forme de noms chimiques.

B : Comparaisons des aires sous les courbes ROC multivariées de chaque top liste de métabolite (comprenant respectivement 15, 10, 7, 5, 3 et 2 métabolites) pour la comparaison du groupe légèrement atteint à un groupe comprenant les sujets ayant nécessité une hospitalisation

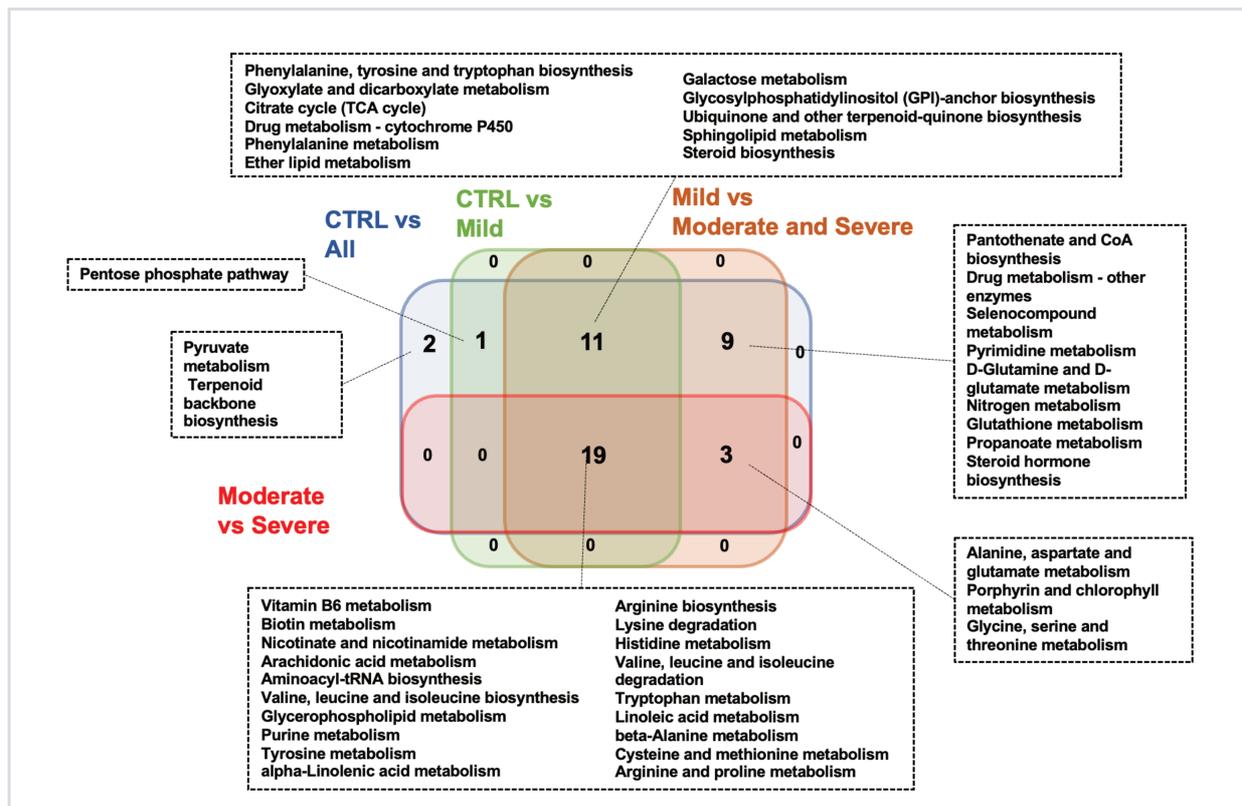
C : Diagramme représentant la top liste de 15 métabolites sélectionnées sur la moyenne des cinq premières composantes du modèle PLS-DA pour la comparaison du groupe modérément atteint au groupe sévèrement atteint. Les métabolites identifiés ont été validés par des analyses MS2 et sont représentés sous forme de noms chimiques.

D : Comparaisons des aires sous les courbes ROC multivariées de chaque top liste de métabolite (comprenant respectivement 15, 10, 7, 5, 3 et 2 métabolites) pour la comparaison du groupe modérément atteint au groupe sévèrement atteint.

▪ **Analyse des voies métaboliques impactées**

Afin de mieux comprendre les modifications métaboliques entraînées par l'infection à SARS-CoV-2 nous avons ensuite procédé à l'analyse des voies métaboliques impactées dans chacune des comparaisons réalisées, à l'aide du logiciel MetaboAnalyst version 5.0<sup>®</sup>. Un diagramme de Venn a ensuite été construit (Figure 23) à l'aide d'un outil en ligne en libre accès (<http://bioinformatics.psb.ugent.be/webtools/Venn/>).

Nous avons pu mettre en avant que l'infection à SARS-CoV-2 impactait les voies de biosynthèse des acides aminés, le métabolisme des lipides et le métabolisme de l'énergie.



**Figure 23. Diagramme de Venn des voies métaboliques plasmatiques significativement modifiées par l'infection à SARS-CoV-2 et selon le niveau de gravité de l'atteinte de la COVID-19**

En bleu : Comparaison du métabolome plasmatique des volontaires sains avec celui des sujets infectés par le SARS-CoV-2 (CTRL vs all)

En vert : Comparaison du métabolome plasmatique des volontaires sains avec celui des sujets légèrement atteints (CTRL vs Mild)

En orange : Comparaison du métabolome plasmatique des sujets légèrement atteints avec celui des sujets ayant nécessité une hospitalisation tous services confondus (Mild vs Moderate & Severe)

En rouge : Comparaison du métabolome plasmatique des sujets modérément atteints avec celui des sujets ayant nécessité une hospitalisation en service de soins critiques (Moderate vs Severe)

Une valeur de  $p < 0,05$  était considérée comme significative

# Article 1. L'empreinte métabolomique non ciblée du plasma met en évidence plusieurs biomarqueurs pour le diagnostic et le pronostic de la maladie à coronavirus 19 (Occelli et al. 2022)



## OPEN ACCESS

EDITED BY  
Evangelos Gikas,  
National and Kapodistrian University  
of Athens, Greece

REVIEWED BY  
George Anderson,  
CRC Scotland and London,  
United Kingdom  
Liang Cui,  
Singapore-MIT Alliance for Research  
and Technology (SMART), Singapore

\*CORRESPONDENCE  
Thierry Pourcher  
thierry.pourcher@univ-cotedazur.fr

SPECIALTY SECTION  
This article was submitted to  
Infectious Diseases – Surveillance,  
Prevention and Treatment,  
a section of the journal  
Frontiers in Medicine

RECEIVED 18 July 2022  
ACCEPTED 01 September 2022  
PUBLISHED 29 September 2022

CITATION  
Occelli C, Guignonis J-M, Lindenthal S,  
Cagnard A, Graslin F, Brglez V,  
Seitz-Polski B, Dellamonica J, Levraut J  
and Pourcher T (2022) Untargeted  
plasma metabolomic fingerprinting  
highlights several biomarkers for the  
diagnosis and prognosis of coronavirus  
disease 19.  
*Front. Med.* 9:995069.  
doi: 10.3389/fmed.2022.995069

COPYRIGHT  
© 2022 Occelli, Guignonis, Lindenthal,  
Cagnard, Graslin, Brglez, Seitz-Polski,  
Dellamonica, Levraut and Pourcher.  
This is an open-access article  
distributed under the terms of the  
[Creative Commons Attribution License  
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or  
reproduction in other forums is  
permitted, provided the original  
author(s) and the copyright owner(s)  
are credited and that the original  
publication in this journal is cited, in  
accordance with accepted academic  
practice. No use, distribution or  
reproduction is permitted which does  
not comply with these terms.

## Untargeted plasma metabolomic fingerprinting highlights several biomarkers for the diagnosis and prognosis of coronavirus disease 19

Céline Occelli<sup>1,2,3</sup>, Jean-Marie Guignonis<sup>1</sup>, Sabine Lindenthal<sup>1</sup>,  
Alexandre Cagnard<sup>1</sup>, Fanny Graslin<sup>1</sup>, Vesna Brglez<sup>4,5</sup>,  
Barbara Seitz-Polski<sup>3,4,5</sup>, Jean Dellamonica<sup>3,4,6</sup>,  
Jacques Levraut<sup>2,3</sup> and Thierry Pourcher<sup>1\*</sup>

<sup>1</sup>Transporter in Imaging and Radiotherapy in Oncology Laboratory (TIRO), Direction de la Recherche Fondamentale (DRF), Institut des Sciences du Vivant Frederic Joliot, Commissariat à l'Energie Atomique et aux Energies Alternatives (CEA), Université Côte d'Azur, School of Medicine, Nice, France, <sup>2</sup>Department of Emergency, University Hospital, Nice, France, <sup>3</sup>School of Medicine, Université Côte d'Azur, Nice, France, <sup>4</sup>Unité de Recherche Clinique Côte d'Azur (UR2CA), Université Côte d'Azur, Nice, France, <sup>5</sup>Department of Immunology, University Hospital, Nice, France, <sup>6</sup>Medical Intensive Care Unit, University Hospital, Nice, France

**Objectives:** The COVID-19 pandemic has been a serious worldwide public health crisis since 2020 and is still challenging healthcare systems. New tools for the prognosis and diagnosis of COVID-19 patients remain important issues.

**Design:** Here, we studied the metabolome of plasma samples of COVID-19 patients for the identification of prognosis biomarkers.

**Patients:** Plasma samples of eighty-six SARS-CoV-2-infected subjects and 24 healthy controls were collected during the first peak of the COVID-19 pandemic in France in 2020.

**Main results:** Plasma metabolome fingerprinting allowed the successful discrimination of healthy controls, mild SARS-CoV-2 subjects, and moderate and severe COVID-19 patients at hospital admission. We found a strong effect of SARS-CoV-2 infection on the plasma metabolome in mild cases. Our results revealed that plasma lipids and alterations in their saturation level are important biomarkers for the detection of the infection. We also identified deoxy-fructosyl-amino acids as new putative plasma biomarkers for SARS-CoV-2 infection and COVID-19 severity. Finally, our results highlight a key role for plasma levels of tryptophan and kynurenine in the symptoms of COVID-19 patients.

**Conclusion:** Our results showed that plasma metabolome profiling is an efficient tool for the diagnosis and prognosis of SARS-CoV-2 infection.

### KEYWORDS

biomarkers, COVID-19, metabolomics (OMICS), diagnostic, prognostic

## Introduction

Since 2020, we have been facing the global pandemic of coronavirus disease 19 (COVID-19), caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). The main difficulties associated with COVID-19 are the lack of general specific symptoms, ranging from asymptomatic forms to acute respiratory distress syndrome (ARDS), its contagiousness and its morbidity and mortality rate (1). Despite rapid vaccine developments, COVID-19 is still a major health threat. COVID-19 vaccines are efficient tools for limiting the spread and severity of the disease but in many countries vaccination rates are still low for economic or ideological reasons. The development of efficient drugs for COVID-19 treatment (2) may improve the outcome of SARS-CoV-2 infection in the future. However, efficient global medical management of the COVID-19 pandemic requires a better understanding of the pathophysiological mechanisms of the disease. More tools are needed to help the clinician to establish a reliable prognostic and thus, to rapidly apply adequate treatment.

SARS-CoV-2 infections are commonly detected by RT-PCR and antigen testing on nose-throat swabs. RT-PCR tests are based on viral gene detection and reliable results can be obtained after several hours (3). Antigen tests based on viral antigen detection are more rapid with results available after approximately 20 min. The diagnosis of moderately and severely affected patients is now codified with combined clinical, biological and CT scan assessments. However, progress is still required in terms of prognostic and screening evaluation. Therefore, new approaches using molecular analyses are needed. Metabolomics, which had already been used to study several viral infections such as influenza (4) and EBOLA (5), was recently also used for studies with COVID-19 patients (6, 7). These metabolomic studies were mainly used with patients suffering moderate to severe symptoms with the aim to develop tools for screening severe patients and for predicting their outcome. The studied cohorts included few patients with mild or asymptomatic forms of COVID-19 (8, 9).

The aims of this study were to identify plasma metabolome signatures that enable the robust classification of asymptomatic, mildly, moderately, and severely affected COVID-19 patients, to predict the course of the patient's medical condition, and to assess new insights in the underlying biological mechanisms of the disease.

## Materials and methods

### Patient selection and procedure

The analyses were conducted on a cohort of plasma samples previously used in the study of Ruetsch et al. (10).

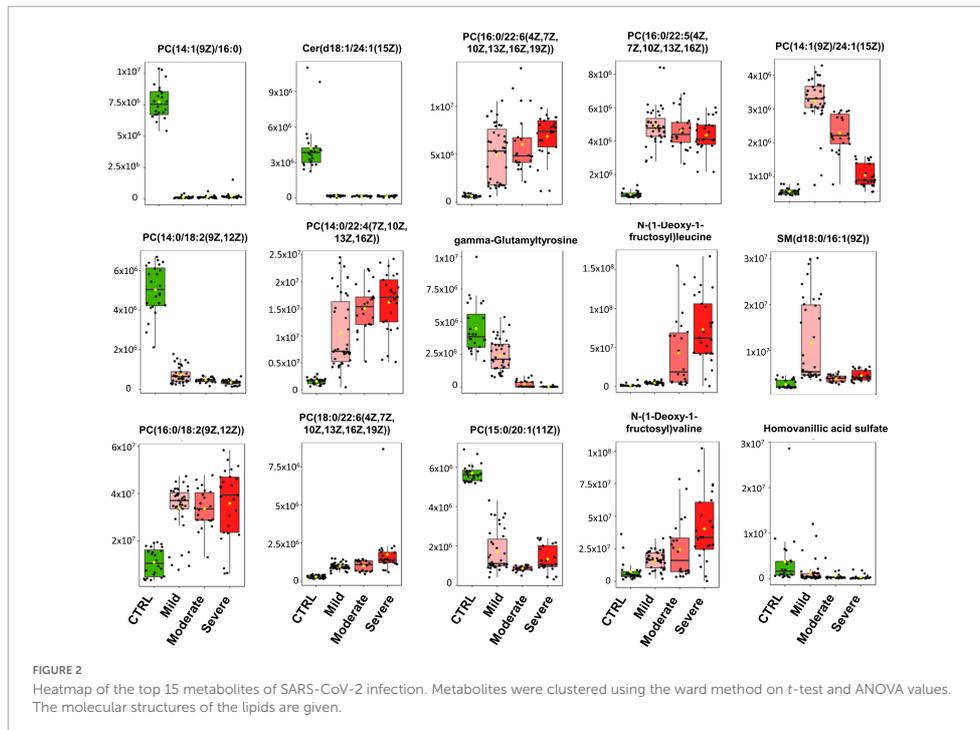
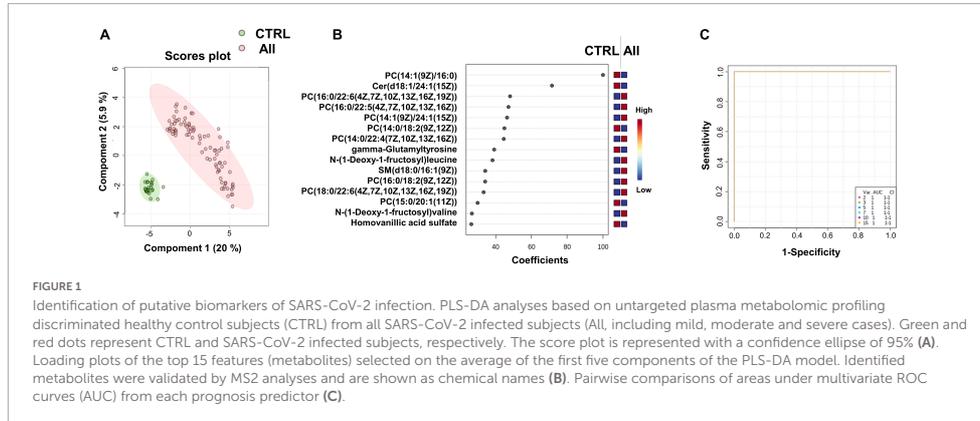
These blood samples were obtained from 86 SARS-CoV-2-infected and 24 healthy control volunteers. All samples were collected during the first peak of the COVID-19 pandemic in France in 2020. Patients were considered to be COVID-19 cases according to the WHO classification: presence of chills or CT scan characteristic of COVID-19 or two consecutive positive RT-PCR tests for SARS-CoV-2 virus or positive serological tests. To evaluate whether the metabolomic approach is suitable for discriminating subjects in terms of diagnosis and/or prognosis, patients were separated into three groups according to the severity of the infection: patients with mild symptoms of COVID-19 (chills or flu-like symptoms not requiring hospital monitoring), patients with moderate symptoms (patients hospitalized in infectious disease units) and patients with a severe form of COVID-19 (requiring hospitalization in an intensive care unit). Plasma samples of the Moderate and Severe groups were obtained from patients at hospital admission. Non-infected subjects with a negative serological test for SARS-CoV-2 comprised the control group. An informed consent was obtained for all participants. The study protocol conformed to the ethical guidelines of the 1975 Declaration of Helsinki and was approved by the appropriate institutional review committee (NCT04355351).

### Collection and processing of plasma samples for omic analyses

A volume of 100  $\mu$ L plasma was mixed with 100  $\mu$ L H<sub>2</sub>O (HPLC grade, Merck Millipore, USA), and 600  $\mu$ L methanol (HPLC grade, Merck Millipore, USA) was added. Samples were vortexed and incubated overnight at  $-20^{\circ}\text{C}$  for protein precipitation. After centrifugation ( $13,000 \times g$ , 15 min,  $4^{\circ}\text{C}$ ), the supernatant was removed, dried using a SpeedVAC concentrator (SVC100H, SAVANT, Thermo Fisher Scientific, Illkirch, France), resuspended in 80  $\mu$ L 20:80 acetonitrile-H<sub>2</sub>O mixture (HPLC grade, Merck Millipore) and stored at  $-20^{\circ}\text{C}$  until use for metabolomic analyses.

### Metabolomic analyses

Chromatographic analyses were performed with the DIONEX Ultimate 3000 HPLC system coupled to a chromatographic column (Phenomenex Synergi 4  $\mu$  Hydro-RP 80A 250  $\times$  3.0 mm) set at  $40^{\circ}\text{C}$  and a flow rate of 0.9 mL/min. Gradients of mobile phases (mobile phase A: 0.1% formic acid in water and mobile phase B: 0.1% formic acid in acetonitrile) were performed over a total of 25 min. MS analyses were carried out on a Thermo Fisher Scientific Exactive Plus Benchtop Orbitrap mass spectrometer. The heated electrospray ionization source (HESI II) was used in both positive and negative ion modes. The instrument was operated in full scan mode from  $m/z$  60 to  $m/z$



900. Data post-treatment was performed using the MZmine2 version 2.39<sup>1</sup> (11). Typical Total Ion Chromatograms for

positive and negative modes are illustrated in [Supplementary Data 1C](#). Gap filling was performed using the “Same RT and m/z range gap filler” method of the MZmine software. The m/z tolerance was set up at 0.001 m/z or 10.0 ppm. Metabolites were putatively annotated [corresponding to level 2 of the

1 <http://mzmine.github.io/>

Metabolomics Standard Initiative (MSI)<sup>2</sup> using the Human Metabolome Database version 5.0<sup>3</sup> (12). Only ions identified as  $[M + H]^+$  adducts in the positive mode and  $[M - H]^-$  adducts in the negative mode and ions found in all the samples after gap filling were included. Metabolites identified as drugs or exogenous compounds were omitted. Full data sets (raw data) are available upon request.

## Statistical analyses

Statistical analyses of the untargeted metabolomic studies were processed using statistical analysis modules (one factor) proposed by MetaboAnalyst 5.0<sup>4</sup> (12). For each comparison, all samples of the data sets were median normalized and peak intensities were Log transformed. Using Partial Least Squares-Discriminant Analysis (PLS-DA), we analyzed loadings (average) of features selected by the PLS-DA model to identify discriminative metabolites. PLS-DA is considered particularly effective for biological systems to select discriminative features. The data identities of the top features shown in **Figures 1–4, 6** were verified by the comparison of the obtained MS2 spectra with those of online data banks (Human Metabolome, MassBank Europe)<sup>5</sup>. MS2 matches are illustrated in **Supplementary Data 1D, 2C, 3C, 4C**. However, most of MS2 spectra that we found in the HMDB were predicted spectra. Therefore, MS2 matches are not corresponding to the level 1 of the MSI. Statistical tests to measure the association between major metabolites and clinical outcome such as Receiver-Operating Characteristic curves (ROC) were performed using the biomarker analysis module of MetaboAnalyst. The datasets of the four comparisons (see below) of the different groups of this study are given in the **Supplementary Data 1A–4A**. For each comparison, the data are ranked according to the PLS-DA average score (as per the loading plots in **Figures 1B, 3B, 4A,C**). Values and graphs are illustrated in the **Supplementary Data 1B–4B**. Several identified ions are putative biomarkers. To simplify the study, we mainly describe and discuss the top 15 metabolites from the PLS models.

## Pathway analyses

Analyses of related pathways were processed using pathway analysis modules proposed by MetaboAnalyst 5.0 (see text footnote 4) (12). Venn diagrams were performed to calculate and draw the intersections of pathway lists using a webtool.<sup>6</sup>

<sup>2</sup> <https://github.com/MSI-Metabolomics-Standards-Initiative/CIMR>

<sup>3</sup> <http://www.hmdb.ca>

<sup>4</sup> <https://www.metaboanalyst.ca>

<sup>5</sup> <https://massbank.eu/MassBank/>

<sup>6</sup> <http://bioinformatics.psb.ugent.be/webtools/Venn/>

## Results

### Group characteristics

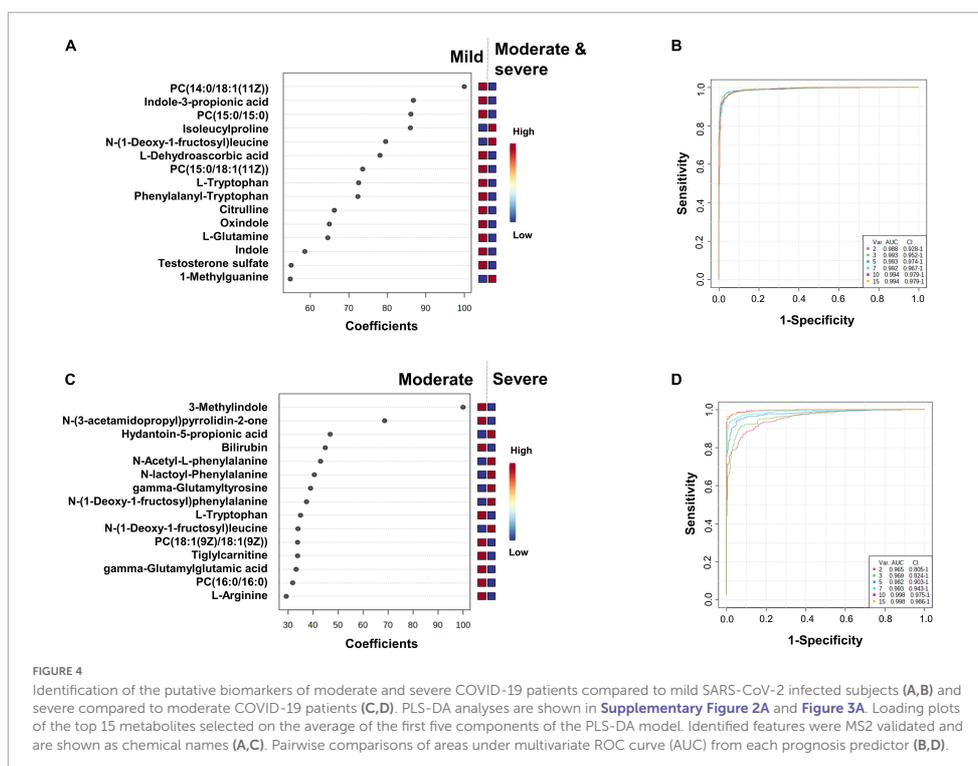
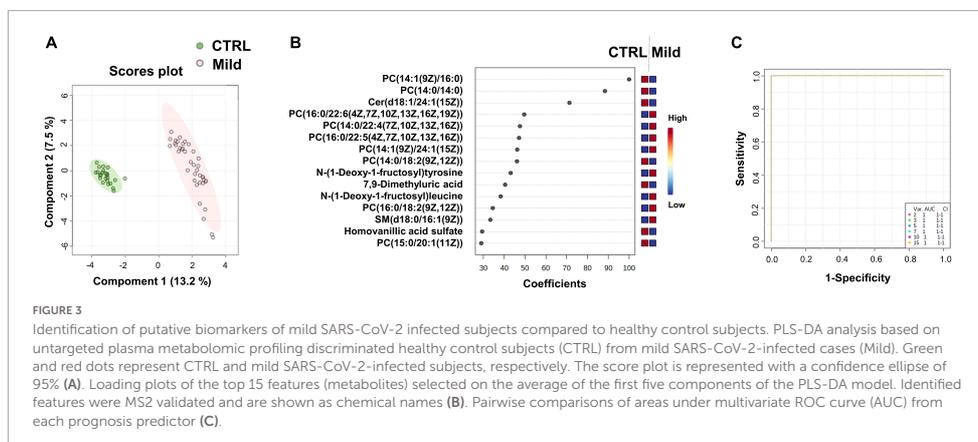
In this study, plasma metabolomes from 110 RT-PCR-tested subjects were analyzed. Epidemiological and clinical data are reported in **Supplementary Table 1**. The group (named “Mild”) of patients with only mild symptoms comprised 39 subjects; the group (named “Moderate”) of patients with moderate symptoms comprised 22 subjects; and the group (named “Severe”) of patients with severe symptoms comprised 25 subjects. Plasma samples from 24 health-care workers with additional negative serological test results and no apparent symptoms constitute the healthy control group (named “CTRL”). The Mild COVID-19 patients did not suffer any respiratory syndrome, however, one patient was treated by oxygen therapy (see **Supplementary Table 1**). Aiming the identification of biomarkers for the prognosis of COVID-19 severity, plasma samples of the Moderate and Severe groups were obtained from patients at hospital admission for respiratory illness and COVID-19 diagnosis *via* a positive RT-PCR test result. These patients were then divided into the two subgroups, i.e., Moderate or Severe through a follow-up, with those whose medical condition progressed to acute respiratory distress being assigned to the Severe group. All samples of the healthy control donors with a negative serological test for SARS-CoV-2 were obtained at the beginning of 2020 and were therefore from subjects supposedly never exposed to SARS-CoV-2 infection.

Gender ( $p = 0.004$ ), mean age ( $p < 0.001$ ), presence of cardiovascular history ( $p < 0.0001$ ), diabetes ( $p = 0.0003$ ), and cancer history ( $p = 0.04$ ) were statistically different between the groups.

### Comparison of untargeted plasma metabolomic fingerprintings of healthy controls (CTRL) and severe acute respiratory syndrome coronavirus 2-infected (mild, moderate, and severe groups) subjects revealed that important metabolome changes occur in all infected patients (All)

The plasma metabolomes of non-infected control subjects were compared to those of the infected volunteers. After data post-treatment of the LC-MS analyses (see section “Materials and methods”), 605 metabolites were selected from the ions ( $m/z$ ) obtained in both, positive and negative electrospray ionization modes. Metabolomics data (ranked according to the loading, see below) are available in **Supplementary Data 1A**.

We performed supervised multivariate analyses [Partial Least Squares-Discriminant Analysis (PLS-DA)] on these data



using the MetaboAnalyst website. As illustrated in the score plot of [Figure 1A](#), the plasma metabolomes of the control group (CTRL) could clearly be discriminated from those

of the SARS-CoV-2-infected groups (All). The diagnostic performance of the model was evaluated by cross-validation. High accuracy values (1.0) for the PLS-DA model were

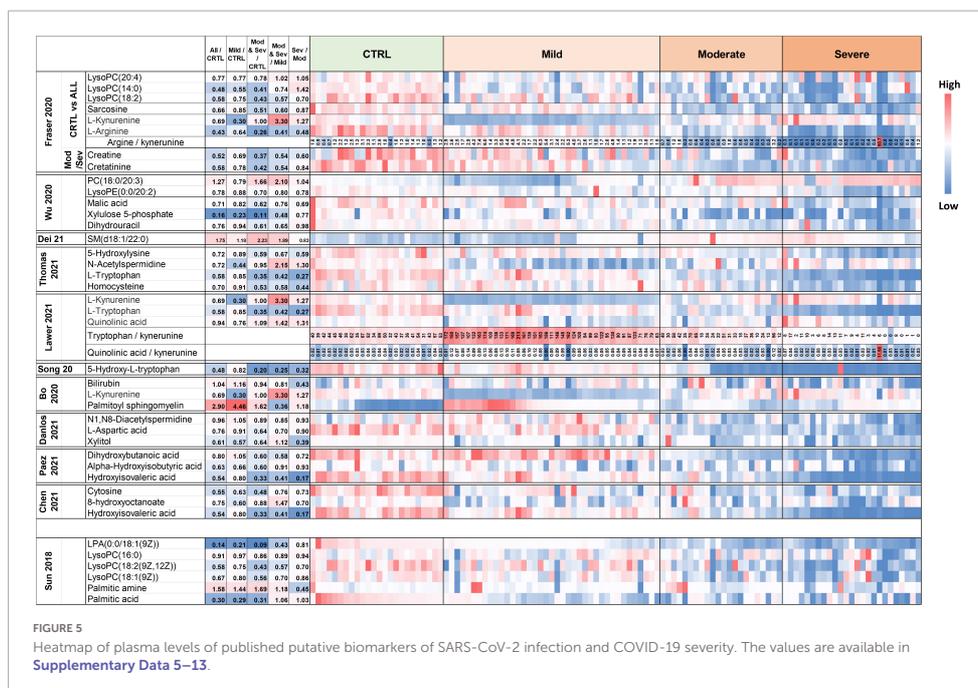


FIGURE 5 Heatmap of plasma levels of published putative biomarkers of SARS-CoV-2 infection and COVID-19 severity. The values are available in [Supplementary Data 5–13](#).

obtained with 2 or more components. The most relevant model was obtained with five components (Q2 value of 0.99, for details see [Supplementary Data 1B](#)). [Figure 1B](#) shows the loading plots of the top 15 features ranked according to their overall coefficient scores of the PLS-DA model. This feature top list of 15 metabolites with highlighted lipids as the main contributors of the PLS-DA model.

As illustrated by the loading plots ([Figure 1B](#)) and the heatmap of these top 15 features ([Figure 2](#)), the levels of saturated lipids were lower in the plasma of SARS-CoV-2-infected subjects when compared to the control samples, whereas the levels of unsaturated lipids were higher. In addition, two N-(1-deoxy-1-fructosyl) amino acids (leucine and valine) were found among the top 15 features. The levels of these two amino acids were higher in the plasma of COVID-19 infected subjects compared to the control samples ([Figures 1B, 2](#)). [Figure 1C](#) shows that the area under the combined Receiving Operating Characteristic curve (ROC) for the prediction of SARS-CoV-2 infection of the top 15 features identified by the PLS-DA model that reached the score 1. This score remained unchanged even when considering only two metabolites from this top list [AUC = 1, CI 95% (1–1)].

### Comparison of untargeted plasma metabolomic fingerprintings of the control group samples and mild group samples revealed important differences

The main goal of this comparison was to determine if a metabolomic signature could be identified and used for the diagnosis of mild forms of COVID-19 infection. After post-treatment of the LC-MS/MS analyses data, 574 metabolites were selected ([Supplementary Data 2A](#)).

The score plot in [Figure 3A](#) shows the accuracy of the plasma-metabolome PLS-DA model to differentiate between the control and the mild group of SARS-CoV-2-infected patients. High accuracy values (1.0) of the PLS-DA model are obtained with 1 or more components. The highest scores were obtained with five components (Q2 value of 0.99) (details in [Supplementary Data 2B](#)). The top 15 metabolites identified by the PLS-DA model are shown in [Figure 3B](#) and in the heatmap of [Supplementary Figure 1](#). Similar to the CTRL vs. All analysis, the metabolic profile shows variations in the levels of several lipids and two N-(1-deoxy-1-fructosyl)

amino acids [N-(1-deoxy-1-fructosyl)tyrosine and N-(1-deoxy-1-fructosyl)leucine]. As expected, most metabolites of the top lists of the two comparisons were identical (see also CTRL vs. All in [Figure 1B](#)). Three metabolites of the CTRL vs. All top list, i.e., N-(1-deoxy-1-fructosyl)valine, the phosphatidylcholine PC[18:0/22:6(4Z,7Z,10Z,13Z,16Z,19Z)] and gamma-glutamyltyrosine were not found in the CTRL vs. Mild top list, which instead included 7,9-dimethyluric acid, PC (14:0/14:0) and N-(1-deoxy-1-fructosyl)tyrosine. As found for saturated lipids of the CTRL vs. All top list, the levels of PC (14:0/14:0) were significantly lower in the plasma of SARS-CoV-2-infected patients with mild symptoms compared to samples from the healthy control subjects. Similar to the findings of the amino acids of the CTRL vs. All top list, N-(1-deoxy-1-fructosyl)tyrosine levels were higher in samples of the mild group compared to those of the control group. [Figure 3C](#) shows the areas under the ROC curve to predict the presence of SARS-CoV-2 infection in asymptomatic subjects or patients with mild symptoms. The AUC values are 1 [95% CI (1–1)] when using two or more metabolites from the top list.

### Comparison of untargeted plasma metabolomic fingerprintings from subjects with mild symptoms (mild) with those of subjects of the moderate and severe group (moderate and severe)

This comparison aimed to determine if metabolomic signatures could be identified and used for the diagnosis and prognosis of moderate and severe forms of COVID-19. After post-treatment of the LC-MS/MS analyses data, 547 metabolites were selected ([Supplementary Data 3A](#)). The most relevant plasma metabolome PLS-DA model for the robust differentiation of samples of the mild group from those of the moderate and severe groups was obtained with five components (AUC = 0.99; Q2 value of 0.86) (see score plot in [Supplementary Figure 2A](#) and details in [Supplementary Data 3B](#)).

The top 15 metabolites identified by the PLS-DA model are shown in [Figure 4A](#) (heatmap in [Supplementary Figure 2B](#)). We observed that the levels of unsaturated lipids were lower in the plasmas of the moderate and severe group compared to those of the mild group. The N-(1-deoxy-1-fructosyl) leucine level was found to be higher in the plasmas of the moderate and severe group compared to those of the mild group. Indole, oxindole, indole-3-propionic acid, 3-methylindole levels were lower in the plasmas of the moderate and severe group compared to those of the mild group, as well as the L-dehydroascorbic acid, phenylalanyl-tryptophan, and citrulline levels. The areas under the ROC curve to predict the moderate or severe forms in SARS-CoV-2-infected patients are shown in [Figure 4B](#). The

AUC value is 0.988 [95% CI (0.928–1)] when using only two metabolites from the top list and the AUC values are 0.994 [95% CI (0.979–1)] when using the 10 or 15 top metabolites.

### Comparison of untargeted metabolomic fingerprintings of plasma from subjects with moderate symptoms (moderate) with those of subjects with severe symptoms (severe)

This comparison aimed to determine if metabolomic signatures could be used for the diagnosis and prognosis of severe forms of COVID-19. After post-treatment of the LC-MS/MS analyses data, 527 metabolites were selected ([Supplementary Data 4A](#)). The most relevant plasma-metabolome PLS-DA model to discriminate the moderate group from the severely affected patient group was obtained with four components (AUC = 0.94; Q2 value of 0.72) (score plot in [Supplementary Figure 3A](#) and details in [Supplementary Data 4B](#)). The top 15 metabolites identified by the PLS-DA model are shown in [Figure 4C](#) (heatmap in [Supplementary Figure 3B](#)).

We observed that the levels of two saturated lipids [PC (18:1(9Z)/18:1(9Z)) and PC (16:0/16:0)] were lower in the plasmas of the severe group compared to those of the moderate group. The levels of the two amino acids, N-(1-deoxy-1-fructosyl)alanine and N-(1-deoxy-1-fructosyl)leucine were higher in the plasmas of the severe group compared to those of the moderate group. In contrast, 3-methylindole levels were lower in the plasmas of the severe group compared to those of the moderate group, as were L-tryptophan, tiglylcarnitine, gamma-glutamylglutamic acid and L-arginine levels. N-acetyl-L-phenylalanine, N-lactoyl-phenylalanine and gamma-glutamyltyrosine levels were higher in plasmas of the severe group. The areas under the ROC curve to predict plasma from patients with moderate or severe forms of COVID-19 are shown in [Figure 4D](#). The AUC values are 0.965 [95% CI (0.805–1)] when using only two metabolites from the top list, and 0.998 [95% CI (0.986–1)] when using 15 metabolites.

### Comparisons of the pathway analyses of untargeted plasma metabolomic fingerprintings for CTRL vs. all, CTRL vs. mild, and mild vs. moderate and severe

Using the MetaboAnalyst website, we studied the pathways related to the three comparative analyses described above. Only

significant pathways were considered ( $p$ -value < 0.05). All the results are shown in the [Supplementary Data 5–10](#).

A Venn diagram showing the intersections of the pathway lists for three comparisons (CTRL vs. All, CTRL vs. Mild and Mild vs. Moderate and Severe) is illustrated in [Supplementary Figure 4](#) shows the Venn diagram for all four comparisons including Moderate vs. Severe. The Venn diagram ([Supplementary Figure 4](#)) highlights the major effects of SARS-CoV-2 infection that were already detected in the plasma samples from COVID-19 patients with mild symptoms compared to that of healthy control subjects. The major pathways found to be altered in mild COVID-19 patients included amino acid biosynthetic pathways, and lipid and energy metabolism.

### Correlation analyses between plasma untargeted metabolomic fingerprintings and IL-6 levels

In a previous study with blood samples from the same patient cohort as those used for the present analyses, Ruetsch et al. (10) revealed a correlation between plasma IL6 levels and the patient outcome. As illustrated in [Supplementary Figure 5](#), our results showed that amino acid (tryptophan, glutamine, and threonine) levels are inversely correlated to plasma IL6 levels (sample correlation coefficient values of  $-0.616$ ,  $-0.585$ , and  $-0.0551$ , respectively). In addition, six unidentified metabolites appeared to be positively correlated to plasma IL6 levels (sample correlation coefficient values of  $0.779$ ,  $0.739$ ,  $0.725$ ,  $0.721$ ,  $0.717$ , and  $0.705$ ) ([Supplementary Figure 5](#)).

## Discussion

This study aims to identify plasma metabolomic signatures for the diagnosis of SARS-CoV-2 infection and the prognosis of COVID-19 disease severity. The identification of new biomarkers may also offer unique insights into the pathophysiology of the disease.

All blood samples from this study were obtained during the first wave of the SARS-CoV-2 infections in France in 2020. Therefore, all samples of the infected subjects were most probably primary infections and the healthy control subjects had never been exposed to the SARS-CoV-2 virus. No vaccine was available at that time. Plasma samples of COVID-19 patients with moderate and severe symptoms were collected at hospital admission. At that point, patients had not yet benefited from specific drug administration or respiratory treatments.

Our results showed that untargeted metabolomic profiling of plasma samples from healthy control and SARS-CoV-2-infected subjects enabled the identification of new biomarkers and new metabolomic signatures for the diagnosis and

prognosis of COVID-19. In the following section, we discuss the potential biological origins of the observed changes in the plasma levels of the identified metabolites due to SARS-CoV-2 infection.

Healthy control subjects (CTRL,  $n = 25$ ) were hospital health-care workers with negative serological test results and no apparent symptoms. Mild subjects ( $n = 39$ ) with positive serological test were patients with chills or flu-like symptoms that did not require hospital monitoring. The subjects of these two groups had no underlying diseases or chronic conditions, so we were able to conduct the analysis without having to include additional interfering factors. Therefore, we assumed that the changes in the plasma metabolite levels observed in the comparison CTRL vs. Mild (and, to a lesser extent in CTRL vs. All) were mainly evoked by SARS-CoV-2 infection. The baseline characteristics of the Moderate COVID-19 patients were not significantly different from those of the Severe subjects. Furthermore, the number of samples was low, 22 Moderate cases and 25 Severe cases, respectively. Therefore, we could not form subgroups. However, we assumed that the changes in the plasma metabolite levels found in the comparison Moderate vs. Severe were mainly caused by the severity of COVID-19 symptoms. In contrast, the baseline characteristics of the subjects in the Mild group differed from those in the Moderate and Severe groups. Corrections for the patient's baseline characteristics were not performed as this would require a larger number of subjects in each group and subgroup. Therefore, the results of this comparison should be treated with caution. Thus, we have them carefully discussed here, focusing mainly on metabolites that have already been described in other publications.

We found that metabolome fingerprinting of control subjects were significantly different from those of all SARS-CoV-2-infected subjects independent of the severity of the disease (CTRL vs. All, illustrated by the PLS-DA score plot and multivariate ROC curves in [Figures 1A,C](#), respectively and CTRL vs. Mild, illustrated by the PLS-DA score plot and multivariate ROC curves in [Figures 3A,C](#), respectively). For both comparisons CTRL vs. All and CTRL vs. Mild, we identified a top list of metabolites that allow one to assign each subject to a specific class, i.e., CTRL, ALL, or Mild (illustrated in [Figures 1B, 2, 3B](#) and [Supplementary Figure 1](#)). The most frequent metabolites of both lists were lipids. In the top metabolite lists of the PLS models of the comparisons between infected subjects suffering from the disease at different severities (Mild vs. Moderate and Severe and Moderate vs. Severe), the elevated lipids were not found (see loading plots in [Figures 4A,C](#)). This result shows that the alteration of lipid metabolism is an early event in the SARS-CoV-2 infection. Fraser and collaborators performed targeted LC-MS/MS analyses with plasma samples of COVID-19 patients (13) and report the presence of LysoPCs in the top metabolite list.

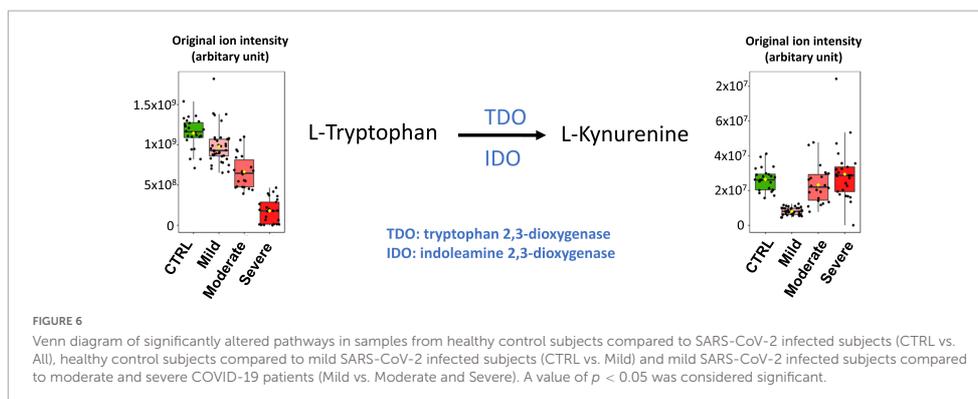


Figure 5 (top lines) shows the heat map of the peak intensities of some of these LysoPCs of our dataset. Although these lipids are not part of our top list, we also detected higher levels in plasma from infected subjects. Wu and collaborators (7) report variations of the plasma lipid concentrations in COVID-19 patients. We detected similar changes in the plasma concentration of two of these lipids in our dataset (see Figure 5 and Supplementary Figure 6). Dei Cas and collaborators report a correlation between lipid signatures and prognostic factors for SARS-CoV-2 infected subjects (14). Only one of the lipids described in this study was also identified in our dataset (variation illustrated in Figure 5). Thomas and collaborators describe changes in free fatty acids and acylcarnitine levels in the plasma of COVID-19 patients when compared to control subjects (15). In addition, Bruzzone et al. show altered regulation of lipoproteins in the plasma of COVID-19 patients, with a reduction in HDL size, an increase in LDL size, and an increase in the level of intermediate VLDL subclasses (16). Song and collaborators report that plasma levels of sphingomyelins (SMs) are enhanced, and GM3s and those of diacylglycerols (DAGs) are reduced in COVID-19 patients (17). Several other groups suggest that certain fatty acids, including arachidonic acid and other unsaturated fatty acids play a key role in the cytokine storm during the initial phase of the infection and later in the inactivation and inhibition of SARS-CoV-2 proliferation (18, 19). Altered plasma lipid levels were also found in recovered COVID-19 patients (20) and during COVID-19 progression as shown by longitudinal metabolomic analyses (21). It will be interesting to assess if the levels of the lipids identified in this study also remain altered in recovered patients.

We performed LC-MS/MS experiments for untargeted metabolomic analyses of polar but also of hydrophobic compounds leading to the identification of new lipid biomarkers for SARS-CoV-2 infection. Interestingly, the top lists of our classification models revealed that the majority of lipids with

lower levels in the plasma of SARS-CoV-2 infected subjects, i.e., CTRL vs. All and CTRL vs. Mild were mainly saturated lipids (as illustrated in Figures 2, 5 and Supplementary Figures 1, 6). In contrast, the majority of lipids with higher levels in the plasma of SARS-CoV-2 infected patients vs. healthy control subjects were unsaturated lipids. This finding could be related to an increased level of oxidative stress in SARS-CoV-2 infected subjects as has been suggested by other metabolomic studies (14, 15). Lipid metabolic remodeling of the host cell is a key feature of virus infection (22) and has been described and discussed for SARS-CoV-2 infection in several reviews (18, 23, 24). Recent study also indicates that expression of Spike protein impairs lipid metabolism of the host cells (25). This remodeling could evoke most variations in the endogenous lipid levels that we reported here. But, we also found that 2 odd-chain PCs were putative biomarkers of the SARS-CoV-2 infection and severity of the disease: PC[15/0/20:1(11Z)] for the comparison of CTRL vs. Mild (Figure 3B) and PC[15:0/18:1(11Z)] for the comparison Mild vs. Moderate and Severe (Figure 4A). Odd-chain lipids are considered as exogenous lipids mainly from milk consumption. The presence of odd-chain lipids could be related to exogenous (ingested) compounds and, in this case, should be removed from datasets in metabolomics studies. But, PC[15/0/20:1(11Z)] was detected at lower level in plasma samples of Mild COVID-19 patients compared to CTRL subjects. These 2 groups have many similarities (age and sex). The subjects of these groups were healthy. It is difficult to guess that in the Mild subjects had mainly diet with low odd-chain lipids. In addition, low levels of PC[15/0/20:1(11Z)] was also detected in plasma samples of COVID-19 patients (Moderate and Severe). These results indicate that PC[15/0/20:1(11Z)] is a putative biomarker of the SARS-CoV-2 infection. More experiments are needed to confirm this finding. PC[15:0/18:1(11Z)] was also detected at lower level in plasma samples of Moderate and Severe COVID-19 patients compared to Mild subjects.

For this comparison Mild vs. Moderate and Severe, baseline characteristics of the subjects are different ([Supplementary Table 1](#)). The Moderate and Severe patients had COVID-19 symptoms and were hospitalized. But plasma samples were collected at hospital admission. All the subjects are living in the same area and they could have a similar diet. However, we cannot exclude an effect of the baseline characteristics when Mild (and CTRL) subjects are compared to Moderate and Severe patients. Therefore, further experiments are required to confirm that PC[15:0/18:1(11Z)] is a biomarker of disease severity.

Our study showed that N-(1-deoxy-1-fructosyl)amino acids were also putative biomarkers for SARS-CoV-2 infection (see [Supplementary Figure 6](#)). Leucine, valine and tyrosine derivatives were found in the top lists of our PLS models for the comparisons of CTRL vs. Mild and CTRL vs. ALL. N-(1-deoxy-1-fructosyl)leucine was detected in the top list of all the PLS models of this study and N-(1-deoxy-1-fructosyl)phenylalanine, as well as other phenylalanine derivatives (N-acetyl-L-phenylalanine and N-lactoyl-phenylalanine), were found in the top list of the PLS model of the comparison Moderate vs. Severe. To our knowledge, N-(1-deoxy-1-fructosyl)amino acids are poorly described in the literature and further studies are needed to understand the altered plasma levels of these amino acids in SARS-CoV-2 infected subjects. As discussed for odd-chain lipids, the presence of N-(1-deoxy-1-fructosyl)amino acids could be related to exogenous (ingested) compounds and, in this case, should be removed from datasets in metabolomics studies. Here, we included the N-(1-deoxy-1-fructosyl)amino acids in our analyses because these compounds were identified in all subject classes allowing us to assume that if they were exogenous, they were ingested randomly by all volunteers. Four N-(1-deoxy-1-fructosyl)amino acids were detected at a higher level in plasma samples of mild COVID-19 patients compared to healthy control subjects and were even higher with increasing severity of the disease.

In the CTRL vs. Mild comparison, lower levels of gamma-glutamyltyrosine and homovanillic acid sulfate were found for samples of infected subjects ([Figure 1B](#) and [Figure 2](#)). The levels of these compounds decreased even further in samples from the moderate group and especially in those from the severe group (see heat map in [Figure 2](#)). Gamma-glutamyltyrosine is produced when protein digestion or protein catabolism is incomplete. Homovanillic acid is a putative marker of metabolic stress (26). The role of these compounds during SARS-CoV-2 infection is still unclear and remains to be investigated. Likewise, the levels of 7,9-dimethyluric acid, a methyl derivative of uric acid, were found to be lower in the plasma of infected subjects than in the plasma of healthy control subjects ([Figure 3B](#) and [Supplementary Figure 1](#)).

The plasma levels of the amino acids, glutamine, arginine, citrulline, and tryptophan and their derivatives, indole-3-propionic acid and phenylalanyl-tryptophan, were found to be

lower in samples of the moderate and severe groups than in samples of the mild group ([Figure 3A](#) and [Supplementary Figure 2](#)) suggesting an effect of disease severity on amino acid metabolism. Indole and indole-3-propionic acid are microbial metabolites and the observed effects on these metabolite levels could result from COVID-19-associated digestive disorders. The effects on dehydroascorbic acid levels could be linked to oxidative stress.

For the comparison of the Mild vs. Moderate and Severe groups ([Figure 4C](#) and [Supplementary Figure 3](#)), the metabolite top list of the PLS model contained N-(3-acetamidopropyl)pyrrolidin-2-one, an intermediate of spermidine catabolism and bilirubin, a degradation product of heme. Bilirubin degradation products are also found by Shen et al. (6) in the sera of COVID-19 patients. Similar to findings by Fraser and collaborators (13), we detected lower levels of arginine in samples of the severe group when compared to those of the moderate group. Kynurenine is a potential biomarker for COVID-19 reported in several studies (6, 13, 27–30) and Fraser et al. propose the use of an arginine vs. kynurenine ratio to predict the severity of the COVID-19 disease.

In our study, kynurenine was not found among the top lists of 15 metabolites of our four PLS models. However, kynurenine occupied the 29th, 24th, and 60th rank of the PLS models of the comparisons CTRL vs. All, CTRL vs. Mild and Mild vs. Moderate and Severe, respectively (see [Supplementary Data 1A–4A](#)). Although kynurenine did not appear in our top lists, this metabolite could still be considered as an interesting biomarker of SARS-CoV-2 infection.

