



**HAL**  
open science

# Utilisation de l'approche métabolomique sans a priori des métabolites urinaires issus de composés électrophiles pour comprendre la relation viande rouge et cancer colorectal

Loic Mervant

## ► To cite this version:

Loic Mervant. Utilisation de l'approche métabolomique sans a priori des métabolites urinaires issus de composés électrophiles pour comprendre la relation viande rouge et cancer colorectal. Cancer. Université Paul Sabatier - Toulouse III, 2022. Français. NNT : 2022TOU30284 . tel-04508385

**HAL Id: tel-04508385**

**<https://theses.hal.science/tel-04508385>**

Submitted on 18 Mar 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

En vue de l'obtention du  
**DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE**  
Délivré par l'Université Toulouse 3 - Paul Sabatier

---

Présentée et soutenue par  
**Loïc MERVANT**

Le 17 mars 2022

**Utilisation de l'approche métabolomique sans a priori des  
métabolites urinaires issus de composés électrophiles pour  
comprendre la relation viande rouge et cancer colorectal**

---

Ecole doctorale : **SEVAB - Sciences Ecologiques, Vétérinaires, Agronomiques et  
Bioingénieries**

Spécialité : **Infectiologie, Physiopathologie, Toxicologie, Génétique et  
Nutrition**

Unité de recherche :  
**TOXALIM - Laboratoire de Toxicologie Alimentaire**

Thèse dirigée par  
**Laurent DEBRAUWER et Françoise GUERAUD**

Jury

**M. Olivier Laprévote, Rapporteur**  
**M. Jean-Charles Martin, Rapporteur**  
**Mme Véronique Santé-Lhoutellier, Examinatrice**  
**M. Serge Rudaz, Examineur**  
**Mme Anne Nègre-Salvayre, Examinatrice**  
**M. Laurent Debrauwer, Co-directeur de thèse**  
**Mme Françoise Guéraud, Co-directrice de thèse**

## Remerciements

## **Remerciements**

Je tiens tout d'abord à remercier l'ensemble du jury d'avoir accepté d'évaluer mes travaux de thèse. En particulier Mr Olivier Laprèvote et Mr Jean-Charles Martin, tous deux rapporteurs de ce travail, que je remercie pour leur lecture attentive du manuscrit. Je remercie également les membres examinateurs de ce jury : Mme Véronique Santé-Lhoutellier, Mr Serge Rudaz et Mme Anne Nègre-Salvayre.

Je remercie également Nicolas Cabaton, François Fenaille et Jean-Philippe Antignac d'avoir accepté de faire partie du comité de suivi de thèse et pour leur regard bienveillant sur l'avancement du projet.

Je remercie tout particulièrement mes directeurs de thèse, Françoise Guéraud et Laurent Debrauwer. Je vous remercie pour toute l'aide que vous m'avez apportée au cours de ces trois années, pour vos conseils, pour les nombreuses virgules que vous avez ajoutées à ce manuscrit et bien sûr pour m'avoir donné la possibilité de laisser libre cours à ma volonté d'entreprendre et d'expérimenter. C'était un vrai plaisir d'apprendre à vos côtés et j'espère que nos échanges perdureront bien après cette thèse. Je remercie également Emilien Jamin pour son encadrement tout au long de ce doctorat et pour la confiance qu'il m'a accordée.

Mes remerciements sont également adressés à Fabrice Pierre, pour son implication dans ce projet de thèse et pour m'amener à réfléchir sur l'impact de nos travaux en tant que chercheurs.

Je tiens à remercier l'ensemble des membres des équipes AXIOM et PPCA pour m'avoir accueilli chaleureusement et d'avoir participé à rendre ces 3 ans de thèse particulièrement agréables. Merci à Alyssa, Justine, Elodie, Sylvie, Cécile, Roselyne, et Jean-François. Merci aux membres de mon bureau Kevin, Carla, Robin et Justin de m'avoir supporté. Un merci particulier à Marie grâce à qui j'ai pu mieux comprendre les stats et avec qui j'ai beaucoup apprécié collaborer malgré le gap générationnel nous séparant. Tu vas avoir beaucoup plus de temps maintenant que tu n'as plus à répondre à toutes mes questions.

Merci également à Cécile, Claire, Nathalie, Pascale, Elodie et Jacques avec qui j'ai pu m'initier aux expérimentations animales, culture cellulaire et autres joyeusetés qui m'étaient parfaitement inconnues avant le début de cette thèse. Merci pour votre aide pour ces nombreuses manips. Je remercie également mes deux stagiaires, Fanny et Emma, pour leur aide précieuse et pour la qualité de leur travail. Merci pour ces premières expériences d'encadrement j'aurais difficilement pu rêver mieux.

Merci également aux autres personnes qui ont été impliquées dans ce travail, en particulier Maiwenn pour ton aide sur le microbiote, Florence ainsi que tous animaliers d'avoir pris bien soin de mes petits rats.

## Remerciements

Je remercie également ma presque troisième équipe, MeX, qui malgré mon insistance ne m'a toujours pas convié au repas de Noël. Merci à Nicolas, François, Cyndel, Louisa, Clément, Maxime, Elodie (N'Guyen), Nathalie, Florence et Daniel d'avoir largement participé à l'animation des pauses café. Je remercie également Fabien, qui en plus de sa qualité de consultant international, fut un partenaire de vélo de qualité. Merci beaucoup pour l'aide que tu m'as apporté et je te souhaite d'enfin arriver à monter cette côte de Pujaudran...

Un grand merci à la belle bande de tuches composée de Vincent, Charline, Edwin et Anaïs d'avoir été de supers collègues et amis. Merci pour tous ces moments au labo ou en dehors ! Charline je te souhaite bien du courage pour supporter les deux affreux à qui je souhaite aussi beaucoup de courage pour endurer tes goûts musicaux plus que discutables. Anaïs, je te souhaite une belle fin de thèse et je te fais largement confiance pour préparer la fête d'après-thèse au moins aussi sérieusement que la soutenance.

Je n'oublie pas non plus les anciens thésards que j'ai eu plaisir à côtoyer, merci en particulier à Marie, Jasper et Adèle pour ces soirées et ces discussions sur les tracas du quotidien de doctorant.

Merci également à toutes les personnes croisées au cours de mon parcours, merci en particulier à Floriant et Maud pour m'avoir donné le goût de la métabolomique lors de mon stage de master. Merci aussi à tous mes anciens collègues de Medday pour cette superbe expérience et en particulier à Sandrine qui m'a encouragé à suivre cette voie.

Je remercie enfin tous mes proches, mes amis, qui sans le savoir ont été d'une grande aide durant ces 3 années, merci à Sylvain, Frank, Yannick\*2, Robin, Loic, Bazet, Andréas, Maxime et tous les autres. Merci pour toutes ces années passées et pour celles à venir.

Merci Christelle, sans qui cette thèse serait probablement bien différente. Tu as été d'un soutien immense durant toute cette période et je ne saurais te remercier assez pour tout. J'ai énormément de chance de t'avoir à mes côtés et je me réjouis de notre future vie à Londres.

Pour terminer, un grand merci à ma famille, à mes parents, à mon frère qui ont toujours été là pour moi et c'est notamment grâce à vous que je suis arrivé jusqu'ici et que je peux écrire aujourd'hui cette page de remerciements de thèse. Merci pour tout.