As illustrated in the heatmap in [Figure 5](#) and the box plot in [Figure 6](#), lower plasma kynurenine levels were found for mild COVID-19 patients when compared to healthy control subjects and to moderate and severe COVID-19 patients. Kynurenine is described as a suppressor of the tumor-induced immune response in cancer patients (31, 32). The lower levels of kynurenine in the plasma of patients with mild COVID-19 symptoms, when compared to CTRL, could stimulate the immune response. Consequently, SARS-CoV-2 infection and symptoms are reduced. In contrast, kynurenine levels in the plasmas of moderate and severe COVID-19 patients are like those in healthy control subjects and may contribute to limit the “cytokine storm” induced by SARS-CoV-2 infection and thus prevent inflammation and the appearance of more severe symptoms. The kynurenine precursor is tryptophan, and two enzymes catalyze the transformation of tryptophan to kynurenine [see (33) for review], indoleamine-2,3-dioxygenase (IDO) and tryptophan-2,3-dioxygenase (TDO). The plasma levels of tryptophan were decreased in the mild group compared to CTRL, but, in contrast to kynurenine, tryptophan levels were lower in the moderate group and lowest in the severe group ([Figure 6](#)). Our tryptophan level data corroborate the findings of Thomas et al. (15) and Lawler et al. (27). The two groups also report altered plasma tryptophan levels in the plasma of

COVID-19 patients. We found that the levels of tryptophan decreased with the severity of the symptoms. Our results also showed that plasma kynurenine levels are only decreased in mild COVID-19. Therefore, we propose that the lower levels of kynurenine in mild COVID-19 patients compared to healthy control subjects could be a consequence of lower tryptophan levels in mild COVID-19 patients without activation of IDO or TDO. In contrast, increased transformation of tryptophan to kynurenine should be evoked in moderate and severe COVID-19 patients. This transformation stimulation should restore kynurenine levels in the presence of lower tryptophan levels but also contribute to lower the tryptophan level. Higher IL-6 levels (as reported in plasma of moderate and severe COVID-19 patients, see [Supplementary Figure 5](#)) could be responsible for the activation of IDO (34). Therefore, we propose that IL-6 could mediate an induction of tryptophan to kynurenine transformation for these patients *via* IDO activation. The raised levels of pro-inflammatory cytokines evident in SARS-CoV-2 infected patients have a number of important consequences, including the induction of IDO and the conversion of tryptophan to kynurenine, which can then activate the aryl hydrocarbon receptor (AhR), thereby dysregulating the immune response, including suppressing natural killer cell activation (35). Pro-inflammatory cytokines also increase gut permeability and gut dysbiosis, leading heightened levels of circulating LPS and suppressed levels of the short-chain fatty acid, butyrate (35). The suppression of butyrate heightens histone acetylation-driven epigenetic regulation, thereby altering the nature of patterned gene expression, including as driven by kynurenine activation of the AhR (36). Given the clinical utility of melatonin in the management of SARS-CoV-2 infection severity and mortality (37), it is important to note that the conversion of tryptophan to kynurenine will suppress serotonin levels, and therefore serotonin as the necessary precursor for the induction of the melatonergic pathway in body cells, including immune cells (38). Given that melatonin production and autocrine effects are necessary to shift macrophages and microglia from an M1-like pro-inflammatory phenotype to an M2-like pro-phagocytic phenotype (39, 40), such alterations in tryptophan availability will significantly dysregulate the patterned immune response. The data in the current study may therefore link to wider systemic aspects of immune regulation involving the gut microbiome/permeability, AhR activation and suppressed mitochondrial melatonergic pathway activity, as recently proposed (41).

We also studied the pathways related to the obtained metabolome fingerprintings. As shown in [Supplementary Data 5–8](#), we used MetaboAnalyst to identify significant pathways ( $p$ -values > 0.05) by performing two-class comparisons (45 pathways for the comparison CTRL vs. All; 31 pathways for CTRL vs. Mild; 42 pathways for Mild vs. Moderate and Severe; 22 pathways for Moderate vs. Severe). We then established Venn diagrams (see [Supplementary Data 9, 10](#)) to

identify common pathways among the different comparisons ([Supplementary Figure 4](#)). The Venn diagram obtained for the three comparisons, CTRL vs. All, CTRL vs. Mild and Mild vs. Moderate and Severe, is given in [Supplementary Data 10](#). It showed that the majority of pathways that were altered by SARS-CoV-2 infection were detected in samples from subjects with only mild symptoms. These pathways mainly belonged to lipid and amino acid metabolism ([Supplementary Figure 4](#)). These results clearly indicate that SARS-CoV-2 infection caused strong alterations of the plasma metabolome even though the infected subjects suffered only very mild symptoms. Further studies are now needed to obtain a better understanding of the physiopathology of the mild form of COVID-19 and possible side effects the virus infection. In this context, it will be interesting to elucidate the effect of COVID-19 vaccination on the plasma metabolome response to SARS-CoV-2 infection.

All samples of our cohort were collected in the hospital of Nice during the first COVID-19 wave of the pandemic. This means there was a restricted coverage area for sample collection but also guaranteed that all subjects that participated in the study (healthy controls and COVID-19 patients) were never exposed to SARS-CoV-2 infection or vaccination prior to sample collection. The limited number of samples of each class did not allow for both a training set and a validation set. However, as described above and illustrated in [Figure 6](#), our results are supported by metabolomic analyses published by other groups.

Based on our studies and those of the other groups, several robust metabolome plasma biomarkers for COVID-19 diagnosis and prognosis have been identified. These plasma biomarkers could be used for targeted LC-MS/MS analyses which are fast, low-cost, and suitable for routine clinical testing. Metabolomic studies of other biological fluids such as urine (42) or exhaled breath (43) already gave interesting results and could also be very useful for rapid COVID-19 diagnosis and prognosis. However, further studies are now needed to identify biomarkers for the diagnosis of specific SARS-CoV-2 variants. It would also be beneficial to compare the plasma metabolome fingerprintings of SARS-CoV-2-infected subjects and patients infected by other viruses that evoke flu-like or COVID-like symptoms. In this context, Sun and collaborators studied the levels of serum fatty acids related to disease severity after infection with H7N9, an avian-origin influenza A virus that can lead to severe lung damage (44). As illustrated in [Figure 6](#), and similar to the findings of Sun et al., the plasma levels of the identified biomarkers for H7N9 decrease with increasing COVID-19 severity. Here, aiming the identification of biomarkers for the prognosis of COVID-19 severity, we studied plasma samples of COVID-19 patients with moderate and severe symptoms that were collected at hospital admission. For better understanding of the physiopathology of COVID-19, further studies will focus on metabolome analyses of samples that were collected repeatedly over the whole period in which our patient cohort stayed at the hospital. Finally, this study

will be complemented by metabolomic analyses on samples from subjects infected with other SARS-CoV-2 variants (delta, omicron, etc.) and with samples from infected subjects who were previously vaccinated.

In conclusion, our results showed that plasma metabolome profiling is an efficient tool for the diagnosis and prognosis of SARS-CoV-2 infection. We found that modifications in the plasma levels of lipids and amino acids were the main features of these metabolomic signatures. In particular, high levels of unsaturated lipids and fructosyl amino acids, and low levels of saturated lipids were identified in the plasma of infected subjects and serve as robust biomarkers for the diagnosis of SARS-CoV-2 infection and the prognosis of COVID-19 severity. Our results also show that major alterations of metabolite levels due to SARS-CoV-2 infection occurred primarily in the plasma of subjects with only mild symptoms and that the kynurenate pathways play a key role in the symptoms of COVID-19.

## Data availability statement

All datasets can be found in the [Supplementary material](#).

## Ethics statement

The studies involving human participants were reviewed and approved by the Comité de Protection des Personnes Sud-Ouest et Outre-Mer 1. The patients/participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

## Author contributions

CO, SL, JL, and TP designed the study. CO, VB, BS-P, and JD collected clinical data. CO, JL, and TP verified the underlying data. CO and FG prepared the samples. J-MG performed the mass spectrometry analyses. CO, AC, and TP performed post treatment of the raw data, statistical analysis, and pathway analysis. CO, SL, and TP analyzed and interpreted the results. CO and TP wrote the manuscript. SL revised critically the report

## References

1. Majumder J, Minko T. Recent developments on therapeutic and diagnostic approaches for Covid-19. *AAPS J.* (2021) 23:14. doi: 10.1208/s12248-020-00532-2
2. Poti F, Pozzoli C, Adami M, Poli E, Costa LG. Treatments for Covid-19: emerging drugs against the coronavirus.

for important intellectual content. All authors revised critically the report and gave final approval of the version for publication and agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

## Funding

Equipment for this study was purchased through grants from the Recherche en Matières de Sécurité Nucléaire et Radioprotection program from the French National Research Agency and the Conseil Départemental 06.

## Acknowledgments

We thank Abby Cuttriss for editorial correction of the manuscript.

## 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.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmed.2022.995069/full#supplementary-material>

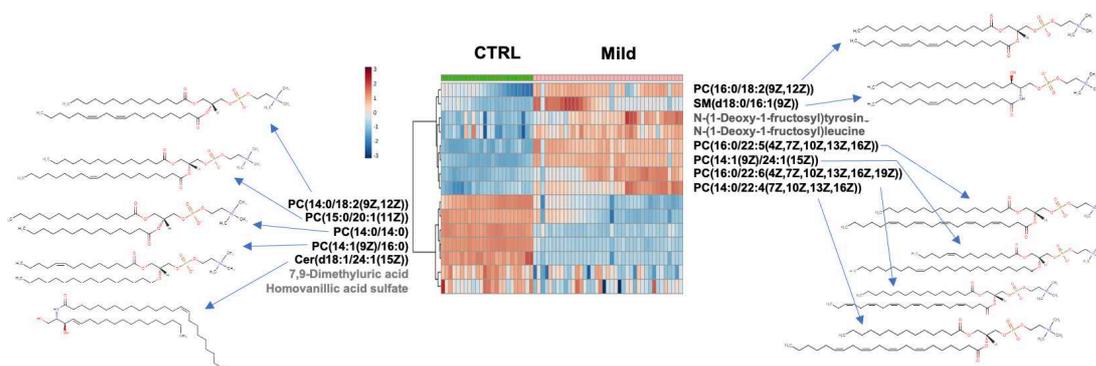
*Acta Biomed.* (2020) 91:118–36. doi: 10.23750/abm.v91i2.9639

3. Yuce M, Filiztekin E, Ozkaya KG. Covid-19 diagnosis -a review of current methods. *Biosens Bioelectron.* (2021) 172:112752. doi: 10.1016/j.bios.2020.112752

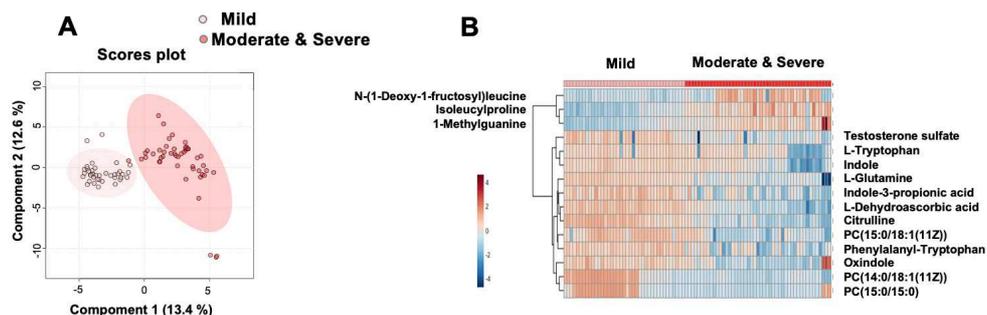
4. Keshavarz M, Soleymani-Mohammadi F, Namdari H, Arjeini Y, Mousavi MJ, Rezaei F. Metabolic host response and therapeutic approaches to influenza infection. *Cell Mol Biol Lett.* (2020) 25:15. doi: 10.1186/s11658-020-00211-2
5. Eisfeld AJ, Halfmann PJ, Wendler JP, Kyle JE, Burnum-Johnson KE, Peralta Z, et al. Multi-platform 'omics analysis of human ebola virus disease pathogenesis. *Cell Host Microbe.* (2017) 22:817–29.e8. doi: 10.1016/j.chom.2017.10.011
6. Shen B, Yi X, Sun Y, Bi X, Du J, Zhang C, et al. Proteomic and metabolomic characterization of Covid-19 patient sera. *Cell.* (2020) 182:59–72e15. doi: 10.1016/j.cell.2020.05.032
7. Wu D, Shu T, Yang X, Song J-X, Zhang M, Yao C, et al. Plasma metabolomic and lipidomic alterations associated with Covid-19. *Natl Sci Rev.* (2020) 7:1157–68. doi: 10.1093/nsr/nwaa086
8. Atzrodt CL, Maknojia I, McCarthy RDP, Oldfield TM, Po J, Ta KTL, et al. A guide to Covid-19: a global pandemic caused by the novel coronavirus Sars-Cov-2. *FEBS J.* (2020) 287:3633–50. doi: 10.1111/febs.15375
9. Gao Z, Xu Y, Sun C, Wang X, Guo Y, Qiu S, et al. A systematic review of asymptomatic infections with Covid-19. *J Microbiol Immunol Infect.* (2021) 54:12–6. doi: 10.1016/j.jmii.2020.05.001
10. Ruetsch C, Brglez V, Cremonini M, Zorzi K, Fernandez C, Boyer-Suavet S, et al. Functional exhaustion of type I and II interferons production in severe Covid-19 patients. *Front Med (Lausanne).* (2020) 7:603961. doi: 10.3389/fmed.2020.603961
11. Pluskal T, Castillo S, Villar-Briones A, Oresic M. Mzmine 2: modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinformatics.* (2010) 11:395. doi: 10.1186/1471-2105-11-395
12. Pang Z, Chong J, Zhou G, de Lima Morais DA, Chang L, Barrette M, et al. MetaboAnalyst 5.0: narrowing the gap between raw spectra and functional insights. *Nucleic Acids Res.* (2021) 49:W388–96. doi: 10.1093/nar/gkab382
13. Fraser DD, Slessarev M, Martin CM, Daley M, Patel MA, Miller MR, et al. Metabolomics profiling of critically ill coronavirus disease 2019 patients: identification of diagnostic and prognostic biomarkers. *Crit Care Explor.* (2020) 2:e0272. doi: 10.1097/CCE.0000000000000272
14. Dei Cas M, Ottolenghi S, Morano C, Rinaldo R, Roda G, Chiumello D, et al. Link between serum lipid signature and prognostic factors in Covid-19 patients. *Sci Rep.* (2021) 11:21633. doi: 10.1038/s41598-021-00755-z
15. Thomas T, Stefanoni D, Reisz JA, Nemkov T, Bertolone L, Francis RO, et al. Covid-19 infection alters kynurenine and fatty acid metabolism, correlating with IL-6 levels and renal status. *JCI Insight.* (2020) 5:e140327. doi: 10.1172/jci.insight.140327
16. Bruzzzone C, Bizkarguenaga M, Gil-Redondo R, Diercks T, Arana E, Garcia de Vicuna A, et al. Sars-Cov-2 infection dysregulates the metabolomic and lipidomic profiles of serum. *iScience.* (2020) 23:101645. doi: 10.1016/j.isci.2020.101645
17. Song JW, Lam SM, Fan X, Cao WJ, Wang SY, Tian H, et al. Omics-driven systems interrogation of metabolic dysregulation in Covid-19 pathogenesis. *Cell Metab.* (2020) 32:188–202e5. doi: 10.1016/j.cmet.2020.06.016
18. Casari I, Manfredi M, Metharom P, Falasca M. Dissecting lipid metabolism alterations in Sars-Cov-2. *Prog Lipid Res.* (2021) 82:101092. doi: 10.1016/j.plipres.2021.101092
19. Das UN. Bioactive lipids in Covid-19—further evidence. *Arch Med Res.* (2021) 52:107–20. doi: 10.1016/j.arcmed.2020.09.006
20. Bai Y, Huang W, Li Y, Lai C, Huang S, Wang G, et al. Lipidomic alteration of plasma in cured Covid-19 patients using ultra high-performance liquid chromatography with high-resolution mass spectrometry. *Biosci Rep.* (2021) 41:BSR20204305. doi: 10.1042/BSR20204305
21. Sindelar M, Stancliffe E, Schwaiger-Haber M, Anbukumar DS, Adkins-Travis K, Goss CW, et al. Longitudinal metabolomics of human plasma reveals prognostic markers of Covid-19 disease severity. *Cell Rep Med.* (2021) 2:100369. doi: 10.1016/j.xcrm.2021.100369
22. Ketter E, Randall G. Virus impact on lipids and membranes. *Annu Rev Virol.* (2019) 6:319–40. doi: 10.1146/annurev-virology-092818-015748
23. Ebrahimi KH, McCullagh JSO. A lipidomic view of Sars-Cov-2. *Biosci Rep.* (2021) 41:BSR20210953. doi: 10.1042/BSR20210953
24. Theken KN, Tang SY, Sengupta S, FitzGerald GA. The roles of lipids in Sars-Cov-2 viral replication and the host immune response. *J Lipid Res.* (2021) 62:100129. doi: 10.1016/j.jlcr.2021.100129
25. Nguyen V, Zhang Y, Gao C, Cao X, Tian Y, Carver W, et al. The spike protein of Sars-Cov-2 impairs lipid metabolism and increases susceptibility to lipotoxicity: implication for a role of Nrf2. *Cells.* (2022) 11:1916. doi: 10.3390/cells11121916
26. Breier A, Davis OR, Buchanan RW, Moricle LA, Munson RC. Effects of metabolic perturbation on plasma homovanillic acid in schizophrenia: relationship to prefrontal cortex volume. *Arch Gen Psychiatry.* (1993) 50:541–50. doi: 10.1001/archpsyc.1993.01820190043005
27. Lawler NG, Gray N, Kimhofer T, Boughton B, Gay M, Yang R, et al. Systemic perturbations in amine and kynurenine metabolism associated with acute Sars-Cov-2 infection and inflammatory cytokine responses. *J Proteome Res.* (2021) 20:2796–811. doi: 10.1021/acs.jproteome.1c00052
28. Mangge H, Herrmann M, Meinitzer A, Pailer S, Curcic P, Sloup Z, et al. Increased kynurenine indicates a fatal course of Covid-19. *Antioxidants (Basel).* (2021) 10:1960–71. doi: 10.3390/antiox10121960
29. Cai Y, Kim DJ, Takahashi T, Broadhurst DL, Yan H, Ma S, et al. Kynurenine acid may underlie sex-specific immune responses to Covid-19. *Sci Signal.* (2021) 14:eabf8483. doi: 10.1126/scisignal.abf8483
30. Marin-Corral J, Rodriguez-Morato J, Gomez-Gomez A, Pascual-Guardia S, Munoz-Bermudez R, Salazar-Degracia A, et al. Metabolic signatures associated with severity in hospitalized Covid-19 patients. *Int J Mol Sci.* (2021) 22:4794–808. doi: 10.3390/ijms22094794
31. Samanic CM, Yue Y, Cote DJ, Stampfer MJ, Wang M, McCann A, et al. A prospective study of pre-diagnostic circulating tryptophan and kynurenine, and the kynurenine/tryptophan ratio and risk of glioma. *Cancer Epidemiol.* (2021) 76:102075. doi: 10.1016/j.canep.2021.102075
32. Huang JY, Larose TL, Luu HN, Wang R, Fanidi A, Alcalá K, et al. Circulating markers of cellular immune activation in prediagnostic blood sample and lung cancer risk in the lung cancer cohort consortium (Lc3). *Int J Cancer J Int Cancer.* (2020) 146:2394–405. doi: 10.1002/ijc.32555
33. Ye Z, Yue L, Shi J, Shao M, Wu T. Role of ido and tdo in cancers and related diseases and the therapeutic implications. *J Cancer.* (2019) 10:2771–82. doi: 10.7150/jca.31727
34. Barreto FS, Chaves Filho AJM, de Araujo M, de Moraes MO, de Moraes MEA, Maes M, et al. Tryptophan catabolites along the indoleamine 2,3-dioxygenase pathway as a biological link between depression and cancer. *Behav Pharmacol.* (2018) 29:165–80. doi: 10.1097/FBP.0000000000000384
35. Anderson G, Reiter RJ. Melatonin: roles in influenza, Covid-19, and other viral infections. *Rev Med Virol.* (2020) 30:e2109. doi: 10.1002/rmv.2109
36. Modoux M, Rolhion N, Lefevre JH, Ouevray C, Nadvornik P, Illes P, et al. Butyrate acts through hdac inhibition to enhance aryl hydrocarbon receptor activation by gut microbiota-derived ligands. *Gut Microbes.* (2022) 14:2105637. doi: 10.1080/19490976.2022.2105637
37. Hasan ZT, Atrakji D, Mehuaiden DAK. The effect of melatonin on thrombosis, sepsis and mortality rate in Covid-19 patients. *Int J Infect Dis.* (2022) 114:79–84. doi: 10.1016/j.ijid.2021.10.012
38. Anderson G, Maes M. Local melatonin regulates inflammation resolution: a common factor in neurodegenerative, psychiatric and systemic inflammatory disorders. *CNS Neurol Disord Drug Targets.* (2014) 13:817–27. doi: 10.2174/1871527313666140711091400
39. Markus RP, Fernandes PA, Kinker GS, da Silveira Cruz-Machado S, Marcola M. Immune-pineal axis - acute inflammatory responses coordinate melatonin synthesis by pinealocytes and phagocytes. *Br J Pharmacol.* (2018) 175:3239–50. doi: 10.1111/bph.14083
40. Muxel SM, Pires-Lapa MA, Monteiro AW, Cecon E, Tamura EK, Floeter-Winter LM, et al. Nf-kappab drives the synthesis of melatonin in raw 264.7 macrophages by inducing the transcription of the arylalkylamine-n-acetyltransferase (Aa-Nat) gene. *PLoS One.* (2012) 7:e52010. doi: 10.1371/journal.pone.0052010
41. Anderson G, Carbone A, Mazzoccoli G. Tryptophan metabolites and aryl hydrocarbon receptor in severe acute respiratory syndrome, Coronavirus-2 (Sars-Cov-2) pathophysiology. *Int J Mol Sci.* (2021) 22:1597. doi: 10.3390/ijms22041597
42. Bi X, Liu W, Ding X, Liang S, Zheng Y, Zhu X, et al. Proteomic and metabolomic profiling of urine uncovers immune responses in patients with Covid-19. *Cell Rep.* (2022) 38:110271. doi: 10.1016/j.celrep.2021.110271
43. Grassin-Delyle S, Roquencourt C, Moine P, Saffroy G, Carn S, Heming N, et al. Metabolomics of exhaled breath in critically ill Covid-19 patients: a pilot study. *EBioMedicine.* (2021) 63:103154. doi: 10.1016/j.ebiom.2020.103154
44. Sun X, Song L, Feng S, Li L, Yu H, Wang Q, et al. Fatty acid metabolism is associated with disease severity after H7n9 infection. *EBioMedicine.* (2018) 33:218–29. doi: 10.1016/j.ebiom.2018.06.019

## Appendix

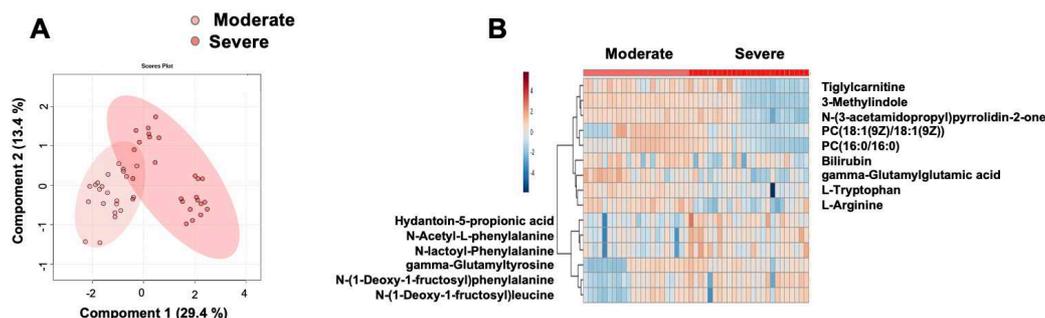
### Supplementary Figures



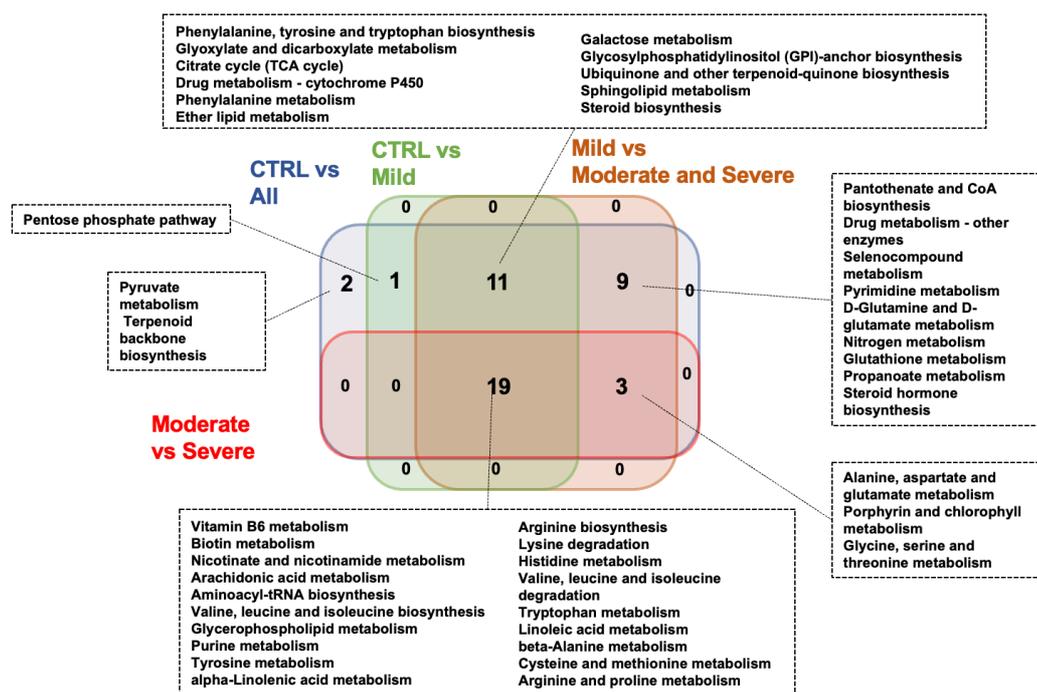
**Fig. S1.** Heatmap of the top 15 metabolites of mild SARS-CoV-2 infected subjects (Mild) compared to non-infected subjects (CTRL). Metabolites were clustered using the Ward method on T-test and ANOVA values. The molecular structures of the lipids are shown.



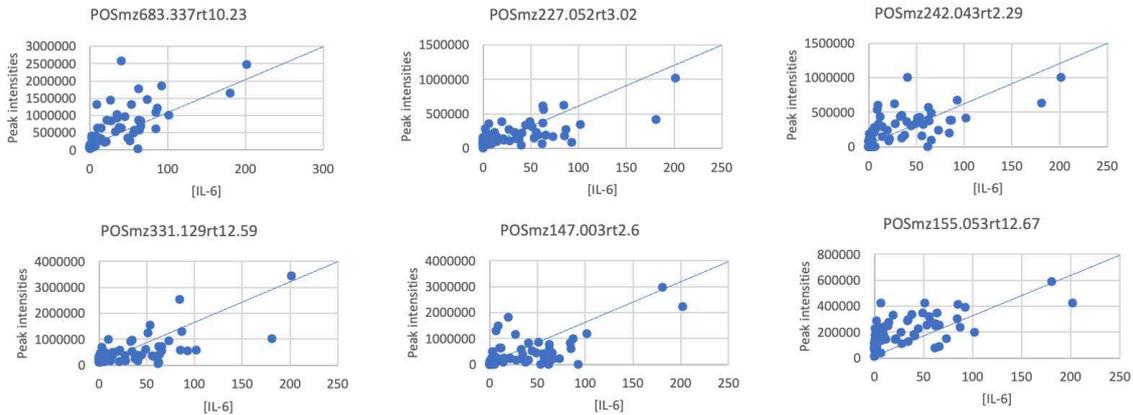
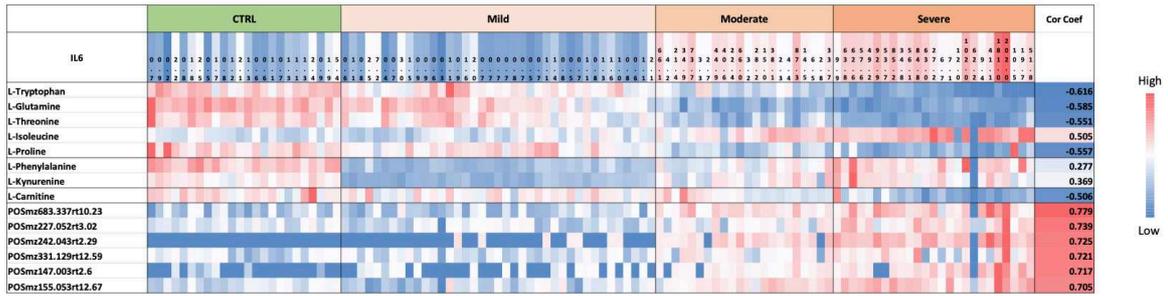
**Fig. S2.** Identification of the putative biomarkers of moderate and severe COVID-19 patients (Moderate & Severe) compared to mild SARS-CoV-2 infected subjects (Mild). PLS-DA analysis based on untargeted plasma metabolomic profiling discriminated mild SARS-CoV-2 infected patients (Mild) from Moderate and Severe COVID-19 patients (Moderate & Severe). Light red and darker red dots represent Mild and Moderate & Severe SARS-CoV-2 infected subjects, respectively. The score plot is represented with a confidence ellipse of 95 % (A). Heatmap of the top 15 putative biomarkers of moderate and severe COVID-19 patients (Moderate & Severe) compared to mild SARS-CoV-2 infected subjects (Mild) (B). Metabolites were clustered using the Ward method on T-test and ANOVA values. Identified features were MS2 validated and are shown as chemical names.



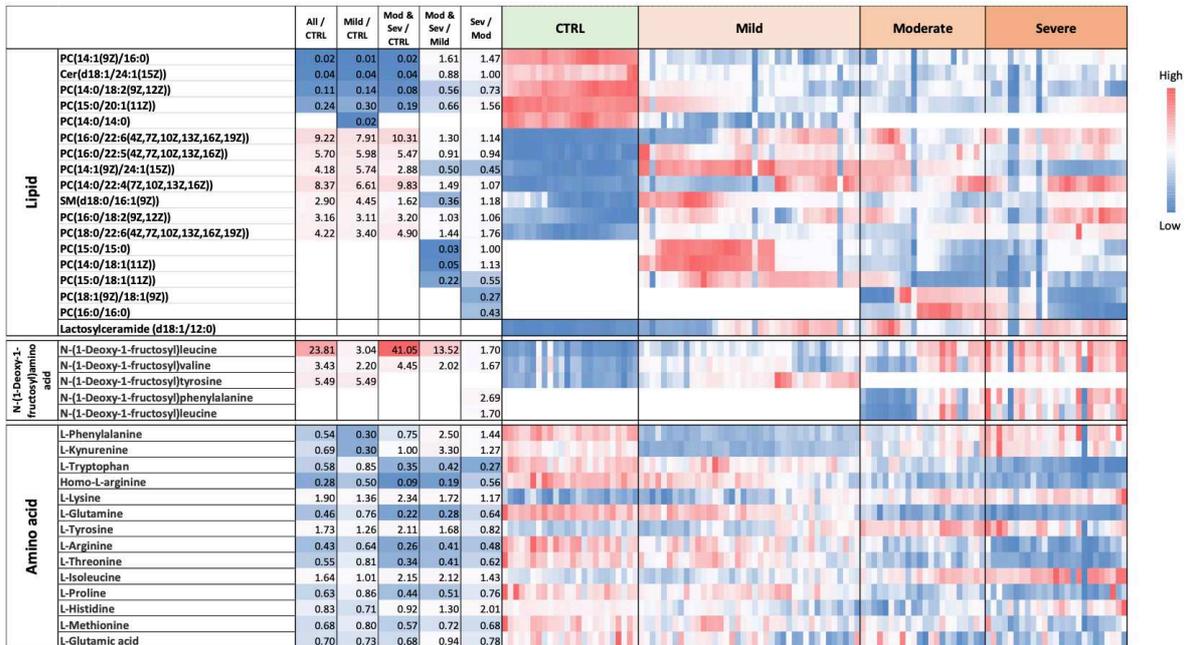
**Fig. S3. Identification of the putative biomarkers of severe (Severe) compared to moderate COVID-19 patients (Moderate).** PLS-DA analysis based on untargeted plasma metabolomic profiling discriminated severe from moderate COVID-19 patients (Severe, Moderate, respectively). Lighter red and darker red dots represent moderate and severe COVID-19 patients, respectively. The score plot is represented with a confidence ellipse of 95 % (A). Heatmap of the top 15 putative biomarkers of moderate COVID-19 patients (Moderate) compared to severe COVID-19 patients (Severe) (B). Metabolites were clustered using the Ward method on T-test and ANOVA values. Identified features were MS2 validated and are shown as chemical named.



**Fig. S4. Venn diagram of significantly altered pathways of healthy control subjects compared to SARS-CoV-2 infected subjects (CTRL vs All), healthy control subjects compared to mild SARS-CoV-2 infected subjects (CTRL vs Mild), mild SARS-CoV-2 infected subjects compared to moderate and severe COVID-19 patients (Mild vs Moderate & Severe) and moderate COVID-19 patients compared to severe COVID-19 patients (Moderate vs Severe).** A value of  $p < 0.05$  was considered significant.



**Fig. S5.** Heatmap of plasma levels of metabolites putatively related to plasma IL-6 concentrations (upper lines) and of unidentified ions showing higher correlation coefficients with IL-6 concentrations. The values are available in the Supplemental Excel File: Supplemental\_data/Supp Fig 5.



**Fig. S6.** Heatmap of plasma levels of various lipids, N-(1-deoxy-1-fructosyl)amino acids and amino acids. The values are available in Supplemental Excel File: Supplemental\_data/Supp Fig 6.

## Supplementary Table

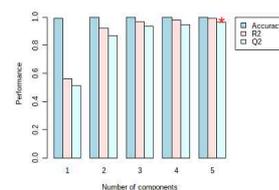
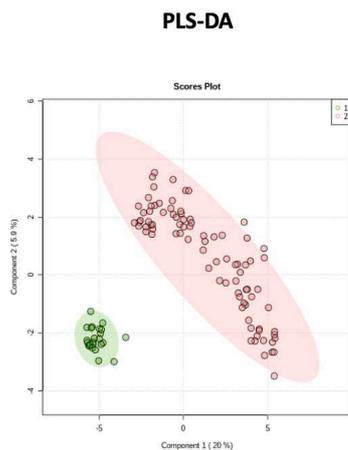
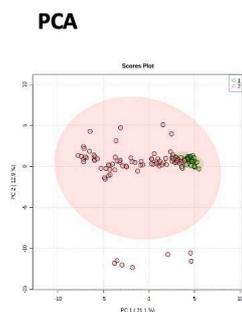
	CTRL n=24	Mild cases n=39	Moderate cases n=22	Severe Cases n=25	<u>p value</u>
<b>Demographic and baseline characteristics</b>					
Age, (years) mean $\pm$ SD	46 $\pm$ 11	35 $\pm$ 16	60 $\pm$ 15	63 $\pm$ 14	<0.001
Male, n (%)	5 (21)	6 (15)	11 (50)	18 (72)	0.0037
Obesity, n (%)	NA	1 (3)	2 (9)	5 (20)	0.06
Cardiovascular history, n (%)	NA	1 (3)	9 (41)	12 (48)	< 0.0001
Chronic respiratory failure, n (%)	NA	0	1 (5)	0	0.24
Diabetes, n (%)	NA	0	4 (18)	9 (36)	0.0003
Malignancy, n (%)	NA	0	2 (9)	4 (16)	0.04
<b>Severity of COVID-19 symptoms</b>					
Days after first signs of COVID-19 (days), median(Q1-Q3)	NA	14 (9-21)	5 (3-8)	5 (3-8)	< 0.00001
Oxygen therapy, n (%)	NA	1 (3)	15 (68)	19 (76)	< 0.0001
Mechanical ventilation or NIV, n (%)	NA	0	0	12 (48)	< 0.0001
In hospital mortality, n (%)	NA	0	0	5 (20)	0.0005

**Table S1: Demographics, baseline characteristics and severity factors of healthy subjects and COVID-19 patients.** Data are presented as median (interquartile range: Q1-Q3) or n (%) and p values are reported. NA: not available.

## Supplementary data

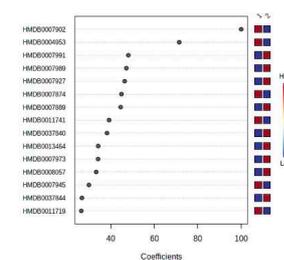
### data S1B

MetaboAnalyst statistical analysis  
CTRL vs COVID  
Sample normalization by median  
Log transformation of the data  
No data scaling



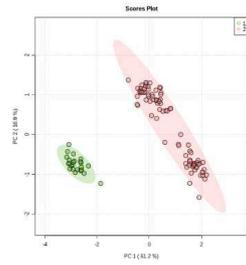
PLS-DA cross validation details:

Measure	1 comps	2 comps	3 comps	4 comps	5 comps
Accuracy	0.992	1.0	1.0	1.0	1.0
R2	0.56303	0.92123	0.96482	0.97802	0.98875
Q2	0.51249	0.65698	0.9342	0.94564	0.95028

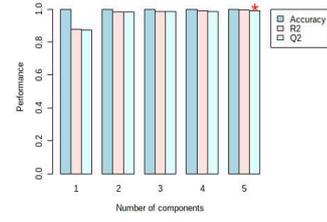
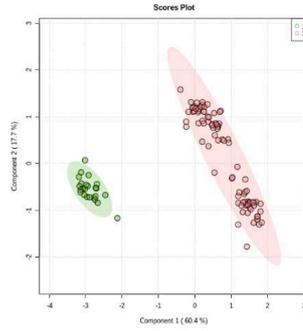


Top 15 metabolites only  
 Log transformation of the data  
 No data scaling

**PCA**

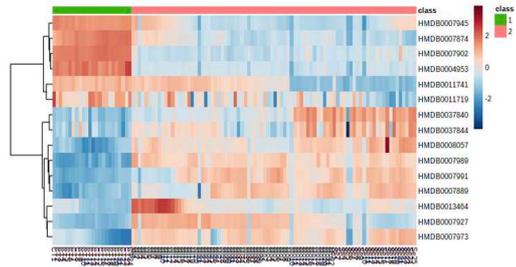


**PLS-DA**

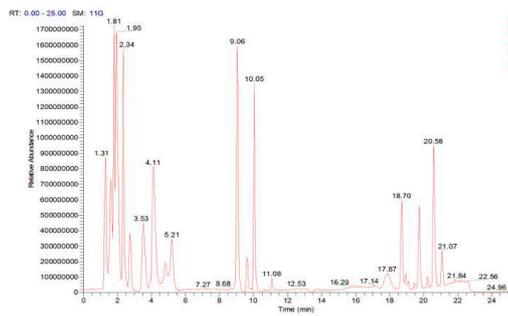
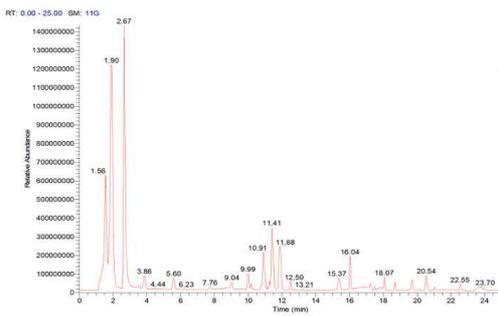


**PLS-DA cross validation details:**

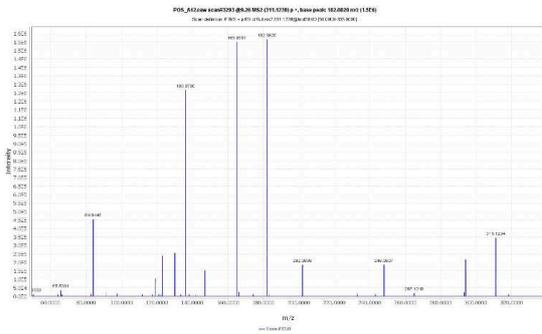
Measure	1 comps	2 comps	3 comps	4 comps	5 comps
Accuracy	1.0	1.0	1.0	1.0	1.0
R2	0.87694	0.98306	0.98796	0.99077	0.99306
Q2	0.87304	0.98159	0.98525	0.98772	0.99104



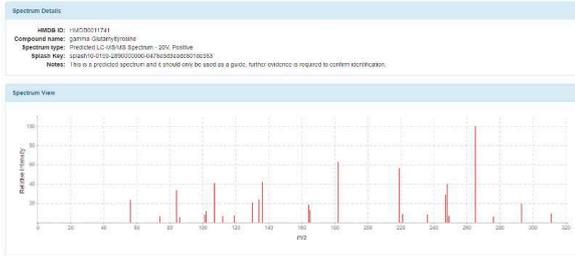
**data S1C**  
**Typical Total Ion Chromatograms for positive and negative modes**



**data S1D**  
**MS2 spectra and matches of top features**  
**HMDB0011741-POSmz311.124rt8.63**  
**gamma-Glutamyltyrosine**

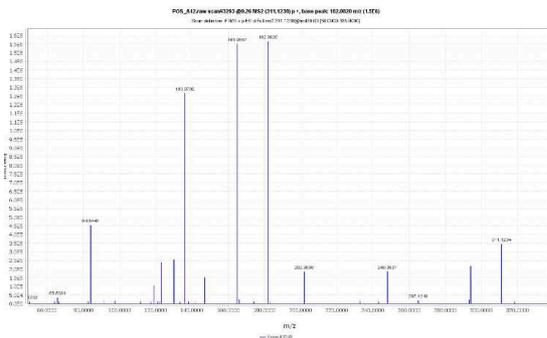


- 84.0448 45
- 119.0488 10
- 123.0488 23
- 136.0768 121
- 165.0578 150
- 182.0819 151
- 202.0868 18
- 248.0907 18
- 265.1242 1
- 294.0960 21
- 311.1233 34



- 56.04947561 3.090659601
- 74.02365479 0.9090021143
- 84.04439023 4.38217896
- 86.0600403 0.759500479
- 101.0709393 1.104857016
- 102.0549549 1.554568153
- 107.0491413 5.35318951
- 112.0393049 0.917253027
- 119.0491413 0.9697036191
- 130.0498695 2.719952128
- 134.0600403 3.164745997
- 136.0756904 5.503739826
- 164.070605 2.463239316
- 165.0546206 1.705453906
- 182.0811697 6.232617036
- 219.1128041 7.412227403
- 221.1284542 1.148004939
- 236.0917343 1.11965838
- 247.1077188 3.778279339
- 248.0917343 5.222418184
- 249.1233688 0.9430655035
- 265.1182834 13.06699515
- 276.089649 0.8137506157
- 293.1131981 2.586864046
- 311.1237627 1.234464109

**HMDB0011741-POSmz311.124rt4.91**  
**gamma-Glutamyltyrosine**



- 84.0448 45
- 119.0488 10
- 123.0488 23
- 136.0768 121
- 165.0578 150
- 182.0819 151
- 202.0868 18
- 248.0907 18
- 265.1242 1
- 294.0960 21
- 311.1233 34

**Spectrum FiehnHILIC000378 for gamma-Glutamyltyrosine** CLEAN

Rating: ★★★★★  
 Originally submitted to the Fiehn Lab HILIC Library

[Fiehn HILIC](#)

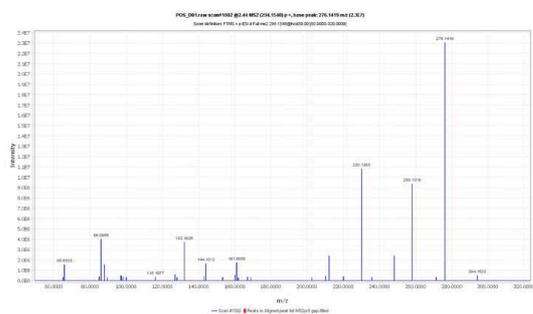
[LC-MS](#)

SPLASH: [splash10-00m0-0910000000-aa63925629601d435f15](#)

Submitter: [qMegan Showalter](#)

**Mass Spectrum**

# HMDB0037840-POSsmz294.155mz2.79 N-(1-Deoxy-1-fructosyl)leucine



- 276.141 230
- 230.138 108
- 258.131 93
- 86.096 40
- 132.102 37
- 212.128 24
- 248.148 24
- 161.063 17
- 344.101 16
- 88.039 15

Predicted LC-MS/MS Spectrum - 20V, Positive (HMDB0037840)

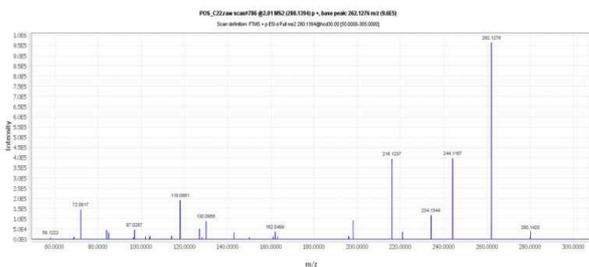
**Spectrum Details**

HMDB ID: HMDB0037840  
 Compound name: N-(1-Deoxy-1-fructosyl)leucine  
 Spectrum type: Predicted LC-MS/MS Spectrum - 20V, Positive  
 Splash Key: splash10.009.6990000000.880344214040470930  
 Notes: This is a predicted spectrum and it should only be used as a guide, further evidence is required to confirm identification.

**Spectrum View**

- 55.01838972 1.50532292
- 57.07042529 5.959208626
- 59.06075335 1.13205916
- 69.07042529 2.68518918
- 71.01330434 1.949568318
- 73.0289544 1.990834972
- 75.04460446 1.001278029
- 86.09697438 1.934403141
- 98.09997438 6.623638131
- 100.1126244 2.119050135
- 115.0759046 1.916500006
- 126.091889 1.195048632
- 128.1075391 0.8836452797
- 144.1024537 11.54874896
- 145.0500838 2.45161707
- 146.1161038 1.135501995
- 162.0768329 1.983468541
- 163.0606485 2.759072922
- 180.0871975 2.894284227
- 188.1286684 1.061536523
- 230.1392331 8.218472294
- 232.1184977 0.8812166813
- 232.1548832 0.8090613845
- 236.0770268 0.7872064206
- 248.1497978 12.0871219
- 258.1341477 2.879183237
- 276.1447124 8.591842219
- 294.1552771 5.45444658

# HMDB37844-POSsmz280.139rt2.33 N-(1-Deoxy-1-fructosyl)valine



- 262.127 964
- 244.116 396
- 216.123 393
- 118.086 190
- 72.081 144
- 234.134 117
- 158.112 91
- 130.086 68
- 127.039 50
- 97.028 45

Predicted LC-MS/MS Spectrum - 20V, Positive (HMDB0037844)

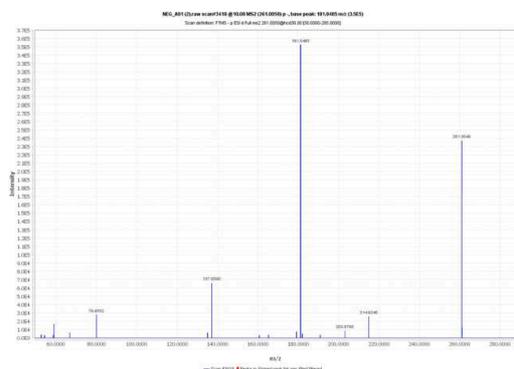
**Spectrum Details**

HMDB ID: HMDB0037844  
 Compound name: N-(1-Deoxy-1-fructosyl)valine  
 Spectrum type: Predicted LC-MS/MS Spectrum - 20V, Positive  
 Splash Key: splash10.009.6990000000.880344214040470930  
 Notes: This is a predicted spectrum and it should only be used as a guide, further evidence is required to confirm identification.