## Table des matières

REMERCIEMENTS	3
TABLE DES MATIERES	5
LISTE DES FIGURES	9
LISTE DES TABLEAUX	13
LISTE DES ABBREVIATIONS	14
CHAPITRE I : INTRODUCTION GENERALE	17
<b>A. Généralités sur le cancer du colon</b>	<b>17</b>
A.1. Incidence et mortalité	17
A.2. Etiologie des cancers colorectaux	18
<b>B. Viande rouge, charcuteries, fer héminique et cancer</b>	<b>21</b>
B.1. Bases épidémiologiques	21
B.2. Autres pathologies	24
B.3. Modèles d'étude de la carcinogénèse colorectale	24
B.4. Hypothèses sur la relation viande rouge et cancer colorectal	29
B.5. Le fer héminique	35
B.6. Carcinogénicité de l'hème	40
<b>C. L'analyse métabolomique pour la toxicologie</b>	<b>44</b>
C.1. Préambule sur les approches omiques	44
C.2. Introduction générale et définitions	45
C.3. Les outils analytiques en métabolomique	46
C.4. La spectrométrie de masse couplée à la chromatographie liquide.	48
<b>D. Le processus d'analyse métabolomique par LC-MS</b>	<b>52</b>
D.1. Planification du schéma expérimental	53
D.2. Préparation des échantillons	53
D.3. Acquisition des données	55
D.4. Traitement du signal	55
D.5. Stratégies de normalisation	57

## Table des matières

D.6. Annotation des données	58
D.7. Analyses statistiques	59
D.8. Interprétation biologique	61
<b>E. La métabolomique au service de la toxicologie</b>	<b>62</b>
E.1. La voie des acides mercapturiques comme marqueur d'exposition	64
<b>Objectifs de la thèse</b>	<b>68</b>
<b>CHAPITRE II : DEVELOPPEMENT ET OPTIMISATION DE LA METHODE D'ANALYSE</b>	<b>72</b>
<b>A. Objectifs</b>	<b>72</b>
A.1. Introduction à l'article	72
A.2. Principaux résultats et conclusion de l'article	88
<b>B. Matériel et méthodes</b>	<b>88</b>
B.1. Réactifs et tampons utilisés	88
B.2. Préparation des échantillons urinaires	88
B.3. Analyse LC-MS	89
B.4. Traitement des données	89
B.5. Détermination de la dilution de travail	89
B.6. Evaluation de la justesse et sensibilité de l'acquisition MS <sup>E</sup>	90
B.7. Optimisations supplémentaires du traitement des données MSE	90
B.8. Comparaison de l'approche MS <sup>E</sup> avec une méthode fast-DDA	91
<b>C. Résultats</b>	<b>92</b>
C.1. Optimisation et évaluation des performances de la méthode	92
C.2. Annotation des données	97
C.3. Etude de méthodes de normalisation	98
C.4. Optimisations supplémentaires mode MS <sup>E</sup>	124
C.5. Comparaison MS <sup>E</sup> avec une méthode fast-DDA (PASEF)	126
<b>D. Discussion et perspectives</b>	<b>126</b>
<b>CHAPITRE III : ETUDE DE L'IMPACT DE LA CONSOMMATION DE VIANDE ROUGE ET DE CHARCUTERIE</b>	<b>129</b>
<b>A. Introduction à l'article</b>	<b>129</b>
A.1. Objectifs de l'étude	129
A.2. Méthodes	130

<b>B. Principaux résultats, discussion et conclusions de l'article</b>	<b>159</b>
<b>CHAPITRE IV : ETUDE DE L'IMPACT DE LA TENEUR EN FER HEMINIQUE, DU CONTEXTE LIPIDIQUE ET DU STATUT EN ANTIOXYDANT</b>	<b>161</b>
<b>A. Introduction</b>	<b>161</b>
<b>B. Matériel et méthodes</b>	<b>162</b>
B.1. Réactifs et tampons utilisés	162
B.2. Analyses LC-MS d'échantillons urinaires	162
B.3. Analyses LC-MS d'échantillons d'eaux fécales	163
B.4. Expérimentations animales et régimes	165
B.5. Dosages du DHN-MA et 8-OH-2-deoxyguanosine par immunodosage	166
B.6. Dosage des TBARS dans les eaux fécales	167
B.7. Cytotoxicité des eaux fécales	167
B.8. Expression génique dans la muqueuse colique	168
B.9. Comptage histologique des lésions prénéoplasiques	168
B.10. Dosage de l'hème dans les eaux fécales	169
B.11. Détermination du taux de phosphorylation de l'histone H2AX	169
B.12. Imagerie par spectrométrie de masse	170
<b>C. Résultats</b>	<b>171</b>
C.1. Les régimes riches en viande rouge induisent une augmentation de la peroxydation lipidique	172
C.2. Les régimes riches en viande rouge induisent une activation transcriptionnelle de gènes de détoxification	178
C.3. Impact sur le métabolome urinaire	180
C.4. Impact d'un régime riche en viande rouge/ fer héminique sur l'épithélium colique	182
C.5. Réponse toxicologique associée	183
<b>D. Discussion et conclusion</b>	<b>188</b>
<b>CONCLUSIONS ET PERSPECTIVES</b>	<b>193</b>
<b>ANNEXES</b>	<b>199</b>
<b>REFERENCES</b>	<b>232</b>