**Spectrum View**

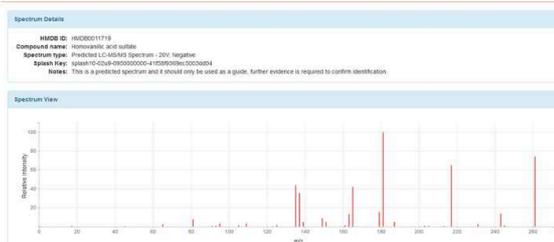
- 55.01838972 1.488384202
- 71.01330434 1.900965494
- 72.08132432 2.270569157
- 73.0289544 2.020548924
- 75.04460446 1.018399304
- 84.08132432 7.113817082
- 85.06997438 2.490919876
- 112.0762389 1.258484153
- 114.091889 0.9688745841
- 130.0868036 11.61913207
- 132.1024537 1.116026522
- 144.1024537 0.8584703423
- 145.0500838 2.450300598
- 163.0606485 2.777968922
- 174.1130194 1.455391413
- 194.1028476 0.8805613357
- 216.1235831 9.202143376
- 218.1362331 1.063277178
- 220.0821122 0.9741092507
- 222.1341477 0.761047631
- 234.1341477 13.71020442
- 236.0770268 0.7635335577
- 236.1497978 0.8021429827
- 244.1184977 2.757984894
- 262.1290624 8.32777654
- 280.139627 5.182260798

# HMDB0011719-NEGmz261.005rt9.45 Homovanillic acid sulfate



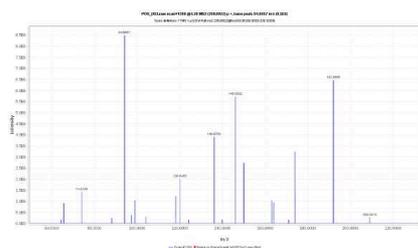
- 181.048 352
- 261.004 237
- 137.059 66
- 79.955 27
- 214.924 25
- 59.011 16
- 203.078 8
- 179.032 7
- 135.043 6
- 66.753 5

Predicted LC-MS/MS Spectrum - 20V, Negative (HMDB0011719)



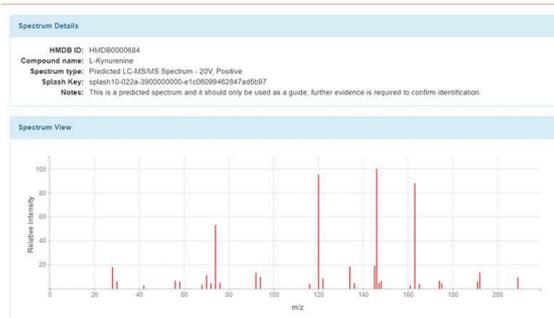
- 65.03912516 0.5354413059
- 80.96463989 1.681135423
- 91.01838972 0.1556255371
- 92.99765427 0.2341283559
- 95.04968984 0.7152760139
- 105.0340398 0.2911365253
- 109.0289544 0.700118041
- 123.008219 0.0936759559
- 125.023869 0.3559112347
- 135.0446045 9.535519405
- 137.0602545 7.745068138
- 139.0395191 1.10229885
- 149.023869 1.964196002
- 151.0395191 1.06437631
- 160.9908546 0.2036359653
- 161.023869 0.252076245
- 163.0395191 2.916976165
- 165.0187836 9.210287938
- 179.0344337 3.326849638
- 181.0500838 21.82551292
- 187.0065047 1.072214743
- 202.9650338 0.2204171099
- 204.9806839 0.1400208218
- 212.9857693 0.1223364614
- 217.0170694 14.22164936
- 230.9963339 0.5753821738
- 242.9963339 2.97281213
- 244.9755985 0.2530962078
- 261.0068986 16.22631087

# HMDB0000684-POSmz209.092rt4.65 L-Kynurenine



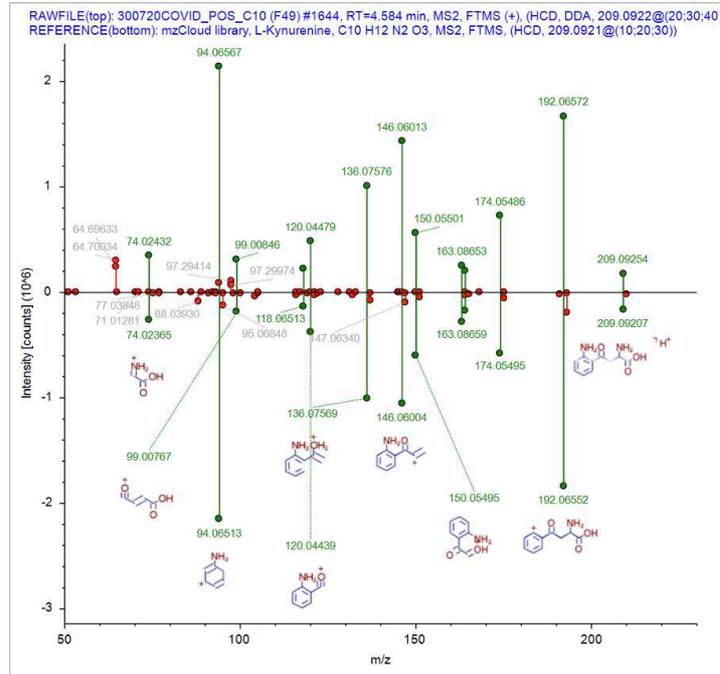
- 94.005 847
- 192.065 645
- 146.067 570
- 136.075 390
- 174.055 393
- 150.054 272
- 120.044 200
- 74.024 141
- 118.065 122
- 99.008 103

Predicted LC-MS/MS Spectrum - 20V, Positive (HMDB0000684)



- 56.01363688 1.238362
- 58.02828875 1.077138969
- 68.01363688 0.5568134019
- 70.02928875 2.056380748
- 72.0085533 0.7229290319
- 72.04493881 0.8056927466
- 74.02420337 9.833847873
- 76.03985343 0.8947303303
- 92.05002419 2.481265362
- 94.06567426 1.815377291
- 116.0347881 0.5748675028
- 116.0500242 0.7208143943
- 120.0449388 17.57411567
- 122.0605889 1.56487464
- 134.0605889 3.446180532
- 136.0762389 0.8397125823
- 145.0765733 3.524973773
- 146.0605889 18.4537027
- 147.0922234 0.9059414802
- 148.0762389 1.199529612
- 161.0714879 0.4751485555
- 163.087138 16.24160595
- 165.102788 0.7351195281
- 174.0555035 1.211394306
- 175.0395191 0.8319393504
- 191.0820526 1.123548881
- 192.0660682 2.462030142
- 209.0926173 1.729167802

HMDB0000684-POS<sub>m/z</sub>209.092<sub>rt</sub>4.65  
L-Kynurenine



data S2B

MetaboAnalyst statistical analysis

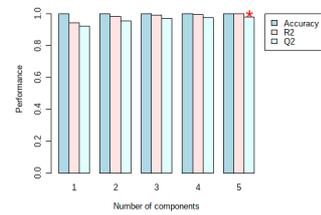
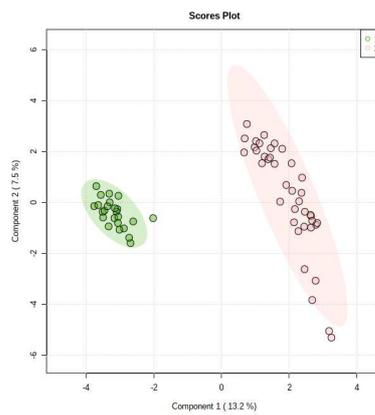
CTRL vs Mild

Sample normalization by median

Log transformation of the data

No data scaling

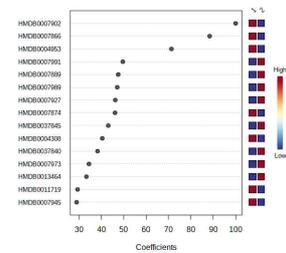
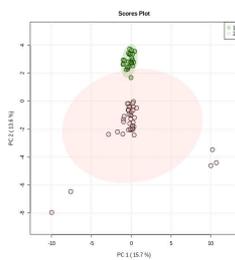
PLS-DA



PLS-DA cross validation details:

Measure	1 comps	2 comps	3 comps	4 comps	5 comps
Accuracy	1.0	1.0	1.0	1.0	1.0
R2	0.94267	0.98228	0.99174	0.99512	0.99571
Q2	0.91978	0.95414	0.96895	0.9736	0.9763

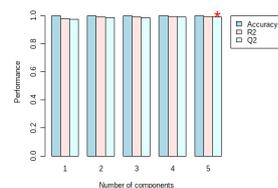
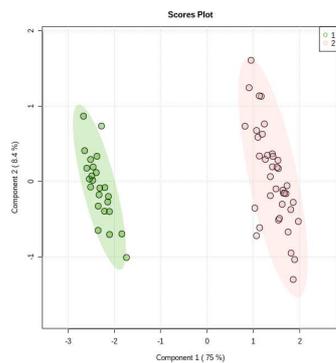
PCA



Top 15 metabolites only

Log transformation of the data  
No data scaling

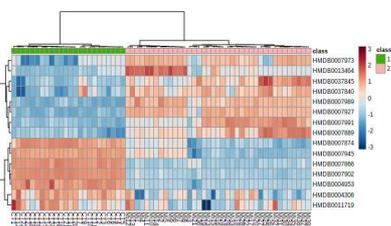
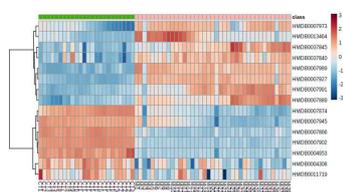
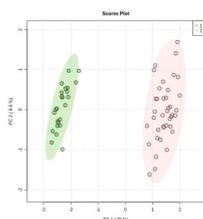
### PLS-DA



PLS-DA cross validation details:

Measure	1 comp	2 comp	3 comp	4 comp	5 comp
Accuracy	1.0	1.0	1.0	1.0	1.0
R2	0.97742	0.98908	0.99225	0.99428	0.99474
Q2	0.97532	0.98612	0.98799	0.98958	0.99032

### PCA

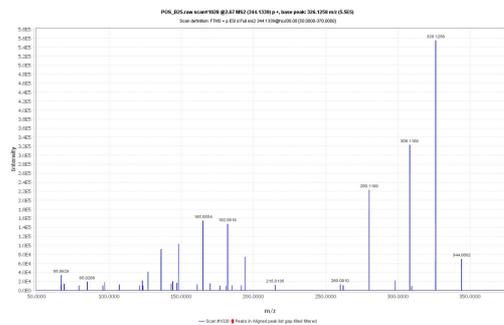
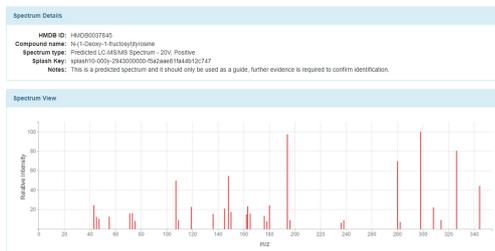


### data S2C

### MS2 spectra and matches of top features

HMDB0037845-POSmz344.134rt3.02  
N-(1-Deoxy-1-fructosyl)tyrosine

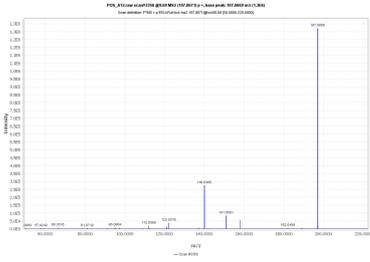
Predicted LC-MS/MS Spectrum - 20V, Positive (HMDB0037845)



- 326.125 554
- 308.11 323
- 280.119 223
- 155.055 154
- 182.081 147
- 148.075 103
- 136.075 91
- 194.082 74
- 344.06 69
- 127.038 40

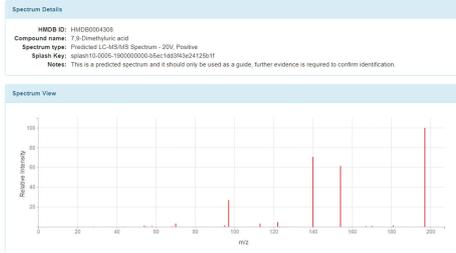
- 55.01838972 1.497858729
- 71.01330424 1.878927837
- 73.0289544 1.916631634
- 75.04460446 0.9882643725
- 107.0496898 5.880743001
- 109.0653389 1.073659441
- 119.0496898 2.685011975
- 136.0762389 1.816198552
- 145.0500838 2.493071531
- 148.0762389 6.439186022
- 150.091889 2.075101556
- 162.0766329 1.788241547
- 163.0606485 2.751536382
- 165.0551691 1.870906774
- 176.0711536 1.583445991
- 178.0868036 0.9435752261
- 180.0871975 2.871935768
- 194.0817182 11.52197555
- 196.0973683 1.079203539
- 236.0770268 0.7766168198
- 238.107933 1.057778759
- 280.1184977 8.25411801
- 282.1341477 0.865398677
- 298.1290624 11.81823654
- 308.1134123 2.632567953
- 314.123977 1.097479187
- 326.123977 5.517102337
- 344.1345417 5.216917417

HMDB0004308-POSmz197.065rt9.57  
7,9-Dimethyluric acid



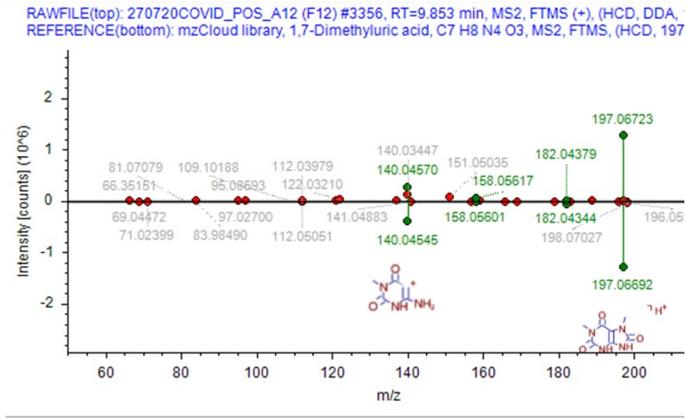
- 197.066 127
- 140.046 27
- 151.05 8
- 158.057 5
- 122.031 3
- 197.122 2
- 112.030 1
- 121.029 1
- 66.351 0
- 95.086 0

Predicted LC-MS/MS Spectrum - 20V, Positive (HMDB0004308)



- 68.01363668 0.1231298854
- 69.04527316 0.0358604038
- 70.02928675 1.116774217
- 72.0085533 0.0213864805
- 85.04018778 0.0302180788
- 86.02420337 0.026703597
- 95.02453772 0.482125039
- 96.0085533 0.004657773
- 97.04018778 9.690383604
- 100.0398534 0.0260304691
- 112.0510868 0.0019540812
- 113.0351024 1.10891938
- 122.0354367 1.733356212
- 123.0194523 0.0621447514
- 125.0351024 0.0330551477
- 127.0507525 0.0418671892
- 140.0460014 25.44208353
- 154.025266 0.0052210304
- 154.0616515 22.09635331
- 167.0569005 0.2120558916
- 170.0565661 0.3168925976
- 181.036165 0.4774295929
- 197.0674651 35.99223359

HMDB0004308-POSmz197.065rt9.57  
7,9-Dimethyluric acid

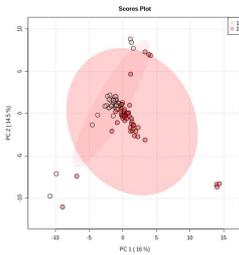


## data S3B

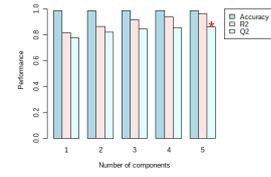
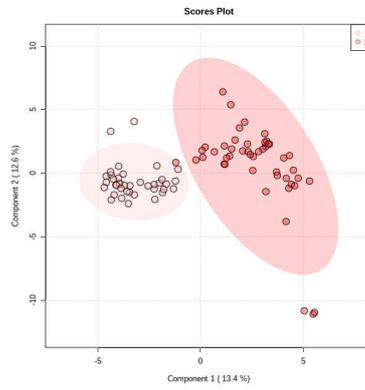
MetaboAnalyst statistical analysis

Mild vs Moderate&Severe  
 Sample normalization by median  
 Log transformation of the data  
 No data scaling

### PCA

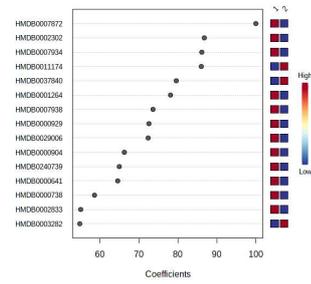


### PLS-DA



#### PLS-DA cross validation details:

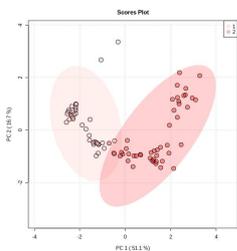
Measure	1 comps	2 comps	3 comps	4 comps	5 comps
Accuracy	0.98667	0.98667	0.98667	0.98667	0.98667
R2	0.81588	0.86521	0.91862	0.93812	0.96189
Q2	0.77761	0.81877	0.84563	0.85468	0.86092



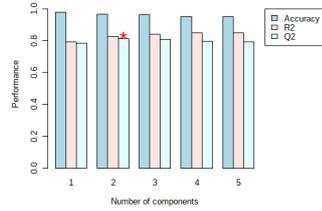
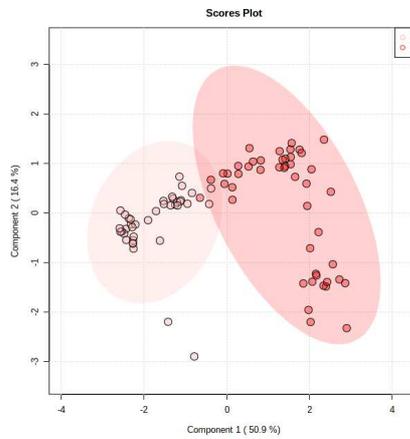
Top 15 metabolites only

Log transformation of the data  
 No data scaling

### PCA

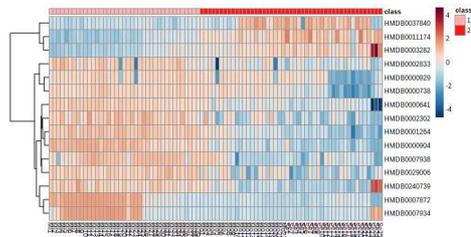


### PLS-DA



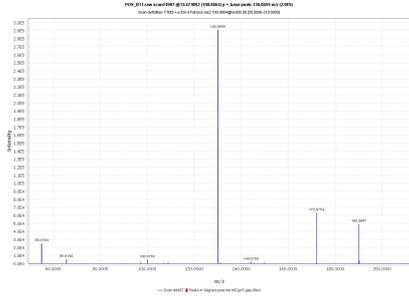
#### PLS-DA cross validation details:

Measure	1 comps	2 comps	3 comps	4 comps	5 comps
Accuracy	0.97667	0.96667	0.96333	0.95	0.95
R2	0.79376	0.82604	0.84043	0.84673	0.84955
Q2	0.78312	0.81163	0.8066	0.79425	0.79024



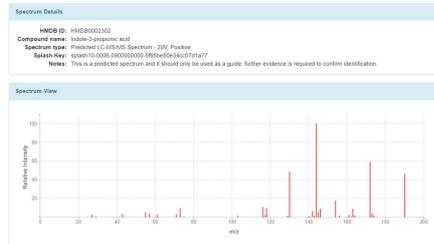
**data S3C**  
**MS2 spectra and matches of top features**

HMDB0002302-POSmz190.086mz13.34  
 Indole-3-propionic acid



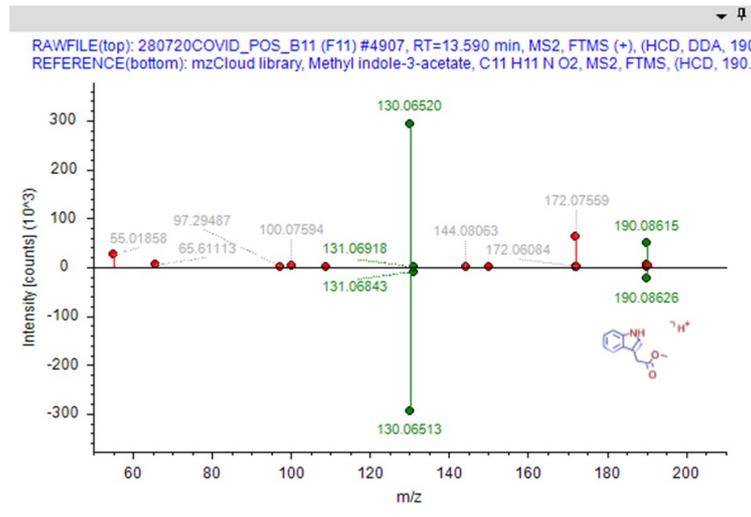
- 130.065 290
- 131.068 43
- 132.071 57
- 133.074 71
- 134.077 85
- 135.080 99
- 136.084 13
- 137.087 27
- 138.090 41
- 139.093 55
- 140.096 69
- 141.100 00
- 142.103 14
- 143.106 28
- 144.109 42
- 145.112 56
- 146.115 70
- 147.118 84
- 148.121 98
- 149.125 12
- 150.128 26
- 151.131 40
- 152.134 54
- 153.137 68
- 154.140 82
- 155.143 96
- 156.147 10
- 157.150 24
- 158.153 38
- 159.156 52
- 160.159 66
- 161.162 80
- 162.165 94
- 163.169 08
- 164.172 22
- 165.175 36
- 166.178 50
- 167.181 64
- 168.184 78
- 169.187 92
- 170.191 06
- 171.194 20
- 172.197 34
- 173.200 48
- 174.203 62
- 175.206 76
- 176.209 90
- 177.213 04
- 178.216 18
- 179.219 32
- 180.222 46
- 181.225 60
- 182.228 74
- 183.231 88
- 184.235 02
- 185.238 16
- 186.241 30
- 187.244 44
- 188.247 58
- 189.250 72
- 190.253 86
- 191.257 00
- 192.260 14
- 193.263 28
- 194.266 42
- 195.269 56
- 196.272 70
- 197.275 84
- 198.278 98
- 199.282 12
- 200.285 26

Predicted LC-MS/MS Spectrum - 20V, Positive (HMDB0002302)

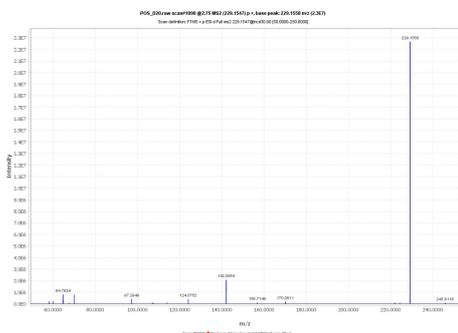


- 55.01838972 1.409913163
- 57.03403978 0.9651209774
- 61.0289544 0.8606350433
- 71.01330434 0.779286998
- 73.0289544 2.589704048
- 75.04460446 0.2287804966
- 103.0547752 0.3622066604
- 116.0500242 2.895068925
- 117.0704253 0.6677135185
- 118.0656743 2.623190231
- 129.0704253 0.4264316881
- 130.0656743 13.08734414
- 140.0500242 0.2531632039
- 142.0656743 1.710712003
- 143.0496898 0.4497103395
- 144.0813243 26.80581219
- 145.0653399 1.395002893
- 146.0969744 2.403789133
- 154.0856743 4.741043673
- 156.0449388 0.4117148874
- 161.0602545 0.6261725518
- 162.0555035 0.2096355267
- 163.0759046 2.362466551
- 164.0711536 0.5095295954
- 172.0762389 15.7742866
- 173.0602545 1.078927751
- 174.0555035 0.547305101
- 190.0868036 12.31183111

HMDB0002302-POSmz190.086mz13.34  
 Indole-3-propionic acid

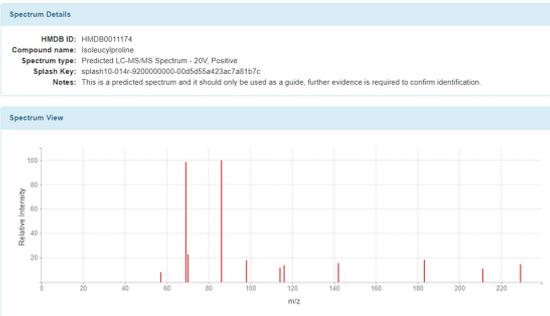


## HMDB0011174-POSsmz229.155mz2.6 Isoleucylproline



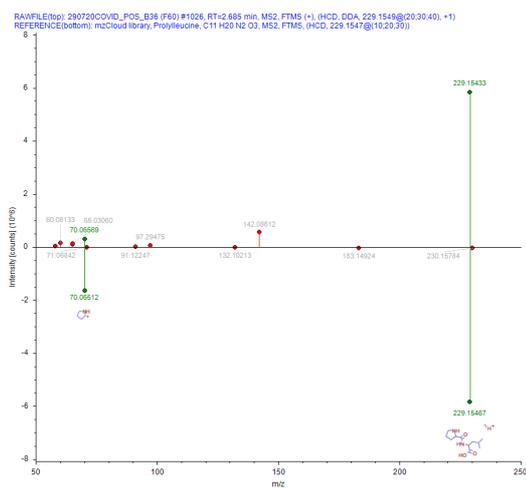
229.155 2.26  
142.085 2.0  
64.783 1.7  
70.065 1.5  
97.294 1.4  
124.075 1.3  
64.779 1.2  
60.081 1.1  
58.065 1.0  
170.081 0.9

Predicted LC-MS/MS Spectrum - 20V, Positive (HMDB0011174)



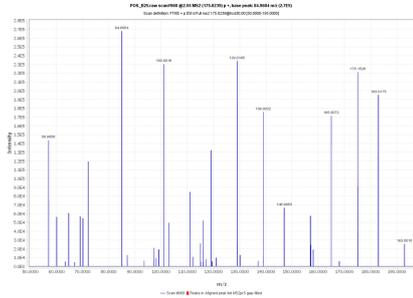
57.06987671 1.974451661  
69.06987671 23.98881387  
70.06512568 5.582992249  
86.0964258 24.35447209  
98.0600403 4.394366336  
114.0913404 2.919581151  
116.070605 3.393502732  
142.0498695 3.819780625  
183.1491896 4.458673705  
211.1441043 2.667689359  
229.154669 3.573182623

## HMDB0011174-POSsmz229.155mz2.6 Isoleucylproline



# HMDB0001264-NEGmz173.008rt2.58

## L-Dehydroascorbic acid

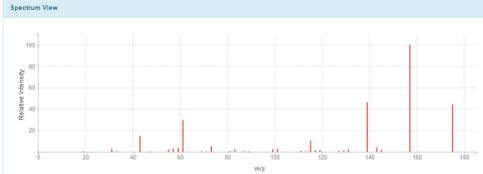


- 84.96 267
- 129.018 233
- 100.991 230
- 175.152 221
- 183.017 195
- 139.002 175
- 165.007 171
- 56.965 143
- 119.001 132
- 72.081 119

Predicted LC-MS/MS Spectrum - 20V, Positive (HMDB0001264)

**Spectrum Details**

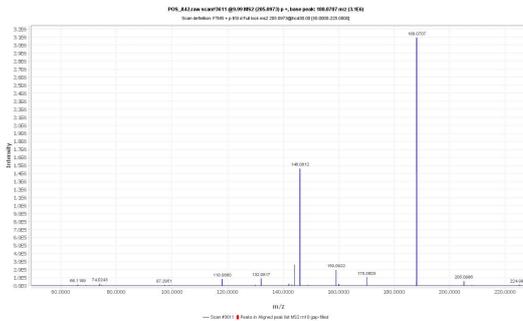
HMDB ID: HMDB0001264  
 Compound name: Dehydroascorbic acid  
 Spectrum type: Predicted LC-MS/MS Spectrum - 20V, Positive  
 Splash Key: splash10-0a6-290000000-524a15574c114929  
 Notes: This is a predicted spectrum and it should only be used as a guide, further evidence is required to confirm identification.



- 55.01838972 0.7787940315
- 56.99765427 1.102540023
- 59.01330434 1.23008822
- 61.0289544 10.23880935
- 68.99765427 0.2328747303
- 71.01330434 0.1394152505
- 73.0289544 1.796313195
- 80.99765427 0.27766387537
- 83.01330434 0.7752793353
- 87.00821896 0.2287000459
- 89.02386902 0.2036007188
- 99.00821896 0.8068907823
- 101.023869 0.995427972
- 111.008219 0.3192095557
- 113.023869 0.2463485281
- 115.0031336 3.539203352
- 117.0187836 0.5081702762
- 119.0344337 0.628379459
- 127.0031336 0.3525113333
- 129.0187836 0.4799402857
- 131.0344337 0.8721453689
- 139.0031336 15.90536086
- 142.9980482 1.46063754
- 145.0136983 0.6493595196
- 157.0136983 34.27149066
- 175.0242629 15.19863601

# HMDB0000929-POSmz205.097rt9.64

## L-Tryptophan

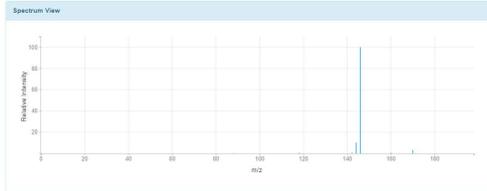


- 74.0243 2
- 97.2951 1
- 118.066 8
- 132.061 8
- 143.079 1
- 146.061 146
- 159.092 19
- 170.063 10
- 188.070 2 309

LC-MS/MS Spectrum - LC-ESI-ITFT (LTQ Orbitrap XL, Thermo Scientific), Positive (HMDB0000929)

**Spectrum Details**

HMDB ID: HMDB0000929  
 Compound name: L-Tryptophan  
 Spectrum type: LC-MS/MS Spectrum - LC-ESI-ITFT (LTQ Orbitrap XL, Thermo Scientific), Positive  
 Splash Key: splash10-0002-060000000-417e0b6e18e5bd1 View in MolNA C



**Experimental Conditions**

Instrument Type: LC-ESI-ITFT (LTQ Orbitrap XL, Thermo Scientific)  
 Ionization Mode: Positive

- 88.021561 386.503540 3
- 118.079071 1045.226196 8
- 132.163635 83.285217 1
- 142.171295 1017.944397 8
- 144.087158 13159.058594 103
- 146.004166 127765.289063 999
- 146.892166 79.571304 1
- 159.992203 785.907166 6
- 169.999130 4282.784668 33
- 173.174606 108.019554 1
- 188.026962 68.911499 1





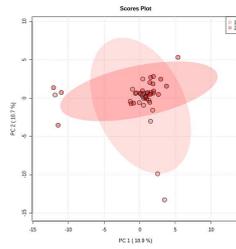


## data S4B

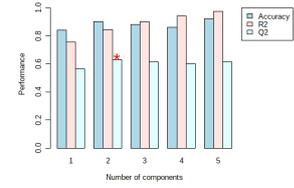
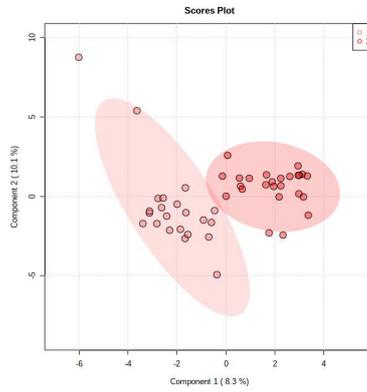
MetaboAnalyst statistical analysis

Moderate vs Severe  
 Sample normalization by median  
 Log transformation of the data  
 No data scaling

### PCA

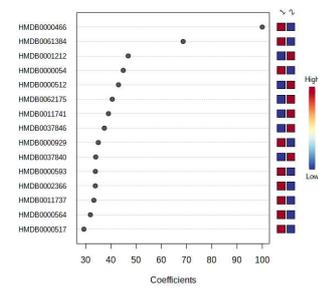


### PLS-DA



PLS-DA cross validation details:

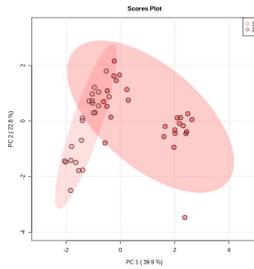
Measure	1 comps	2 comps	3 comps	4 comps	5 comps
Accuracy	0.84	0.9	0.88	0.86	0.92
R2	0.7539	0.84636	0.90034	0.94242	0.97226
Q2	0.56566	0.62846	0.61342	0.60199	0.61466



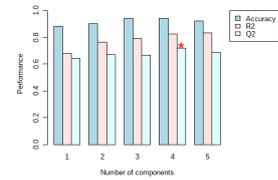
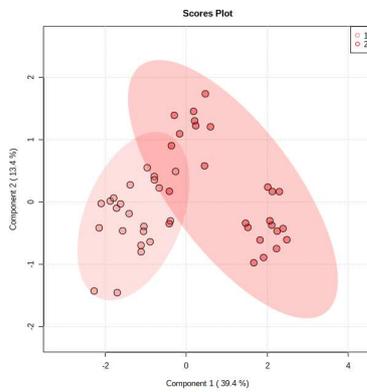
Top 15 metabolites only

Log transformation of the data  
 No data scaling

### PCA

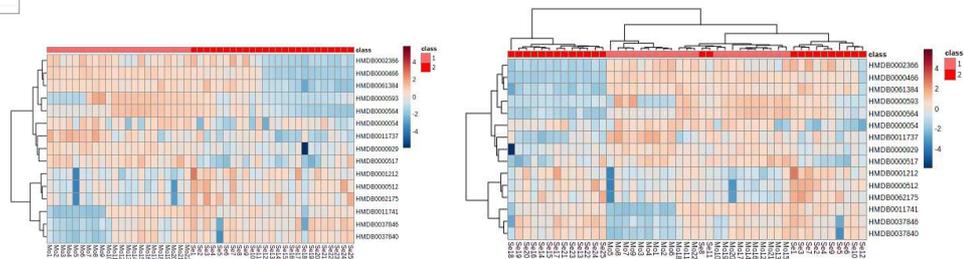


### PLS-DA



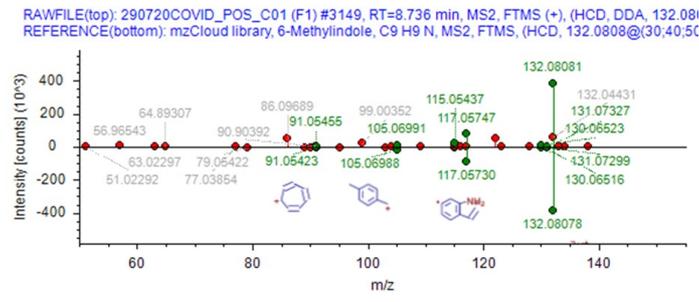
PLS-DA cross validation details:

Measure	1 comps	2 comps	3 comps	4 comps	5 comps
Accuracy	0.88	0.9	0.94	0.94	0.92
R2	0.67744	0.76268	0.79184	0.82604	0.83376
Q2	0.64264	0.67002	0.66618	0.71738	0.68723



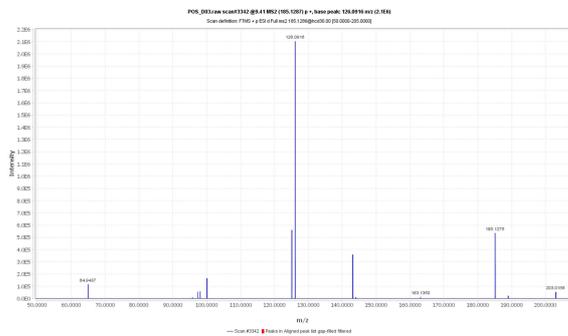
**data S3C**  
**MS2 spectra and matches of top features**

HMDB0000466-POSmz132.081rt8.49  
 3-Methylindole



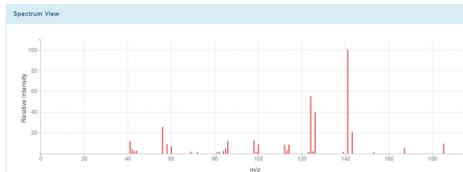
HMDB0061384-POSmz185.129rt9  
 N-(3-acetamidopropyl)pyrrolidin-2-one

Predicted LC-MS/MS Spectrum - 20V, Positive (HMDB0061384)



**Spectrum Details**

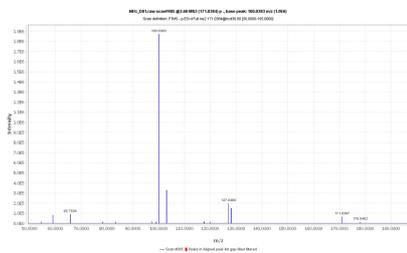
HMDB ID: HMDB0061384  
 Compound name: N-(3-acetamidopropyl)pyrrolidin-2-one  
 Spectrum type: Predicted LC-MS/MS Spectrum - 20V, Positive  
 Splash Key: splash10-0006-2900000000-5e9f92409d99f01d1f0e  
 Note: This is a predicted spectrum and it should only be used as a guide, further evidence is required to confirm identification.



56.05002419	6.898370438
58.06567426	2.420976219
60.04493881	1.848258931
69.03403978	0.5052863095
72.04493881	0.3871005466
81.04527316	0.3551390824
82.06567426	0.3976606191
84.04493881	0.6776481714
85.07657329	1.3083923179
86.06568888	3.238666363
98.06568888	3.377358138
99.09222335	0.51225926
100.0763389	2.439171754
112.0762389	2.339108982
113.1078734	0.8730230093
114.001880	2.323611894
123.0922234	0.458598054
124.0762389	14.83369747
125.0714879	0.411432862
125.1078734	0.6746253031
126.091889	10.70129979
139.087138	0.4418423997
141.102788	26.81337499
143.1184381	5.527297036
153.102788	0.3546220564
167.1184381	1.50597344
185.1290028	2.589464572

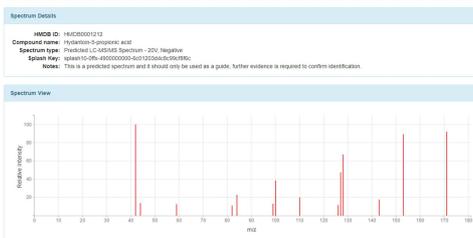
# HMDB0001212-NEGmz171.039rt2.39

## Hydantoin-5-propionic acid



- 100.038187
- 103.03732
- 127.04819
- 128.03215
- 65.75219
- 59.01118
- 171.03816
- 97.29212
- 115.83111
- 117.63811

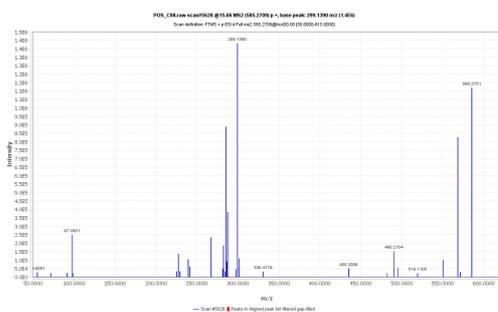
Predicted LC-MS/MS Spectrum - 20V, Negative (HMDB0001212)



- 59.01385292 1.819974028
- 82.02982733 1.611981911
- 84.04548730 3.302670239
- 99.02000092 1.886333118
- 100.040402 5.611316455
- 110.0247510 2.894353399
- 126.0196666 1.68715387
- 127.051301 6.921131283
- 128.0353166 9.77109072
- 143.0462157 2.566834977
- 153.0305656 13.01557598
- 171.0411303 13.41256178

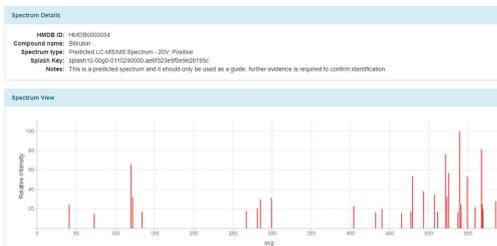
# HMDB0000054-POSmz585.271rt14.75

## Bilirubin



- 299.139138
- 585.275111
- 285.121188
- 568.237182
- 287.14138
- 97.29515
- 267.112123
- 282.11318
- 490.2115
- 227.11913

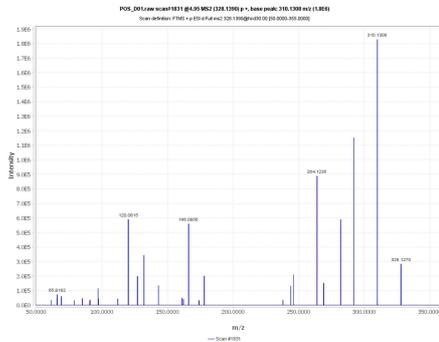
Predicted LC-MS/MS Spectrum - 20V, Positive (HMDB0000054)



- 41.03912516 2.327945209
- 73.0289544 1.421901312
- 120.0449388 6.332804827
- 122.0605889 3.082177032
- 134.0605889 1.636067137
- 267.1133527 1.681941678
- 281.1290028 1.944678077
- 285.1239174 2.8601024
- 299.1395675 3.002364181
- 404.1974167 2.191624709
- 432.192313 1.566453224
- 440.2185461 1.905533855
- 465.2290512 1.493834998
- 477.2290512 1.665604476
- 479.2447012 5.154289726
- 493.2603513 3.674131936
- 507.2396159 3.335040444
- 511.2345305 1.637988539
- 521.2552659 7.390015375
- 523.2709163 1.129430937
- 525.2501805 5.491383245
- 537.2501805 1.628432361
- 539.2658306 9.649654545
- 541.2450952 2.369693882
- 541.2814807 1.419233059
- 549.2501805 5.117863407
- 559.2556598 2.094052817
- 567.2607452 7.8820827
- 568.2447608 2.394338764
- 569.2400098 1.840951828
- 585.2713099 7.688113589



HMDB0037846-POS<sub>mz</sub>328.139<sub>rt</sub>4.79  
N-(1-Deoxy-1-fructosyl)phenylalanine



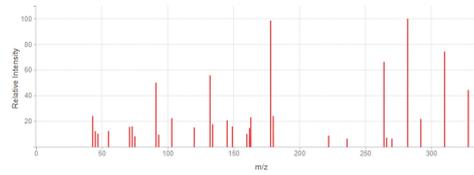
310.13 182  
292.119 115  
264.113 88  
120.081 59  
282.125 58  
166.085 56  
132.08 34  
328.126 28  
246.111 21  
178.086 20

Predicted LC-MS/MS Spectrum - 20V, Positive (HMDB0037846)

Spectrum Details

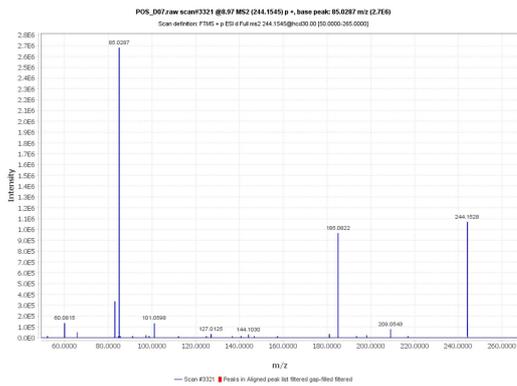
HMDB ID: HMDB0037846  
Compound: N-(1-Deoxy-1-fructosyl)phenylalanine  
name:  
Spectrum type: Predicted LC-MS/MS Spectrum - 20V, Positive  
Splash Key: splash10-01sa-495300000-18a2037532e5b4c4429  
Notes: This is a predicted spectrum and it should only be used as a guide, further evidence is required to confirm identification.

Spectrum View



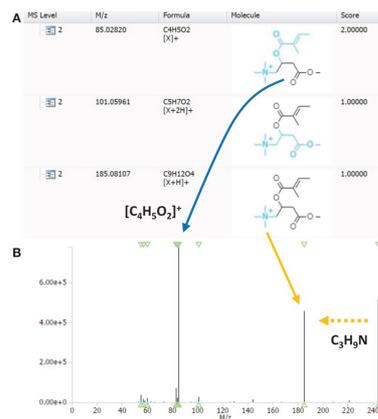
55.01838972 1.50024729  
71.01330434 1.892259885  
73.0289544 1.93572651  
75.04460446 0.9977764487  
91.05477522 6.029727629  
93.07042529 1.149824094  
103.0547752 2.707829167  
120.0813243 1.838858277  
132.0813243 6.731759554  
134.0989744 2.135085195  
145.050838 2.496701671  
149.0802545 1.922319228  
160.0762389 1.218741042  
162.0765329 1.78382153  
162.091889 0.890149337  
163.0606485 2.778900709  
178.0868036 11.85830979  
180.0871975 2.899724219  
180.1024537 1.135229133  
222.1120184 1.068074901  
236.0770268 0.781950618  
264.125831 7.982967531  
266.1392331 0.866797833  
270.1341477 0.7799847623  
282.1341477 12.03771295  
292.1184977 2.63763461  
310.1290624 8.951266829  
328.139627 5.346718615

HMDB0002366-POS<sub>mz</sub>244.154<sub>rt</sub>9.99  
Tiglylcarnitine



85.028 268  
244.152 107  
183.082 96  
83.05 33  
60.081 13  
101.059 13  
209.054 7  
65.79 4  
181.06 3  
127.012 3

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4353373/figure/F2/>





## data S6

MetaboAnalyst pathway analysis

CTRL vs Mild

Sample normalization by median

Log transformation of the data

No data scaling

Pathway Name	Match Status	p	-log(p)	Holm p	FDR	Impact	Details
<a href="#">Arachidonic acid metabolism</a>	2 over 36	3.25E-48	47.488	1.66E-46	1.50E-46	0	KEGG SMP
<a href="#">Linoleic acid metabolism</a>	3 over 5	7.53E-48	47.123	3.77E-46	1.50E-46	1	KEGG SMP
<a href="#">alpha-Linolenic acid metabolism</a>	2 over 13	8.84E-48	47.053	4.33E-46	1.50E-46	0.33333	KEGG SMP
<a href="#">Glycerophospholipid metabolism</a>	6 over 36	1.65E-42	41.781	7.94E-41	2.11E-41	0.31146	KEGG
<a href="#">Sphingolipid metabolism</a>	4 over 21	2.51E-42	41.6	1.18E-40	2.56E-41	0.29818	KEGG SMP
<a href="#">Phenylalanine, tyrosine and tryptophan biosynthesis</a>	3 over 4	2.94E-28	27.532	1.35E-26	2.50E-27	1	KEGG SMP
<a href="#">Aminoacyl-tRNA biosynthesis</a>	12 over 48	1.64E-25	24.785	7.38E-24	1.19E-24	0	KEGG
<a href="#">Phenylalanine metabolism</a>	4 over 10	2.05E-16	15.688	9.02E-15	1.31E-15	0.61904	KEGG SMP
<a href="#">Pentose phosphate pathway</a>	1 over 22	6.47E-10	9.1888	2.78E-08	3.67E-09	0.00262	KEGG SMP
<a href="#">Lysine degradation</a>	5 over 25	6.41E-09	8.193	2.69E-07	3.27E-08	0.2817	KEGG SMP
<a href="#">Biotin metabolism</a>	2 over 10	1.37E-07	6.8636	5.61E-06	6.35E-07	0.2	KEGG SMP
<a href="#">Vitamin B6 metabolism</a>	2 over 9	8.72E-07	6.0596	3.49E-05	3.71E-06	0.07843	KEGG SMP
<a href="#">Citrate cycle (TCA cycle)</a>	3 over 20	2.13E-06	5.6707	8.33E-05	8.37E-06	0.10859	KEGG SMP
<a href="#">Tryptophan metabolism</a>	9 over 41	5.74E-05	4.2408	0.0021824	2.09E-04	0.31659	KEGG SMP
<a href="#">Drug metabolism - cytochrome P450</a>	1 over 55	2.86E-04	3.5439	0.010577	9.11E-04	0	KEGG
<a href="#">Arginine biosynthesis</a>	7 over 14	2.86E-04	3.5437	0.010577	9.11E-04	0.48223	KEGG
<a href="#">Valine, leucine and isoleucine degradation</a>	2 over 40	3.94E-04	3.4046	0.013786	0.0011192	0.01084	KEGG SMP
<a href="#">Ubiquinone and other terpenoid-quinone biosynthesis</a>	1 over 9	3.95E-04	3.4034	0.013786	0.0011192	0	KEGG SMP
<a href="#">Galactose metabolism</a>	1 over 27	4.69E-04	3.3286	0.015487	0.0012597	0	KEGG SMP
<a href="#">Purine metabolism</a>	7 over 65	7.25E-04	3.14	0.023184	0.0018475	0.14876	KEGG SMP
<a href="#">Cytosine and methionine metabolism</a>	4 over 33	9.40E-04	3.027	0.029132	0.0022822	0.26448	KEGG SMP SMP
<a href="#">Nicotinate and nicotinamide metabolism</a>	3 over 15	0.0017192	2.7647	0.051576	0.003911	0.38509	KEGG SMP
<a href="#">Glycosylphosphatidylinositol ( GPI)-anchor biosynthesis</a>	1 over 14	0.0017638	2.7536	0.051576	0.003911	0.00399	KEGG
<a href="#">Histidine metabolism</a>	6 over 16	0.0042739	2.3692	0.11967	0.0090821	0.62294	KEGG SMP
<a href="#">Tryptophan metabolism</a>	9 over 42	0.0086924	2.0609	0.23469	0.017732	0.30158	KEGG SMP SMP
<a href="#">Steroid biosynthesis</a>	4 over 42	0.01197	1.9219	0.31121	0.023479	0.0392	KEGG
<a href="#">Valine, leucine and isoleucine biosynthesis</a>	3 over 8	0.016168	1.7913	0.40421	0.03054	0	KEGG SMP
<a href="#">Arginine and proline metabolism</a>	7 over 38	0.022965	1.6389	0.55115	0.041828	0.42886	KEGG SMP
<a href="#">Ether lipid metabolism</a>	1 over 20	0.031396	1.5031	0.7221	0.055213	0	KEGG
<a href="#">Glycerate and dicarboxylate metabolism</a>	3 over 32	0.042091	1.3758	0.926	0.071555	0.02381	KEGG
<a href="#">beta-Alanine metabolism</a>	5 over 21	0.044314	1.3535	0.9306	0.072904	0.60821	KEGG SMP

## data S7

MetaboAnalyst pathway analysis

Mild vs Moderate&Severe

Sample normalization by median

Log transformation of the data

No data scaling

Pathway Name	Match Status	p	-log(p)	Holm p	FDR	Impact	Details
<a href="#">beta-Alanine metabolism</a>	5 over 21	1.80E-33	32.744	9.19E-32	9.19E-32	0.60821	KEGG SMP
<a href="#">Aminoacyl-tRNA biosynthesis</a>	12 over 48	1.67E-26	25.778	8.33E-25	4.25E-25	0	KEGG
<a href="#">Arginine biosynthesis</a>	7 over 14	1.85E-25	24.733	9.06E-24	3.14E-24	0.48223	KEGG
<a href="#">Phenylalanine metabolism</a>	4 over 10	3.25E-24	23.489	1.56E-22	4.14E-23	0.61904	KEGG SMP
<a href="#">Valine, leucine and isoleucine biosynthesis</a>	3 over 8	9.55E-24	23.02	4.49E-22	9.74E-23	0	KEGG
<a href="#">D-Glutamine and D-glutamate metabolism</a>	3 over 6	3.82E-21	20.418	1.76E-19	2.52E-20	0.5	KEGG SMP
<a href="#">Alanine, aspartate and glutamate metabolism</a>	3 over 28	3.82E-21	20.418	1.76E-19	2.52E-20	0.35898	KEGG SMP SMP SMP
<a href="#">Glycerate and dicarboxylate metabolism</a>	2 over 32	4.44E-21	20.353	1.95E-19	2.52E-20	0	KEGG
<a href="#">Nitrogen metabolism</a>	2 over 6	4.44E-21	20.353	1.95E-19	2.52E-20	0	KEGG SMP
<a href="#">Tryptophan metabolism</a>	9 over 41	2.83E-19	18.549	1.19E-17	1.44E-18	0.31659	KEGG SMP
<a href="#">Valine metabolism</a>	11 over 42	4.16E-17	16.381	1.71E-15	1.93E-16	0.403	KEGG SMP SMP
<a href="#">Valine, leucine and isoleucine degradation</a>	2 over 40	1.68E-16	15.775	6.71E-15	7.13E-16	0.01084	KEGG SMP
<a href="#">Glycerophospholipid metabolism</a>	5 over 36	2.24E-16	15.65	8.73E-15	8.79E-16	0.26445	KEGG
<a href="#">Vitamin B6 metabolism</a>	1 over 9	5.36E-16	15.271	2.04E-14	1.95E-15	0	KEGG SMP
<a href="#">Arachidonic acid metabolism</a>	2 over 36	4.03E-14	13.395	1.49E-12	1.37E-13	0	KEGG SMP
<a href="#">alpha-Linolenic acid metabolism</a>	2 over 13	1.36E-13	12.866	4.90E-12	4.34E-13	0.33333	KEGG SMP
<a href="#">Linoleic acid metabolism</a>	3 over 5	1.81E-13	12.743	6.33E-12	5.43E-13	1	KEGG SMP
<a href="#">Lysine degradation</a>	6 over 25	9.36E-13	12.029	3.18E-11	2.65E-12	0.2817	KEGG SMP
<a href="#">Phenylalanine, tyrosine and tryptophan biosynthesis</a>	3 over 4	1.52E-12	11.819	5.01E-11	4.08E-12	1	KEGG SMP
<a href="#">Arginine and proline metabolism</a>	6 over 38	2.39E-11	10.621	7.66E-10	6.10E-11	0.4054	KEGG SMP
<a href="#">Glycine, serine and threonine metabolism</a>	5 over 33	5.06E-11	10.296	1.57E-09	1.23E-10	0.14714	KEGG SMP
<a href="#">Glutathione metabolism</a>	3 over 28	5.08E-11	10.246	1.70E-09	1.32E-10	0.02679	KEGG SMP
<a href="#">Pantothenate and CoA biosynthesis</a>	4 over 19	1.04E-09	8.9824	3.02E-08	2.31E-09	0.04286	KEGG SMP
<a href="#">Biotin metabolism</a>	2 over 10	4.04E-09	8.3941	1.13E-07	8.58E-09	0.2	KEGG SMP
<a href="#">Ether lipid metabolism</a>	1 over 20	4.33E-09	8.3634	1.17E-07	8.83E-09	0	KEGG
<a href="#">Propanoate metabolism</a>	1 over 23	3.02E-08	7.5202	7.85E-07	5.92E-08	0	KEGG SMP
<a href="#">Purine metabolism</a>	6 over 65	3.52E-08	7.4537	8.79E-07	6.64E-08	0.13659	KEGG SMP
<a href="#">Cysteine and methionine metabolism</a>	4 over 33	7.46E-07	6.127	1.79E-05	1.36E-06	0.26448	KEGG SMP SMP
<a href="#">Galactose metabolism</a>	1 over 27	1.00E-06	5.9985	2.31E-05	1.76E-06	0	KEGG SMP
<a href="#">Porphyrin and chlorophyll metabolism</a>	5 over 30	1.38E-06	5.8593	3.04E-05	2.35E-06	0.12753	KEGG
<a href="#">Ubiquinone and other terpenoid-quinone biosynthesis</a>	1 over 9	2.20E-06	5.658	4.62E-05	3.62E-06	0	KEGG SMP
<a href="#">Steroid biosynthesis</a>	2 over 42	4.98E-06	5.3028	9.96E-05	7.94E-06	0.0392	KEGG
<a href="#">Pyrimidine metabolism</a>	6 over 39	6.41E-06	5.1932	1.22E-04	9.91E-06	0.15301	KEGG SMP
<a href="#">Sphingolipid metabolism</a>	5 over 21	3.09E-05	4.5103	5.56E-04	4.63E-05	0.07302	KEGG SMP
<a href="#">Steroid hormone biosynthesis</a>	4 over 85	6.21E-05	4.2069	0.0010557	9.05E-05	0.07915	KEGG SMP
<a href="#">Histidine metabolism</a>	6 over 16	6.85E-05	4.1646	0.0010952	9.70E-05	0.62294	KEGG SMP
<a href="#">Nicotinate and nicotinamide metabolism</a>	3 over 15	0.0019076	2.7195	0.028614	0.0026294	0.38509	KEGG SMP
<a href="#">Glycosylphosphatidylinositol (GPI)-anchor biosynthesis</a>	1 over 14	0.0067439	2.1711	0.094415	0.0090511	0.00399	KEGG
<a href="#">Citrate cycle (TCA cycle)</a>	1 over 20	0.010614	1.9741	0.13798	0.01388	0.08586	KEGG SMP
<a href="#">Drug metabolism - other enzymes</a>	1 over 39	0.016958	1.7706	0.2035	0.021622	0.03804	KEGG
<a href="#">Drug metabolism - cytochrome P450</a>	1 over 55	0.028217	1.5495	0.31039	0.0351	0	KEGG
<a href="#">Selenocompound metabolism</a>	1 over 20	0.04082	1.3891	0.4082	0.049567	0.15809	KEGG SMP

## data S8

MetaboAnalyst pathway analysis

Moderate vs Severe

Sample normalization by median

Log transformation of the data

No data scaling

Pathway Name	Match Status	p	-log(p)	Holm p	FDR	Impact	Details
<a href="#">Arginine biosynthesis</a>	6 over 14	1.38E-07	6.8602	6.62E-06	6.62E-06	0.48223	<a href="#">KEGG</a>
<a href="#">Aminoacyl-tRNA biosynthesis</a>	12 over 48	1.25E-06	5.903	5.88E-05	3.00E-05	0	<a href="#">KEGG</a>
<a href="#">beta-Alanine metabolism</a>	4 over 21	2.56E-05	4.5921	0.0011767	3.89E-04	0.55224	<a href="#">KEGG SMP</a>
<a href="#">Arginine and proline metabolism</a>	6 over 38	3.24E-05	4.4889	0.0014598	3.89E-04	0.4054	<a href="#">KEGG SMP</a>
<a href="#">Valine, leucine and isoleucine biosynthesis</a>	4 over 8	8.62E-05	4.0646	0.0037921	8.27E-04	0	<a href="#">KEGG SMP</a>
<a href="#">Arachidonic acid metabolism</a>	2 over 36	7.29E-04	3.1374	0.031334	0.0058296	0	<a href="#">KEGG SMP</a>
<a href="#">Glycine, serine and threonine metabolism</a>	6 over 33	8.51E-04	3.0699	0.035758	0.0055381	0.14714	<a href="#">KEGG SMP</a>
<a href="#">alpha-Linolenic acid metabolism</a>	2 over 13	0.0010929	2.9614	0.044808	0.0065572	0.33333	<a href="#">KEGG SMP</a>
<a href="#">Vitamin B6 metabolism</a>	2 over 9	0.0012417	2.906	0.049667	0.0066222	0.07843	<a href="#">KEGG SMP</a>
<a href="#">Lysine degradation</a>	6 over 25	0.0023467	2.6295	0.09152	0.01019	0.2817	<a href="#">KEGG SMP</a>
<a href="#">Tyrosine metabolism</a>	12 over 42	0.0024199	2.6162	0.091956	0.01019	0.47287	<a href="#">KEGG SMP SMP</a>
<a href="#">Tryptophan metabolism</a>	9 over 41	0.0026368	2.5789	0.097561	0.01019	0.31659	<a href="#">KEGG SMP</a>
<a href="#">Valine, leucine and isoleucine degradation</a>	2 over 40	0.0027598	2.5591	0.099353	0.01019	0.01084	<a href="#">KEGG SMP</a>
<a href="#">Linoleic acid metabolism</a>	3 over 5	0.0046781	2.3299	0.16373	0.016039	1	<a href="#">KEGG SMP</a>
<a href="#">Glycerophospholipid metabolism</a>	5 over 36	0.0087429	2.0583	0.29726	0.027977	0.26445	<a href="#">KEGG</a>
<a href="#">Porphyrin and chlorophyll metabolism</a>	5 over 30	0.025193	1.5987	0.83136	0.075578	0.12753	<a href="#">KEGG</a>
<a href="#">Alanine, aspartate and glutamate metabolism</a>	3 over 28	0.033799	1.4711	1	0.080198	0.39744	<a href="#">KEGG SMP SMP 5 MP</a>
<a href="#">Cysteine and methionine metabolism</a>	4 over 33	0.035008	1.4558	1	0.080198	0.18613	<a href="#">KEGG SMP SMP</a>
<a href="#">Nicotinate and nicotinamide metabolism</a>	3 over 15	0.03575	1.4467	1	0.080198	0.38509	<a href="#">KEGG SMP</a>
<a href="#">Histidine metabolism</a>	5 over 16	0.036208	1.4412	1	0.080198	0.53278	<a href="#">KEGG SMP</a>
<a href="#">Purine metabolism</a>	7 over 65	0.036721	1.4351	1	0.080198	0.14876	<a href="#">KEGG SMP</a>
<a href="#">Biotin metabolism</a>	1 over 10	0.036757	1.4347	1	0.080198	0	<a href="#">KEGG SMP</a>

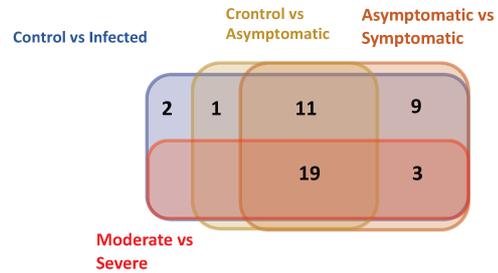
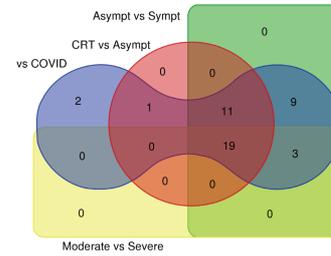
## data S9

Venn diagrams

(VIB, Ghent University, Belgium)

4 comparisons

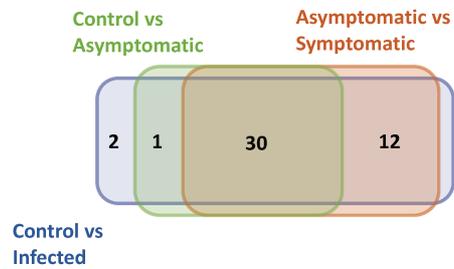
Names	total	elements
Asympt vs Sympt CRT vs Asympt Asympt CRT vs COVID Moderate vs Severe	19	Vitamin B6 metabolism; Biotin metabolism; Nicotinate and nicotinamide metabolism; Arachidonic acid metabolism; Aminoacyl-tRNA biosynthesis; Valine, leucine and isoleucine biosynthesis; Glycerophospholipid metabolism; Purine metabolism; Tyrosine metabolism; alpha-Linolenic acid metabolism; Arginine biosynthesis; Lysine degradation; Histidine metabolism; Valine, leucine and isoleucine degradation; Tryptophan metabolism; Linoleic acid metabolism; beta-Alanine metabolism; Cysteine and methionine metabolism; Arginine and proline metabolism
Asympt vs Sympt CRT vs Asympt Asympt CRT vs COVID	11	Phenylalanine, tyrosine and tryptophan biosynthesis; Glyoxylate and dicarboxylate metabolism; Citrate cycle (TCA cycle); Drug metabolism - cytochrome P450; Phenylalanine metabolism; Ether lipid metabolism; Galactose metabolism; Glycosylphosphatidylinositol (GPI)-anchor biosynthesis; Ubiquinone and other terpenoid-quinone biosynthesis; Sphingolipid metabolism; Steroid biosynthesis
Asympt vs Sympt CRT vs COVID Moderate vs Severe	3	Alanine, aspartate and glutamate metabolism; Porphyrin and chlorophyll metabolism; Glycine, serine and threonine metabolism
CRT vs Asympt CRT vs COVID	1	Pentose phosphate pathway
Asympt vs Sympt CRT vs COVID	9	Pantothenate and CoA biosynthesis; Drug metabolism - other enzymes; Selenocompound metabolism; Pyrimidine metabolism; D-Glutamine and D-glutamate metabolism; Nitrogen metabolism; Glutathione metabolism; Propanoate metabolism; Steroid hormone biosynthesis
CRT vs COVID	2	Pyruvate metabolism; Terpenoid backbone biosynthesis



## data S10

Venn diagrams  
(VIB, Ghent University, Belgium)  
3 comparisons

Names	total	elements
Asymptomatic vs Symptomatic Control vs Asymptomatic Control vs COVID	30	Phenylalanine, tyrosine and tryptophan biosynthesis Vitamin B6 metabolism Biotin metabolism Glyoxylate and dicarboxylate metabolism Nicotinate and nicotinamide metabolism Citrate cycle (TCA cycle) Arachidonic acid metabolism Drug metabolism - cytochrome P450 Aminoacyl-tRNA biosynthesis Valine, leucine and isoleucine biosynthesis Phenylalanine metabolism Glycerophospholipid metabolism Ether lipid metabolism Purine metabolism Galactose metabolism Tyrosine metabolism Glycosylphosphatidylinositol (GPI)-anchor biosynthesis alpha-Linolenic acid metabolism Arginine biosynthesis Ubiquinone and other terpenoid-quinone biosynthesis Lysine degradation Sphingolipid metabolism Histidine metabolism Valine, leucine and isoleucine degradation Tryptophan metabolism Linoleic acid metabolism beta-Alanine metabolism Cysteine and methionine metabolism Steroid biosynthesis Arginine and proline metabolism
Control vs Asymptomatic Control vs COVID	1	Pentose phosphate pathway
Asymptomatic vs Symptomatic Control vs COVID	12	Pantothenate and CoA biosynthesis Alanine, aspartate and glutamate metabolism Drug metabolism - other enzymes Porphyrin and chlorophyll metabolism Selenocompound metabolism Pyrimidine metabolism D-Glutamine and D-glutamate metabolism Nitrogen metabolism Glutathione metabolism Propanoate metabolism Steroid hormone biosynthesis Glycine, serine and threonine metabolism
Control vs COVID	2	Pyruvate metabolism Terpenoid backbone biosynthesis



### Chapitre 3 : Identification d'une signature métabolomique urinaire diagnostic et pronostic de la COVID-19

La cohorte SignCov a été constituée par des échantillons d'urine de 152 patients admis dans le service des urgences du CHU de Nice pour suspicion d'infection à SARS COV 2 ou hospitalisés en service de Médecine Intensive et Réanimation pour COVID-19, et de 11 volontaires sains. Les patients ont été considérés comme atteints de la COVID-19 lorsqu'ils présentaient des symptômes compatibles avec la définition de l'OMS et qu'un test RT-PCR réalisé à l'hôpital était positif pour le virus SARS-CoV-2. Tous les échantillons d'urine ont été prélevés lors de l'admission à l'hôpital.

Afin d'évaluer si l'approche métabolomique sur les échantillons d'urine peut discriminer les sujets en termes de diagnostic et/ou de pronostic, les patients ont été séparés en 4 groupes (Figure 24) :

- Groupe « Non Covid » (n=35) : patients symptomatiques dont la RT-PCR était négative pour le SARS-CoV-2 et pour les autres virus identifiables
- Groupe « légèrement atteint » (n=29) : patients présentant des symptômes légers de COVID-19 qui n'ont pas nécessité d'hospitalisation après leur admission aux urgences
- Groupe « modérément atteint » (n=36) : patients présentant des symptômes modérés nécessitant une hospitalisation en dehors d'une unité de soins critiques
- Groupe « sévèrement atteint » (n=52) : patients présentant une forme sévère de COVID-19 nécessitant une hospitalisation dans une unité de soins intensifs ou de réanimation.

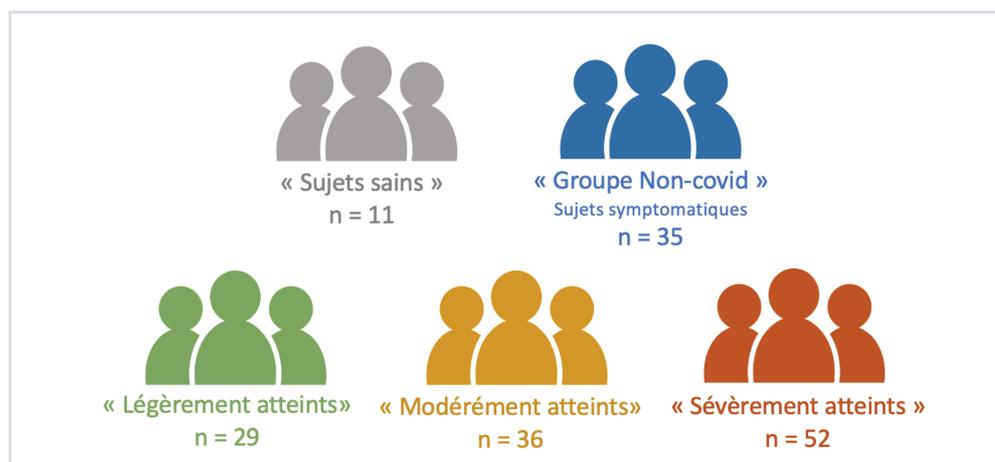


Figure 24. Description des groupes de patients et volontaires sains dans la cohorte SignCov

## 1) Préparation des échantillons et analyse LC-MS

Un volume de 50 µl d'urine a été mélangé à 50 µl de H<sub>2</sub>O (qualité HPLC, Merck Millipore, États-Unis) et 400 µl de méthanol (qualité HPLC, Merck Millipore, États-Unis) ont été ajoutés. Les échantillons ont été vortexés et incubés pendant une nuit à -20°C pour la précipitation des protéines. Après centrifugation (13.000 x g, 15 min, 4°C), le surnageant a été retiré, séché à l'aide d'un concentrateur SpeedVAC (SVC100H, SAVANT, Thermo Fisher Scientific, Illkirch, France), remis en suspension dans 50 µL de mélange acétonitrile-H<sub>2</sub>O 20:80 (qualité HPLC, Merck Millipore) et stocké à -20°C jusqu'à son utilisation pour les analyses métabolomiques.

Tout comme dans notre précédent travail, les analyses chromatographiques ont été effectuées avec le système HPLC Ultimate 3000 de DIONEX couplé à une colonne chromatographique (Phenomenex Synergi 4 u Hydro-RP80A 250 3,0 mm) réglée à 40 °C et à un débit de 0,9 ml/min. Les gradients des phases mobiles (phase mobile A : 0,1 % d'acide formique dans l'eau et phase mobile B : 0,1 % d'acide formique dans l'acétonitrile) ont été effectués sur une durée totale de 25 min. Les analyses MS ont été effectuées sur un spectromètre de masse Thermo Fisher Scientific Exactive Plus Benchtop Orbitrap. La source d'ionisation électrospray chauffée (HESI II) a été utilisée en mode ionique positif et négatif. L'instrument a été utilisé en mode de balayage complet de m/z 60 à m/z.

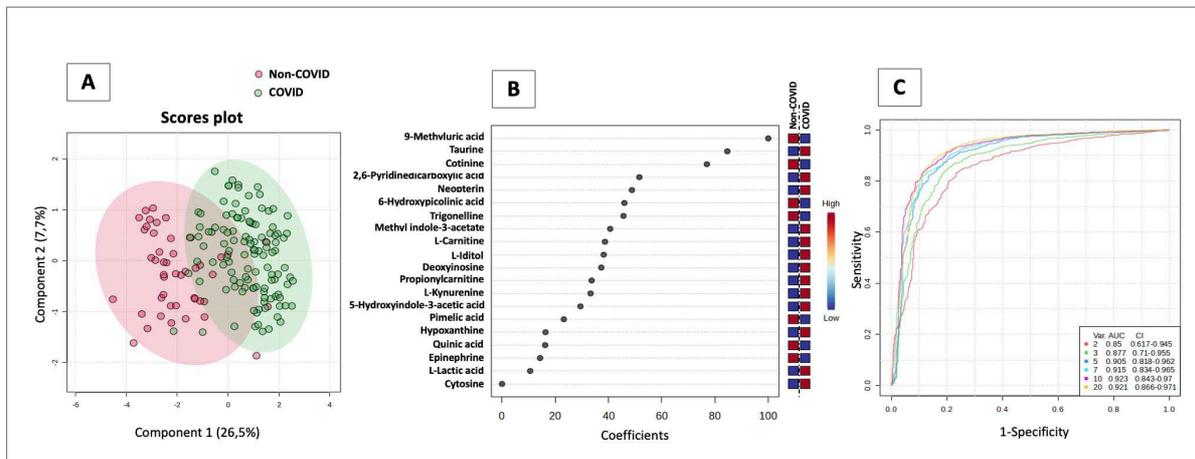
Le post-traitement des données a été effectué selon le même procédé que dans le travail précédent. Afin de faciliter les analyses, une étape préliminaire de sélection des métabolites endogènes a été réalisée à l'aide de la base de données du logiciel Compound Discoverer™.

## 2) Présentation des principaux résultats

### ▪ Signature métabolomique diagnostique

Notre première analyse s'est intéressée à la comparaison des groupes non-COVID, comprenant les sujets sains et les sujets symptomatiques non-COVID, au groupe de patients atteints de la COVID-19 (légèrement, modérément et sévèrement atteints). Après post traitement des données brutes issues de la LC-MS, nous avons sélectionnés 250 métabolites.

Cette comparaison nous a permis d'identifier une top liste de 20 métabolites. Il s'agissait du modèle diagnostique le plus performant avec une AUC à 0,92 (IC 95%, 0,87 – 0,97) (Figure 25). Dans cette signature, nous constatons, une augmentation de la *taurine*, de la *néoptérine* (métabolite produit dans les macrophages à la suite d'une stimulation par l'interferon  $\gamma$ ), de la *kynurenine*, de la *L-carnitine* et du *propionylcarnitine*, de la base purique *hypoxanthine*, et de la *cytosine*. Nous constatons également une baisse du *9-acide méthylurique* et de l'*épinephrine*.



### Figure 25. Identification d'une signature métabolomique urinaire diagnostique de la COVID-19

Les points rouges représentent les échantillons urinaires du groupe de patients non-COVID. Les points verts représentent les échantillons urinaires de tous les patients atteints de la COVID-19.

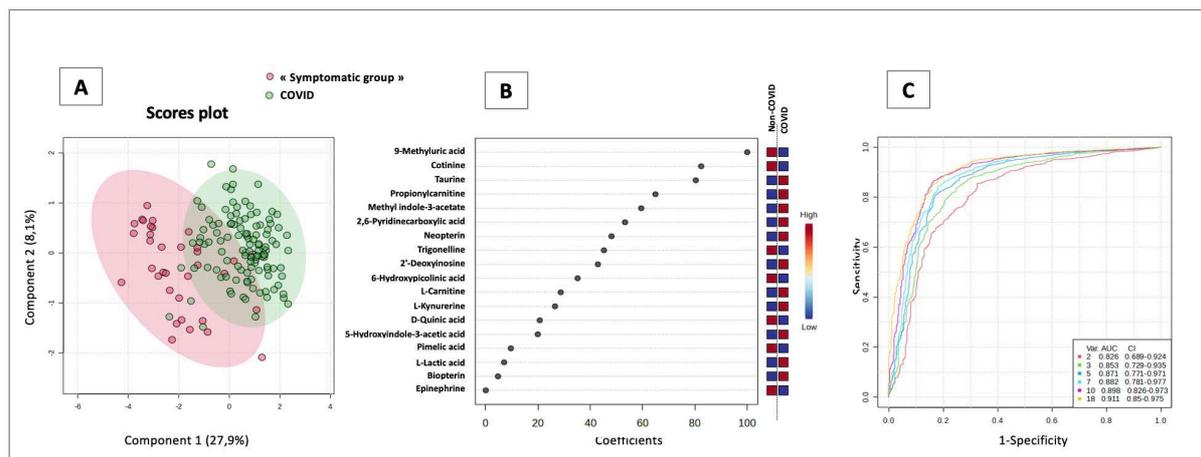
A : Analyse PLS-DA. Le diagramme de score est représenté par une ellipse de confiance de 95 %

B : Diagramme représentant la top liste de 20 métabolites sélectionnées sur la moyenne des cinq premières composantes du modèle PLS-DA. Les métabolites identifiés ont été validés par des analyses MS2 et sont représentés sous forme de noms chimiques.

C : Comparaisons des aires sous les courbes ROC multivariées de chaque top liste de métabolite (comprenant respectivement 20, 15, 10, 7, 5, 3 et 2 métabolites)

Afin de mieux évaluer les pouvoirs discriminants d'une signature métabolomique urinaire, nous avons souhaité comparer le métabolome du groupe « symptomatique non-COVID » au groupe de patients atteints de la COVID-19. Après post-traitement des données, nous avons obtenu 251 métabolites.

Nous nous sommes, une nouvelle fois, intéressé à une top liste de 20 métabolites. Dans cette comparaison, c'est un modèle diagnostique à 18 métabolites qui possédait les meilleures performances avec une AUC à 0,91 (IC 95%, 0,85 – 0,98) (Figure 26). Cette top liste, très proche de la précédente, confirme l'augmentation de la *taurine*, de la *néoptérine*, du *propionylcartinine* et de la *L-cartinite*, de l'*acide lactique*, et de la *kynurenine*. Il apparait également une augmentation de la *bioptérine* (cofacteur dans la synthèse du NO et métabolite entrant en jeu dans la conversion de la phénylalanine en tyrosine et dans la conversion du tryptophane en 5-Hydroxytryptophane). La baisse du *9-acide methylurique* et de l'*épinephrine* est aussi confirmée dans cette signature.



**Figure 26. Comparaison du métabolome urinaire des sujets symptomatiques non-COVID avec celui des patients atteints de la COVID-19**

Les points rouges représentent les échantillons urinaires des patients symptomatiques non-COVID. Les points verts représentent les échantillons urinaires de tous les patients atteints de la COVID-19.

A : Analyse PLS-DA. Le diagramme de score est représenté par une ellipse de confiance de 95 %

B : Diagramme représentant la top liste de 18 métabolites sélectionnées sur la moyenne des cinq premières composantes du modèle PLS-DA. Les métabolites identifiés ont été validés par des analyses MS2 et sont représentés sous forme de noms chimiques.

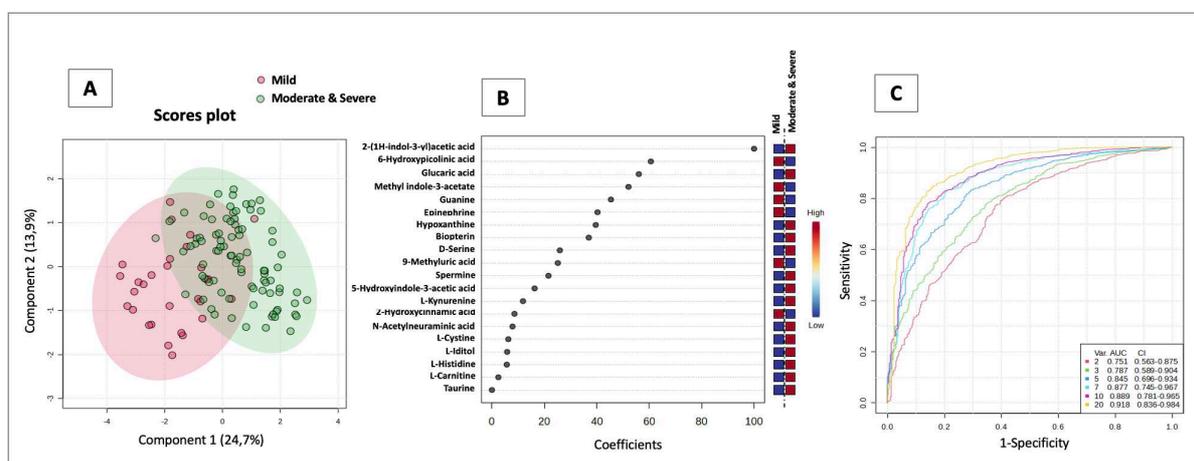
C : Comparaisons des aires sous les courbes ROC multivariées de chaque top liste de métabolite (comprenant respectivement 18, 10, 7, 5, 3 et 2 métabolites)

- **Signature métabolomique pronostique**

Nous avons souhaité par la suite chercher à identifier si une signature métabolomique urinaire pourrait permettre de prédire le risque d'hospitalisation des patients atteints de la COVID-19. Pour ce faire, nous avons comparé, parmi les patients infectés par le SARS-CoV-2, le groupe légèrement atteint au groupe de patients hospitalisés, que ce soit en service d'hospitalisation traditionnelle ou en service de soins critiques (modérément et sévèrement atteints). Lors de cette comparaison, un total de 251 métabolites a été analysé.

Dans cette comparaison, c'est une top liste de 20 métabolites qui présentait les performances pronostiques les plus intéressantes avec une AUC à 0,92 (IC 95%, 0,84 – 0,98) (Figure 27).

Dans cette signature, nous constatons dans le groupe de patient hospitalisés, une augmentation de certains acides aminés tels que la *serine*, la *cystine*, la *taurine*, alors que nous constatons une baisse de la *guanine*. Nous retrouvons également l'augmentation de l'*hypoxanthine*, de la *kynurenine* et de la *cartinine*. Il apparait également augmentation de la *spermine*. La baisse de l'*épinephrine* est confirmée et se situe plus haut dans cette top liste pronostique.



**Figure 27. Identification d'une signature métabolomique urinaire pronostique du risque d'hospitalisation**

Les points rouges représentent les échantillons urinaires du groupe de patients légèrement atteints par la COVID-19. Les points verts représentent les échantillons urinaires de tous les patients hospitalisés (modérément et sévèrement atteints)

A : Analyse PLS-DA. Le diagramme de score est représenté par une ellipse de confiance de 95 %

B : Diagramme représentant la top liste de 20 métabolites sélectionnées sur la moyenne des cinq premières composantes du modèle PLS-DA. Les métabolites identifiés ont été validés par des analyses MS2 et sont représentés sous forme de noms chimiques.

C : Comparaisons des aires sous les courbes ROC multivariées de chaque top liste de métabolite (comprenant respectivement 20, 10, 7, 5, 3 et 2 métabolites)

Les analyses statistiques MetaboAnalyst® (*annexes 1,3,5*) ainsi que les validations de métabolites des top listes pour chaque comparaison (*annexes 2,4,6*) sont présentées en annexes.

## Chapitre 4 : Discussion - Place de la métabolomique dans la prise en charge et la compréhension de la COVID-19

---

Depuis le début de nos travaux, l'apport de la métabolomique dans l'aide à la compréhension physiopathologique et dans l'identification de nouveaux biomarqueurs n'a cessé de croître. Son point fort réside dans le fait qu'elle permet une analyse de l'interaction hôte-virus, au plus proche du phénotype ; et qu'il est possible, par le biais d'analyses répétées, de suivre les réponses biochimiques de l'hôte tout au long de la maladie. A l'heure actuelle, la principale limite à son extrapolation en pratique de routine semble être la difficulté de standardiser les approches analytiques. Nous faisons encore face à une hétérogénéité dans les méthodes utilisées, que ce soit dans la préparation des échantillons, dans les analyses chromatographiques mais également dans le post traitement des données (Bourgin, Durand, and Kroemer 2023). Ces limites rendent donc difficiles la généralisation des résultats et la réalisation de méta-analyses. Les études révèlent toutefois souvent des effets similaires sur certains métabolites.

Cependant, le partage est surtout la mutualisation des résultats obtenus par les différentes équipes est nécessaire. C'est d'ailleurs dans cette idée que certaines initiatives telles que la cartographie de la COVID-19 (C19DMap) d'Ostaszewski et son équipe *COVID-19 Disease Map* ont été créés (Ostaszewski et al. 2021). L'objectif étant d'établir un référentiel de connaissances sur les mécanismes moléculaires du COVID-19 et étant libre d'accès ([https://covid.pages.uni.lu/map\\_contents](https://covid.pages.uni.lu/map_contents)).

Un autre point fort de l'approche métabolique, comme nous avons souhaité l'aborder par le biais de notre étude SignCov, est la possibilité d'analyser plusieurs types d'échantillons (tissus, plasma, serum, urines, air expiré...). L'analyse des différents tissus et biofluides pourrait permettre une meilleure compréhension de la dysrégulation métabolique secondaire à l'infection à SARS-CoV-2 mais également de recherche des outils diagnostiques et pronostiques peu invasifs et mieux acceptables par la population.

### *Le métabolome sanguin*

Très rapidement après la découverte du SARS-CoV-2, les équipes scientifiques travaillant dans le domaine de la métabolomique se sont intéressées à l'analyse du métabolome plasmatique ou sérique des patients atteints de la COVID-19. Comme expliqué plus haut, cette approche semblait prometteuse pour améliorer la compréhension des mécanismes physiopathologiques ainsi également des signatures diagnostiques et pronostiques. Et les bio-banques d'échantillons, en majeure partie issues des prélèvements sanguins de patients atteints, se multipliaient.

Naturellement, notre première étude s'est attachée à ces objectifs. La force de cette dernière est qu'elle concernait des patients atteints par la COVID-19 lors de la première vague et donc une cohorte probablement quasi représentée que par des patients « primo-infectés » par le SARS-CoV-2. Ce point semble important à soulever car une partie de nos recherches étant ciblée sur la compréhension des modifications métaboliques entraînées par le SARS-CoV-2, il était important de ne pas avoir de modifications liées à une récurrence d'infection ou à un COVID long. Dans le même ordre d'idée, ces cohortes (comprenant celle des volontaires sains) n'avaient pas encore bénéficié de la vaccination puisqu'aucun vaccin n'était disponible à cette époque. Nos résultats ont montré que le profilage métabolomique non ciblé d'échantillons de plasma provenant de volontaires sains et de sujets infectés par le SRAS-CoV-2 a permis d'identifier de nouvelles signatures métabolomiques plasmatiques et de nouveaux biomarqueurs d'intérêt pour le diagnostic et le pronostic du COVID-19 mais également pour les mécanismes biologiques sous-jacents.

### *L'atteinte du métabolisme lipidique*

Lors de nos analyses sur échantillons plasmatiques, nous avons constaté que des variations de taux de lipides étaient un signal fort de nos signatures métabolomiques diagnostiques (comparaisons des sujets sains versus tous les sujets infectés par le SARS-CoV-2 et sujets sains versus sujets légèrement atteints). Les lipides étant très représentés dans les top-listes de 15 métabolites de ces deux comparaisons.

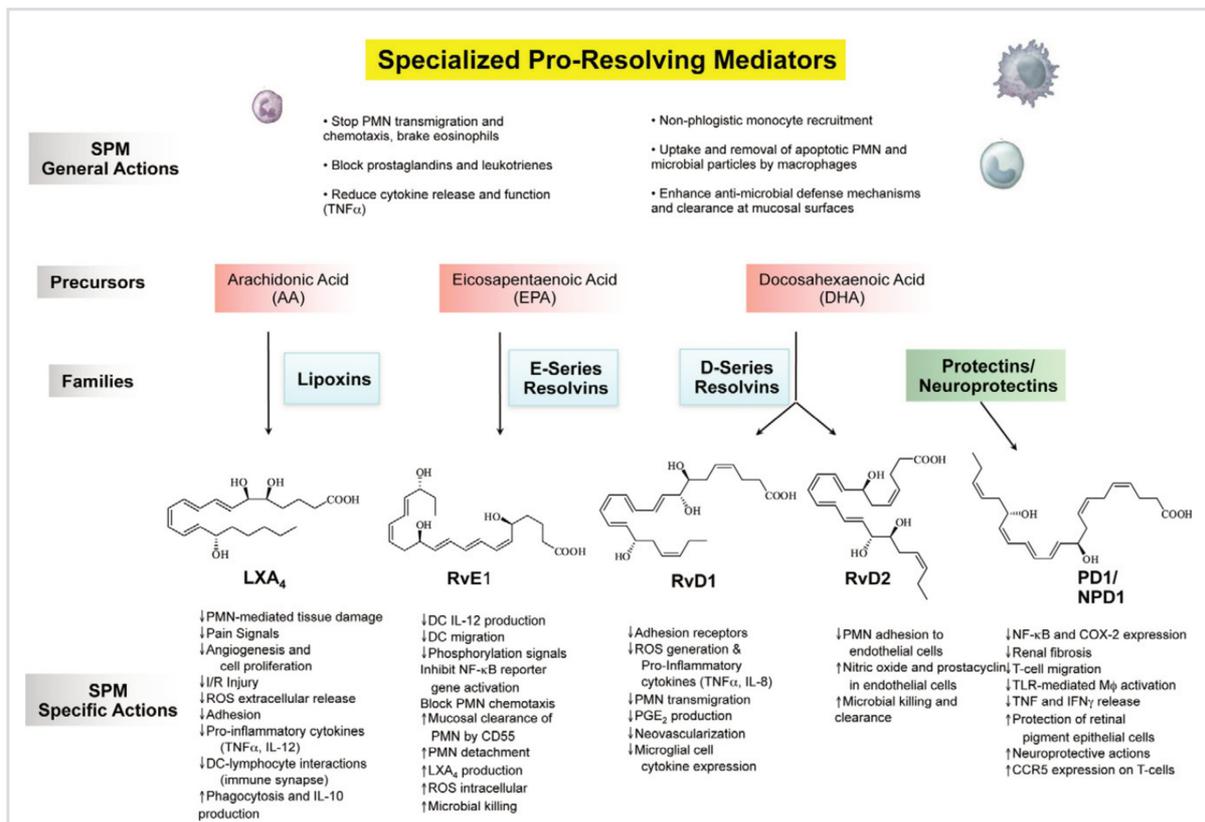
D'autres équipes ont mis en avant une atteinte du métabolisme lipidique chez les patients atteints de la COVID-19.

L'étude princeps, de Wu et al., basée sur une analyse métabolomique et lipidomique, met en avant une régulation à la hausse des lipides composant leur signature plasmatique. Cette signature comprenait, entre autres, des diglycérides, des acides gras libres et des triglycérides. Dans leur cohorte, ces modifications étaient d'autant plus importantes que le patient était gravement atteint. Dans leurs résultats, les phosphatidylcholines étaient quant-à-elles régulées à la baisse. Les auteurs concluent à une potentielle accélération de la lipolyse entraînée par l'infection. Pour finir, ils ont également mis en avant une régulation à la baisse du *Xylulose 5-phosphate* chez les patients gravement atteints. Ce métabolite étant un signal de coordination activant la phosphofructokinase dans la glycolyse et favorisant la transcription des gènes de la lipogénèse et de la glycolyse, les auteurs suggèrent qu'il pourrait exister une altération du métabolisme glucidique (que nous discuterons par la suite) et lipidique pouvant traduire une atteinte hépatique (D. Wu et al. 2020).

Nos résultats sont également en accord avec l'analyse du métabolome lipidique de patients présentant une pneumopathie secondaire à la COVID-19 réalisée par Nguyen et son équipe objectivant une augmentation du taux d'acides gras libres poly-insaturés (principalement les *acides linoléique* et *arachidonique*) ainsi qu'une baisse des HDL-cholestérol (Nguyen et al. 2021). Dans le même sens Thomas et al. montraient également une baisse des acylcarnitines à chaîne courte et moyenne et une augmentation du taux d'acides gras libres dans le métabolome plasmatique des sujets infectés par le SARS-CoV-2. Ceci pourrait être le reflet d'une activité accrue de la Phospholipase A2 entraînant, entre autres, une surproduction d'*acide arachidonique*, précurseur d' eicosanoïdes (leucotriènes et prostaglandines) médiateurs pro-inflammatoires et donc impliqué dans la tempête cytokinique mais également impliqué dans la synthèse du facteur d'activation plaquettaire et donc dans la coagulopathie (T. Thomas et al. 2020).

Schwarz et al., ont également constaté une signature mettant en avant une variation de taux de lipides. Ils observaient des changements dans les sphingomyélines, les lyso-phospholipides, les esters de cholestérol et de multiples classes de glycérophospholipides. En s'intéressant plus précisément aux variations des eicosanoïdes et docosanoïdes, lipides reconnus comme immunomodulateurs ainsi qu'aux lipides précurseurs d'acides gras polyinsaturés dans le lipidome sérique, ils mettaient en avant une dysrégulation de ces médiateurs lipidiques selon le niveau de gravité du tableau clinique (Schwarz et al. 2021).

Ces données laissent penser qu'il existe une dysrégulation du métabolisme des lipides pro-inflammatoires et des lipides pro-résolutifs de l'inflammation (Specialized Proresolving Mediators = SPM) lors de l'infection à SARS-CoV-2. Quatre familles de SPM sont actuellement identifiées : les lipoxines, les résolvines, les protectines et les marésines. Les SPM proviennent du métabolisme de certains acides gras polyinsaturés précurseurs appartenant aux familles des omégas-3 et omégas-6 et entrent en jeu, selon diverses actions schématisées dans la figure 28, dans la résolution de l'inflammation (Calvino 2021; Serhan 2010). Une partie des modifications du métabolome plasmatique des sujets atteints de la COVID-19 constatées pourraient donc être expliquées par une mise en jeu accrue de lipides pro-inflammatoires et par une synthèse de ces SPM lors de l'infection à SARS-CoV-2.



**Figure 28. Les lipides pro-résolutifs de l'inflammation**

Extrait de Serhan, Charles N. 2010. "Novel Lipid Mediators and Resolution Mechanisms in Acute Inflammation: To Resolve or Not?" American Journal of Pathology 177(4):1576–91.

Ainsi, le métabolisme lipidique jouerait un rôle central dans l'infection à SARS-CoV-2. Il interviendrait à de multiples étapes de la physiopathologie virale (Casari et al. 2021):

- Interaction virus-hôte par reprogrammation du métabolisme des lipides
- Endocytose et exocytose par utilisation des radeaux lipidiques membranaires

- Participation à la formation d'autophagosome et des exosomes qui facilitent la dissémination virale, la cytotoxicité et l'activité pro-inflammatoire
- Activation des voies immuno-modulatrices
- Participation à la coagulopathie

### *L'atteinte du métabolisme des acides aminés*

En accord avec nos résultats, de nombreuses études analysant le métabolome de patients infectés par le SARS-CoV-2 ont retrouvé une forte altération du métabolisme des acides aminés, ciblant notamment la *valine*, la *L-citrulline*, la *L-isoleucine*, l'*asparagine*, l'*aspartate*, l'*arginine*, la *proline*, la *glycine*, la *glutamine* et le *tryptophane* (Bourgin, Durand, and Kroemer 2023).

Pour exemple, l'équipe de Shen ont mis en avant plus de 100 métabolites, comprenant des acides aminés et leurs dérivés, fortement régulés à la baisse dans le plasma de patients atteints de COVID-19 et dont 10 sont directement impliqués dans le métabolisme de l'arginine (Shen et al. 2020).

Dans la cohorte de Thomas et al., il a également été constaté une baisse des acides aminés gluconéogènes, de l'*ornithine* et de la *citruline* dans le serum des patients atteints de COVID-19, ceci pouvant traduire une altération du cycle de l'urée. Les polyamines du métabolisme de l'azote (*spermidine*, *acétyl-spermidine*) étaient aussi fortement augmenté (T. Thomas et al. 2020).

Ces polyamines joueraient un rôle important dans les interactions virus-hôte et dans la réplication virale, notamment pour les phases de polymérisation de l'ADN et de l'ARN, d'encapsulation des acides nucléiques et de synthèse des protéines (Mounce et al. 2017)

D'autres travaux vont dans le sens d'une altération du catabolisme des acides aminés qui pourrait être favorisée par l'hypoxie. Ainsi avec augmentation de trois acides  $\alpha$ -hydroxylés selon la sévérité de la maladie constatée Páez-Franco et al. pourrait témoigner de la contribution du métabolisme cétonique au maintien du recyclage du Nicotinamide Adénine Dinucléotide (NAD) malgré les conditions d'hypoxie (Páez-Franco et al. 2021).

Certains auteurs mettent en avant l'impact physiopathologique de la baisse du taux d'arginine puisque cet acide aminé serait le substrat utilisé dans la production d' *oxyde nitrique* (NO) molécule de signalisation impliquée dans les réponses inflammatoires, dans la vasodilatation et

la bronchodilatation mais surtout agissant comme agent antithrombotique et cytoprotecteur (Adebayo et al. 2021; Rees et al. 2021)

Il a également été constaté dans plusieurs travaux, une diminution de la *glutamine* au profit de sa transformation en *glutamate*. La *glutamine*, nutriment essentiel pour la prolifération des lymphocytes et la production de cytokines, les activités phagocytaires et sécrétoires des macrophages, et l'élimination des bactéries par les macrophages, pourrait alors être une cible à compléter (Cruzat et al. 2018; Matsuyama et al. 2021). Les travaux longitudinaux de Lewis et al., évaluant le métabolome de sujets atteints de la COVID-19 sur plus long terme (7 mois après l'infection), a d'ailleurs montré que l'altération du métabolisme de la *glutamine* persistait bien après la guérison clinique (Lewis et al. 2022).

### *L'implication de la voie de la kynurénine*

Nos données suggèrent que le catabolisme du *tryptophane* est fortement impacté chez les patients sévèrement atteints de la COVID-19. Ceci se traduisant par des signatures métabolomique objectivant une diminution du *tryptophane* et une augmentation de *kynurénine*, de *kynurénate* voire d'*acide anthranilique*. Des résultats similaires ont également été décrits par de nombreuses équipes (Blasco et al. 2020; Bourgin, Durand, and Kroemer 2023; T. Thomas et al. 2020; D. Wu et al. 2020b).

Le *tryptophane* est un acide aminé essentiel qui est catabolisé par le foie via la voie de la *kynurénine* afin de produire plusieurs composés, dont du  $\text{NAD}^+$ , rentrant en jeu dans des processus biologiques fondamentaux. Le *tryptophane* non utilisé par le catabolisme hépatique va être utilisé par les tissus périphériques où l'*indoléamine 2,3-dioxygénase 1 (IDO 1)*. Au cours des processus inflammatoire, et notamment suite à la stimulation par l'interféron  $\gamma$ , il est constaté une régulation à la hausse de l'IDO 1. Selon le type de cellule et le milieu cytokinique, l'expression de l'IDO peut être également modulée par les interleukines, le TGF- $\alpha$ , ou les prostaglandines. Les métabolites obtenus par ce catabolisme du *tryptophane* médié par IDO 1 sont des médiateurs importants des réponses immuno-inflammatoires dans l'inflammation, dans le stress oxydatif et dans le dysfonctionnement endothélial. Le métabolisme du *tryptophane* serait également un élément clef de la modulation du microbiote intestinal. Cependant, il a déjà été démontré que la surexpression de l'IDO peut atténuer les réponses immunitaires aux antigènes (Wang et al. 2015). Elle agirait comme un régulateur négatif de l'inflammation et de l'activation immunitaire voire déclencherait des processus de tolérance immunitaire pouvant avoir un impact désastreux dans certains mécanismes infectiologiques ou cancérologiques.

Elle limiterait la fonction des lymphocytes T (Mangge et al. 2021). Le ratio kynurénine/tryptophane pourrait constituer un biomarqueur intéressant de la sévérité de la maladie, reflétant l'activité de l'IDO (T. Thomas et al. 2020)

### *L'atteinte du métabolisme glucidique et énergétique*

Wu et al. ont également mis en avant dans leur signature métabolomique plasmatique que l'acide malique, un intermédiaire du cycle de Krebs (ou cycle des acides tricarboxyliques = TCA), présentait le métabolisme le plus significativement modifié lors de leur comparaison entre les sujets sains et les patients décédés de la COVID-19, montrant une forte baisse d'abondance relative dans le groupe de patients décédés. Ils postulent ainsi que l'altération du métabolisme énergétique serait secondaire à la réplication du SARS-CoV-2, argumentant que le métabolisme énergétique des cellules hôtes pourrait être privilégié afin de garantir cette réplication. Dans leurs travaux, le *glycerol-3-phosphate* est également retrouvé moins abondant chez les patients les plus graves. Métabolite impliqué dans la glycolyse mais également reflet de l'immunité acquise, il pourrait, selon les auteurs, être un biomarqueur d'intérêt pour dépister un dysfonctionnement immunitaire (D. Wu et al. 2020a).

Pang et al. constatent, eux aussi, une diminution du métabolisme du TCA probablement en lien avec la forte consommation en énergie du SARS-CoV-2 et discutent les dangers engendrés par cette baisse. En effet, cette altération métabolique pourrait entraîner une diminution de l'effet anti-oxydant et engendrer des dommages inflammatoires (Pang, Zhou, et al. 2021).

### *Métabolisme de l'hème*

Nos résultats objectivant une augmentation de la *bilirubine* dans nos signatures pronostiques sont en accord avec la littérature sur le métabolome plasmatique des patients infectés à SARS-CoV-2. L'augmentation de ce métabolite ou des produits de dégradation de la *bilirubine* pourraient être le reflet de l'atteinte hépatique (Shen et al. 2020).

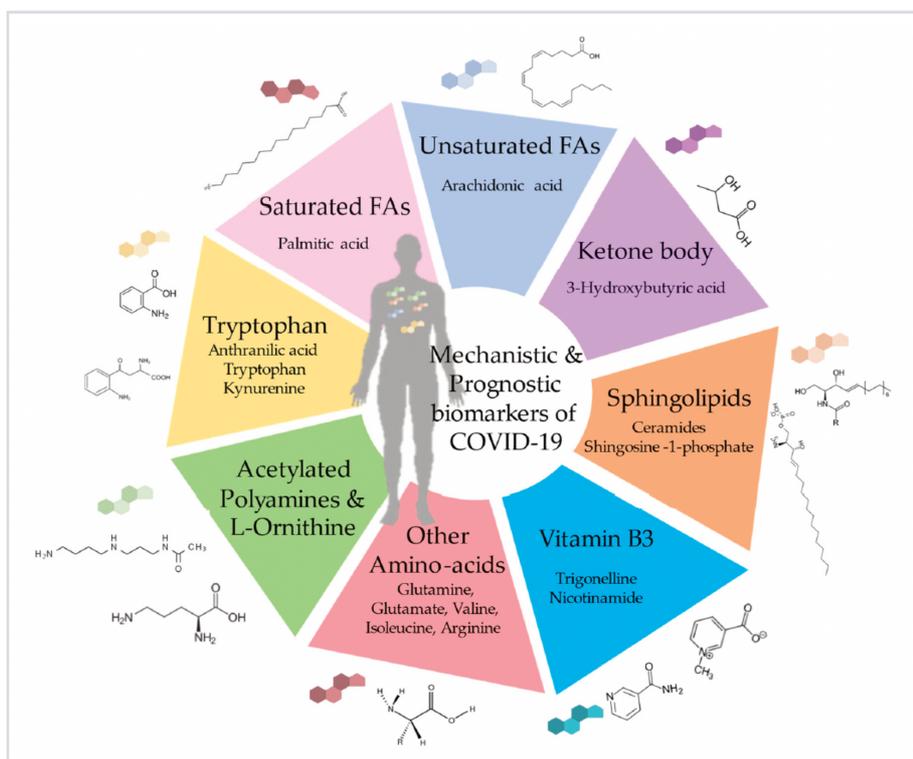
### *L'atteinte du métabolisme des purines*

Wu et al. ont retrouvé une régulation à la baisse du *Guanosine-Monophosphate* chez les patients les plus graves (D. Wu et al. 2020a).

Blasco et al. ont, quant à eux, mis en avant que la *xanthine* et l'*adénosine*, impliquées dans le métabolisme des purines, faisait partie de leur signature pronostique de 15 métabolites. Ils expliquent que l'inflammation induite par l'infection pourrait être en partie contrôlée par la libération de purines secondaires à la lyse cellulaire et que ce phénomène pourrait être régulé par l'hypoxie. Les purines possèderaient, elles-mêmes, des activités immuno-modulatrices. Certaines, comme le NAD, entraîneraient une réponse pro inflammatoire alors que d'autres, telles que l'*adénosine*, seraient plutôt des effecteurs anti-inflammatoires sur les cellules immunitaires (Blasco et al. 2020).

Les différentes revues de la littérature et méta-analyses sur le sujet vont dans le sens des résultats discutés ci-dessus (Bourgin, Durand, and Kroemer 2023; Hasan, Suleiman, and Pérez-López 2021; Pang, Zhou, et al. 2021).

Ainsi, selon Bourgin et al., certains métabolites issus de ces voies métaboliques impactées pourraient constituer de nouveaux biomarqueurs (Figure 29). Parmi eux, le *tryptophane*, les acides aminés retrouvés comme modifiés (par exemple *Glutamine*, *Valine*, *Isoleucine*, *Arginine*), les polyamines tels que la *spermine* ou la *spermidine*, certains acides gras tels que l'*acide palmitique* ou l'*acide arachidonique*, la *vitamine B3*, ou encore le *céramide* et la *sphingosine-1-phosphate*, ou le  *$\beta$ -hydroxybutyrate* seraient de bons candidats (Bourgin, Durand, and Kroemer 2023).



**Figure 29 : Schéma des potentiels biomarqueurs d'intérêt diagnostiques et pronostiques de la COVID-19 selon Bourgin et al.**

Extrait de : Bourgin et al. 2023. "Diagnostic, Prognostic and Mechanistic Biomarkers of COVID-19 Identified by Mass Spectrometric Metabolomics." *Metabolites* 13(3).

### **Le métabolome urinaire**

Notre étude prospective SignCov s'est intéressée à l'analyse du métabolome urinaire de patients atteints de la COVID-19. L'analyse de ce bio-fluide est intéressante en plusieurs points. Tout d'abord, le prélèvement urinaire est une technique souvent non invasive et donc indolore pour le patient, ne nécessitant pas de cathétérisme dans la majeure partie des cas. Il s'agit donc d'une technique tout à fait acceptable pour le patient notamment quand les prélèvements sont répétés où lorsqu'un dépistage de masse est nécessaire. Ensuite, l'analyse du métabolome urinaire pour compléter une analyse plasmatique nous a semblé évidente dans l'objectif d'améliorer la compréhension des phénomènes métaboliques impactés par l'infection à SARS-CoV-2.

D'autres équipes se sont intéressées à l'analyse de ce métabolome urinaire et mettent en avant des données métabolomiques en lien avec nos résultats.