## Liste des figures

Figure 1 Incidence du cancer colorectal en 2018 (Bray <i>et al.</i> , 2018)	17
Figure 2 Premier modèle adénome-carcinome proposé par Fearon et Vogelstein (d'après Fearon and Vogelstein, 1990)	19
Figure 3 Séquence adénome-carcinome suivant la modèle d'instabilité chromosomique ou d'instabilité des microsatellites (Walther <i>et al.</i> , 2009)	21
Figure 4 Foyer de cryptes aberrantes (ACF) de 5 cryptes observé au microscope après coloration au bleu de méthylène	27
Figure 5 Foyer déplété en mucine observé au microscope après coloration HIDAB	28
Figure 6 Formation des amines aromatiques hétérocycliques d'après Jägerstad <i>et al.</i>	30
Figure 7 Absorption et métabolisme de l'hème dans le tractus intestinal (Hooda <i>et al.</i> , 2014)	37
Figure 8 Mécanisme général de la réaction de Fenton (A) Mécanisme proposé de la réaction Fenton-like par le fer héminique (B) Mécanisme proposé de l'oxydation par le fer héminique hypervalent. D'après Carlsen <i>et al.</i>	38
Figure 9 Mécanisme de la peroxydation lipidique	39
Figure 10 Schéma représentatif de la cascade des approches omiques	45
Figure 11 Nombre de publications dans PubMed pour les termes « NMR + Metabolomics » et « MS + metabolomics » entre 2001 et 2020	47
Figure 12 Représentation schématique de la structure d'un spectromètre de masse	48
Figure 13 Représentation schématique du déroulement d'une analyse métabolomique non-ciblée par LC-MS	52
Figure 14 Voie de synthèse des acides mercapturiques. R correspond à un composé électrophile. Adapté de Mathias <i>et al.</i>	65
Figure 15 Schéma général du devenir des composés électrophiles au sein de l'organisme	70
Figure 16 Schéma général de la stratégie analytique utilisée pour l'analyse du métabolome urinaire et du screening des mercapturates par MS <sup>E</sup>	72
Figure 17 Représentation schématique du Synapt-G2-Si (Waters)	92
Figure 18 Comparaison de l'intensité du signal de deux standards de mercapturates en fonction des paramètres de tune. (A) Intensité du signal de l'ion parent du NAC-Phe-Cys (B) Ratio entre l'aire du signal de l'ion fragment et de l'ion parent (C) Intensité du signal de l'ion parent du DHN-MA (D) Ratio entre l'aire du signal de l'ion fragment et de l'ion parent	93
Figure 19 Nombre de mercapturates potentiels détectés (A) Somme des candidats détectés (B) Candidats retrouvés dans chaque triplicat	94
Figure 20 Evolution du signal en fonction de la dilution et du nombre d'injection (A) Signal relatif en % par rapport à la valeur maximale détectée pour chaque dilution (B) Signal observé pour 200 injections d'urine diluée au ¼	95