Ainsi, Rosolanka et al. ont mis en avant, sur leur étude longitudinale du métabolome urinaire de patients atteints de la COVID-19, une augmentation du taux d'*acétone* ainsi qu'une diminution du taux de *citrate* urinaires, à la phase initiale de la maladie, pouvant corroborer les résultats plasmatiques et sériques déjà discutés auparavant et étant en faveur d'une altération du métabolisme énergétique au profit du métabolisme cétonique. Nous constatons d'ailleurs de notre côté, une diminution significative de citrate (nous ne pouvons pas observer l'acétone) pour nos trois comparaisons (COVID-19 versus non-COVID-19, COVID-19 versus symptomatiques non-COVID, COVID-19 modérés ou sévères versus peu sévères) mais également une augmentation du taux de l'*acide lactique* urinaire dans notre cohorte de patient COVID comparativement au groupe contrôle. Parallèlement à ce résultat, et comme le montre également nos résultats, une augmentation du taux de *carnitine* urinaire était également observée dans leur cohorte, témoin de l'augmentation de la beta-oxydation des acides gras et donc témoin indirect de l'augmentation du métabolisme des corps cétonique.

De manière similaire aux résultats que nous constatons dans SignCov, l'équipe de Rosolanka constate une augmentation de l'abondance d'*hypoxanthine* (Rosolanka et al. 2023). Cette augmentation est significative pour nos comparaisons COVID-19 versus non-COVID-19 et COVID-19 versus symptomatiques non COVID-19. L'*hypoxanthine*, nucléoside purique libéré lors de l'hydrolyse de l'ATP et de l'ADP en AMP et pouvant circuler librement dans le sang, est déjà identifiée comme un métabolite d'intérêt, biomarqueur potentiel des phénomènes d'ischémie ou d'hypoxie, dans la littérature. (Fisher, Benson, and Imray 2018).

Dans SignCov, nous remarquons une baisse du taux d'*épinéphrine* urinaire pour nos 3 comparaisons. Rosolanka et al. arrivent aux mêmes conclusions en constatant une augmentation de la *tyrosine* urinaire (Rosolanka et al. 2023) et en concluant à une potentielle sous-utilisation de la *tyrosine*, pouvant entraîner une sous-production d'hormones thyroïdiennes et de neurotransmetteurs dérivés de la *tyrosine*, tels que la *dopamine* et la *noradrénaline* (Liptak et al. 2022).

Les travaux de Jing et al. mettent en avant dans leur cohorte que le métabolome urinaire proviendrait principalement du métabolisme associé au microbiome (33,3 %) et de la voie du métabolisme du *tryptophane* (12,8 %) confirmant ainsi une altération du microbiote par la COVID-19 (Jing et al. 2022). Les preuves de l'impact de la COVID-19 sur le métabolisme des acides aminés et de la voie de la *kynurénine* étant déjà précédemment établies, Dewulf et al ont mené une analyse métabolomique ciblée sur échantillon urinaire dans l'objectif de quantifier

ces changements. Ils confirment ainsi la présence d'une augmentation importante de concentration urinaires des métabolites du tryptophane (*kynurénine*, *3-hydroxykynurénine* et *3-hydroxyanthranilate*) fortement corrélée à l'augmentation de la CRP, biomarqueur historique de l'inflammation, et à la gravité de la maladie. Parallèlement à cela, ils objectivent également une augmentation de l'acidourie notamment sur les acides aminés neutres (Dewulf et al. 2022). Lonati et al. confirment l'importance de la voie de la *kynurénine* dans les métabolomes urinaires de patients infectés par le SARS-CoV-2 et éclairent le lien entre ces dysrégulations, celles constatées sur la biosynthèse du NAD<sup>+</sup> et leur impact sur l'immunomodulation. Ils remarquent dans leurs analyses portant sur les métabolomes plasmatiques et urinaires des niveaux sanguins de *Sirtuine 1* (SIRT1) plus faibles. Dans leur cohorte, les patients dont la SIRT1 était altérée présentaient des concentrations plus élevées de biomarqueurs d'inflammation systémique, de biomarqueurs reflet du dysfonctionnement épithélial, ainsi qu'une libération accrue de médiateurs inflammatoires. A noter que Lonati et son équipe retrouve, tout comme nous, une augmentation de la *néoptérine* et de la *taurine* ainsi qu'une diminution de la *trigonelline* (Lonati et al. 2023). Comme expliqué précédemment, la voie de la *kynurénine* conduit à la production d'*acide quinolinique*, précurseur dans la biosynthèse du NAD<sup>+</sup>. Les sirtuines (SIRT), enzymes dont l'action est NAD<sup>+</sup> dépendante, sont des régulatrices de nombreux processus biochimiques et biologiques indispensables à l'homéostasie comme la sénescence cellulaire, l'apoptose, les métabolismes glucidiques et lipidiques, le stress oxydatif et l'inflammation ...De plus, plusieurs voies de signalisation provoquées par l'activation des cellules immunitaires sont étroitement liées à l'activité des SIRT (Yang et al. 2022).

Cette étude, comprenant l'analyse du métabolome urinaire, permet une avancée majeure dans la compréhension du lien physiopathologique entre atteinte de la voie de la *kynurénine* et du métabolisme des purines et l'inflammation secondaire à la COVID-19.

Nos résultats dans SignCov montrent des augmentations significatives des concentrations urinaires de tryptophane, de kynurénine et de kynurénate en comparant COVID-19 versus non COVID-19 et COVID-19 versus symptomatiques non-COVID-19 pour le diagnostic de l'infection. Elle est aussi significative pour la concentration de kynurénine en comparant COVID-19 versus symptomatiques non-COVID-19 pour le pronostic de gravité.

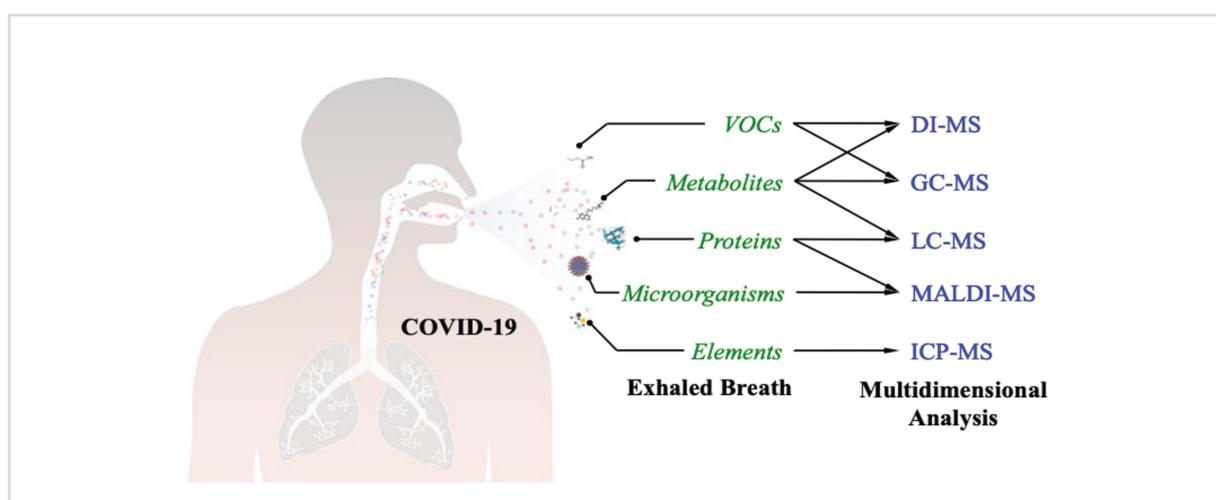
### Le métabolome de l'air expiré

Le SARS-CoV-2 touchant principalement les voies aériennes supérieures et inférieures, les équipes de recherche ont jugé pertinent d'analyser le métabolome de l'air expiré. D'autant que l'analyse métabolomique de l'air expiré a pour principaux avantages d'être non invasive, de permettre une analyse multidimensionnelle présentant de bonnes performances analytiques et de potentiellement compléter l'analyse des fluides du corps humain (Yuan and Hu 2021).

Elle permet de détecter plusieurs classes de composés et particules parmi lesquelles (Figure 30)

:

- les composés organiques volatils (COV)
- les métabolites non volatils
- les protéines
- les micro-organismes
- les oligo-éléments



**Figure 30. Analyse multidimensionnelle de l'air expiré basée sur la spectrométrie de masse**

Extrait de Yuan, Zi Cheng and Bin Hu. 2021. "Mass Spectrometry-Based Human Breath Analysis: Towards COVID-19 Diagnosis and Research." *Journal of Analysis and Testing* 5(4):287-97.

DI-MS : Spectrométrie de masse avec ionisation par Désorption-Ionisation

GC-MS : Chromatographie en phase gazeuse-spectrométrie de masse

LC-MS : Chromatographie en phase liquide-spectrométrie de masse

MALDI-MS : Spectrométrie de masse avec désorption/ionisation laser assistée par matrice

ICP-MS : spectrométrie de masse à plasma à couplage inductif

De plus, les méthodes de prélèvements et d'analyses étaient déjà standardisées (Walker and Burrell 2020) et avaient déjà fait leurs preuves dans le Syndrome de Détresse Respiratoire Aiguë (SDRA) (Bos et al. 2014; S. Zhang et al. 2024) et d'autres pathologies respiratoires (Belizário, Faintuch, and Malpartida 2021; Bos et al. 2016; Jareño-Esteban et al. 2017).

Concernant l'analyse des Composés Organiques Volatils ou COV, Ibrahim et al. ont pu mettre en avant, dans leur étude pilote, que l'air expiré des patients atteints de COVID-19 contenait des concentrations plus élevées de *1-propanol* (COV d'intérêt déjà identifié dans la pneumopathie bactérienne, la néoplasie pulmonaire et l'asthme), et le 3,6-diméthylundécane, COV dérivé probable de la peroxydation lipidique (Ibrahim et al. 2021).

Dans l'objectif de permettre de différencier l'atteinte pulmonaire de la COVID-19 des autres affections respiratoires aiguës, les travaux de Grassin-Delyle et al. ont permis de prouver que 4 COV (*méthylpent-2-éanal*, *2,4-octadiène*, *1-chloroheptane* et *nonanal*) pourraient discriminer le SDRA secondaire à la COVID-19 des SDRA non COVID (Grassin-Delyle et al. 2021). Toujours dans cet objectif, Ruszkiewicz et al. ont pu construire un « score COVID-19 » basé sur l'abondance relative d'un panel de composés organiques volatils. Ce score, évalué sur deux cohortes provenant de 2 pays différents, permettait, avec une précision de 80% (AUC : 0,87 (IC à 95 % 0,67–1)) et 81,5% (AUC : 0,91 (IC à 95 % 0,87–1)), de différencier les patients atteints de COVID-19 de ceux souffrant d'une autre atteinte respiratoire (asthme, BPCO, pneumonie bactérienne et décompensation cardiaque). La signature permettant la construction de ce score comprenait des aldéhydes (*éthanal*, *octanal*), des cétones (*acétone*, *butanone*) et du *méthanol*. Selon les auteurs, la baisse constatée du *méthanol* pourrait être le reflet de la réduction de l'activité du microbiote intestinal. Les aldéhydes pourraient, quant-à-eux, refléter une peroxydation lipidique accrue et donc la présence de processus pro-inflammatoires (Ruszkiewicz et al. 2020).

L'analyse métabolomique non ciblée sur air expiré de Barberis et son équipe retrouve une signature métabolomique composée de 26 métabolites dont 8 sont des monoglycérides d'acides gras. Parmi eux, 3 étaient des monoglycérides saturés (*monopalmitine*, *monomyristine* et *monolaurine*) et étaient régulés à la baisse chez les patients atteints de COVID-19 comparativement aux sujets sains (Barberis et al. 2021). Ces résultats objectivant une baisse des lipides saturés viennent corroborer nos résultats sur le métabolome plasmatique des patients COVID-19.

## Le métabolome salivaire

Rapidement, les équipes scientifiques se sont également intéressés à l'analyse du métabolome salivaire. En effet, la présence du SARS-CoV-2 dans la salive est bien établie et semble secondaire à plusieurs mécanismes (Sapkota et al. 2021):

- Entrée du virus à partir des voies aériennes supérieures et/ou inférieures
- Entrée via le liquide crévulaire gingival
- Libération des particules virales par les canaux salivaires via les glandes salivaires infectées

Tout ceci conduit à penser qu'un changement du métabolisme salivaire à la suite d'une infection à SARS-CoV-2 est possible et que l'analyse du métabolome salivaire des patients infectés pourrait être utile pour le diagnostic, le suivi ou le pronostic de la COVID-19. De plus, la collecte d'un échantillon salivaire est facilement réalisable, indolore et ne nécessite pas d'équipement spécifique (seul un récipient stérile semble nécessaire). Il est cependant important de prendre en compte que le contenu salivaire peut être influencé par l'état bucco-dentaire (variations inter-individuelles), par le rythme circadien et les habitudes alimentaires, par la méthode de prélèvement et il est important de noter que certaines protéines salivaires peuvent être dégradées par les enzymes protéolytiques (Costa Dos Santos Junior et al. 2020; Sapkota et al. 2021).

En s'intéressant à l'analyse du métabolome salivaire, Pozzi et al mettent en évidence une signature métabolique de 9 métabolites qui serait capable de prédire le recours à une hospitalisation chez les patients infectés par le SARS-CoV-2 (*acide 2-pyrrolidine-acétique, 1-3-diaminopropane, 3-hydroxypyridine-acétate, 1-3-diaminopropane, 3-hydroxypyridine, cyclo(leu-pro), myo-inositol, N,N-diméthyl-5-aminovalérate, 3-(3-hydroxyphényl)propionate, pantothénate et mannonate*). Ils ont également pu identifier que le microbiote salivaire des patients atteints de COVID-19 différait de celui des sujets non infectés. (Pozzi et al. 2022). A noter que dans le même ordre d'idée, l'analyse métabolomique des selles a permis de mettre en évidence des changements dans le microbiote intestinal reflétant l'inflammation intestinale et la mauvaise absorption de certains nutriments (Lv et al. 2021). La métagénomique a également prouvé des modifications du microbiotes nasopharyngé (Ventero et al. 2021) et des voies respiratoires inférieures (I. Sulaiman et al. 2021).

Saheb Sharif-Askari et al. ont mis également en évidence, sur des échantillons plasmatiques et salivaires de 62 patients atteints de COVID-19 et 18 patients non infectés, que la dysrégulation métabolique entraînée par la maladie pourrait être plus efficacement détectée dans les échantillons salivaires comparativement aux échantillons plasmatiques et que ces métabolites diffèrent selon la gravité du tableau clinique. Des biomarqueurs, tels que *l'acide linoléique* et *l'alpha-cétoisovalérique* ressortent comme intéressants dans le diagnostic de l'infection à SARS-CoV-2 alors que la *sphingosine* et la *kynurénine* ou le *tryptophane* pourraient être des biomarqueurs salivaires pronostiques (Saheb Sharif-Askari et al. 2022).

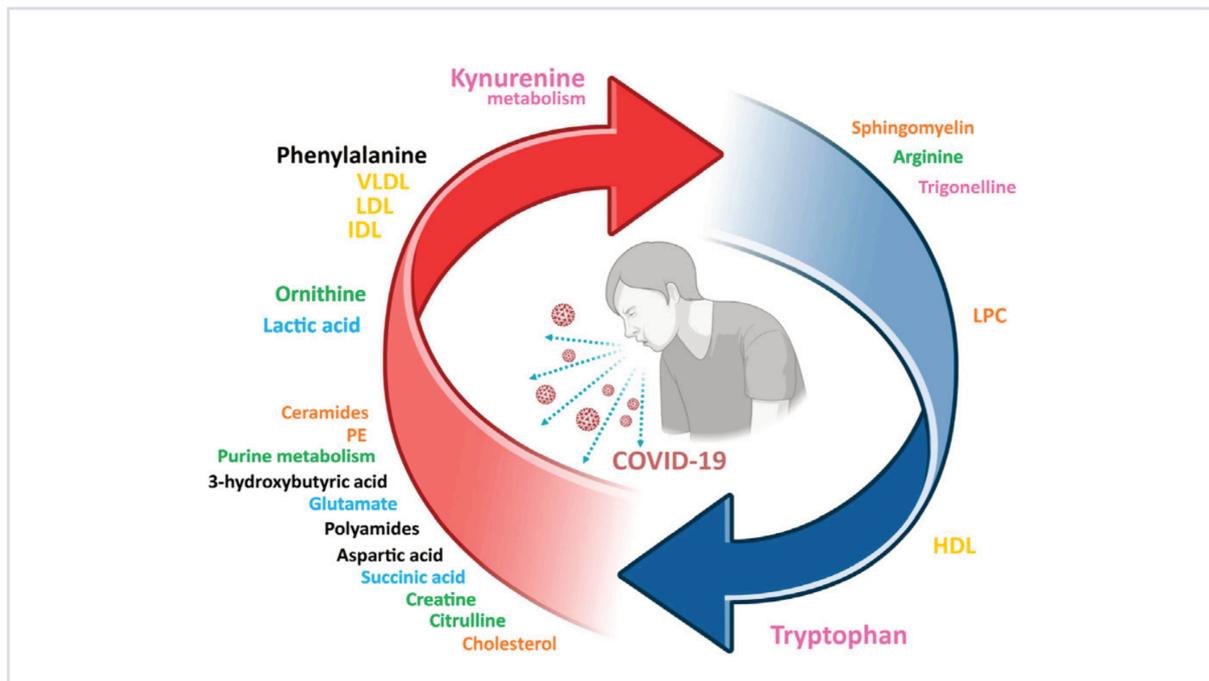
Comme discuté auparavant, la dysrégulation de la voie des acides aminés a un réel impact dans les métabolites plasmatiques et urinaires des patients atteints de la COVID-19. Dans le même sens, Frampas et al ont mis en avant une signature métabolomique salivaire composée de 6 métabolites permettant de prédire la sévérité de l'atteinte avec une AUC de 1. Sur ces 6 métabolites, 5 étaient des acides aminés (*valine, leucine, phénylalanine, tyrosine et proline*) et tous montraient des taux plus faibles dans le métabolome salivaire des patients les plus graves (Frampas et al. 2022).

L'équipe de Da Silva Fidalgo s'est, elle aussi, intéressée à l'analyse du métabolisme salivaire de 35 patients (dont 9 avec PCR COVID négative, 9 avec PCR COVID positive et 18 présentant une infection ancienne et considérée comme guérie). Les résultats mettent en avant une diminution des taux d'acides aminés majeurs (*alanine, glutamine, histidine, lysine, phénylalanine, proline*) dans le groupe PCR positive comparativement aux non infectés. D'autres métabolites tels que le *saccharose* et le *butyrate* étaient plus élevés chez les patients du groupe post-infection que chez les patients PCR négative. Bien que les scores de PLS-DA soient discutables dans cette étude (accuracy = 0,69,  $R^2 = 0,39$ ,  $Q^2 = 0,08$ ), ce travail va dans le sens de la modification du métabolome salivaire pouvant impacter les acides aminés et ouvrir la voie sur la présence de modifications potentielles de ce métabolome persistantes même après la guérison clinique (da Silva Fidalgo et al. 2023).

Bien que souffrant de limites, majoritairement dues à l'absence de standardisation des méthodes de recueil et d'analyse des échantillons entre les différents travaux et à la taille des cohortes, ces données suggèrent que l'approche par analyse du métabolome salivaire est prometteuse et que les efforts pour développer cette approche doivent continuer.

A ce jour, la majorité des études en métabolomiques dans le cadre de la COVID-19 se sont intéressées à ces différents bio-fluides, mais d'autres sources d'échantillons semblent prometteuses, comme l'analyse du sébum par exemple (Spick et al. 2022).

La mise en commun des signatures métabolomiques (Figures 31 et 32) identifiées dans la littérature a déjà permis de grandes avancées dans la lutte contre la COVID-19 (Bruzzone et al. 2023). La poursuite de ces recherches scientifiques, combinant les différentes approches « omiques » permettra de mieux comprendre le phénotype des patients atteints de la COVID-19 et donc de perfectionner la compréhension des mécanismes physiopathologiques engendrés par l'infection à SARS-CoV-2.



**Figure 31 : Consensus sur la signature métabolomique des patients COVID-19, selon Bruzzone et al.**

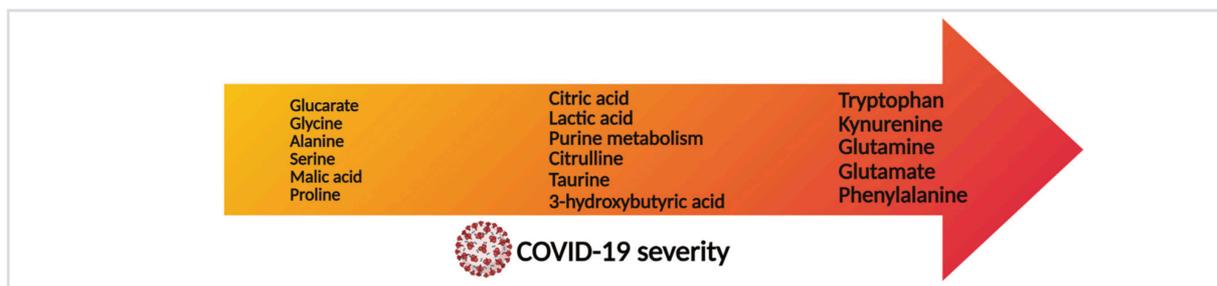
Extrait de Bruzzone C et al. "Metabolomics as a Powerful Tool for Diagnostic, Prognostic and Drug Intervention Analysis in COVID-19." *Frontiers in Molecular Biosciences* 10(Février):1–11

La couleur des flèches indique si les métabolites sont régulés à la hausse (rouge) ou à la baisse (bleu).

La taille des lettres est proportionnelle au niveau de preuve expérimentale.

Légende des couleurs pour les voies métaboliques : métabolisme du tryptophane (rose), métabolisme central (bleu), métabolisme des purines et cycle de l'urée (vert), lipides (orange), lipoprotéines (jaune), autres (noir).

VLDL : lipoprotéines de très basse densité ; LDL : lipoprotéines de basse densité ; HDL : lipoprotéines de haute densité ; IDL : lipoprotéines de densité intermédiaire ; PE : phosphatidyléthanolamine.



### Figure 32 : Métabolites associés à la gravité du pronostic de COVID-19

Extrait de Bruzzone C et al. "Metabolomics as a Powerful Tool for Diagnostic, Prognostic and Drug Intervention Analysis in COVID-19." *Frontiers in Molecular Biosciences* 10(February):1–11

L'identification et surtout la validation des différents « métabotypes » spécifiques de la maladie, de son évolution et de sa gravité pourrait permettre, en plus de développer des tests diagnostiques robustes et moins invasifs, de prédire le risque de mauvaise évolution clinique d'une infection comme le COVID-19 ou de COVID long (Lonati et al. 2023; Saito et al. 2024; Whiley et al. 2024). Ces connaissances permettront de poursuivre les efforts dans le développement de thérapies et de vaccins efficaces (Guo et al. 2023) et cet exemple pourrait être utile lors de potentielles futures pandémies.

**Partie IV**  
**Perspectives en médecine d'urgence**

Au cours des dernières années, les progrès technologiques et de post-traitement des données ont permis une couverture de plus en plus complète du métabolome humain (Bourgin, Durand, and Kroemer 2023). A l'heure de la médecine personnalisée et de précision, la métabolomique est un outil séduisant permettant la lecture de l'influence des facteurs intrinsèques, tels que le génotype, et exogènes, tels que l'environnement, le régime alimentaire, le microbiote ou encore les maladies sur le métabolisme humain. Certaines réponses biochimiques peuvent ainsi être détectées dans les différents tissus ou biofluides alors même que l'homéostasie cellulaire est conservée et donc que l'expression clinique est nulle (Ellero-Simatos, Claus, and Guillou 2019). Ainsi, la métabolomique pourrait tout d'abord permettre de détecter de manière précoce certains biomarqueurs ou signatures multi-marqueurs en lien avec une pathologie.

L'obtention d'un diagnostic précis et la détection précoce de la gravité du tableau clinique d'un patient sont essentielles en médecine d'urgence. Une prise en charge adaptée et précoce peut influencer directement le pronostic des patients, et ce pour un grand nombre de pathologies rencontrées dans nos structures d'urgence. Dans ce contexte, l'évaluation des paramètres vitaux, l'utilisation de scores d'alerte précoces et de biomarqueurs sont pratique courante. Bien que largement utilisé, les performances de ces outils sont continuellement débattues (Occelli and Contenti 2020). En effet, les paramètres vitaux évalués étant principalement liés au système macrocirculatoire, ils peuvent se détériorer tardivement et rassurer à tort le clinicien. De plus, l'émergence de nouvelles pathologies, comme la COVID-19, ont mis en avant les limites de nos outils diagnostiques et pronostiques habituels.

A l'heure des sciences omiques, il est désormais possible d'évaluer dynamiquement les variations physiopathologiques, et donc d'identifier des signatures moléculaires multi-marqueurs qui reflètent directement les fonctions de la cellule ou de l'organe, et donc de détecter les changements liés à la présence d'une pathologie ou à l'exposition à un agent toxique (S. Claus and Guillou 2019). Le diagnostic et l'évaluation pronostique d'une pathologie ne dépendrait donc plus d'un seul biomarqueur, ou de la combinaison de plusieurs biomarqueurs évalués indépendamment, mais serait basée sur des signatures moléculaires complexes reflet de la réponse physiologique à la maladie.

Dans la littérature, émerge de plus en plus de signatures métabolomiques diagnostiques ou pronostiques de pathologies rencontrées quotidiennement en médecine d'urgence. Ces signatures pourraient dans un futur plus ou moins proche intégrer nos algorithmes décisionnels.

Pour exemple, dans la population de traumatisés crâniens les enjeux diagnostiques et pronostiques sont importants en médecine d'urgence. Il faut pouvoir caractériser les patients à risque de lésions intra cérébrales et donc devant nécessité d'un bilan d'imagerie, et ceux à risque de mauvaise évolution que ce soit sur des lésions primaires (causées par des forces mécaniques directes) ou secondaires (causées par l'ischémie, l'œdème vasogénique, l'œdème cytogénique, l'inflammation ou encore le stress oxydatif). En ce sens, des profilages métaboliques ont permis d'identifier des métabolites d'intérêt, intéressant principalement la famille des acides aminés, des intermédiaires du sucre, des lipides prédicteurs de la présence de lésions à l'Imagerie par Résonance Magnétique (IRM) structurale (Kharouba et al. 2023). Plus précisément, trois métabolites d'intérêt (*acide érythronique* et *myo-inositol* et le *polypeptide léger des neurofilaments*), permettrait de prédire la présence de lésions de la substance blanche et/ou de la substance grise cérébrale à l'IRM avec une aire sous la courbe de la caractéristique opérationnelle du récepteur de 0,85 (spécificité = 0,89, sensibilité = 0,65) (I. Thomas et al. 2020). Dans ce domaine, les avancées scientifiques dans le domaine de la métabolomique pourrait permettre un diagnostic plus rapide notamment dans des centres n'ayant qu'un accès limité à l'imagerie.

Dans l'optique de signatures pronostiques, certains travaux se sont intéressés au profilage métabolomique sur micro-dialyse cérébrale ou sur liquide céphalo rachidien, peu adaptée à la réponse rapide et peu invasive nécessaire en médecine d'urgence. Cependant des travaux s'intéressant sur des signatures métabolomiques pronostiques sériques (impactant les phospholipides de la choline) et urinaires des patients traumatisés crâniens semblent prometteuses (Kharouba et al. 2023).

Dans le même ordre d'idée, des études métabolomiques du LCR et du sérum de patients souffrant d'hémorragie sous-arachnoïdienne consécutive à une rupture d'anévrisme et d'hémorragies intra cérébrales non traumatiques commencent à identifier des signatures diagnostiques et pronostiques à court, moyen et long terme, qu'il est encore nécessaire de confirmer par des cohortes de validation (Kharouba et al. 2023).

Les données expérimentales sont également prometteuses pour l'identification des différents phénomènes physiopathologiques et étiologiques conduisant à l'insuffisance rénale aiguë (Patschan et al. 2023). Une meilleure connaissance des causes conduisant à la défaillance rénale pourrait conduire à une prise en charge thérapeutique plus adaptée.

Dans l'insuffisance hépatique aigue, l'analyse multi-omique plasmatique retrouve que l'*albumine* est plus oxydée et glycosylée et plus souvent liée à des protéines associées à l'inflammation, à des produits de glycation avancée, et à des métabolites liés au métabolisme de l'*arginine* et de la *proline*, à des acides biliaires et à des produits de dégradation mitochondriale chez les patients non survivants. Une signature métabolomique à 5 métabolites (*acide nicotinique*, *l-acétyl carnitine*, *l-carnitine*, *sulfate de prégnénolone* et *N-(3-hydroxybutanoyl)-l-homosérine lactone*) permettrait de distinguer les patients prédisposés à une mortalité précoce ou nécessitant une transplantation hépatique d'urgence (Sharma et al. 2023). L'identification précoce de ces patients à risque de mauvaise évolution clinique pourrait permettre à l'urgentiste d'orienter le patient vers une filière de soins adaptée.

L'analyse du métabolome plasmatique et urinaire a également été réalisée sur une cohorte de patient consultant dans un service d'urgence pour douleur abdominale aigue. Elle met en évidence des signatures diagnostiques mais surtout étiologiques et pronostiques de la pancréatite aiguë (Villaseñor et al. 2014). Bien que ces données doivent encore être validées, l'approche par phénotypage métabolomique dans cette pathologie pourrait ainsi favoriser le recours aux gestes endoscopiques et aux thérapeutiques adaptées.

La métabolomique a également fait ses preuves dans un motif de recours quotidiennement rencontré dans nos services d'urgence : la douleur thoracique aigue. Ce symptôme, parfois difficile à caractériser ou peu spécifique peut traduire la présence de pathologies bénignes comme la présence de pathologies plus graves, telles que la dissection aortique, l'infarctus du myocarde ou l'embolie pulmonaire, pouvant mettre en jeu le pronostic vital. Dans ces pathologies, des algorithmes décisionnels, basés sur des biomarqueurs usuels et validés ou encore sur l'utilisation de score clinique sont recommandés (Byrne et al. 2023; Konstantinides et al. 2020). Cependant, des enjeux persistent dans le diagnostic de syndrome coronarien aigu sans élévation du segment ST à troponine cardiaque initiale normale (qui nécessitent donc des tests supplémentaires 1 à 3 heures après l'admission du patient), dans la stratification du risque pour l'embolie pulmonaire et dans le diagnostic de dissection aortique par exemple. Concernant le syndrome coronarien aigu, des études translationnelles ont retrouvé des signatures métabolomiques plasmatiques reflétant la présence d'une athérosclérose coronarienne non obstructive, d'un angor stable, d'un angor instable ou d'un infarctus du myocarde. Un biomarqueur, le *N-acétylneuraminique*, permettrait à lui seul d'identifier la présence d'une ischémie myocardique. L'inhibition de la neuraminidase-1, l'enzyme qui régule la production

de ce métabolite pourrait par ailleurs être une cible thérapeutique d'intérêt (L. Zhang et al. 2018). Sur un modèle d'ischémie myocardique provoquée par occlusion coronaire chez l'homme, des changements du métabolome plasmatique très précoces impliquant principalement les lipoprotéines, le HDL cholestérol, les triglycérides, deux acides organiques impliqués dans le TCA et deux acides aminés (*sarcosine et glutathion*) ont été constatés alors qu'une augmentation du lactate survenait plus tardivement (Hasselbalch et al. 2023). La métabolomique représente donc un outil prometteur pour améliorer les approches actuelles basées sur un biomarqueur unique en identifiant des signatures métaboliques précoces permettant de mieux caractériser le stade de la maladie coronarienne et donc pourrait, dans le futur, entraîner une prise en charge plus précoce de l'ischémie myocardique (Bodi et al. 2013).

Des données sur le métabolome plasmatique et des globules rouges des patients souffrant de maladie thrombo-embolique veineuse commencent également à émerger. L'une d'entre elle retrouve une signature métabolomique plasmatique à 23 métabolites dont la voie métabolique la plus impactée était celle de la d-glutamine, et une signature issue de l'analyse des hématies objectivant 3 biomarqueurs avec un pouvoir discriminant performants : *l'adénosine 3',5'-diphosphate* (AUC 0,983), *le glutathion* (AUC 0,923) et *l'acide acétylsalicylique* (AUC 0,923) (Febra et al. 2024). Le profilage métabolomique pourrait ainsi permettre de différencier les patients souffrant d'embolie pulmonaire, les patients souffrant d'infarctus du myocarde sans élévation du segment ST et les sujets sains (Xie et al. 2023).

L'analyse du métabolome plasmatique de patient souffrant de douleur thoracique pourrait également, à terme, être une aide diagnostique dans la dissection aortique (Hao et al. 2022; J. Zhang et al. 2023).

Mais c'est également dans le domaine de l'infectiologie que la métabolomique pourrait être une approche très intéressante. A l'heure où la physiopathologie complexe du sepsis continue de faire l'objet de recherches et que des phénotypes de patients septiques commencent à être identifiés (Seymour et al. 2019). Les sciences omiques permettent de poursuivre les efforts dans l'identification et la compréhension des mécanismes physiopathologiques qui sous-tendent les différentes réponses des hôtes à l'infection et pourrait entraîner une prise en charge thérapeutique plus personnalisée, basée notamment sur des thérapies adjuvantes ciblées (Li et al. 2022). Alors que nous sommes toujours à la recherche du score clinique ou du biomarqueur idéal, des signatures métabolomiques pourraient faciliter le diagnostic du sepsis et aider à une

meilleure stratification du risque de mauvaise évolution clinique, voire permettre la découverte de nouveaux biomarqueurs jusqu'alors non identifiés.

Des signatures métabolomiques ont déjà fait leurs preuves pour discriminer l'inflammation stérile du sepsis avec des scores encourageants. Par exemple, chez le traumatisé sévère, pathologie reconnue comme favorisant les états pro-inflammatoires et comme étant à risque de sepsis secondaire, Blaise et al. ont mis en avant, sur une petite cohorte une signature métabolomique plasmatique permettant de discriminer les patients traumatisés à risque de sepsis secondaire lors de leur admission en service de soins critiques ( $R^2 = 0.855$ ,  $Q^2 = 0.384$ , AUC de 0,778). Au sein de cette signature, il existait *des variations importantes de la valine, du citrate, de l'aspartate, de l'allantoïne et de l'hydroxybutyrate* (Blaise et al. 2013). Chez le traumatisé sévère atteint de sepsis, l'analyse métabolomique non ciblée sur sérum de Feng et al. ont identifiés trois biomarqueurs potentiels (*semialdéhyde de l'acide succinique, uracile et uridine*) qui pourraient être utile au diagnostic de sepsis dans cette population de patients (K. Feng et al. 2022).

Sur un modèle animal, puis après validation chez l'homme, Langley et al. mettent en avant un modèle à 4 métabolites, *kynurénine, tauroolithocholate-3-sulfate* (reflet de la dysfonction hépatique), *acyl-carnitine* (reflet de la dysfonction mitochondriale) et *acyl-GPC* (reflet de l'activation plaquettaire) qui permettait de discriminer le sepsis du SIRS (Syndrome de réponse inflammatoire systémique) (Langley et al. 2014). Mickiewicz et al. retrouvent, quant-à-eux, une signature métabolomique à 31 métabolites séparant le groupe de patients atteints de choc septique de ceux présentant un SIRS sans infection avec une AUC de 0,98. Cette signature multi-marqueur comprenait principalement des métabolites impliqués dans le métabolisme énergétique, des acides aminés à chaîne ramifiée et des métabolites reflets de défaillance d'organes (Mickiewicz et al. 2014).

Un des principaux enjeux, dans la prise en charge du patient septique aux urgences, est l'identification des patients à risque de mauvaise évolution clinique, afin de les intégrer à une filière adaptée et de débiter le plus rapidement possible la prise en charge thérapeutique et donc d'en améliorer le pronostic (Seymour et al. 2017). L'approche métabolomique pourrait, une fois de plus, aider à prédire la mortalité (Ferrario et al. 2016), la réponse au traitement (Cambiaghi et al. 2017; Ferrario et al. 2016) ou même à identifier la présence de dysfonctions organiques spécifiques, telles que la dysfonction cardiaque (Aletti et al. 2016; Cao et al. 2023) ou le SDRA (Chang et al. 2023).

Pour exemple, la baisse de concentration des acides aminés à chaîne ramifiée (*valine, leucine, isoleucine*), impliqués dans la synthèse des protéines musculaires, la fonction immunitaire et l'activation des voies mTOR, serait corrélée à la gravité du tableau septique et pourrait constituer une cible thérapeutique intéressante (Reisinger et al. 2021; She et al. 2023).

Le métabolisme des lipides est, lui aussi, fortement impacté par le sepsis et impacterait fortement le système immunitaire inné. Par exemple, les lipoprotéines participent à l'élimination des toxines bactériennes, à la libération de corticostéroïdes, à la protection de l'endothélium et à la réduction de la réponse inflammatoire des monocytes. Aux côtés des HDL, les LDL et les VLDL participent à cette élimination en liant et séquestrant le lipopolysaccharide (LPS). Le sepsis entraînerait des altérations qualitatives et quantitatives de ces lipides et donc une diminution leurs fonctions immunoprotectrices. Ces HDL dysfonctionnants possèderaient même une activité pro-inflammatoire (Barker et al. 2021). Des signatures lipidomiques commencent à voir le jour. Elles permettraient, elles aussi, de caractériser les patients septiques à risque de mauvais pronostic (D. Sulaiman et al. 2024). Comme décrit précédemment dans le contexte de la COVID-19, les médiateurs lipidiques spécialisés pro-résolutifs, semblent également élevés dans le plasma des patients septiques relevant d'une prise en charge de soins critiques et pourraient être corrélés avec l'évolution vers le SDRA (Dalli et al. 2017).

Ainsi, les métabolismes les plus impactés par le choc septique pourraient être ceux du métabolisme énergétique, des acides aminés et des lipides avec notamment une augmentation de la glycolyse anaérobie, de la protéolyse, de la lipolyse et de la néo-glucogenèse (Thooft et al. 2023). Une revue de la littérature comprenant 27 articles s'intéressant à l'approche métabolomique pronostique dans le sepsis semble aller dans le même sens. Elle conclue que les métabolites présents dans les signatures métabolomiques de ces études étaient principalement des molécules lipidiques, des acides organiques et des acides carboxyliques. Les potentiels biomarqueurs les plus fréquemment décrits étaient, par ordre de fréquence, les suivants : *lactate, acetylcartinine, phénylalanine, isoleucine, urée, glutamine, kynurénine, 3-hydroxyisobutyrate, citrate, fumarate, tyrosine, alanine, indolacetate, pyruvate,  $\gamma$ -glutamylphenylalanine, leucine, 3-(4-hydroxyphenyl)lactate, valine, 5-methylthioadenosine, butyrolcarnitine et taurocholentae sulfate*. (S. Feng et al. 2023).

L'identification rapide de changements spécifiques dans l'empreinte métabolique des patients septiques pourrait, à l'avenir, être utilisée pour améliorer le monitoring des patients et développer de nouvelles thérapies ciblées.

Les sciences métabolomiques, qui s'intègrent et se combinent avec toutes les sciences « omiques » ouvrent donc une nouvelle voie dans le diagnostic et la prise en charge du sepsis, avec une approche plus centrée sur la réponse du patient que sur la détection du pathogène. Elles peuvent offrir la perspective d'un traitement personnalisé voire prédire la réponse à un traitement (Eckerle et al. 2018; Isac et al. 2024). Cependant les données de la littérature sont encore hétérogènes. Ceci pouvant en partie être expliqué par des procédures analytiques non encore parfaitement standardisées et par des cohortes hétérogènes, comme par exemple, des patients inclus à différents stades du processus infectieux notamment fréquemment inclus dans des services de soins critiques donc présentant déjà un certain niveau de gravité. De plus, une fois fiabilisées, ces signatures métabolomiques devront être encore validées en pratique clinique.

L'approche est déjà peu coûteuse en particulier si la signature recherchée est déjà établie. Un spectromètre plus simple peut être alors utilisé. Le développement de préparation et d'analyse automatisée et la miniaturisation des outils d'analytiques, ainsi que le développement de l'intelligence artificielle pour le post-traitement des données, devraient permettre, à terme, des analyses métabolomiques ciblées simples et d'obtention rapides, facilitant ainsi l'utilisation des signatures métabolomiques en pratique de routine voire directement dans nos services d'urgences.

# Metabolomics, the future of biomarkers?

Céline Occelli<sup>a,b,c</sup>, Jacques Levraut<sup>a,b</sup> and Thierry Pourcher<sup>c</sup>

European Journal of Emergency Medicine 2024, 31:7–8

<sup>a</sup>Department of Emergency Medicine, University Hospital of Nice, <sup>b</sup>Université Côte d'Azur, School of Medicine and <sup>c</sup>Université Côte d'Azur, Commissariat à l'Énergie Atomique et aux Énergies Alternatives (CEA), Direction de la Recherche Fondamentale (DRF), Institut des Sciences du Vivant Frédéric Joliot, Transporter in Imaging and Radiotherapy in Oncology Laboratory (TIRO), School of Medicine, Nice, France

Correspondence to Céline Occelli, MD, Department of Emergency Medicine (Pôle Urgences), CHU Pasteur 2, 30, avenue de la voie Romaine, F06100 – Nice, France  
Tel: +33 618847104; e-mail: occelli.c@chu-nice.fr

Received 19 September 2023 Accepted 21 September 2023.

Accurate diagnosis and early assessment of the severity of a patient's condition are crucial in emergency medicine. They enable rapid management, leading to improved clinical outcomes in many pathologies, in particular sepsis. Clinicians measure vital parameters and use early warning scores and biomarkers to help with diagnosis and assessment. Although these tools are widely used in current practice, their performance is still often debated [1]. As the clinical parameters assessed are mainly related to the macrocirculatory system, they may deteriorate relatively late in the progression of the pathology giving the clinician false reassurance. Furthermore, new emerging pathologies such as COVID-19 have highlighted the limits of our traditional diagnostic and prognostic tools [2].

Metabolomics concerns the study of metabolites present in a biological sample (cells, tissues or biological fluids). As metabolites are the products of cellular metabolism, an analysis of the metabolites in a sample can indicate the organism's metabolic response to a physiological or pathological stimulus. The human metabolome currently comprises some 42 000 identified metabolites, including lipids, amino acids, organic acids, sugars, hormones and exogenous metabolites [3]. Thanks to technological advances in mass spectrometry, it is now possible to dynamically evaluate variations in this metabolome, and thus identify multi-marker molecular signatures that directly reflect the functions of the cell or organ, and therefore detect changes linked to the presence of a pathology or exposure to a toxic agent. The diagnosis and prognostic evaluation of a pathology need no longer depend on a single biomarker, or on the combination of several independently assessed biomarkers, but could be based on complex molecular signatures that may include variations of several metabolites. In the future, a single targeted metabolomics analysis combining the biomarkers of several pathologies could also enable joint diagnosis and prognostic evaluations. To date, metabolomics studies have already proved their value in the development of new diagnostic and prognostic tools, and have also improved our understanding of

pathophysiology and, consequently, patient management in pathologies of interest to the emergency care domain. For example, metabolomics allowed the identification of new prognostic biomarkers in traumatic brain injury patients that may serve as indicators for optimizing management [4]. Experimental data also showed promise in identifying the various phenomena leading to acute renal failure [5]. Multi-omics analysis also enabled us to identify biomolecules specific to non-survival in acute liver failure [6]. Early identification of these metabolomic signatures could then be used for patient stratification, enabling appropriate therapeutic management and referral to suitable further care.

Furthermore, a non-targeted metabolomics approach also enables novel molecular signatures to be rapidly identified in response to new emerging pathologies, as demonstrated by the COVID-19 experience. Several publications report that variations in plasma lipids and metabolites of the kynurenate pathway are relevant biomarkers for the diagnosis and prognosis of SARS COV-2 infection [7].

In the field of infectiology, metabolomics could allow identification of the early dysfunction of the host response to infection and of the pro-inflammatory processes of immunosuppression. Metabolomic signatures have already proved their worth in discriminating sterile inflammation from sepsis, with encouraging scores (area under the curve 0.778 for the study by Blaise *et al.* [8] and area under the curve 0.78 and 0.82 for the 4-metabolite model by Langley *et al.*, validated on two cohorts [9]). Variations in metabolites involved in energy metabolism, amino acids and lactate are also found in the metabolomes of septic patients [10].

Once the diagnosis is established, it remains important to rapidly determine the potential severity of the septic picture and identify patients at risk of unfavorable evolution. Once again, the metabolomics approach could help predict mortality [11], response to treatment [12] and even identify the presence of specific organ dysfunctions, such as cardiac dysfunction [13] or ARDS [14]. The rapid identification of specific changes in the metabolomic fingerprint of septic patients could, in the future, be used to improve the stratification of patients and to develop new targeted therapies.

Metabolomics could be combined with other 'omics' sciences such as transcriptomics which, for example, has validated a classifier combining 29 host mRNAs that predicts the severity of infected patients more effectively

than the usual biomarkers or clinical scores [15]. These promising omics results open up a new avenue in the diagnosis and management of sepsis, with an approach more focused on patient response than on pathogen detection, that may offer the prospect of personalized treatment. For clinical routine, metabolomics could be a fast and inexpensive assay, although metabolomic signatures must first be validated on larger cohorts. Moreover, new automated liquid chromatography and tandem mass spectrometry systems, coupled with artificial intelligence for data post-processing, should soon enable simple and rapid targeted metabolomic analyses, thus facilitating routine clinical use.

### Acknowledgements

The authors acknowledge Catherine Buchana for linguistic proofreading of the manuscript.

### Conflicts of interest

There are no conflicts of interest.

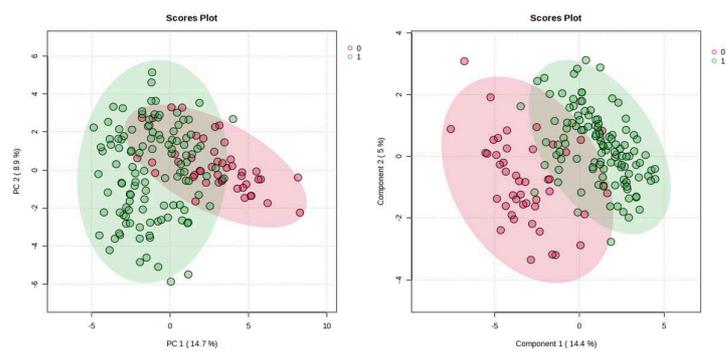
### References

- Ocelli C, Contenti J. Early warning scores: are they clinically relevant? *Eur J Emerg Med* 2020; **27**:325–326.
- Malinverni S, Nuñez M, Cotton F, Martiny D, Collot V, Konopnicki D, et al. Is procalcitonin a reliable marker of bacterial community-acquired pneumonia in adults admitted to the emergency department during SARS-CoV-2 pandemic? *Eur J Emerg Med* 2021; **28**:312–314.
- Qiu S, Cai Y, Yao H, Lin C, Xie Y, Tang S, et al. Small molecule metabolites: discovery of biomarkers and therapeutic targets. *Signal Transduct Target Ther* 2023; **8**:1–37.
- Kharouba M, Patel DD, Jaber RH, Mahmoud SH. Metabolomic analysis in neurocritical care patients. *Metabolites* 2023; **13**:745.
- Patschan D, Patschan S, Matyukhin I, Hoffmeister M, Lauxmann M, Ritter O, et al. Metabolomics in acute kidney injury: the experimental perspective. *J Clin Med Res* 2023; **15**:283–291.
- Sharma N, Pandey S, Yadav M, Mathew B, Bindal V, Sharma N, et al. Bio-molecular map of albumin identifies signatures of severity and early mortality in Acute Liver Failure. *J Hepatol* 2023; **79**:677–691.
- Ocelli C, Guignon JM, Lindenthal S, Cagnard A, Graslin F, Brglez V, et al. Untargeted plasma metabolomic fingerprinting highlights several biomarkers for the diagnosis and prognosis of coronavirus disease 19. *Front Med* 2022; **9**:1–13.
- Blaise BJ, Gouel-Chéron A, Floccard B, Monneret G, Allaouchiche B. Metabolic phenotyping of traumatized patients reveals a susceptibility to sepsis. *Anal Chem* 2013; **85**:10850–10855.
- Langley RJ, Tipper JL, Bruse S, Baron RM, Tsalik EL, Huntley J, et al. Integrative 'omic' analysis of experimental bacteremia identifies a metabolic signature that distinguishes human sepsis from systemic inflammatory response syndromes. *Am J Respir Crit Care Med* 2014; **190**:445–455.
- Mickiewicz B, Duggan GE, Winston BW, Doig C, Kubes P, Vogel HJ; Alberta Sepsis Network. Metabolic profiling of serum samples by 1H nuclear magnetic resonance spectroscopy as a potential diagnostic approach for septic shock. *Crit Care Med* 2014; **42**:1140–1149.
- Ferrario M, Cambiagli A, Brunelli L, Giordano S, Caironi P, Guatteri L, et al. Mortality prediction in patients with severe septic shock: a pilot study using a target metabolomics approach. *Sci Rep* 2016; **6**:1–11.
- Cambiagli A, Pinto BB, Brunelli L, Falcetta F, Aletti F, Bendjelid K, et al. Characterization of a metabolomic profile associated with responsiveness to therapy in the acute phase of septic shock. *Sci Rep* 2017; **7**:1–16.
- Cao Y, Liu Z, Ma W, Fang C, Pei Y, Jing Y, et al. Untargeted metabolomic profiling of sepsis-induced cardiac dysfunction. *Front Endocrinol (Lausanne)* 2023; **14**:1–13.
- Chang Y, Yoo HJ, Kim SJ, Lee K, Lim CM, Hong SB, et al. A targeted metabolomics approach for sepsis-induced ARDS and its subphenotypes. *Crit Care* 2023; **27**:263.
- Galtung N, Diehl-Wiesenecker E, Lehmann D, Markmann N, Bergström WH, Wacker J, et al. Prospective validation of a transcriptomic severity classifier among patients with suspected acute infection and sepsis in the emergency department. *Eur J Emerg Med* 2022; **29**:357–365.

# ANNEXES

## Annexe 1. Analyses statistiques MetaboAnalyst® comparaison métabolomes urinaires des sujets non- COVID (sains et symptomatiques) et des sujets COVID (légèrement, modérément et sévèrement atteints)

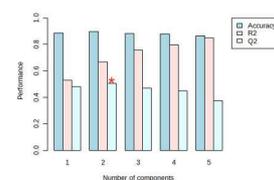
### Analyses statistiques MetaboAnalyst Tous métabolites Log transformation



PCA

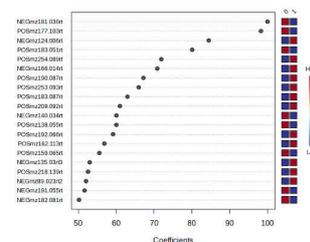
PLS-DA

- Non-COVID (sujets sains et symptomatiques non-COVID)
- COVID (légèrement, modérément et sévèrement atteints)

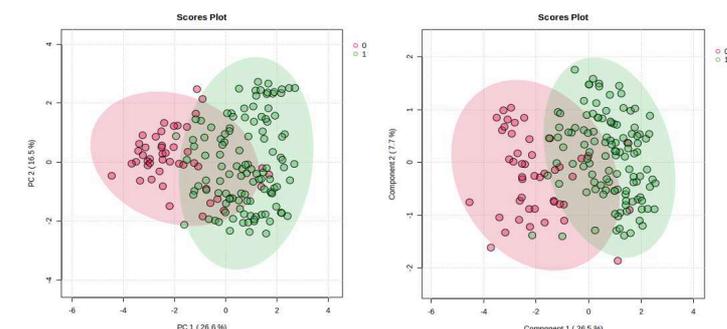


PLS-DA cross validation details:

Measure	1 comps	2 comps	3 comps	4 comps	5 comps
Accuracy	0.88314	0.89527	0.88295	0.87689	0.86477
R2	0.52988	0.66498	0.75571	0.79789	0.84899
Q2	0.47837	0.50563	0.46777	0.44629	0.37497



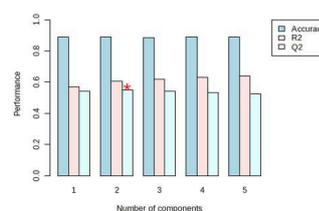
### Analyses statistiques MetaboAnalyst Top liste 20 métabolites Log transformation



PCA

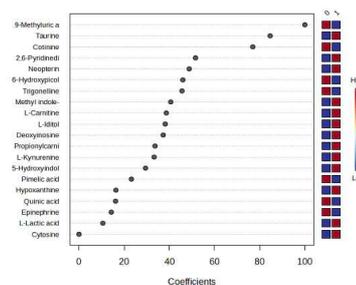
PLS-DA

- Non-COVID (sujets sains et symptomatiques non-COVID)
- COVID (légèrement, modérément et sévèrement atteints)



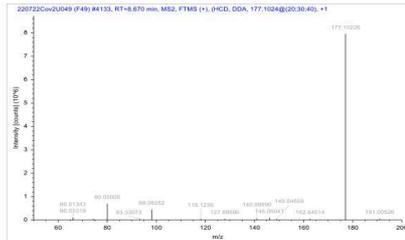
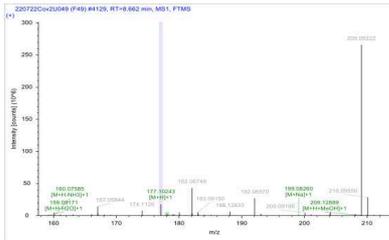
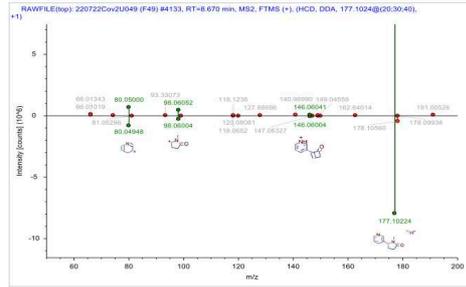
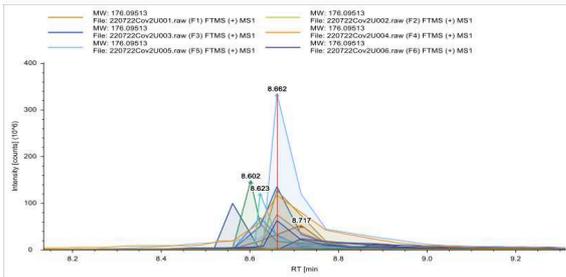
PLS-DA cross validation details:

Measure	1 comps	2 comps	3 comps	4 comps	5 comps
Accuracy	0.88958	0.88977	0.88314	0.88939	0.88958
R2	0.56774	0.60617	0.61863	0.6295	0.63925
Q2	0.54072	0.54698	0.54097	0.53076	0.52576



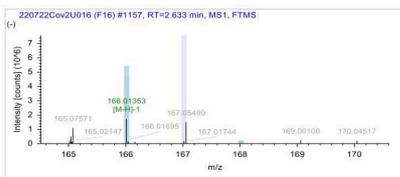
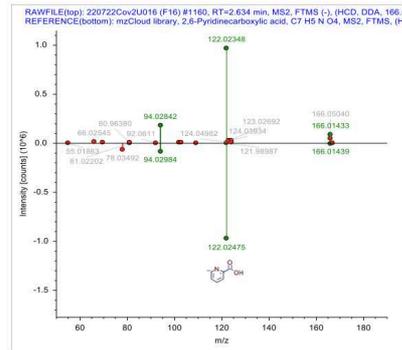
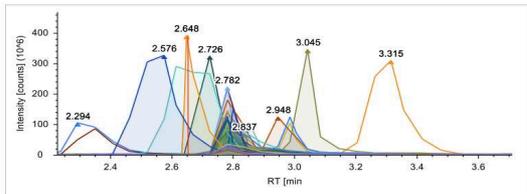


**HMDB001046 – POSmz177.103rt8.91**  
**Cotinine**

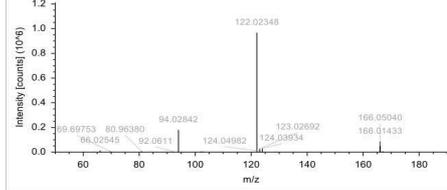


m/z	Intensity (counts)
66.01019	116743
66.01143	85488
74.41127	50470
80.05	700324
83.33073	45476
98.06052	450989
118.12383	50964
123.08996	51124
140.9899	69131
146.09041	93380
149.04559	49542
162.64014	53950
177.10226	792432
191.05028	58911

**HMDB0033161 – NEGmz166.014rt2.77**  
**2,6-Pyridinedicarboxylic acid**

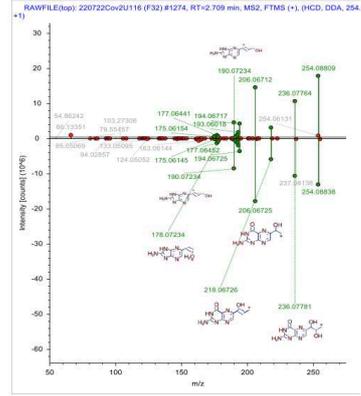
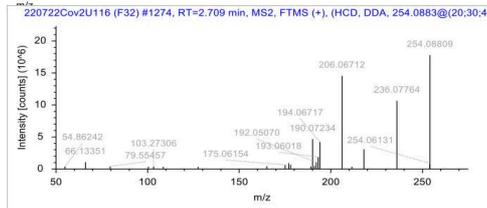
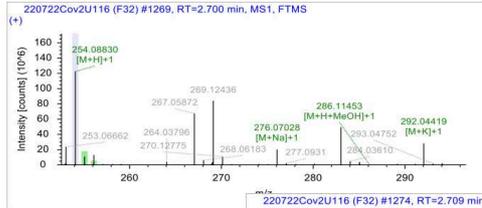
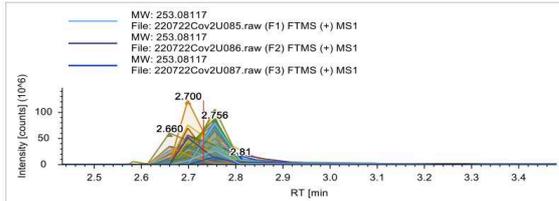


220722Cov2U016 (F16) #1160, RT=2.634 min, MS2, FTMS (-), (HCD, DDA, 166.0135@20:30:4)



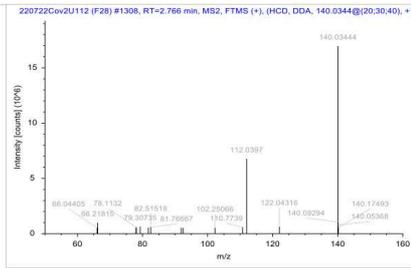
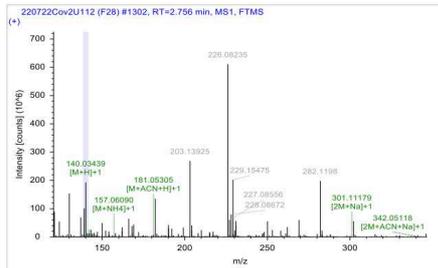
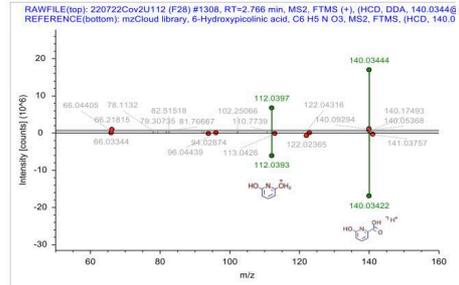
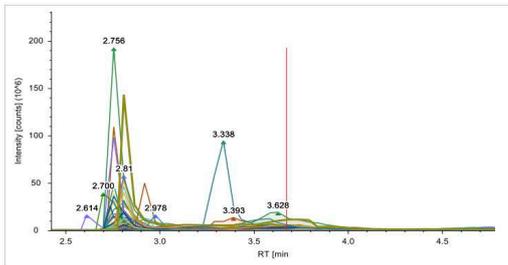
m/z	Intensity (counts)
66.02545	13100
69.69753	5767
80.9638	5921
92.06113	4949
94.02842	37959
102.02588	5645
102.71333	5634
122.02348	968028
123.02692	27235
124.03934	30961
124.04982	7008
146.01433	86691
166.0504	58338
167.03834	4706
191.05041	509

**HMDB000845 – POSmz254.089rt2.72**  
**Neopterin**



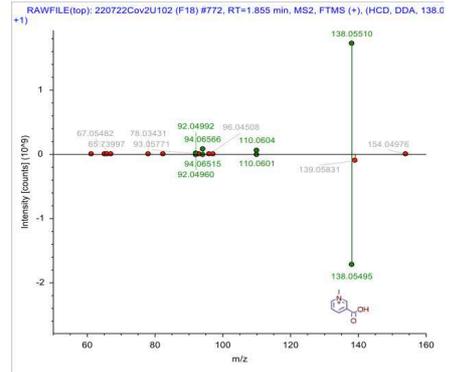
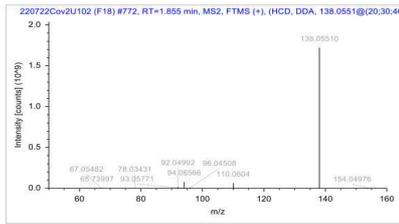
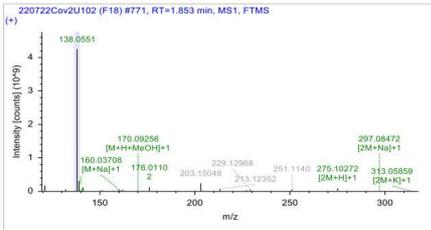
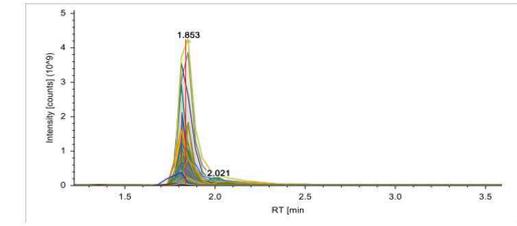
m/z	Intensity [counts]
54.86242	370213
66.13351	307274
79.55457	397739
103.27306	346229
108.59901	344878
127.53368	349459
165.06522	428671
175.06154	648114
177.06441	930054
178.07175	681183
189.08424	398590
190.07234	466193
191.05992	492196
192.05070	1038429
193.06018	187332
194.06717	4198957
206.06712	14571059
211.45938	372735
218.06732	3082424
236.07764	10073890
254.06131	889917
254.08809	17810690

**HMDB0013188 – NEGmz140.034rt2.59**  
**6-Hydroxypicolinic acid**

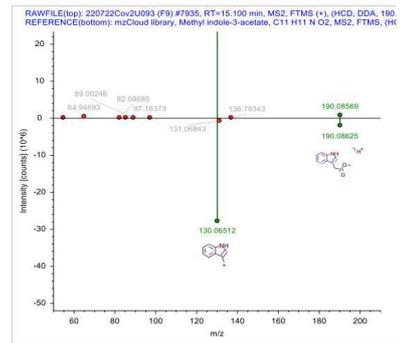
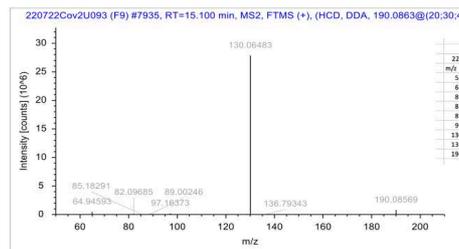
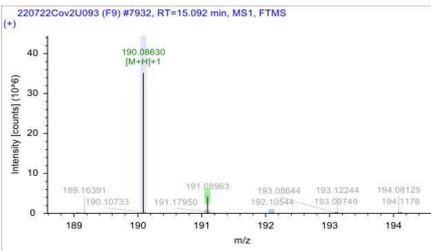
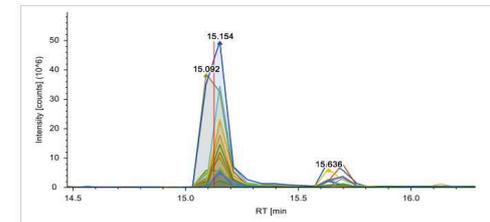


m/z	Intensity [counts]
66.04405	523758
68.12815	92296
77.97491	508454
78.11128	576314
79.30735	630747
81.76667	538511
82.51518	584560
91.91208	528816
92.49226	528236
102.25066	933080
110.77398	585112
112.0397	6766615
122.04316	633784
140.03444	16969580
140.05368	761932
140.09294	1897637
140.17493	540669

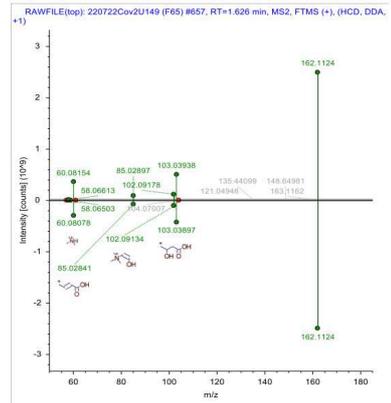
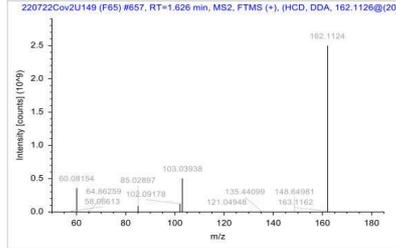
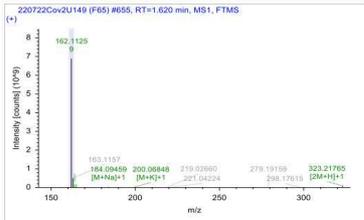
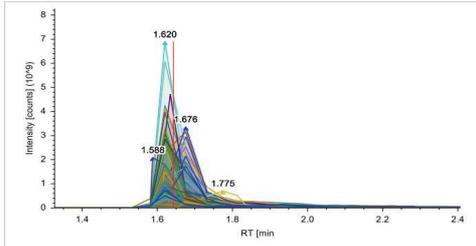
**HMDB000875 – POSmz138.055rt1.83**  
**Trigonelline**



**HMDB0029738 – POSmz190.087rt15.21**  
**Methyl indole-3-acetate**

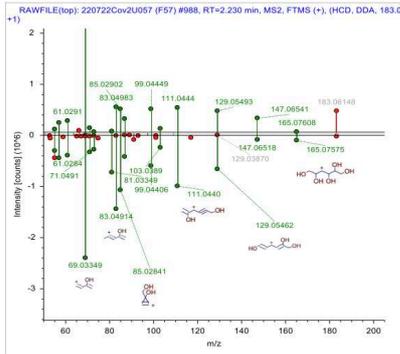
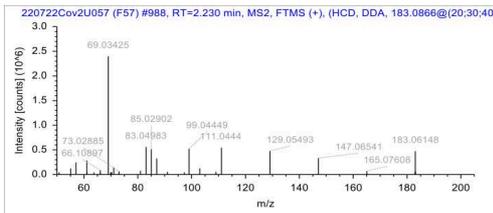
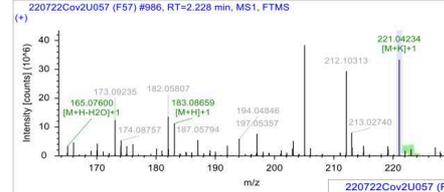
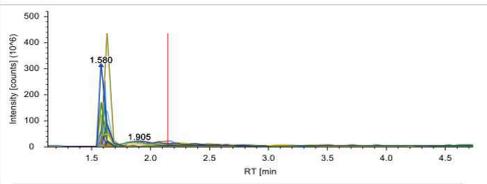


**HMDB000062 – POSmz162.113rt1.64**  
**L-Carnitine**



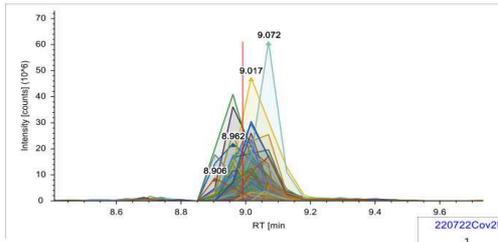
m/z	Intensity [counts]
50.14479	1481583
50.62368	1117094
57.03435	3672875
58.06613	13445860
59.0737	3514031
60.08154	35755000
61.52237	1181451
63.75962	1264901
64.86229	7879713
85.02897	8976950
102.09178	122549800
103.03938	50358500
121.04948	3688017
135.44099	1408648
148.64981	1636055
162.11249	249762800
163.11624	2683111

**HMDB0011632 – POSmz183.087rt1.64**  
**L-Iditol**

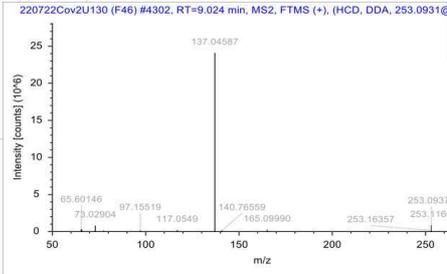
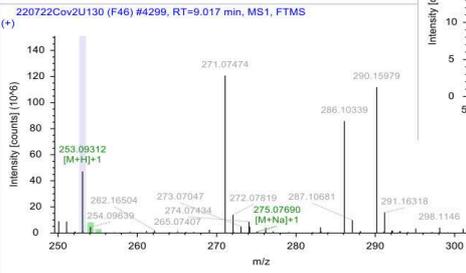
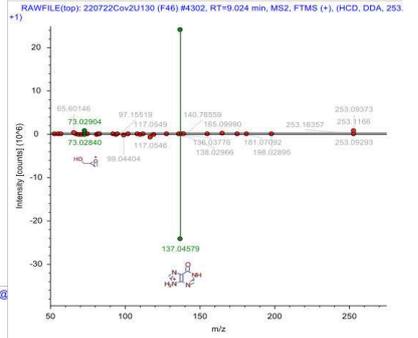


m/z	Intensity [counts]
56.78351	44600
59.05496	129193
59.0737	246893
61.02913	285264
61.0405	59794
63.70705	43575
66.10897	88628
66.1121	55347
69.03425	2399326
69.30755	45333
70.3076	48473
71.04984	140944
73.02885	62393
81.03425	78007
83.04983	528169
85.02902	512921
87.04457	326772
91.03928	52902
97.28601	44789
99.04449	520351
103.03965	125180
109.04835	54470
111.04445	545405
129.05493	477733
147.06541	338804
165.07608	69104
183.06148	475425
188.08722	52841
188.11307	60347

**HMDB000071 – POSmz253.093rt9.04**  
**Deoxyinosine**

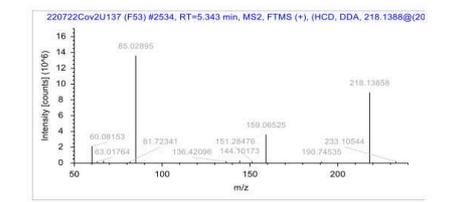
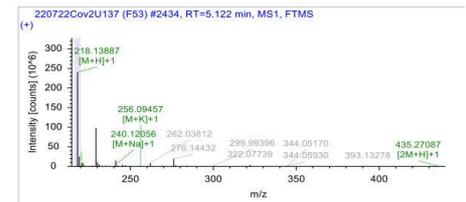
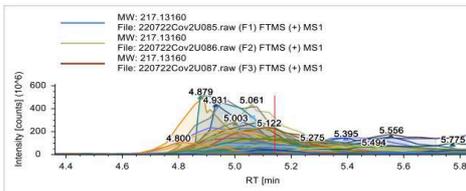


220722Cov2U130 (F46) #4302, RT=9.024 min, MS2, FTMS (+), (HCD, DDA, 253.0931@

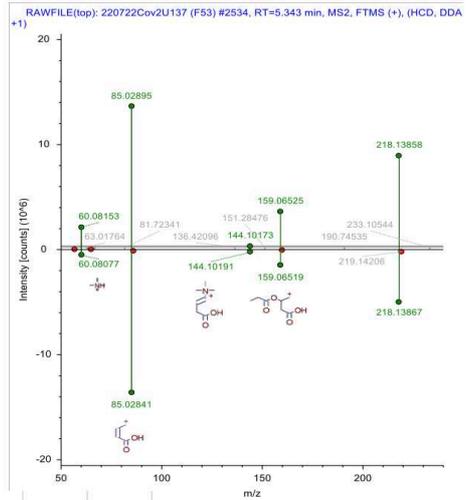


m/z	Intensity [counts]
65.5978	210448
65.60146	318908
65.60813	144000
65.94601	135612
73.02904	805477
77.51139	141670
87.15519	151416
117.05491	203696
137.04587	24101460
140.76559	131159
140.76559	243790
155.0999	243790
253.09373	154488
253.11668	854096
253.16357	149576

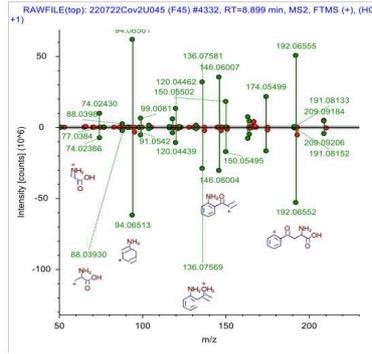
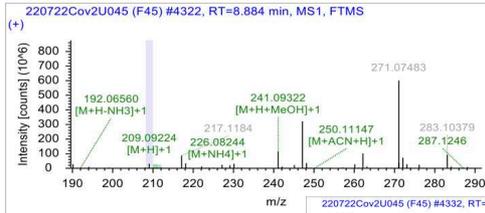
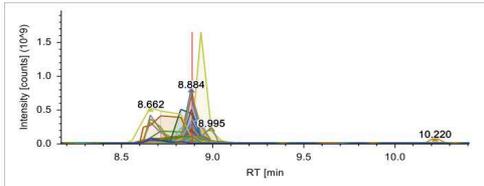
**HMDB0000824 – POSmz218.1388rt5.43**  
**Propionylcarnitine**



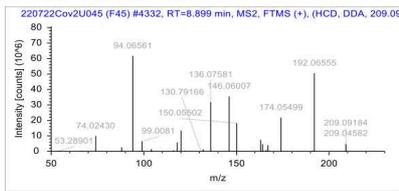
m/z	Intensity [counts]
60.08153	2136699
63.01764	213490
66.55646	271120
81.72341	198507
85.02895	13605610
136.42096	211843
144.10173	322129
151.28476	213026
159.06525	3613359
190.74535	200590
218.13858	8944492
233.10544	196393



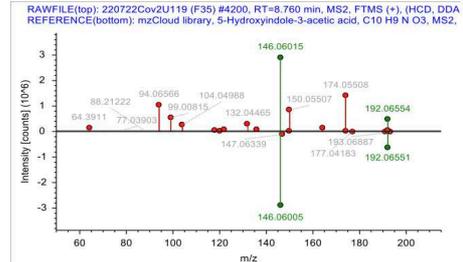
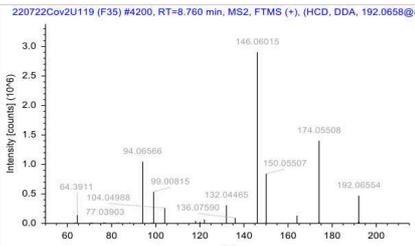
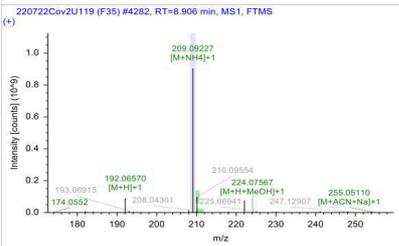
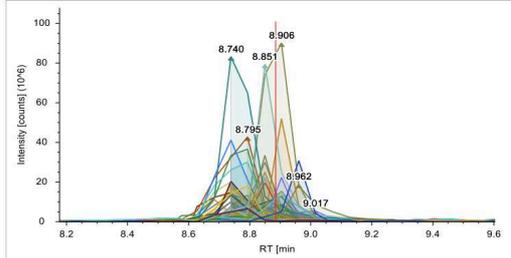
**HMDB0000684 – POSmz209.092rt8.88  
L-Kynurenine**



m/z	Intensity [count]
51.2854	147796
51.2801	125781
59.02749	119821
69.09322	141317
65.98471	287067
67.86361	108861
74.0243	902782
77.03934	156077
85.18137	133181
88.07985	2034326
91.04641	172094
94.06561	6127260
95.12785	106383
98.02625	211967
99.00811	6521022
102.96739	135600
104.04988	1495942
111.91002	131313
116.0495	670201
118.06538	1794712
120.04462	13514390
122.06001	395887
130.79166	113571
132.04465	1303788
136.07581	3176790
146.06007	3950900
148.11148	145783
150.0602	23356
150.0502	1810950
153.0863	748234
154.0761	465545
156.0124	862378
167.0556	4020070
167.0569	16485
168.0197	278545
174.0558	2383550
191.0558	242593
191.08139	48339
192.0655	5066070
209.04582	1049060
209.0184	444180
211.1909	140054

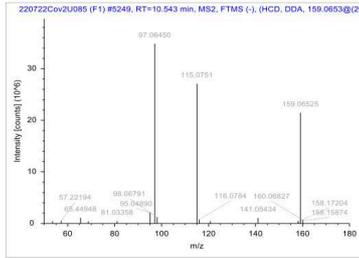
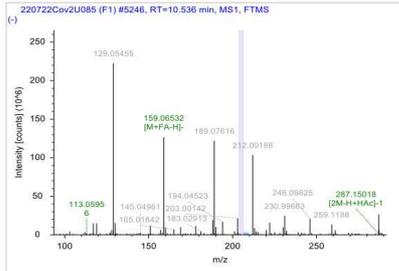
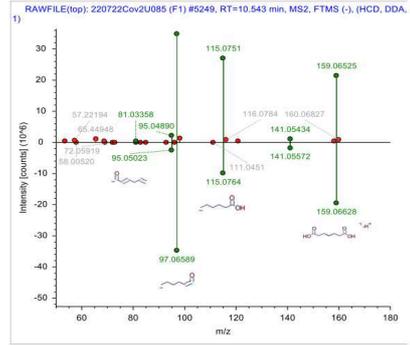
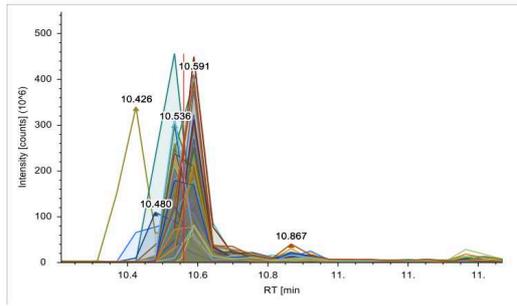


**HMDB0000763 – POSmz192.066rt8.89  
5-Hydroxyindole-3-acetic acid**



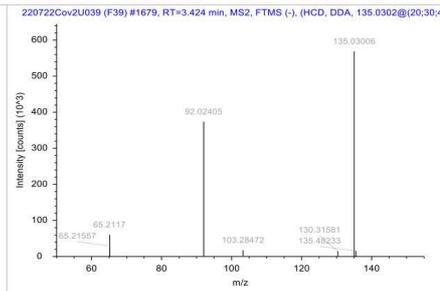
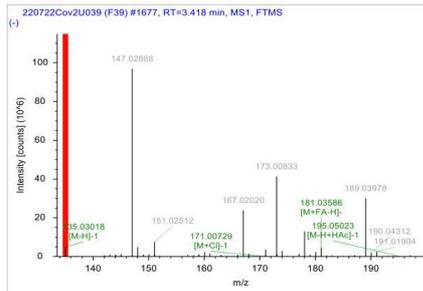
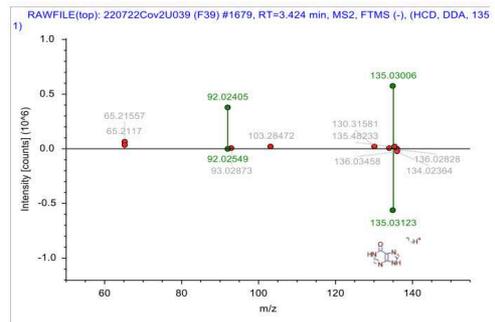
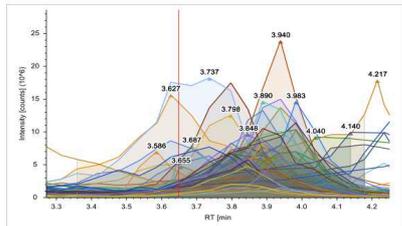
m/z	Intensity [count]
63.52547	14725
64.39111	142936
70.66737	15078
71.48687	14200
76.24927	15544
77.03903	17709
82.16885	13675
83.23363	13105
88.21222	16512
94.06566	1049282
98.20961	14442
99.00815	516083
104.04988	261985
118.06538	45199
120.04469	26458
120.08121	15089
122.06039	66925
132.04465	307931
136.0759	93448
146.06015	2904470
150.0507	844511
150.06729	26143
164.07066	133329
174.05556	1402894
174.14861	24014
192.06554	472650
192.08125	27264
192.1015	25516
192.15967	33116

**HMDB0000857 – POSmz159.065rt10.12**  
**Pimelic acid**



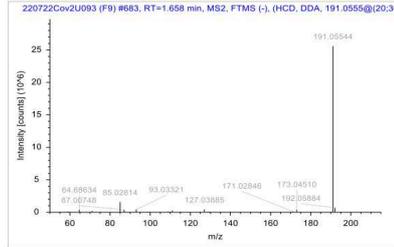
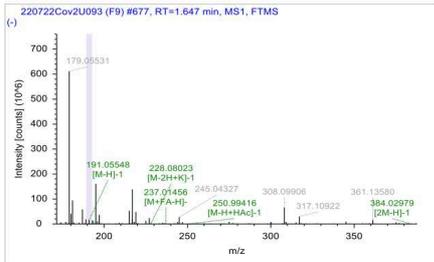
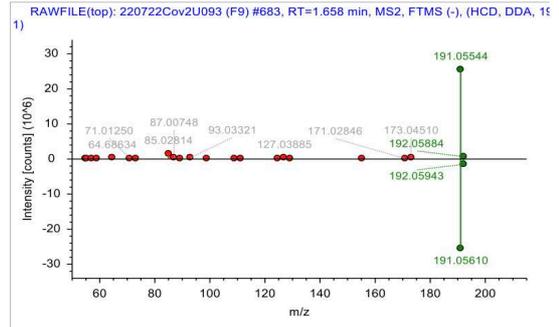
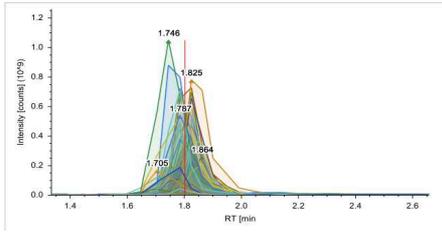
m/z	Intensity [counts]
57.21194	520475
57.2194	520475
65.44948	1108191
68.7539	429713
81.03358	654127
95.0489	2145329
97.0645	34838560
98.06791	1156414
115.07517	27068860
116.07847	791257
120.7849	437559
141.05434	1093884
158.15874	441651
158.17204	40732
159.06525	2145290
160.06827	716571

**POSmz137.046rt2.82 – POSmz183.087rt1.64**  
**Hypoxanthine**



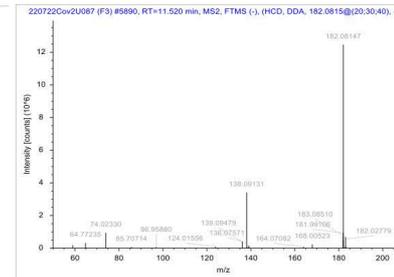
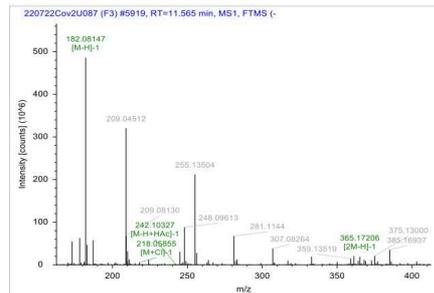
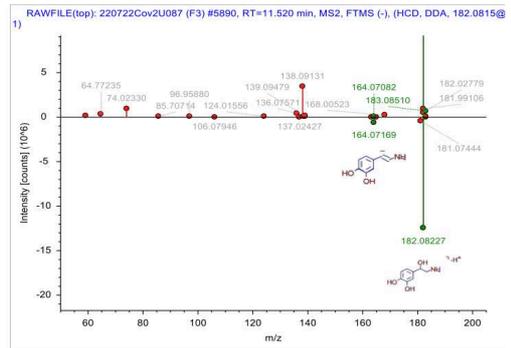
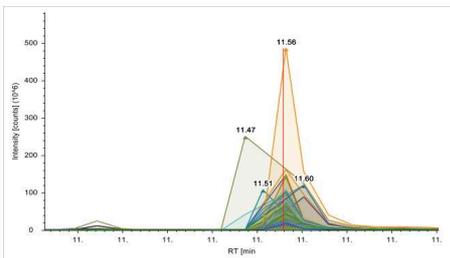
m/z	Intensity [counts]
65,21178	60186
65,21557	29079
92,02405	373806
103,28472	17597
130,31581	15059
135,03006	568630
135,48233	15727

**HMDB0003072 – NEGmz191.055rt1.79**  
**Quinic acid**



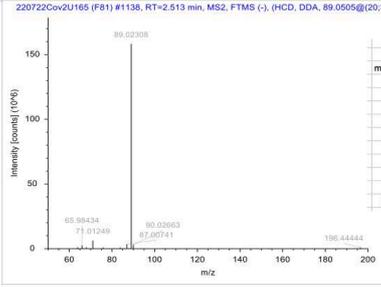
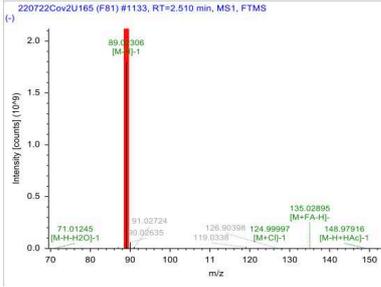
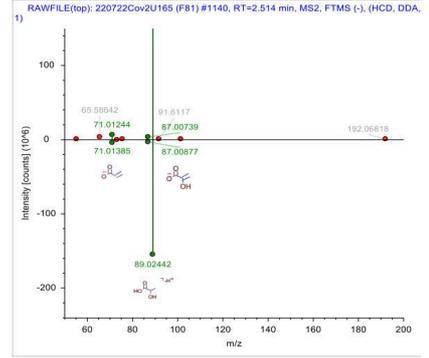
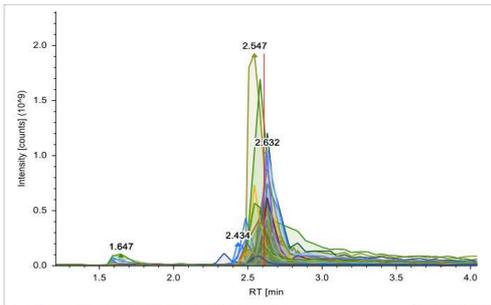
m/z	Intensity [counts]
54.82524	56789
55.09191	48219
57.03318	56619
59.01252	124262
64.68634	361533
71.0125	213503
73.02795	81146
85.02814	573912
87.00748	374209
89.02336	69208
93.03321	402352
99.04404	90004
109.02856	119394
111.04386	245330
111.05065	58134
124.47247	59221
127.03885	416157
129.01828	104504
155.03456	57857
171.02846	113626
173.0451	353153
191.05544	25586860
192.05884	683374

**HMDB0000068 – NEGmz182.081rt11.47**  
**Epinephrine**



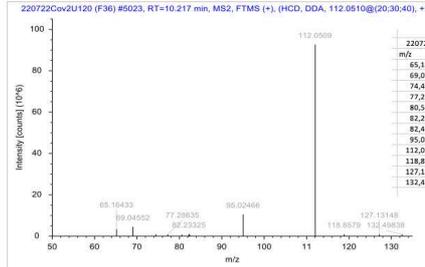
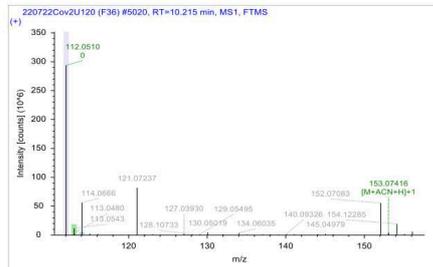
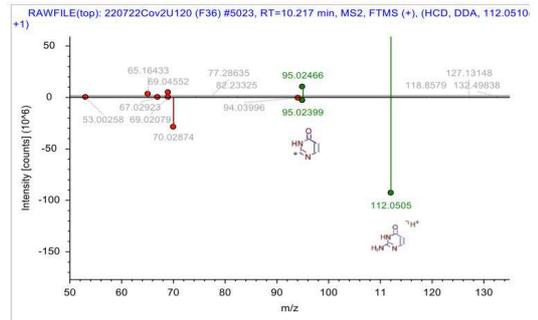
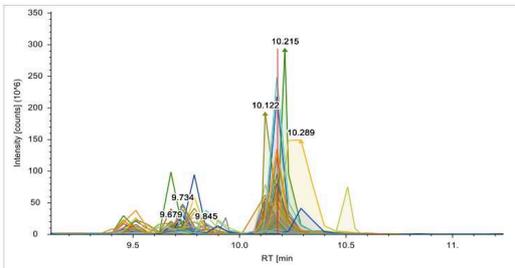
m/z	Intensity [counts]
59.01242	183647
64.77235	321283
74.0223	360442
85.70714	65383
96.9588	54257
124.01556	87195
136.07571	415521
138.09131	3409346
138.33284	55613
139.07555	81852
139.09479	145974
164.07082	74872
168.00523	233140
181.99106	932591
182.02779	539763
182.08147	12454530
183.02801	96552
183.0851	677324
183.10049	94833

**HMDB0000190 – NEGmz89.023rt2.6**  
**L-Lactic acid**



m/z	Intensity (counts)
63.94404	1099759
65.98434	2609985
67.96803	1121697
71.01249	6357533
75.86484	942560
83.85836	917904
87.00741	3496281
89.02308	15820000
90.02663	3198579
196.44444	989395

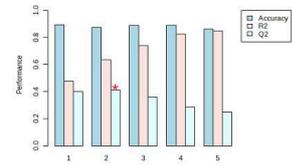
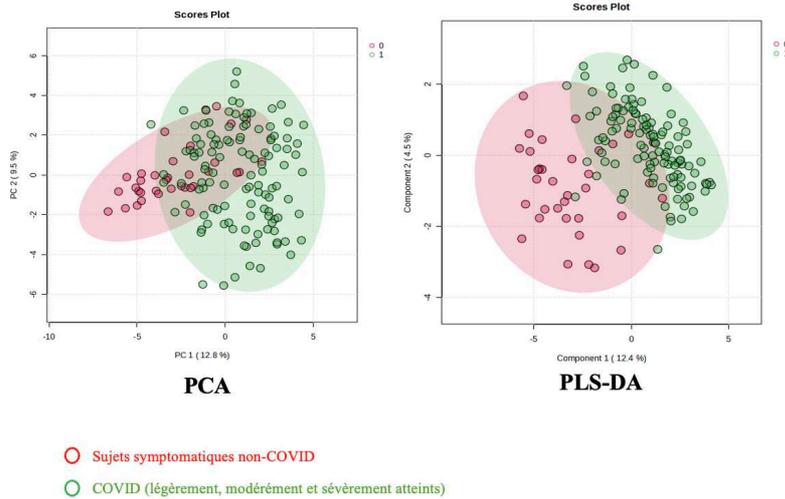
**HMDB0000630 – POSmz112.049rt9.94**  
**Cytosine**



m/z	Intensity (counts)
65.16433	3351468
69.04552	4478039
74.41448	920462
77.28635	920881
80.55111	827415
82.23325	990357
84.41233	761270
95.02466	10538080
112.05095	6273140
118.85792	938185
127.13148	896428
132.49638	933743

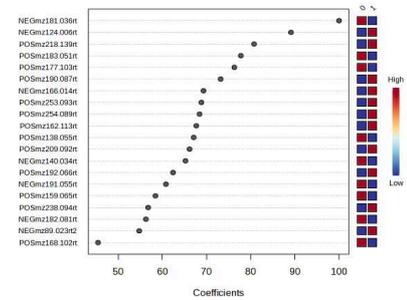
### Annexe 3. Analyses statistiques MetaboAnalyst® comparaison métabolomes urinaires des symptomatiques non-COVID et des sujets COVID (légèrement, modérément et sévèrement atteints)

#### Analyses statistiques MetaboAnalyst Tous métabolites Log transformation

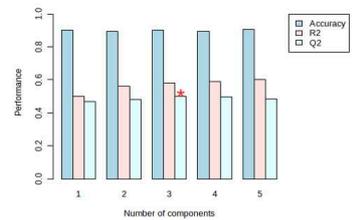
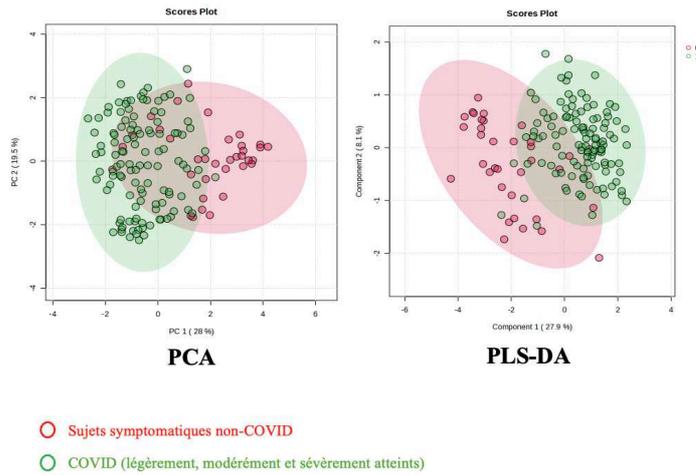


#### PLS-DA cross validation details:

Measure	1 comps	2 comps	3 comps	4 comps	5 comps
Accuracy	0.89441	0.87462	0.86774	0.86753	0.86172
R2	0.47497	0.63593	0.73883	0.82366	0.84809
Q2	0.39742	0.41019	0.35811	0.28476	0.25101

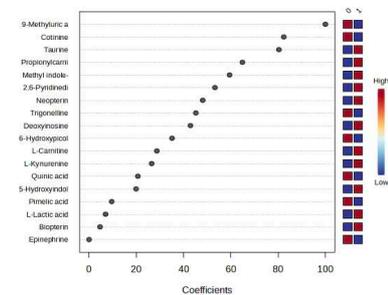


#### Analyses statistiques MetaboAnalyst Top liste 18 métabolites Log transformation



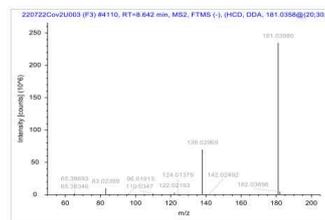
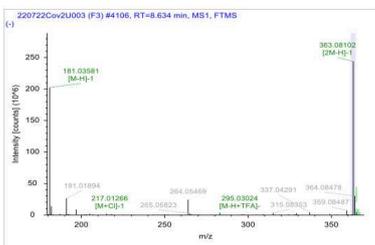
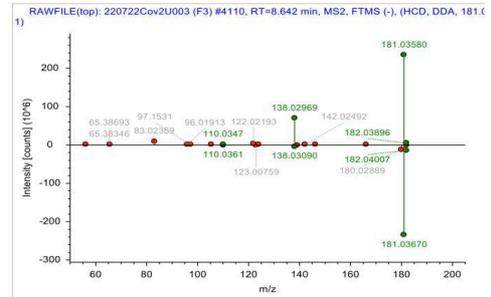
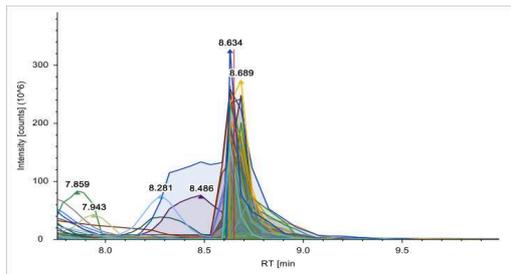
#### PLS-DA cross validation details:

Measure	1 comps	2 comps	3 comps	4 comps	5 comps
Accuracy	0.90086	0.89388	0.90043	0.89388	0.9071
R2	0.50169	0.56207	0.58173	0.58809	0.60278
Q2	0.46883	0.47983	0.49912	0.49695	0.48382



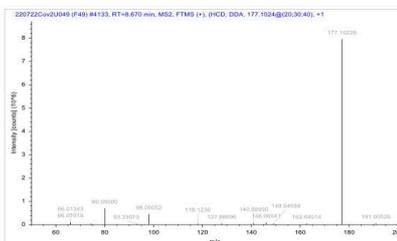
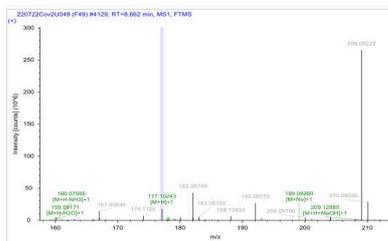
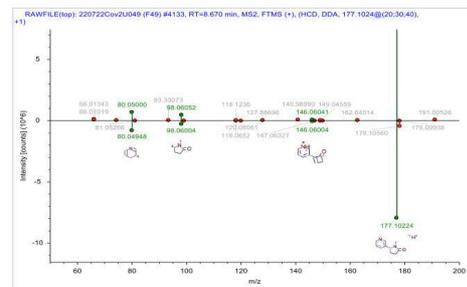
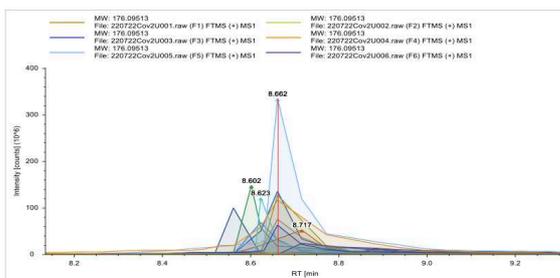
# Annexe 4. Validation top liste 18 métabolites de la comparaison métabolomes urinaires des symptomatiques non-COVID et des sujets COVID (légèrement, modérément et sévèrement atteints)

**HMDB0003099 - NEGmz181.036rt8.75**  
**9-Methyluric acid**



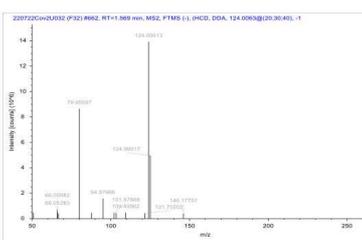
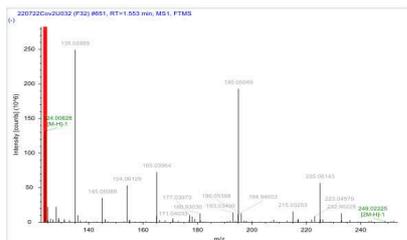
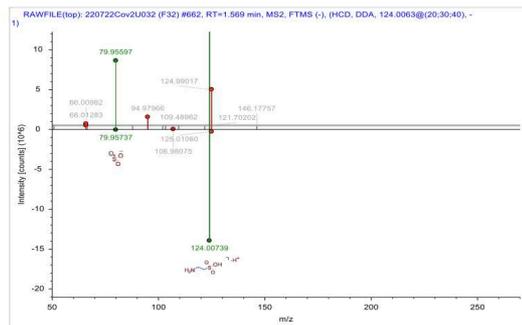
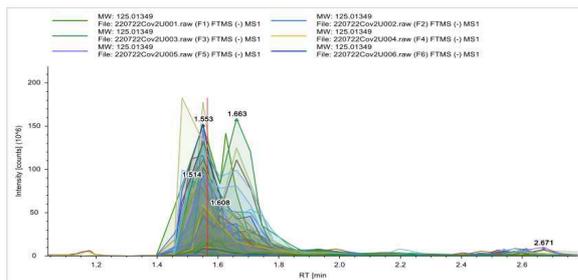
m/z	Intensity [count]	FTMS (+)	[HCD	DDA	181.0368@(-2 -)]
56.17252	686651				
65.38346	1921637				
65.38693	1746004				
83.02359	9991678				
96.01913	1527163				
97.15311	800339				
105.45144	684969				
110.03477	1803321				
122.02193	3239906				
124.01379	982652				
138.02969	6987264				
142.02492	1038634				
146.10616	898900				
166.01276	705650				
181.03558	234787600				
182.03896	4891546				

**HMDB0001046 – POSmz177.103rt8.91**  
**Cotinine**



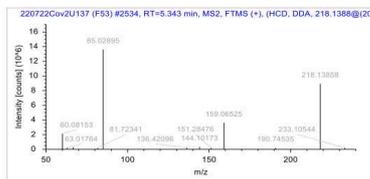
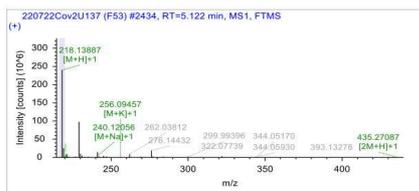
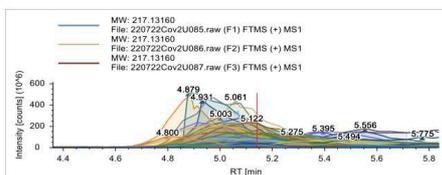
m/z	Intensity [count]
66.01019	116743
66.01311	81488
74.41127	10470
80.051	700234
93.33073	45476
98.09052	450989
118.12383	109644
127.88696	53124
146.06041	49131
146.06041	93180
148.04559	49542
146.06044	33090
177.10226	7912452
191.00526	58911

**HMDB000251 – NEGmz124.006rt1.64**  
**Taurine**

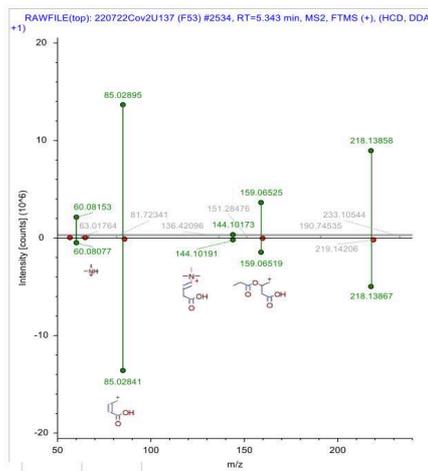


m/z	Intensity [counts]
50.82385	467608
66.00982	518866
66.01283	735505
66.70425	397427
79.95597	854930
87.79192	461966
94.97966	1586186
101.97888	462783
103.29534	450472
109.49862	478841
121.70202	424541
124.00613	13927970
124.99017	4982876
146.17757	395293

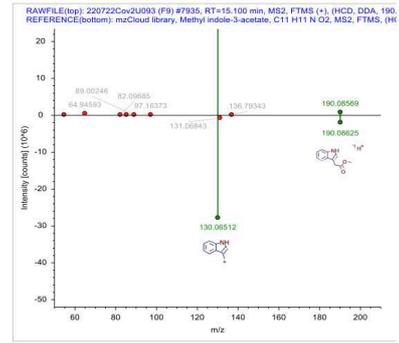
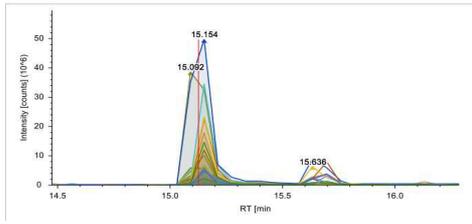
**HMDB000824 – POSmz218.139rt5.43**  
**Propionylcarnitine**



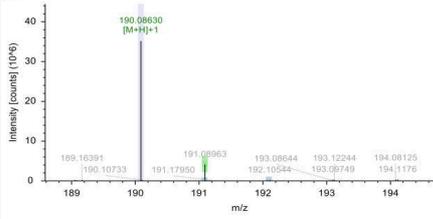
m/z	Intensity [counts]
50.08153	2136699
63.01764	213490
66.65646	271120
81.72341	198507
85.02895	13605610
136.42096	211843
144.10173	322129
151.28476	213026
159.06525	3613359
190.74535	200590
218.13858	8944492
233.10544	196393



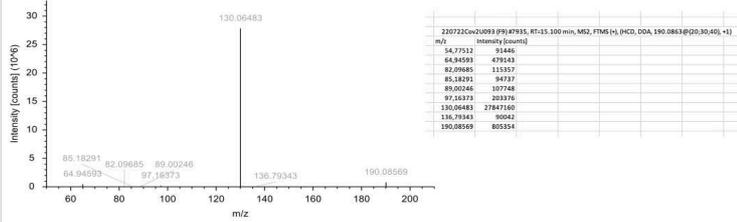
**HMDB0029738 – POSmz190.087rt15.21**  
**Methyl indole-3-acetate**



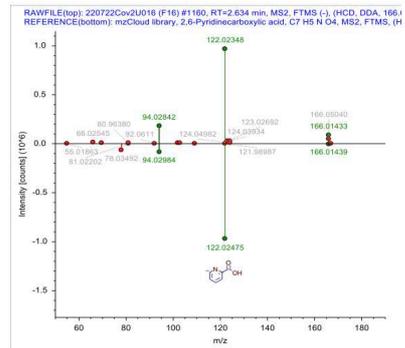
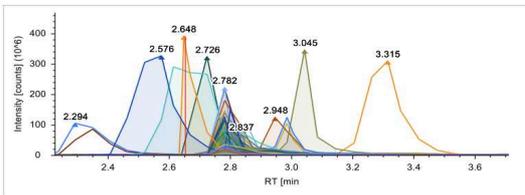
220722Cov2U093 (F9) #7932, RT=15.092 min, MS1, FTMS



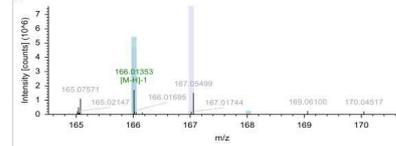
220722Cov2U093 (F9) #7935, RT=15.100 min, MS2, FTMS (+), (HCD, DDA, 190.0863@20:30:4)



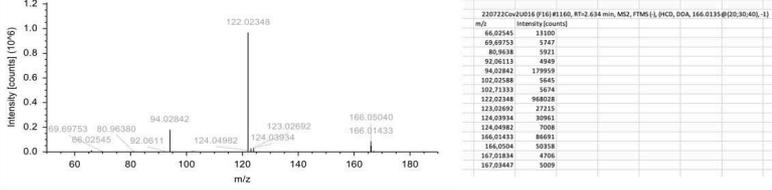
**HMDB0033161 – NEGmz166.014rt2.77**  
**2,6-Pyridinedicarboxylic acid**



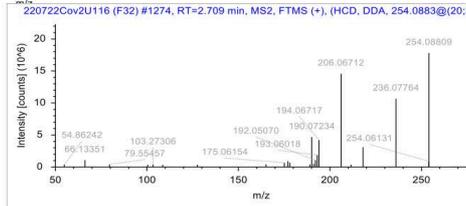
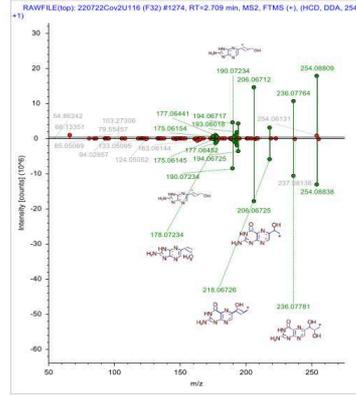
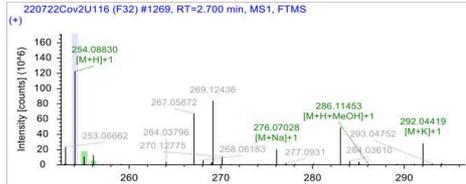
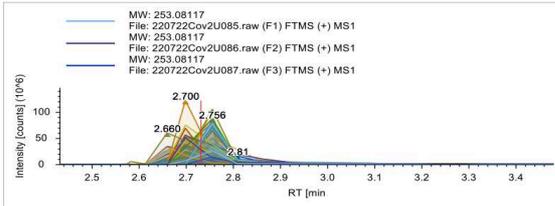
220722Cov2U016 (F16) #1157, RT=2.633 min, MS1, FTMS



220722Cov2U016 (F16) #1160, RT=2.634 min, MS2, FTMS (-), (HCD, DDA, 166.0135@20:30:4)

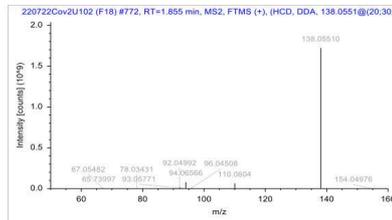
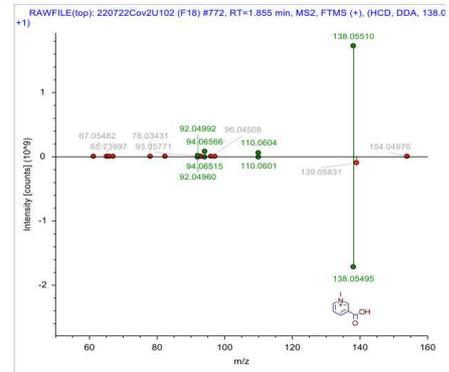
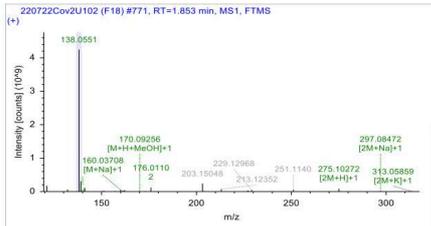
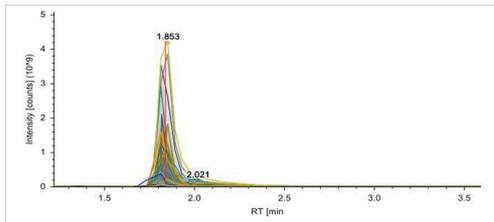


**HMDB000845 – POSmz254.089rt.72  
Neopterin**



m/z	Intensity [counts]
54.86242	370213
66.13351	1072474
68.05502	397739
69.13351	346279
70.55457	383365
73.55457	344878
75.06154	348859
77.06441	428671
78.55457	648114
84.86242	930644
86.05502	681183
87.06154	398590
89.13351	456383
90.23306	492196
92.04992	1038429
93.05771	187382
94.06556	419857
96.04508	1457100
97.24666	372735
98.05502	3082424
99.06154	10671890
100.06712	889917
101.07334	1781090

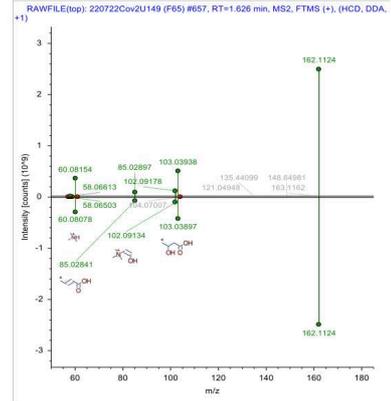
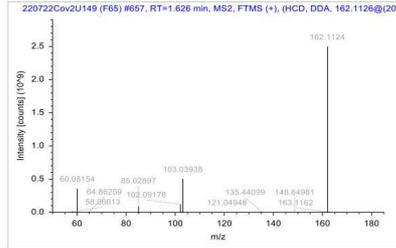
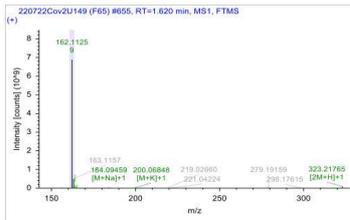
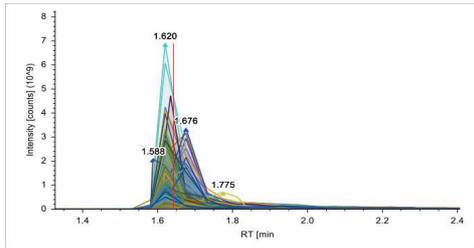
**HMDB000875 – POSmz138.055rt1.83  
Trigonelline**



m/z	Intensity [counts]
61.15607	1137514
65.03928	1457184
67.05482	3684211
78.03431	1842103
82.04992	1948551
83.34281	1081631
92.04992	16265710
93.05771	4564802
94.06556	79871820
96.04508	8583304
97.24666	1277318
110.06041	6412460
138.0551	1721738000
154.04976	2374093

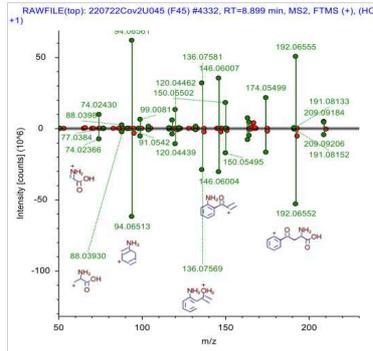
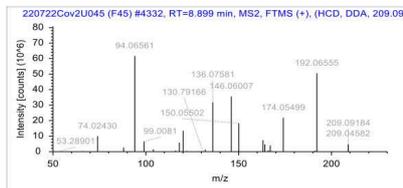
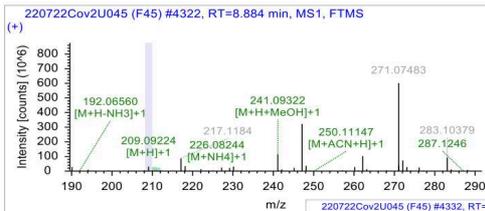
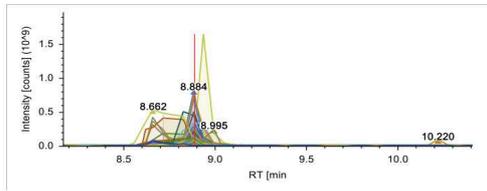


**HMDB000062 – POSmz162.113rt1.64  
L-Carnitine**



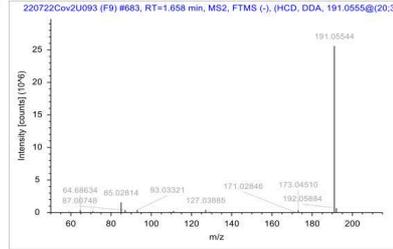
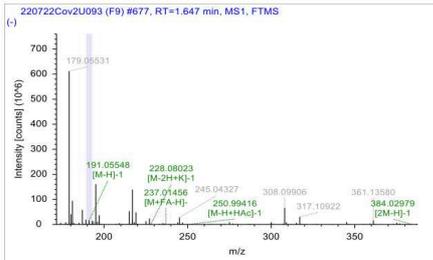
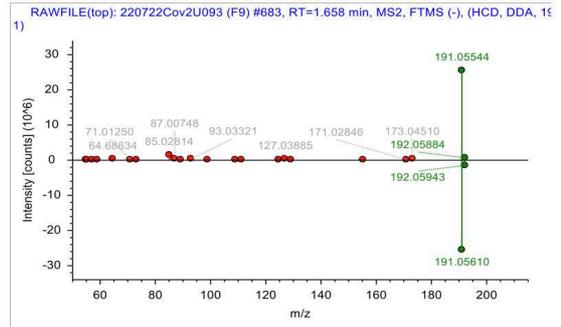
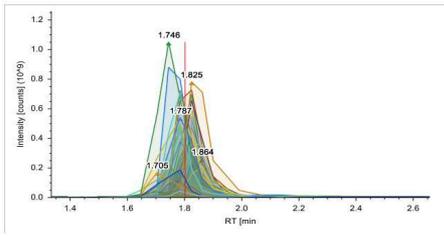
m/z	Intensity [counts]
50.14479	1481683
50.62368	1117054
57.03435	3672875
58.06613	13445660
59.0737	3514611
60.08154	357595000
61.5237	1181451
63.75982	1264901
64.86259	7879073
85.02897	8976950
102.09178	122549800
103.03938	505583500
121.04948	36888017
135.44099	1408648
148.64981	1636055
162.11249	2497628000
163.11624	2983111

**HMDB0000684 – POSmz209.092rt8.88  
L-Kynurenine**



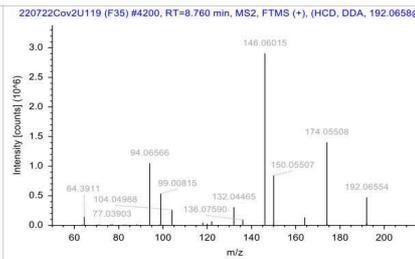
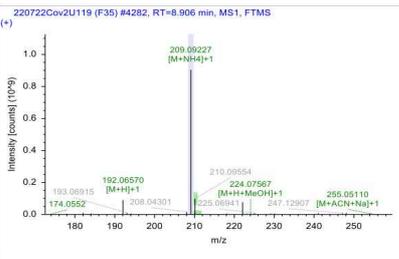
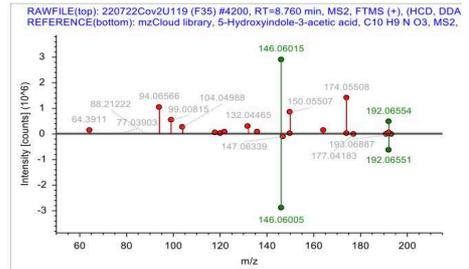
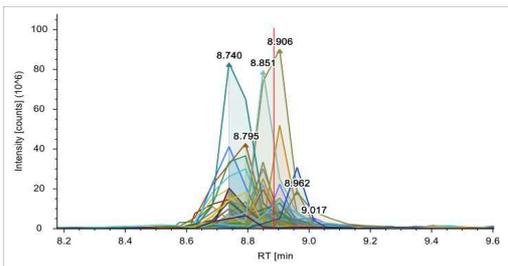
m/z	Intensity [counts]
51.02364	147756
52.0905	128781
59.09743	119821
65.01922	141117
65.08171	207907
67.88261	1200005
74.02430	997340
77.03934	158077
85.18137	123181
88.03985	324326
91.05441	172094
94.06561	61721560
95.29785	106383
98.02425	211917
99.00811	631202
102.96239	135000
104.0498	149961
111.0192	131313
116.0395	670101
118.0638	579672
120.04462	13518280
122.04601	39887
136.07166	118371
137.04443	139789
138.07581	31762760
146.06007	35030400
148.11148	145783
150.0002	253366
150.05502	38320939
151.08019	184824
164.07051	4835545
166.03726	845178
167.03954	4029070
167.05569	164836
168.03917	27805
174.05499	21818250
201.05558	242359
191.08133	484293
192.06555	50360750
200.04562	1040900
209.09184	4442180
211.09009	140054

**HMDB0003072 – NEGmz191.055rt1.79**  
**Quinic acid**



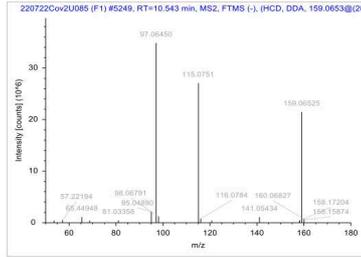
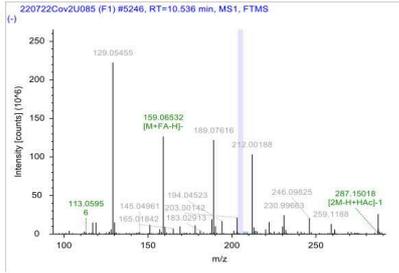
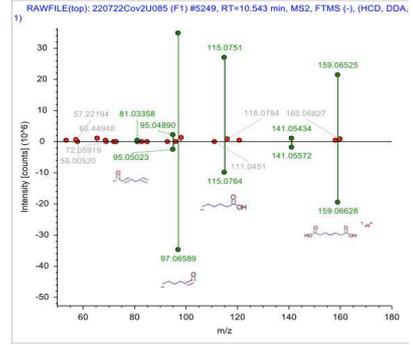
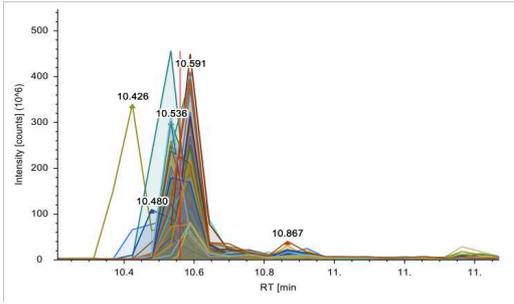
m/z	Intensity [counts]
54.82324	56789
55.09191	48229
57.03318	56619
59.01252	124262
64.08634	361533
71.0125	213503
73.02795	81146
85.02844	157912
87.00748	374209
89.02336	69208
93.03221	402952
99.04404	90004
109.02856	119394
111.04386	245530
111.05065	58134
124.47247	59221
127.03885	416157
129.01828	104054
135.03456	57857
171.02846	113626
173.0451	351513
191.05544	2558680
192.05884	683374

**HMDB0000763 - POSmz192.066rt8.89**  
**5-Hydroxyindole-3-acetic acid**



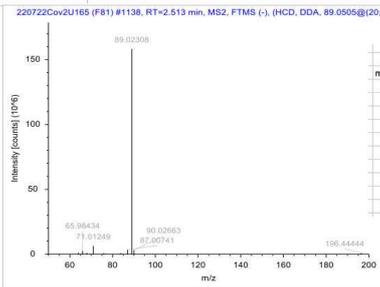
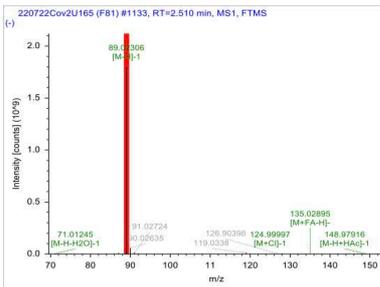
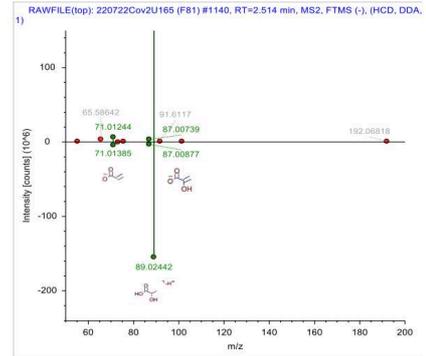
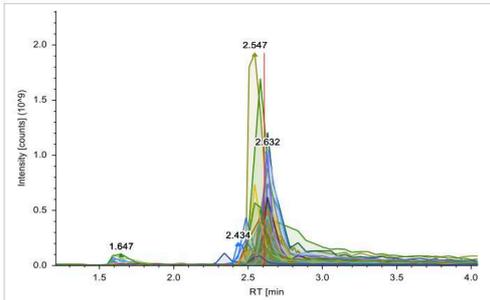
m/z	Intensity [counts]
63.5247	14725
64.3911	142936
70.66737	15078
71.48687	14200
76.36927	15546
77.03903	17709
82.16885	13975
83.23163	13305
88.21222	16512
94.06566	1049282
98.20961	14432
99.00815	536083
104.04988	261985
118.06538	45139
120.04449	26458
120.08121	15089
122.06039	69925
132.04465	307931
136.0759	93448
146.06015	2904470
150.05507	814851
150.06729	26143
164.07066	133329
174.05508	1402894
174.14861	24014
192.06554	472650
192.0825	27264
192.1015	22516
192.15967	33116

**HMDB0000857 – POSmz159.065rt10.12**  
**Pimelic acid**



m/z	Intensity [counts]
53.56002	439629
57.22194	520475
65.44948	1108191
65.7559	423713
81.03358	454127
95.0489	2145329
97.0645	3485890
98.00791	1196414
115.07517	27068860
116.07847	791257
120.7839	437559
141.05434	1083884
158.15874	441651
158.17204	407322
159.06525	2146290
160.06827	716571

**HMDB0000190 – NEGmz89.023rt2.6**  
**L-Lactic acid**

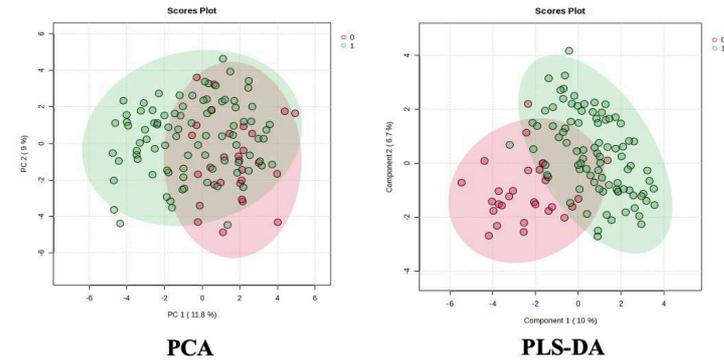


m/z	Intensity [counts]
63.94404	1099759
65.98434	2609985
67.96803	1121697
71.01249	6357533
75.86484	942560
83.85836	917904
87.00741	3496281
89.02308	15820000
90.02663	3198579
196.44444	989395

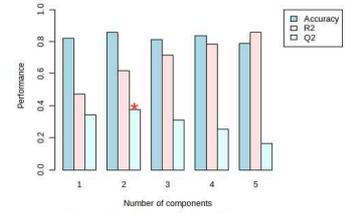


# Annexe 5. Analyses statistiques MetaboAnalyst® comparaison métabolomes urinaires des sujets légèrement atteints par la COVID-19 et des sujets atteints par la COVID-19 et hospitalisés (modérément et sévèrement atteints)

## Analyses statistiques MetaboAnalyst Tous métabolites Log transformation

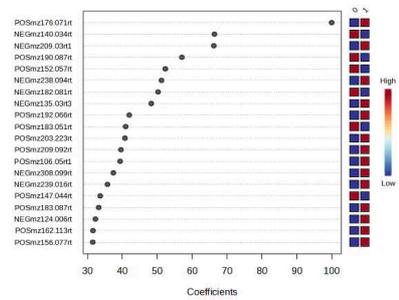


- Sujets légèrement atteints par la COVID-19
- Sujets atteints par la COVID-19 et hospitalisés (modérément et sévèrement atteints)

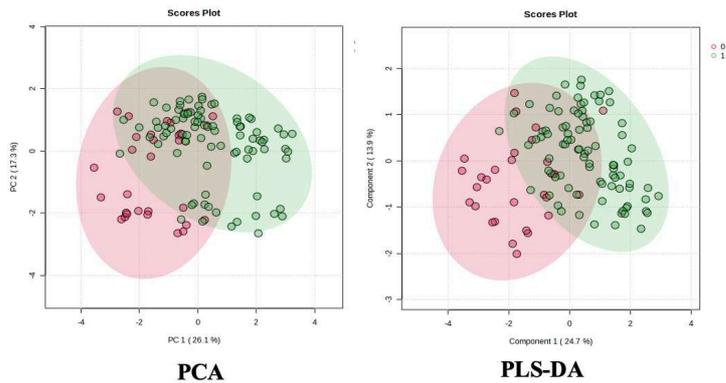


### PLS-DA cross validation details:

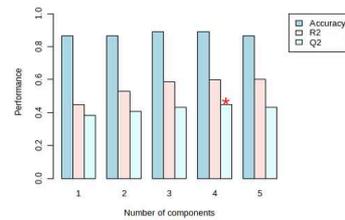
Measure	1 comps	2 comps	3 comps	4 comps	5 comps
Accuracy	0.82138	0.8558	0.81232	0.83841	0.78659
R2	0.47001	0.61675	0.7152	0.78286	0.83812
Q2	0.34125	0.37426	0.3103	0.25379	0.16533



## Analyses statistiques MetaboAnalyst Top liste 20 métabolites Log transformation

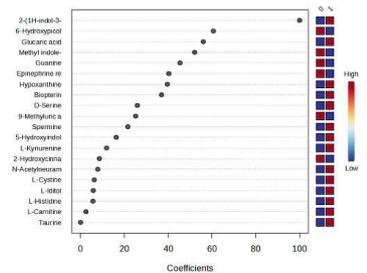


- Sujets légèrement atteints par la COVID-19
- Sujets atteints par la COVID-19 et hospitalisés (modérément et sévèrement atteints)



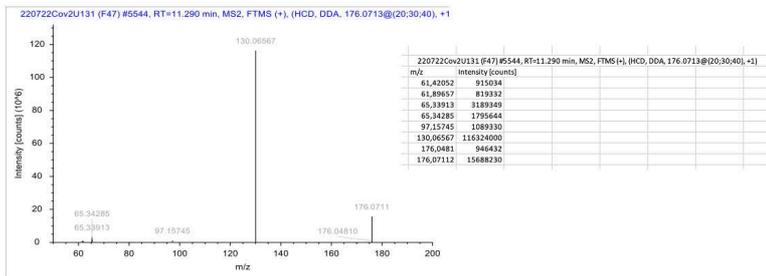
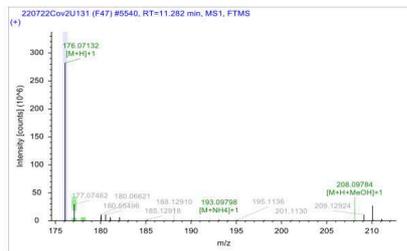
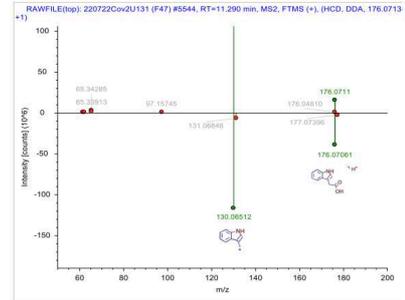
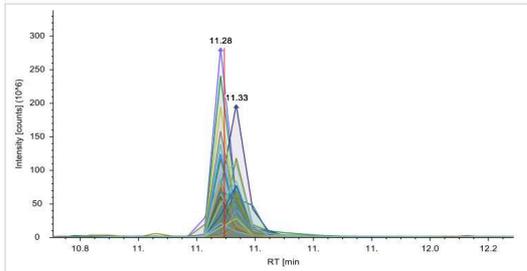
### PLS-DA cross validation details:

Measure	1 comps	2 comps	3 comps	4 comps	5 comps
Accuracy	0.86294	0.86294	0.88946	0.88946	0.8637
R2	0.44604	0.52807	0.58481	0.59707	0.60162
Q2	0.38249	0.40577	0.43077	0.44597	0.43307

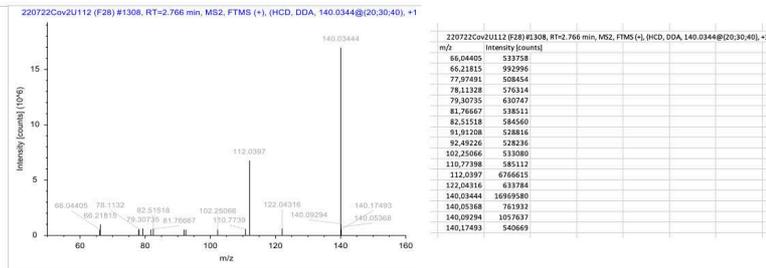
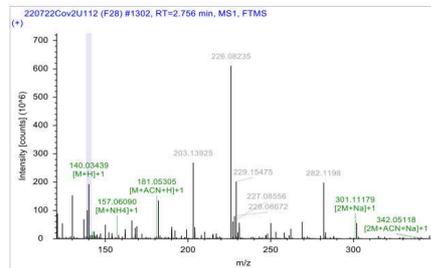
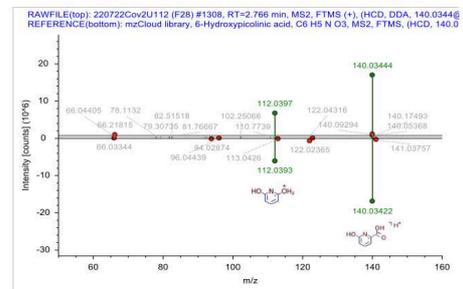
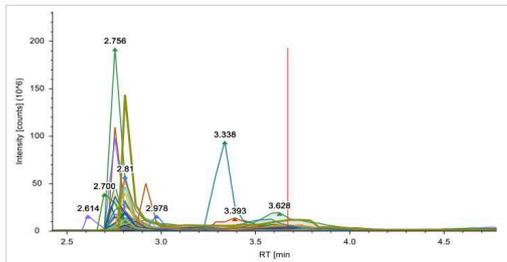


## Annexe 6. Validation top liste 20 métabolites de la comparaison métabolomes urinaires des sujets légèrement atteints par la COVID-19 et des sujets atteints par la COVID-19 et hospitalisés (modérément et sévèrement atteints)

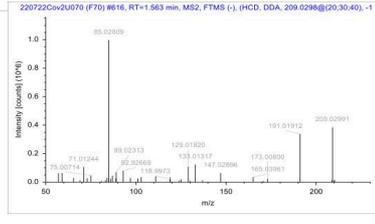
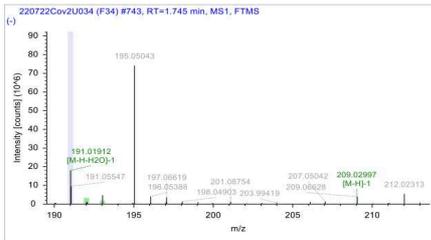
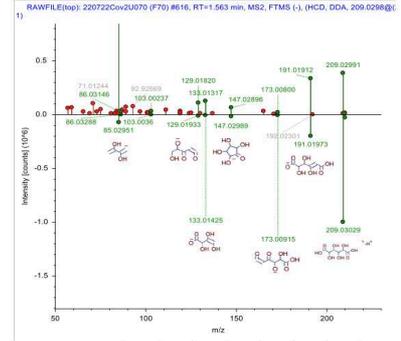
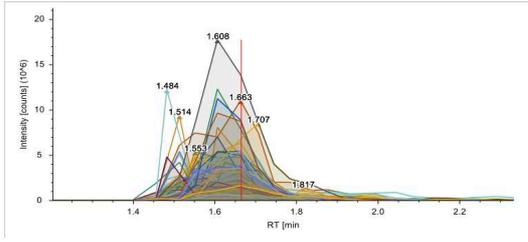
HMDB0000197 – POSmz176.071rt11.27  
2-(1H-indol-3-yl)acetic acid



HMDB0013188 – NEGmz140.034rt2.59  
6-Hydroxypicolinic acid

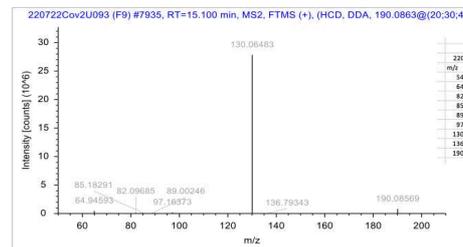
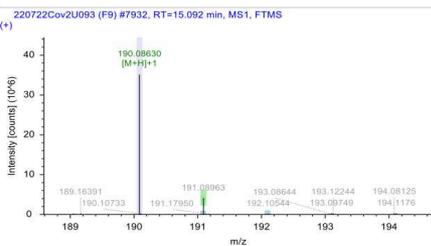
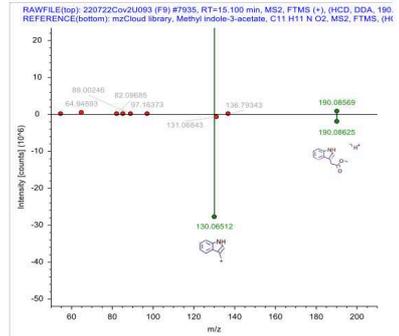
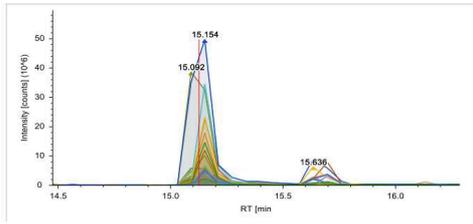


**HMDB0000663 – NEGmz209.03rt1.59**  
**Glucuric acid**



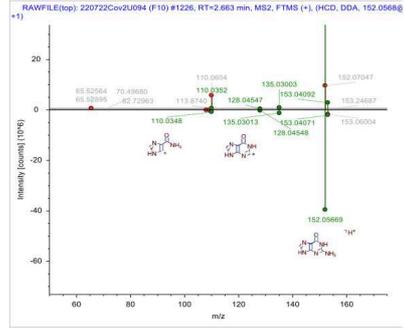
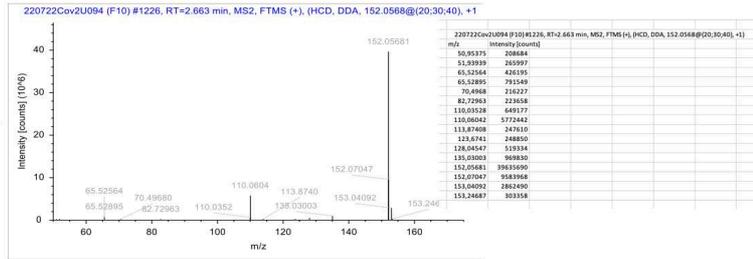
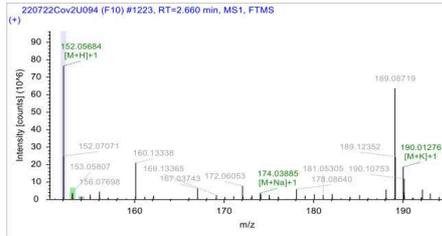
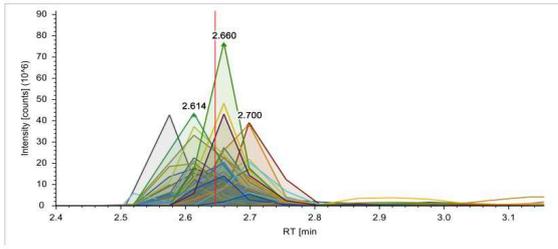
m/z	Intensity [counts]
71.01244	0.05
85.02809	1.00
129.01620	0.15
133.01317	0.10
173.00800	0.10
191.01912	0.40
209.02991	0.15

**HMDB0029738 – POSmz190.087rt15.21**  
**Methyl indole-3-acetate**

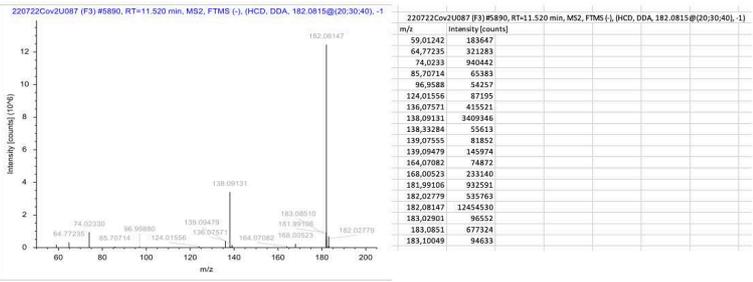
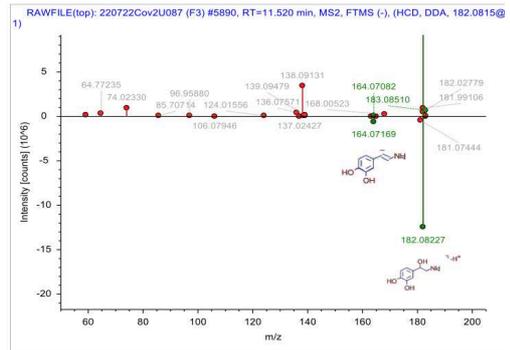
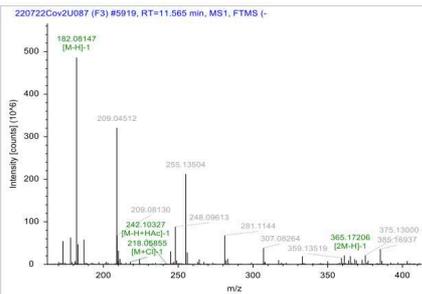
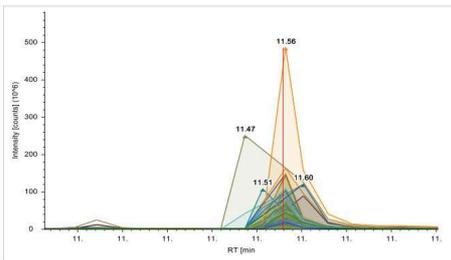


m/z	Intensity [counts]
130.06483	30.00
190.08569	1.00

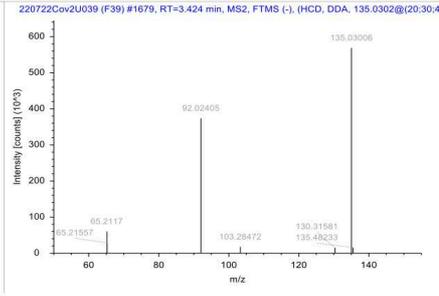
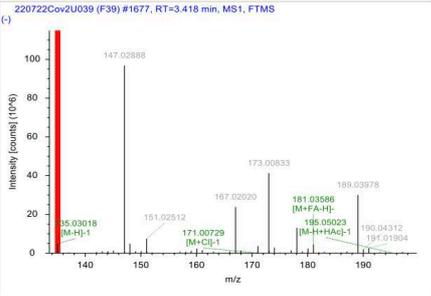
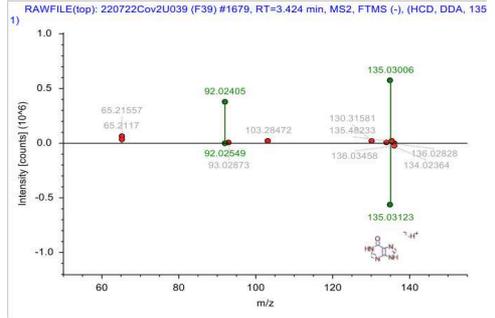
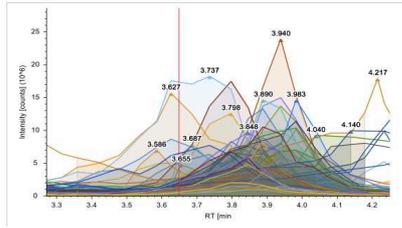
**HMDB0000132 – POSmz152.057rt2.65**  
**Guanine**



**HMDB0000068 – NEGmz182.081rt11.47**  
**Epinephrine**

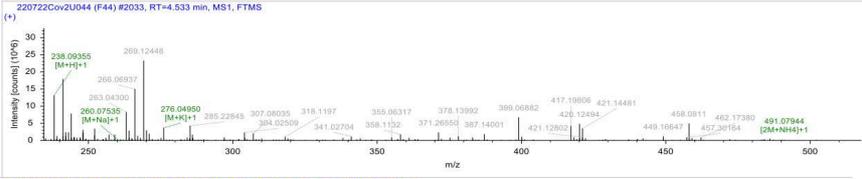
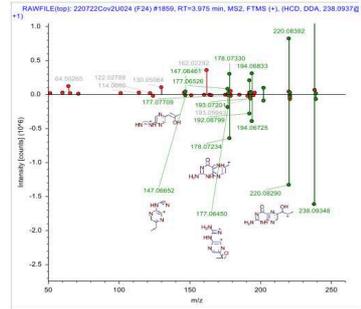
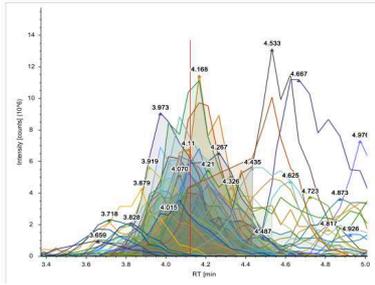


**POSmz137.046rt2.82 – POSmz183.087rt1.64  
Hypoxanthine**

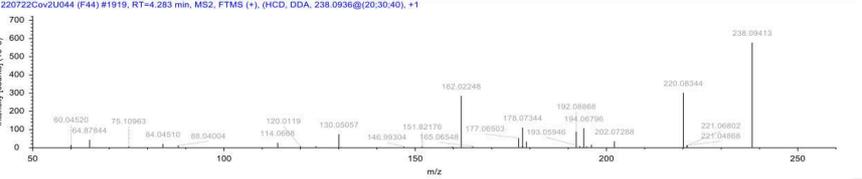


220722Cov RT=3.424 min MS2	
m/z	Intensity [counts]
65,21178	60186
65,21557	29079
92,02405	373806
103,28472	17597
130,31581	15059
135,03006	568630
135,48233	15727

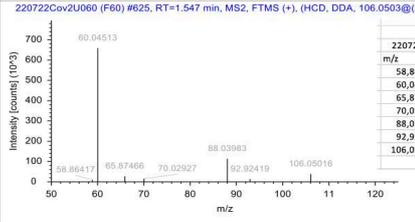
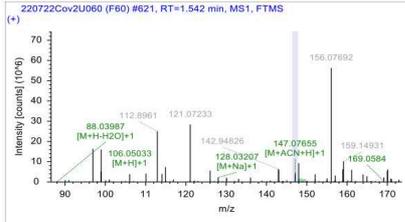
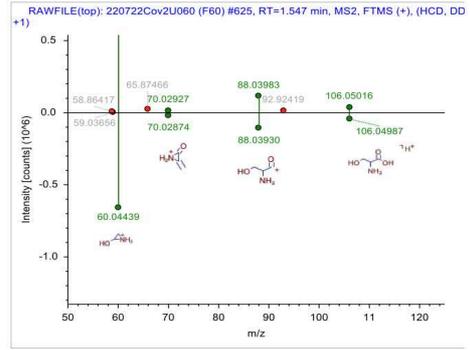
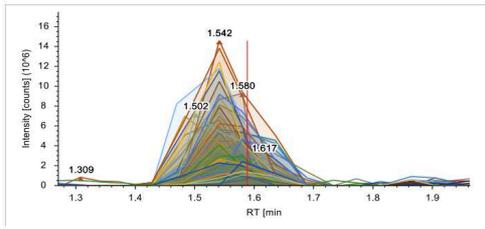
**HMDB0000068 – NEGmz182.081rt1.47  
Biotpterin**



220722Cov2U044 (F44) #1915, RT4.28	
m/z	Intensity [counts]
60.0432	13398
64.8744	43527
75.10963	7735
85.0451	21308
88.04004	12795
114.06868	27321
120.01199	8728
124.05117	8558
130.05057	74175
146.99304	6667
151.82176	7086
156.79795	5919
162.02248	284768
165.06548	7972
177.06503	53021
178.05977	8089
178.07344	131083
179.03818	33424
192.08868	88151
193.05966	8647
194.06796	107018
194.81747	8124
196.06419	16793
202.07288	35988
220.08344	301651
221.04888	13629
221.06802	11005
238.09413	578210

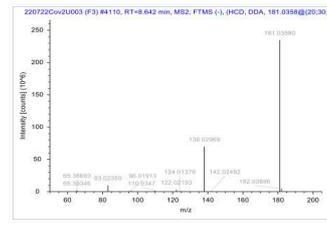
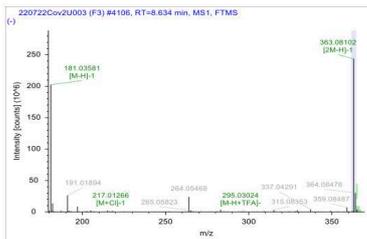
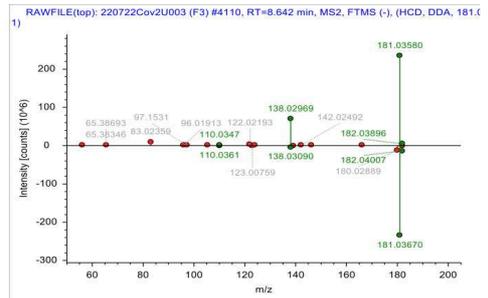
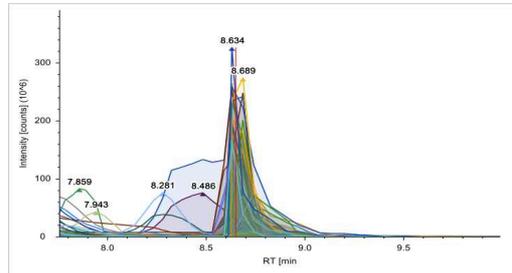


**HMDB0003406 – POSmz106.05rt1.61**  
**D-Serine**



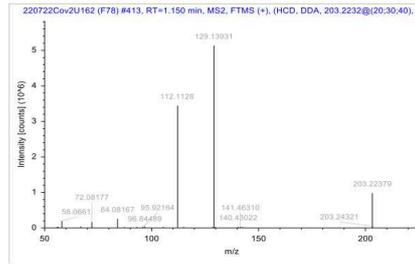
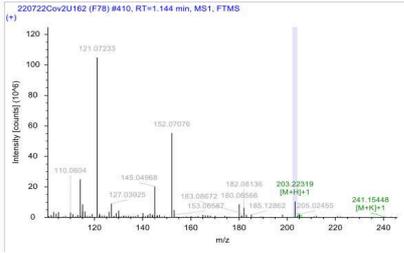
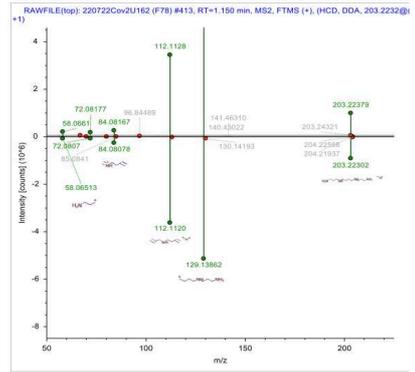
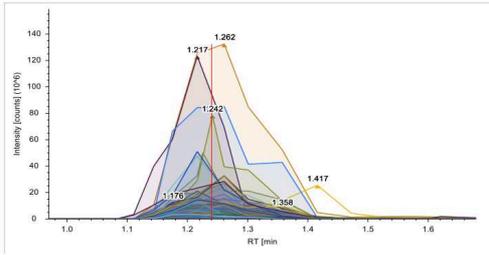
m/z	Intensity (counts)
58.86417	9727
60.04513	658837
65.87466	27530
70.02927	15560
88.03983	113747
92.02419	13439
106.05016	38647

**HMDB0003099 - NEGmz181.036rt8.75**  
**9-Methyluric acid**



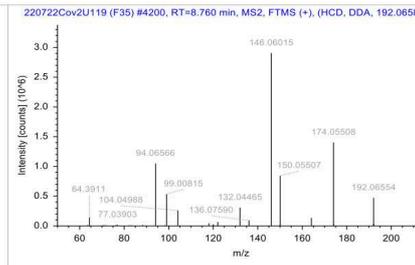
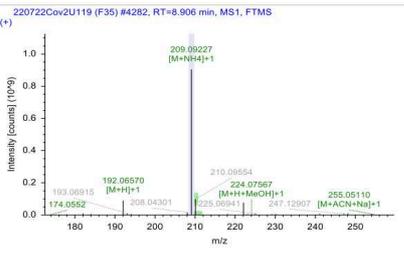
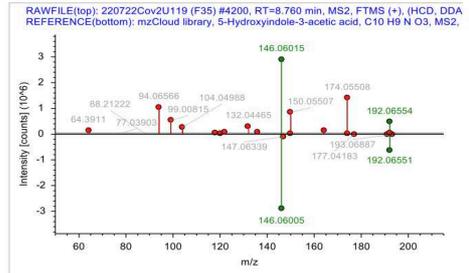
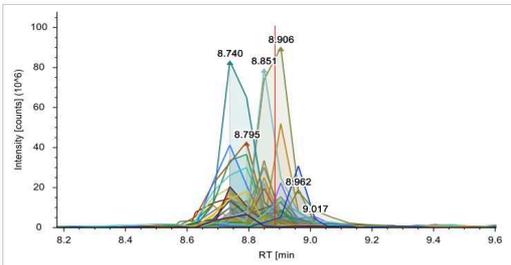
m/z	Intensity (counts)
55.17252	84651
65.38346	1921637
65.38693	1746604
83.02359	2931578
96.01913	1527163
97.15311	800339
105.45144	540369
110.03477	1802321
122.02193	9239906
124.01379	982552
138.02969	69872640
142.02492	1038654
146.10616	890800
166.03276	705650
181.0358	234787600
182.03896	4891648

**HMDB0001256 - POSmz203.223rt1.25**  
**Spermine**



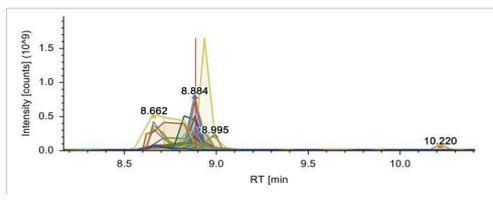
m/z	Intensity [counts]
55.8147	29676
56.2141	30033
58.0611	19668
66.2242	43320
72.08177	168795
72.0852	29014
84.08167	257398
87.0653	28077
93.06102	30062
95.82164	30880
96.84489	31726
105.49768	29267
112.11282	3442581
114.9956	18950
128.66887	28532
129.13931	5134685
129.64735	29693
140.43022	29797
141.4631	29711
142.29963	27070
203.22379	984272
203.24321	55103

**HMDB0000763 - POSmz192.066rt8.89**  
**5-Hydroxyindole-3-acetic acid**

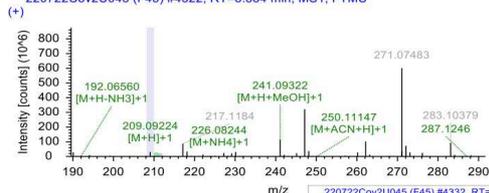


m/z	Intensity [counts]
63.5247	24725
64.39111	142936
70.66737	15078
71.48687	14200
76.89527	35846
77.03903	17709
82.16885	13675
83.23163	13305
88.21222	16512
94.06566	1049282
98.26961	14432
99.00815	536083
104.04988	261985
118.06538	45139
120.04449	26458
120.08121	15089
122.06039	89925
132.04465	307931
136.0759	93448
146.06015	2904470
150.05507	814851
150.06729	26143
164.07066	133329
174.05508	1402894
174.14861	24014
192.06554	472650
192.08135	27264
192.1015	25216
192.15967	33116

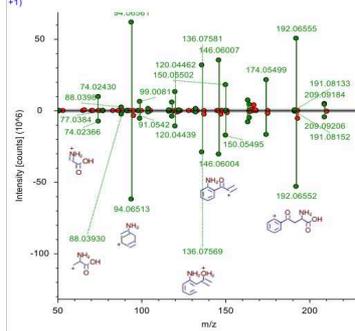
**HMDB0000684 – POSmz209.092rt8.88**  
**L-Kynurenine**



220722Cov2U045 (F45) #4332, RT=8.884 min, MS1, FTMS (+)

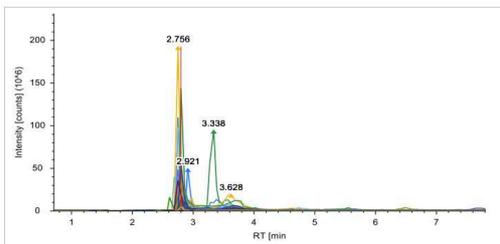


RAWFILE(top): 220722Cov2U045 (F45) #4332, RT=8.899 min, MS2, FTMS (+), (HCl +1)

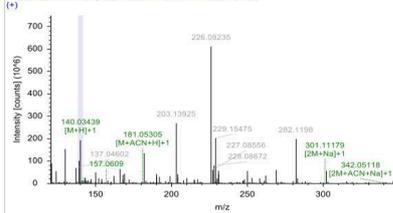


m/z	Intensity [counts]
51.02364	147756
53.28902	128781
59.09765	119821
65.09922	141137
65.06471	282067
67.86261	120605
74.02430	892989
77.03944	158977
85.18127	123181
88.03995	204326
91.05441	172094
94.06664	4122560
96.37995	106181
98.02425	211957
99.00811	623002
102.04239	132000
104.00808	1499942
111.01032	331313
116.04945	670201
118.06398	579612
120.04462	13514280
122.04001	359087
130.78166	115971
132.04443	1392789
136.07581	3170760
146.06007	35053400
148.11448	145783
150.00925	223196
150.05502	18110910
162.08023	748254
164.07051	4855545
166.03224	860278
167.03564	4029700
167.05569	156826
168.03937	278545
174.05499	21818250
191.08598	242599
191.08133	481393
192.06555	10569750
209.09184	500890
209.09184	4444180
231.19099	14054

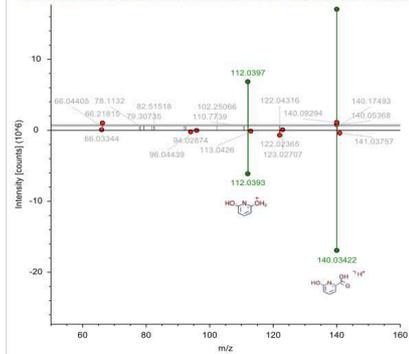
**HMDB0002641 – POSmz147.044rt10.49**  
**2-Hydroxycinnamic acid**



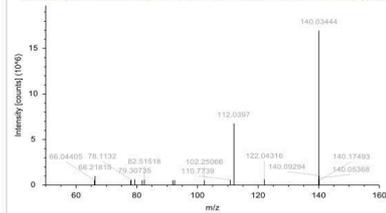
220722Cov2U112 (F28) #1308, RT=2.756 min, MS1, FTMS (+)



RAWFILE(top): 220722Cov2U112 (F28) #1308, RT=2.766 min, MS2, FTMS (+), (HCl, DDA, 140.0344@20:30.40, +1)

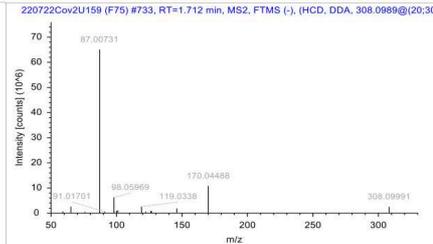
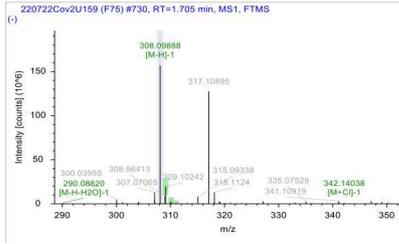
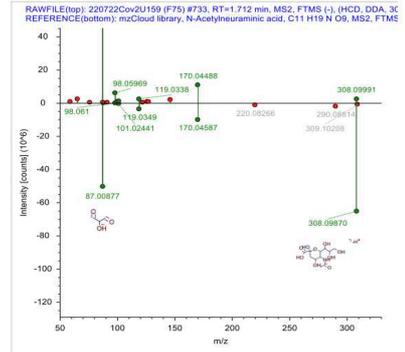
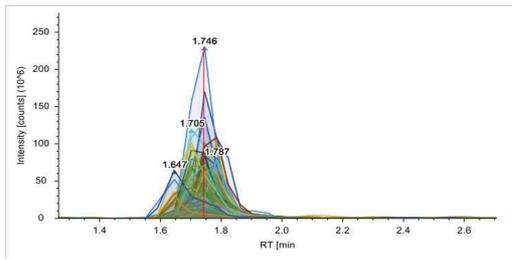


220722Cov2U112 (F28) #1308, RT=2.766 min, MS2, FTMS (+), (HCl, DDA, 140.0344@20:30.40, +1)



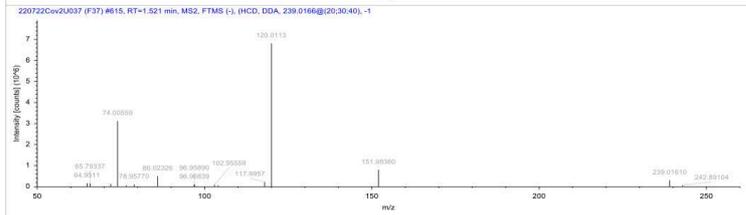
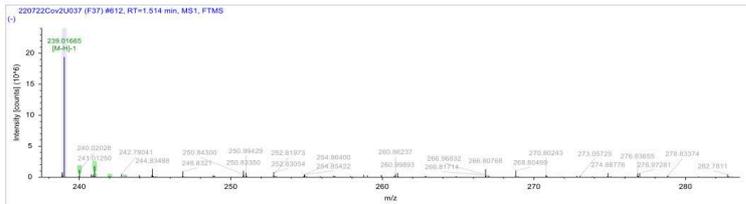
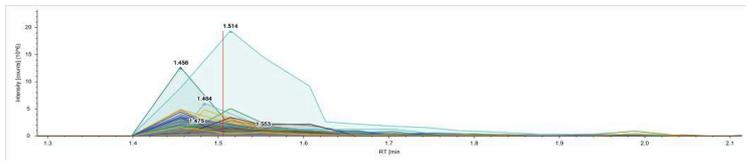
m/z	Intensity [counts]
66.04405	533758
66.21815	992996
77.87491	508454
78.11328	576314
79.30735	690747
81.76567	538511
82.51518	584560
91.91208	528816
93.49226	528736
102.25066	533080
110.77398	585112
112.03897	676615
122.04316	633784
140.03444	16969380
140.03368	763532
140.09294	1057637
140.17493	540669

**HMDB000230 – NEGmz308.099rt1.73**  
**N-Acetylneuraminic acid**



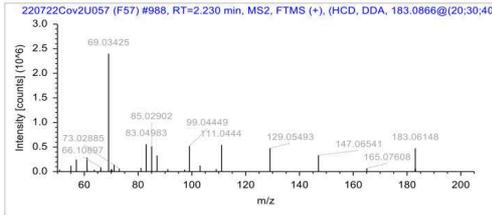
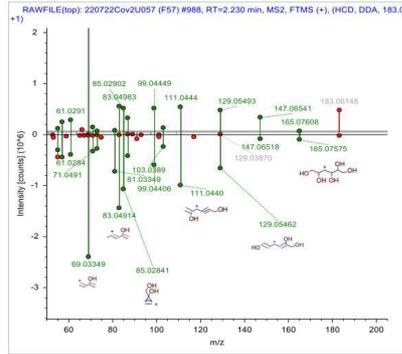
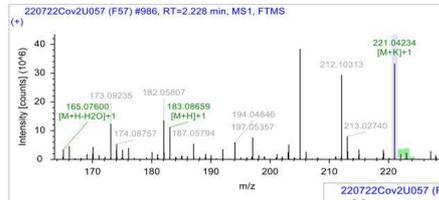
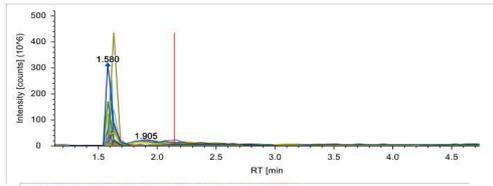
m/z	Intensity [counts]
59.01253	674883
65.18037	257340
75.78641	508765
87.00731	65062500
87.01461	563634
91.01701	581076
98.05969	6239053
100.07314	945881
101.03333	1120731
119.03386	2605758
122.28091	570359
126.05452	916963
126.90383	784120
146.04463	1910259
170.04488	1079380
308.09991	2514000

**HMDB000230 – NEGmz308.099rt1.73**  
**L-Cystine**



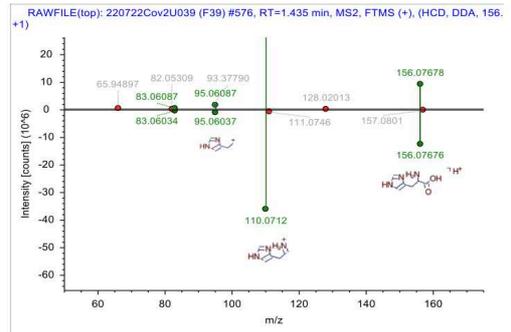
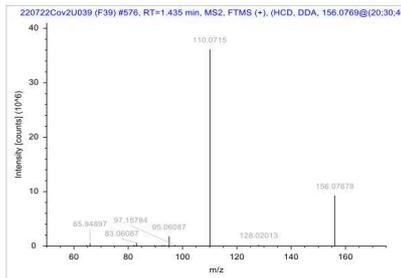
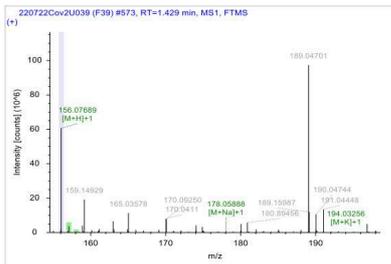
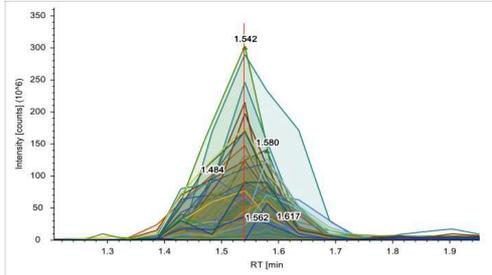
m/z	Intensity [counts]
64,95116	164210
65,79337	157152
71,98994	131268
74,00559	3125089
76,61642	58931
78,9577	105524
86,02326	502950
96,9589	93689
96,96839	104839
102,95559	84331
104,1134	58729
117,99571	227191
120,01131	6815726
151,9836	803247
239,0161	305470
242,89104	61682

**HMDB0011632 – POSmz183.087rt1.64**  
**L-Iditol**



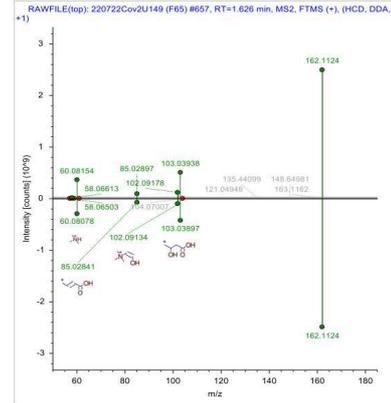
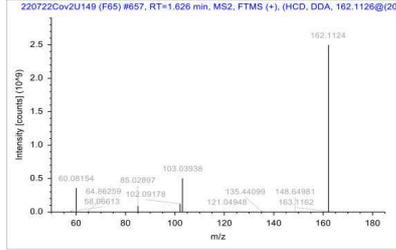
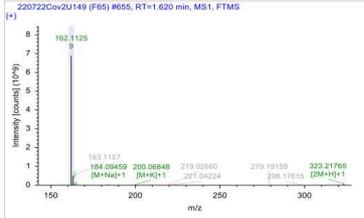
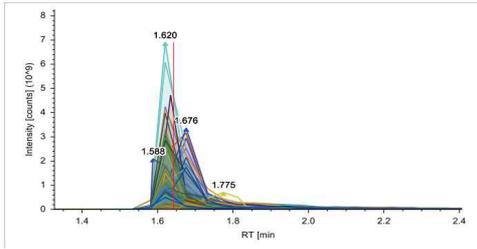
m/z	Intensity [counts]
56.78153	4460
55.05496	32193
57.03434	24893
61.02913	28264
61.0405	59794
63.70705	42575
66.10897	88628
66.1121	55247
68.03425	239936
69.03425	45333
70.3076	48453
71.04984	340544
73.02885	62393
81.03425	78097
83.04983	558269
85.02902	512941
87.04457	326772
91.03926	52902
97.28601	44789
99.04449	520351
103.03905	121380
109.04835	54470
111.04445	545405
129.05493	477733
147.06541	333904
165.07608	89164
183.06148	475425
183.08722	52941
183.11397	60947

**HMDB0000230 – NEGmz308.099rt1.73**  
**L-Histidine**



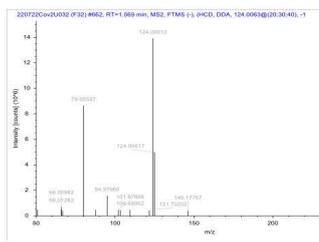
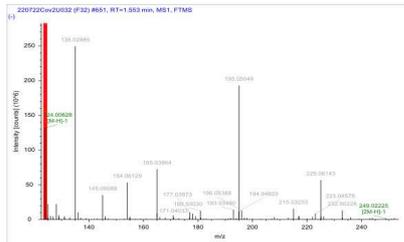
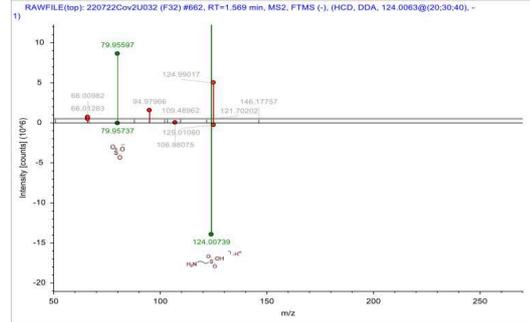
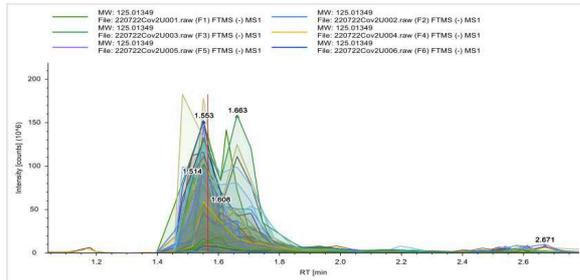
m/z	Intensity [counts]
64.68417	172371
65.94897	510654
82.05309	240808
83.06087	600360
89.66601	169248
92.45215	175068
93.3779	184886
95.06087	1787709
97.15784	183768
110.07156	3614670
128.02013	255721
156.07678	9279467

**HMDB000062 – POSmz162.113rt1.64**  
**L-Carnitine**



m/z	Intensity [counts]
50.14479	1481583
50.62368	1117094
57.03435	3672875
58.06613	13445660
59.0737	3524011
60.08154	35795000
61.52237	1381451
64.79962	1264901
64.86259	787573
85.02897	89766950
102.09178	122549800
103.03938	503583500
109.48862	3688017
121.04948	1408648
135.44099	1408648
148.64981	1636055
161.11249	2497626000
163.11624	2683111

**HMDB0000251 – NEGmz124.006rt1.64**  
**Taurine**



m/z	Intensity [counts]
50.82985	467608
66.00982	518866
66.01283	735505
66.70425	337427
79.95597	8654930
87.79192	461966
94.97966	1586186
101.97888	462783
103.29534	450472
109.48862	476841
121.70202	424541
124.00613	13927970
124.99017	4982876
146.17757	395203

## REFERENCES BIBLIOGRAPHIQUES

---

- Adebayo, Ayobami et al. 2021. "L-Arginine and COVID-19: An Update." *Nutrients* 13(11).
- Aletti, Federico et al. 2016. "ShockOmics: Multiscale Approach to the Identification of Molecular Biomarkers in Acute Heart Failure Induced by Shock." *Scandinavian journal of trauma, resuscitation and emergency medicine* 24: 9.
- Armitage, Emily G, and Coral Barbas. 2014. "Metabolomics in Cancer Biomarker Discovery: Current Trends and Future Perspectives." *Journal of pharmaceutical and biomedical analysis* 87: 1–11.
- Bailleux, Caroline et al. 2023. "Metabolomic Signatures of Scarff-Bloom-Richardson (SBR) Grade in Non-Metastatic Breast Cancer." *Cancers* 15(7).
- Barberis, Elettra et al. 2021. "Metabolomics Diagnosis of Covid-19 from Exhaled Breath Condensate." *Metabolites* 11(12).
- Barker, Grant et al. 2021. "Review Lipid and Lipoprotein Dysregulation in Sepsis: Clinical and Mechanistic Insights into Chronic Critical Illness." *Journal of Clinical Medicine* 10(8): 1–20.
- Beisken, Stephan, Michael Eiden, and Reza M Salek. 2015. "Getting the Right Answers : Understanding Metabolomics Challenges." : 97–109.
- Belizário, José E., Joel Faintuch, and Miguel Garay Malpartida. 2021. "Breath Biopsy and Discovery of Exclusive Volatile Organic Compounds for Diagnosis of Infectious Diseases." *Frontiers in Cellular and Infection Microbiology* 10(January): 1–13.
- Benson, Mikael. 2016. "Clinical Implications of Omics and Systems Medicine: Focus on Predictive and Individualized Treatment." *Journal of Internal Medicine* 279(3): 229–40.
- Benton, Paul H et al. 2015. "An Interactive Cluster Heat Map to Visualize and Explore Multidimensional Metabolomic Data." *Metabolomics : Official journal of the Metabolomic Society* 11(4): 1029–34.
- Blaise, Benjamin J. et al. 2013. "Metabolic Phenotyping of Traumatized Patients Reveals a Susceptibility to Sepsis." *Analytical Chemistry* 85(22): 10850–55.
- Blasco, H. et al. 2020. "The Specific Metabolome Profiling of Patients Infected by SARS-COV-2 Supports the Key Role of Tryptophan-Nicotinamide Pathway and Cytosine Metabolism." *Scientific Reports* 10(1): 1–12. <https://doi.org/10.1038/s41598-020-73966-5>.
- Boccard, Julien, Jean-Luc Veuthey, and Serge Rudaz. 2010. "Knowledge Discovery in Metabolomics : An Overview of MS Data Handling." *Journal of Separation Science* 33: 290–304.
- Bodi, Vicente et al. 2013. "Metabolomics in the Diagnosis of Acute Myocardial Ischemia." *Journal of cardiovascular translational research* 6(5): 808–15.
- Bonny, V et al. 2020. "[COVID-19: Pathogenesis of a Multi-Faceted Disease]. TT - COVID-19 : Physiopathologie d'une Maladie à Plusieurs Visages." *Rev Med Interne* 41(6): 375–89.
- Bos, Lieuwe D J et al. 2014. "Exhaled Breath Metabolomics as a Noninvasive Diagnostic Tool for Acute Respiratory Distress Syndrome." *European Respiratory Journal* 44(1): 188–97.
- Bos, Lieuwe D J, Simone Meinardi, Donald Blake, and Katrine Whiteson. 2016. "Bacteria in the Airways of Patients with Cystic Fibrosis Are Genetically Capable of Producing VOCs in Breath." *Journal of Breath Research* 10(4): 47103.
- Bouatra, Souhaila et al. 2013. "The Human Urine Metabolome." *PLoS ONE* 8(9).
- Bourgin, Mélanie, Sylvère Durand, and Guido Kroemer. 2023. "Diagnostic, Prognostic and Mechanistic Biomarkers of COVID-19 Identified by Mass Spectrometric

- Metabolomics.” *Metabolites* 13(3).
- Bourgueil, Yann et al. 2020. “La Première Vague de Covid-19 En France et Les Soins Primaires.” *Revue Medicale Suisse* 16: 2123–26.
- Broadhurst, David I, and Douglas B Kell. 2007. “Statistical Strategies for Avoiding False Discoveries in Metabolomics and Related Experiments.” 2(4).
- Bruzzone, Chiara et al. 2023. “Metabolomics as a Powerful Tool for Diagnostic, Pronostic and Drug Intervention Analysis in COVID-19.” *Frontiers in Molecular Biosciences* 10(February): 1–11.
- Byrne, Robert A. et al. 2023. “2023 ESC Guidelines for the Management of Acute Coronary Syndromes.” *European Heart Journal* 44(38): 3720–3826.
- Cajka, Tomas, and Oliver Fiehn. 2016. “Toward Merging Untargeted and Targeted Methods in Mass Spectrometry-Based Metabolomics and Lipidomics.” *Analytical Chemistry* 88(1): 524–45.
- Calvino, Bernard. 2021. “Resolution Phase of Inflammation, an Active Phase Orchestrated by the Role of Proresolving Mediators: Resolvins, Protectins, and Maresins.” *Douleurs* 22(5–6): 219–31.
- Cambiaghi, Alice et al. 2017. “Characterization of a Metabolomic Profile Associated with Responsiveness to Therapy in the Acute Phase of Septic Shock.” *Scientific Reports* 7(1): 1–16.
- Cao, Yan et al. 2023. “Untargeted Metabolomic Profiling of Sepsis-Induced Cardiac Dysfunction.” *Frontiers in Endocrinology* 14(February): 1–13.
- Carli, P, and C Telion. 2023. “Première Vague COVID-19, Trois Ans Déjà....” *Journal Européen des Urgences et de Réanimation* 35(1): 1–3.
- Carneiro, Gabriel, Andres Lopez Radcenco, Joseph Evaristo, and Gustavo Monnerat. 2020. “Novel Strategies for Clinical Investigation and Biomarker Discovery: A Guide to Applied Metabolomics.” *Hormone Molecular Biology and Clinical Investigation* 38(3): 1–13.
- Casari, Iliaria, Marcello Manfredi, Pat Metharom, and Marco Falasca. 2021. “Dissecting Lipid Metabolism Alterations in SARS-CoV-2.” *Progress in Lipid Research* 82(October 2020): 101092.
- Chang, Youjin et al. 2023. “A Targeted Metabolomics Approach for Sepsis-Induced ARDS and Its Subphenotypes.” *Critical care (London, England)* 27(1): 263.
- Chardin, David et al. 2023. “Identification of Metabolomic Markers in Frozen or Formalin-Fixed and Paraffin-Embedded Samples of Diffuse Glioma from Adults.” *International journal of molecular sciences* 24(23).
- Chong, Jasmine, David S Wishart, and Jianguo Xia. 2019. “Using MetaboAnalyst 4.0 for Comprehensive and Integrative Metabolomics Data Analysis.” *Current protocols in bioinformatics* 68(1): e86.
- Chu, Derek K et al. 2020. “Articles Physical Distancing , Face Masks , and Eye Protection to Prevent Person-to-Person Transmission of SARS-CoV-2 and COVID-19 : A Systematic Review and Meta-Analysis.” *Lancet* 395: 1973–87.
- Claus, S, and H Guillou. 2019. “La Métabolomique : Applications Médicales.” *Médecine Maladies Métaboliques* 13(3): 263–67.
- Claus, Sandrine P, and Jonathan R Swann. 2013. “Nutrimetabonomics: Applications for Nutritional Sciences, with Specific Reference to Gut Microbial Interactions.” *Annual review of food science and technology* 4: 381–99.
- Clayton, T Andrew et al. 2006. “Pharmaco-Metabonomic Phenotyping and Personalized Drug Treatment.” *Nature* 440(7087): 1073–77.
- Costa Dos Santos Junior, Gilson, Claudia Maria Pereira, Tatiana Kelly Da Silva Fidalgo, and Ana Paula Valente. 2020. “Saliva NMR-Based Metabolomics in the War against

- COVID-19.” *Analytical Chemistry* 92(24): 15688–92.
- Courant, Frédérique, Jean Philippe Antignac, Gaud Dervilly-Pinel, and Bruno Le Bizec. 2014. “Basics of Mass Spectrometry Based Metabolomics.” *Proteomics* 14(21–22): 2369–88.
- Cruzat, Vinicius et al. 2018. “Glutamine: Metabolism and Immune Function, Supplementation and Clinical Translation.” *Nutrients* 10(11): 1–31.
- D’Andréa, Grégoire et al. 2023. “Pilot Study on the Use of Untargeted Metabolomic Fingerprinting of Liquid-Cytology Fluids as a Diagnostic Tool of Malignancy for Thyroid Nodules.” *Metabolites* 13(7).
- Dalli, Jesmond et al. 2017. “Human Sepsis Eicosanoid and Proresolving Lipid Mediator Temporal Profiles: Correlations With Survival and Clinical Outcomes.” *Critical Care Medicine* 45(1).
- Dewulf, Joseph P. et al. 2022. “Urine Metabolomics Links Dysregulation of the Tryptophan-Kynurenine Pathway to Inflammation and Severity of COVID-19.” *Scientific Reports* 12(1): 1–8.
- Doğan, Halef Okan. 2024. “Metabolomics: A Review of Liquid Chromatography Mass Spectrometry-Based Methods and Clinical Applications.” *Turkish Journal of Biochemistry* 0(0): 1–14.
- Dunn, W B et al. 2011. “The Human Serum Metabolome (HUSERMET) Consortium, Procedures for Large-Scale Metabolic Profiling of Serum and Plasma Using Gas Chromatography and Liquid Chromatography Coupled to Mass Spectrometry.” *Nat. Protoc.* 6: 1–24.
- Eckerle, Michelle et al. 2018. “Metabolomics as a Driver in Advancing Precision Medicine in Sepsis.” *Pharmacotherapy* 37(9): 1023–32.
- Ellero-Simatos, S., S. Claus, and H. Guillou. 2019. “Metabolomics: Biomedical Applications.” *Medecine des Maladies Metaboliques* 13(3): 263–67.
- European Commission safety directorate-general for Health and food. 2023. “EU Common List of COVID-19 Antigen Tests.” (Juillet).
- Febra, Cláudia et al. 2024. “Acute Venous Thromboembolism Plasma and Red Blood Cell Metabolomic Profiling Reveals Potential New Early Diagnostic Biomarkers: Observational Clinical Study.” *Journal of Translational Medicine* 22(1): 1–17. <https://doi.org/10.1186/s12967-024-04883-8>.
- Feng, Ke et al. 2022. “Identification of Biomarkers and the Mechanisms of Multiple Trauma Complicated with Sepsis Using Metabolomics.” *Frontiers in Public Health* 10.
- Feng, S et al. 2023. “Prognostic Biomarkers for Sepsis Mortality Based on the Literature and LC-MS-Based Metabolomics of Sepsis Patients.” *American Journal of Translational Research* 15(9): 5757–68.
- Fernandes, Renata Salgado et al. 2022. “Recent Advances in Point of Care Testing for COVID-19 Detection.” *Biomedicine and Pharmacotherapy* 153(July): 113538.
- Ferrario, Manuela et al. 2016. “Mortality Prediction in Patients with Severe Septic Shock: A Pilot Study Using a Target Metabolomics Approach.” *Scientific Reports* 6(August 2015): 1–11.
- Fiehn, Oliver. 2002. “Metabolomics--the Link between Genotypes and Phenotypes.” *Plant molecular biology* 48(1–2): 155–71.
- Fisher, Owain, Ruth A. Benson, and Christopher He Imray. 2018. “The Clinical Application of Purine Nucleosides as Biomarkers of Tissue Ischemia and Hypoxia in Humans in Vivo.” *Biomarkers in Medicine* 13(11): 953–65.
- Frapas, Cecile F et al. 2022. “Untargeted Saliva Metabolomics by Liquid Chromatography-Mass Spectrometry Reveals Markers of COVID-19 Severity.” *PloS one* 17(9): e0274967.
- Gala, Jean-luc et al. 2020. “Méthodes Diagnostiques Du COVID-19.” *Louvain Medical*

- 139(05–06): 228–35.
- Gika, Helen G., Georgios A. Theodoridis, Robert S. Plumb, and Ian D. Wilson. 2014. “Current Practice of Liquid Chromatography-Mass Spectrometry in Metabolomics and Metabonomics.” *Journal of Pharmaceutical and Biomedical Analysis* 87: 12–25.
- Gorrochategui, Eva, Joaquim Jaumot, Silvia Lacorte, and Romà Tauler. 2016. “Trends in Analytical Chemistry Data Analysis Strategies for Targeted and Untargeted LC-MS Metabolomic Studies : Overview and Workflow.” *Trends in Analytical Chemistry* 82: 425–42.
- Grassin-Delyle, Stanislas et al. 2021. “Metabolomics of Exhaled Breath in Critically Ill COVID-19 Patients: A Pilot Study.” *EBioMedicine* 63: 0–6.
- Guo, Mengyu et al. 2023. “Multi-Omics for COVID-19: Driving Development of Therapeutics and Vaccines.” *National Science Review* 10(9). <https://doi.org/10.1093/nsr/nwad161>.
- Gupta, Aakriti et al. 2020. “Extrapulmonary Manifestations of COVID-19.” *Nature Medicine* 26(7): 1017–32.
- Gutiérrez-Gutiérrez, Belén et al. 2021. “Identification and Validation of Clinical Phenotypes with Prognostic Implications in Patients Admitted to Hospital with COVID-19: A Multicentre Cohort Study.” *The Lancet Infectious Diseases* 21(6): 783–92.
- Hantz, Sébastien. 2020. “Biological diagnosis of Sars-CoV-2 infection: strategies and interpretation of results.” *Revue francophone des laboratoires : RFL* 2020(526): 48–56.
- Hao, Xiao Bin et al. 2022. “Potential Metabolomic Biomarkers for the Identification and Diagnosis of Type A Acute Aortic Dissection in Patients with Hypertension.” *Frontiers in Cardiovascular Medicine* 9(November): 1–12.
- Hasan, Mohammad Rubayet, Mohammed Suleiman, and Andrés Pérez-López. 2021. “Metabolomics in the Diagnosis and Prognosis of COVID-19.” *Frontiers in Genetics* 12(May).
- Hasselbalch, Rasmus Bo et al. 2023. “Metabolomics of Early Myocardial Ischemia.” *Metabolomics* 19(4): 1–9.
- Haute Autorité de Santé. 2020. “Avis N° 2020.0020/AC/SEAP Du 6 Mars 2020 Du Collège de La Haute Autorité de Santé Relatif à l’inscription Sur La Liste Des Actes et Prestations Mentionnée à l’article L. 162-1-7 Du Code de La Sécurité Sociale, de La Détection Du Génome Du Coronavirus SARS.” 33(0): 5–7.
- . 2021. *Méta-Analyse de l’intérêt Diagnostique Des Tests RT-PCR Salivaires de Détection Du SARS- CoV-2*. [www.has-sante.fr](http://www.has-sante.fr).
- Holmes, Edward C et al. 2021. “The Origins of SARS-CoV-2 : A Critical Review.” *Cell* 184(19): 4848–56. <https://doi.org/10.1016/j.cell.2021.08.017>.
- Ibrahim, Wadah et al. 2021. “Diagnosis of Covid-19 by Exhaled Breath Analysis Using Gas Chromatography–Mass Spectrometry.” *ERJ Open Research* 7(3). <http://dx.doi.org/10.1183/23120541.00139-2021>.
- Iliescu, Florina Silvia et al. 2021. “Point-of-Care Testing-the Key in the Battle against SARS-CoV-2 Pandemic.” *Micromachines* 12(12): 1–33.
- Isac, Sebastian et al. 2024. “The Omics Complexity in Sepsis: The Limits of the Personalized Medicine Approach.” *Journal of Personalized Medicine* 14(3).
- Izcovich, Ariel et al. 2020. “Prognostic Factors for Severity and Mortality in Patients Infected with COVID-19: A Systematic Review.” *PLoS ONE* 15(11 November): 1–30.
- Jareño-Esteban, José Javier et al. 2017. “Study of 5 Volatile Organic Compounds in Exhaled Breath in Chronic Obstructive Pulmonary Disease.” *Archivos de bronconeumologia* 53(5): 251–56.
- Jing, Yixian et al. 2022. “Alterations of Urinary Microbial Metabolites and Immune Indexes Linked With COVID-19 Infection and Prognosis.” *Frontiers in Immunology* 13(March).

- Kailasam, Srividya. 2021. "Analysis and LC-MS / MS." *Technology Networks (Lc)*.
- Kharouba, Maged, Dimple D. Patel, Rami H. Jaber, and Sherif Hanafy Mahmoud. 2023. "Metabolomic Analysis in Neurocritical Care Patients." *Metabolites* 13(6): 745.
- Klassen, Aline et al. 2017. "Metabolomics: Definitions and Significance in Systems Biology." *Advances in experimental medicine and biology* 965: 3–17.
- Konstantinides, Stavros V. et al. 2020. "2019 ESC Guidelines for the Diagnosis and Management of Acute Pulmonary Embolism Developed in Collaboration with the European Respiratory Society (ERS)." *European Heart Journal* 41(4): 543–603.
- Krastanov, A. 2010. "Metabolomics - The State of Art." *Biotechnology and Biotechnological Equipment* 24(1): 1537–43.
- Langley, Raymond J. et al. 2014. "Integrative 'Omic' Analysis of Experimental Bacteremia Identifies a Metabolic Signature That Distinguishes Human Sepsis from Systemic Inflammatory Response Syndromes." *American Journal of Respiratory and Critical Care Medicine* 190(4): 445–55.
- Lee, Loong Chuen, Choong-Yeun Liong, and Abdul Aziz Jemain. 2018. "Partial Least Squares-Discriminant Analysis (PLS-DA) for Classification of High-Dimensional (HD) Data: A Review of Contemporary Practice Strategies and Knowledge Gaps." *The Analyst* 143(15): 3526–39.
- Lenski, Marie, and Clément Bruno. 2023. "Metabolomics : Concepts and Applications in Biological and Forensic Toxicology." *Toxicologie Analytique et Clinique* 35(2): 96–112.
- Lewis, Holly May et al. 2022. "Metabolomics Markers of COVID-19 Are Dependent on Collection Wave." *Metabolites* 12(8).
- Li, Han et al. 2022. "Methods for Phenotyping Adult Patients in Sepsis and Septic Shock: A Scoping Review." *Critical Care Explorations* 4(4): E0672.
- Lin, Baoxu et al. 2020. "Critical Reviews in Clinical Laboratory Sciences Progress in Understanding COVID-19 : Insights from the Omics Approach." *Critical Reviews in Clinical Laboratory Sciences* 0(0): 1–18.
- Liptak, Peter et al. 2022. "Persistence of Metabolomic Changes in Patients during Post-COVID Phase: A Prospective, Observational Study." *Metabolites* 12(7).
- Lonati, Caterina et al. 2023. "Urinary Phenotyping of SARS-CoV-2 Infection Connects Clinical Diagnostics with Metabolomics and Uncovers Impaired NAD<sup>+</sup> pathway and SIRT1 Activation." *Clinical Chemistry and Laboratory Medicine* 62(4): 770–88.
- Lv, Longxian et al. 2021. "The Faecal Metabolome in COVID-19 Patients Is Altered and Associated with Clinical Features and Gut Microbes." *Analytica Chimica Acta* 1152: 338267.
- Mangge, Harald et al. 2021. "Increased Kynurenine Indicates a Fatal Course of Covid-19." *Antioxidants* 10(12): 1–11.
- Manzoni, Claudia et al. 2018. "Genome , Transcriptome and Proteome : The Rise of Omics Data and Their Integration in Biomedical Sciences." *Briefings in Bioinformatics* 19(2): 286–302.
- Marco-ramell, Anna et al. 2018. "Evaluation and Comparison of Bioinformatic Tools for the Enrichment Analysis of Metabolomics Data." *BMC Bioinformatics* 19(1): 1–11.
- Marty, Francisco M., Kaiwen Chen, and Kelly A. Verrill. 2020. "How to Obtain a Nasopharyngeal Swab Specimen." *New England Journal of Medicine* 382(22): e76.
- Matsuyama, Toshifumi, Steven K. Yoshinaga, Kimitaka Shibue, and Tak W. Mak. 2021. "Comorbidity-Associated Glutamine Deficiency Is a Predisposition to Severe COVID-19." *Cell Death and Differentiation* 28(12): 3199–3213.
- McDonald, Daniel et al. 2018. "American Gut: An Open Platform for Citizen Science Microbiome Research." *mSystems* 3(3): e00031-18.
- Mickiewicz, Beata et al. 2014. "Metabolic Profiling of Serum Samples by 1H Nuclear

- Magnetic Resonance Spectroscopy as a Potential Diagnostic Approach for Septic Shock.” *Critical Care Medicine* 42(5): 1140–49.
- Mounce, Bryan C, Michelle E Olsen, Marco Vignuzzi, and John H Connor. 2017. “Polyamines and Their Role in Virus.” *Microbiology and Molecular Biology Reviews* 81(4): 1–12.
- Nariai, Maika et al. 2022. “Biomarker Profiling of Postmortem Blood for Diabetes Mellitus and Discussion of Possible Applications of Metabolomics for Forensic Casework.” *International journal of legal medicine* 136(4): 1075–90.
- Nguyen, Maxime et al. 2021. “High Plasma Concentration of Non-Esterified Polyunsaturated Fatty Acids Is a Specific Feature of Severe COVID-19 Pneumonia.” *Scientific Reports* 11(1): 1–12.
- Occelli, Céline et al. 2022. “Untargeted Plasma Metabolomic Fingerprinting Highlights Several Biomarkers for the Diagnosis and Prognosis of Coronavirus Disease 19.” *Frontiers in Medicine* 9(September): 1–13.
- Occelli, Céline, and Julie Contenti. 2020. “Early Warning Scores: Are They Clinically Relevant?” *European Journal of Emergency Medicine* 27(5).
- Oliver, Stephen G, Michael K Winson, Douglas B Kell, and Frank Baganz. 1998. “Systematic Functional Analysis of the Yeast Genome.” *Trends in Biotechnology* 16(9): 373–78. [https://doi.org/10.1016/S0167-7799\(98\)01214-1](https://doi.org/10.1016/S0167-7799(98)01214-1).
- Ostaszewski, Marek et al. 2021. “COVID19 Disease Map, a Computational Knowledge Repository of Virus–Host Interaction Mechanisms.” *Molecular Systems Biology* 17(10): 1–22.
- Páez-Franco, José C. et al. 2021. “Metabolomics Analysis Reveals a Modified Amino Acid Metabolism That Correlates with Altered Oxygen Homeostasis in COVID-19 Patients.” *Scientific Reports* 11(1): 1–12.
- Pang, Zhiqiang, Jasmine Chong, et al. 2021. “MetaboAnalyst 5.0: Narrowing the Gap between Raw Spectra and Functional Insights.” *Nucleic Acids Research* 49(W1): W388–96.
- Pang, Zhiqiang, Guangyan Zhou, Jasmine Chong, and Jianguo Xia. 2021. “Comprehensive Meta-Analysis of Covid-19 Global Metabolomics Datasets.” *Metabolites* 11(1): 1–14.
- Patschan, Daniel et al. 2023. “Metabolomics in Acute Kidney Injury: The Experimental Perspective.” *Journal of Clinical Medicine Research* 15(6): 283–91.
- Perrotta, Fabio, Maria Gabriella Matera, Mario Cazzola, and Andrea Bianco. 2020. “Severe Respiratory SARS-CoV2 Infection: Does ACE2 Receptor Matter?” *Respiratory Medicine* 168(March): 105996.
- Powers, Robert, and Eli Riekeberg. 2017. “New Frontiers in Metabolomics: From Measurement to Insight.” *F1000Research* Jul 19(6): 1148.
- Pozzi, Chiara et al. 2022. “A ‘Multiomic’ Approach of Saliva Metabolomics, Microbiota, and Serum Biomarkers to Assess the Need of Hospitalization in Coronavirus Disease 2019.” *Gastro Hep Advances* 1(2): 194–209.
- Ramos-Casals, Manuel, Pilar Brito-Zerón, and Xavier Mariette. 2021. “Systemic and Organ-Specific Immune-Related Manifestations of COVID-19.” *Nature Reviews Rheumatology* 17(6): 315–32.
- Rees, Chris A. et al. 2021. “Altered Amino Acid Profile in Patients with SARS-CoV-2 Infection.” *Proceedings of the National Academy of Sciences of the United States of America* 118(25): 4–6.
- Reisinger, Alexander Christian et al. 2021. “Branched-chain Amino Acids Can Predict Mortality in Icu Sepsis Patients.” *Nutrients* 13(9): 1–22.
- Roberts, Lee D., Amanda L. Souza, Robert E. Gerszten, and Clary B. Clish. 2012. “Targeted Metabolomics.” *Current Protocols in Molecular Biology* 1(SUPPL.98): 1–24.

- Rosolanka, Robert et al. 2023. “Changes in the Urine Metabolomic Profile in Patients Recovering from Severe COVID-19.” *Metabolites* 13(3).
- Roux, Aurélie, Dominique Lison, Christophe Junot, and Jean-françois Heilier. 2011. “Applications of Liquid Chromatography Coupled to Mass Spectrometry-Based Metabolomics in Clinical Chemistry and Toxicology : A Review.” *Clinical Biochemistry* 44(1): 119–35.
- Ruetsch, Caroline et al. 2021. “Functional Exhaustion of Type I and II Interferons Production in Severe COVID-19 Patients.” *Frontiers in Medicine* 7(January): 1–11.
- Ruszkiewicz, Dorota M. et al. 2020. “Diagnosis of COVID-19 by Analysis of Breath with Gas Chromatography-Ion Mobility Spectrometry - a Feasibility Study.” *EClinicalMedicine* 29–30: 100609.
- Saheb Sharif-Askari, Narjes et al. 2022. “Saliva Metabolomic Profile of COVID-19 Patients Associates with Disease Severity.” *Metabolomics* 18(11): 1–16.  
<https://doi.org/10.1007/s11306-022-01936-1>.
- Saito, Suguru et al. 2024. “Metabolomic and Immune Alterations in Long COVID Patients with Chronic Fatigue Syndrome.” *Frontiers in Immunology* 15(January): 1–16.
- Sapkota, Dipak et al. 2021. “COVID-19 Salivary Signature: Diagnostic and Research Opportunities.” *Journal of Clinical Pathology* 74(6): 344–49.
- Schwarz, Benjamin et al. 2021. “Cutting Edge: Severe SARS-CoV-2 Infection in Humans Is Defined by a Shift in the Serum Lipidome, Resulting in Dysregulation of Eicosanoid Immune Mediators.” *Journal of immunology (Baltimore, Md. : 1950)* 206(2): 329–34.
- Serhan, Charles N. 2010. “Novel Lipid Mediators and Resolution Mechanisms in Acute Inflammation: To Resolve or Not?” *American Journal of Pathology* 177(4): 1576–91.
- Seymour, Christopher W. et al. 2017. “Time to Treatment and Mortality during Mandated Emergency Care for Sepsis.” *New England Journal of Medicine* 376(23): 2235–44.
- Seymour, Christopher W et al. 2019. “Derivation, Validation, and Potential Treatment Implications of Novel Clinical Phenotypes for Sepsis.” *JAMA* 321(20): 2003–17.  
<https://doi.org/10.1001/jama.2019.5791>.
- Sharma, Neha et al. 2023. “Bio-Molecular Map of Albumin Identifies Signatures of Severity and Early Mortality in Acute Liver Failure.” *Journal of hepatology*.
- She, Han et al. 2023. “Metabolomics and Machine Learning Approaches for Diagnostic and Prognostic Biomarkers Screening in Sepsis.” *BMC Anesthesiology* 23(1): 1–13.
- Shen, Bo et al. 2020. “Proteomic and Metabolomic Characterization of COVID-19 Patient Sera.” *Cell* 182(1): 59-72.e15.
- da Silva Fidalgo, Tatiana Kelly et al. 2023. “Salivary Metabolomic Analysis Reveals Amino Acid Metabolism Shift in SARS-CoV-2 Virus Activity and Post-Infection Condition.” *Metabolites* 13(2).
- Song, Qi et al. 2021. “Point-of-Care Testing Detection Methods for COVID-19.” *Lab on a Chip* 21(9): 1634–60.
- Spick, Matt et al. 2022. “An Integrated Analysis and Comparison of Serum, Saliva and Sebum for COVID-19 Metabolomics.” *Scientific Reports* 12(1): 1–12.
- Sugimoto, Masahiro, Masato Kawakami, Martin Robert, and Tomoyoshi Soga. 2012. “Bioinformatics Tools for Mass Spectroscopy-Based Metabolomic Data Processing and Analysis.” *Current Bioinformatics* 7(1): 96–108.
- Suissa, Laurent et al. 2020. “Metabolome of Cerebral Thrombi Reveals an Association between High Glycemia at Stroke Onset and Good Clinical Outcome.” *Metabolites* 10(12).
- . 2021. “Combined Omic Analyzes of Cerebral Thrombi: A New Molecular Approach to Identify Cardioembolic Stroke Origin.” *Stroke* 52(9): 2892–2901.
- Sulaiman, Dawoud et al. 2024. “Lipidomic Changes in a Novel Sepsis Outcome-Based

- Analysis Reveals Potent pro-Inflammatory and pro-Resolving Signaling Lipids.” *Clinical and Translational Science* 17(3): 1–11.
- Sulaiman, Imran et al. 2021. “Microbial Signatures in the Lower Airways of Mechanically Ventilated COVID19 Patients Associated with Poor Clinical Outcome.” *Nature microbiology* 6(10): 1245–58.
- Szymanska, Ewa, Edoardo Saccenti, Age K. Smilde, and Johan A. Westerhuis. 2012. “Double-Check : Validation of Diagnostic Statistics for PLS-DA Models in Metabolomics Studies.” *Metabolomics* 8: S3–S16.
- Tebani, A., and S. Bekri. 2022. “The Promise of Omics in the Precision Medicine Era.” *Revue de Medecine Interne* 43(11): 649–60.
- Thomas, Ilias et al. 2020. “Integrative Analysis of Circulating Metabolite Profiles and Magnetic Resonance Imaging Metrics in Patients with Traumatic Brain Injury.” *International Journal of Molecular Sciences* 21(4).
- Thomas, Tiffany et al. 2020. “COVID-19 Infection Alters Kynurenine and Fatty Acid Metabolism, Correlating with IL-6 Levels and Renal Status.” *JCI Insight* 5(14): 1–16.
- Thooft, Aurélie et al. 2023. “Serum Metabolomic Profiles in Critically Ill Patients with Shock on Admission to the Intensive Care Unit.” *Metabolites* 13(4).
- Ventero, Maria Paz et al. 2021. “Nasopharyngeal Microbial Communities of Patients Infected With SARS-CoV-2 That Developed COVID-19.” *Frontiers in microbiology* 12: 637430.
- Villaseñor, Alma et al. 2014. “<sup>1</sup>H NMR Global Metabolic Phenotyping of Acute Pancreatitis in the Emergency Unit.” *Journal of Proteome Research* 13(12): 5362–75.
- Vinaixa, Aria et al. 2016. “Mass Spectral Databases for LC / MS- and GC / MS-Based Metabolomics : State of the Field and Future Prospects.” *Trends in Analytical Chemistry* 78: 23–35.
- Walker, Heather J., and Michael M. Burrell. 2020. “Could Breath Analysis by MS Could Be a Solution to Rapid, Non-Invasive Testing for COVID-19?” *Bioanalysis* 12(17): 1213–17.
- Wang, Qiongxin, Danxia Liu, Ping Song, and Ming Hui Zou. 2015. “Tryptophan-Kynurenine Pathway Is Dysregulated in Inflammation, and Immune Activation.” *Frontiers in Bioscience - Landmark* 20(7): 1116–43.
- Wells, Greg, Harry Prest, Charles William, and Russ Iv. 2011. “Spectrométrie de Masse Note Technique.” *Agilent Technologies*.
- Whiley, Luke et al. 2024. “Cross-Validation of Metabolic Phenotypes in SARS-CoV-2 Infected Subpopulations Using Targeted Liquid Chromatography-Mass Spectrometry (LC-MS).” *Journal of Proteome Research* 23(4): 1313–27.
- Wiener, Matthew C, Jeffrey R Sachs, Ekaterina G Deyanova, and Nathan A Yates. 2004. “Differential Mass Spectrometry : A Label-Free LC - MS Method for Finding Significant Differences in Complex Peptide and Protein Mixtures.” 76(20): 6085–96.
- Wiersinga, W Joost et al. 2020. “Pathophysiology, Transmission, Diagnosis, and Treatment of Coronavirus Disease 2019 (COVID-19) A Review.” *JAMA - Journal of the American Medical Association* 8(324): 782–793.
- Wishart, David S et al. 2022. “HMDB 5 . 0 : The Human Metabolome Database for 2022.” *Nucleic Acids Research* 50(November 2021): 622–31.
- World Health Organization. 2020. “Virtual Press Conference on COVID-19.” (11 March 2020).
- Wu, Di et al. 2020a. “Plasma Metabolomic and Lipidomic Alterations Associated with COVID-19.” *National Science Review* 7(7): 1157–68.
- . 2020b. “Plasma Metabolomic and Lipidomic Alterations Associated with COVID-19.” *National Science Review* 7(7): 1157–1168.
- Wu, Zunyou, and Jennifer M McGoogan. 2020. “Characteristics of and Important Lessons

- From the Coronavirus Disease 2019 (COVID-19) Outbreak in China: Summary of a Report of 72 314 Cases From the Chinese Center for Disease Control and Prevention.” *JAMA* 323(13): 1239–42.
- Wynants, Laure et al. 2020. “Prediction Models for Diagnosis and Prognosis of Covid-19: Systematic Review and Critical Appraisal.” *The BMJ* 369.
- Xie, Ming et al. 2023. “Serum Metabolic Signatures for Acute Pulmonary Embolism Identified by Untargeted Metabolomics.” *Frontiers in Medicine* 10(June).
- Yang, Yunshu et al. 2022. “Regulation of SIRT1 and Its Roles in Inflammation.” *Frontiers in Immunology* 13(March): 1–16.
- Yuan, Zi Cheng, and Bin Hu. 2021. “Mass Spectrometry-Based Human Breath Analysis: Towards COVID-19 Diagnosis and Research.” *Journal of Analysis and Testing* 5(4): 287–97.
- Zhang, Haibo et al. 2020. “Angiotensin-Converting Enzyme 2 (ACE2) as a SARS-CoV-2 Receptor: Molecular Mechanisms and Potential Therapeutic Target.” *Intensive Care Medicine* 46(4): 586–90.
- Zhang, Jinghui et al. 2023. “Metabolomic Analysis Reveals the Metabolic Disturbance in Aortic Dissection: Subtype Difference and Accurate Diagnosis.” *Nutrition, metabolism, and cardiovascular diseases : NMCD* 33(8): 1556–64.
- Zhang, Lei et al. 2018. “Functional Metabolomics Characterizes a Key Role for N-Acetylneuraminic Acid in Coronary Artery Diseases.” *Circulation* 137(13): 1374–90.
- Zhang, Shiqi et al. 2024. “Breath Metabolomics for Diagnosis of Acute Respiratory Distress Syndrome.” *Critical Care* 28(1): 1–11.
- Zhou, Peng et al. 2020. “A Pneumonia Outbreak Associated with a New Coronavirus of Probable Bat Origin.” *Nature* 579(7798): 270–73.