Figure 21 Stratégie d'extraction automatique des systèmes aldéhydes conjugués à l'aide d'un identifiant SMARTS (les « A » dans la structure correspondent à Any atom)	97
Fig. 22 Concentration variations in urine samples. (A) Dilution factors calculated for both creatinine and osmolality. (B) Correlation of dilution factors calculated according to creatinine levels and osmolality levels	113
Fig. 23 Two-dimensional PCA score plot of HR-MS spectra of urine samples. Each symbol represents an individual data point projected onto the first (horizontal axis) and second (vertical axis) PCA components. Normalization methods are shown in different colours: creatinine in green dots (n=40), none in blue squares (n=40), and osmolality in red triangles (n=40). The black ellipse determines the 95% confidence interval, which was drawn using Hotelling's T2 statistic. (A) Preacquisition normalization: A = 2, R2 = 20%. (B) Post-acquisition normalization: A = 2, R2 = 24%	114
Fig. 24 Radar chart of Q <sup>2</sup> of OPLS-DA models generated for all tested normalization strategies (blue line) compared to Q <sup>2</sup> of OPLS-DA generated from non-normalized model (grey line). For detailed values of Q <sup>2</sup> , see Supplementary S.2	115
Fig. 25 Percentage of (A) variability explained by sex. (B) Residual variability determined by AMOPLS for all tested normalizations. The red bar represents the results from data normalized prior to the acquisition based on osmolality levels	116
Figure 26 Fréquence de détection et nombre de mercapturates potentiels détectés au sein du groupe des gros consommateurs de viande rouge et de charcuterie en fonction de la méthode de traitement des données utilisée	124
Figure 27 Recouvrement des mercapturates potentiels détectés entre chaque groupe pour (A) les mercapturates détectés au sein de chaque réplikat (B) les espèces détectées dans un seul des 6 réplikats. Le disque bleu correspond au groupe des gros consommateurs, le vert à celui des non-consommateurs et le rouge aux individus sélectionnés au hasard	125
Figure 28 Odd-ratios of selected metabolites in the non-nutritionally adjusted model (A) Correlation matrix between red meat consumption and metabolite signal intensity for all metabolites selected in every calculated model (B) Indoxyl sulfate signal intensity measured in both animal experiment and human cohort samples (C) Cinnamoylglycine signal intensity measured in both animal experiment and human cohort samples (D) Coumaric acid signal intensity measured in both animal experiment and human cohort samples (E)	146
Figure 29 Log <sub>2</sub> fold change of all 47 mercapturates detected in the animal experiment. « * » correspond to mercapturates which ViP is > 1 in the OPLS-DA model	147
Figure 30 Concentration en hème dans les eaux fécales (A) dans l'expérimentation de carcinogénèse (B) dans l'expérimentation nutritionnelle. Les valeurs correspondent à des moyennes ± écart-type. La significativité des différents facteurs a été déterminée par (A) ANOVA à deux voies et (B) ANOVA à 3 voies	171
Figure 31 Nombre de mercapturates potentiels détectés dans chaque QC de groupe	172

- Figure 32 Fold change des 47 acides mercapturiques (A) entre les groupes BOSA et CON (B) entre les groupes BOSA et PUSA. Les valeurs sont exprimées en Log<sub>2</sub>(Fold change) une valeur positive correspondant à une augmentation dans le groupe BOSA par rapport au groupe comparé. Une « \* » signifie une annotation putative \_\_\_\_\_ 173
- Figure 33 Fold change des 47 acides mercapturiques entre les groupes BOSAR et BOSA. Les valeurs sont exprimées en Log<sub>2</sub>(Fold change), une valeur négative correspondant à une diminution dans le groupe BOSAR par rapport au groupe BOSA. Une « \* » signifie une annotation putative \_\_\_\_\_ 174
- Figure 34 Evaluation de la peroxydation lipidique induite par les régimes alimentaires par dosage de (A) TBARS dans les eaux fécales (B) DHN-MA urinaire. La significativité est déterminée par ANOVA à 3 facteurs \_\_\_\_\_ 175
- Figure 35 Analyse en composante principale à partir de (A) l'ensemble des variables détectées *via* XMCS (R<sup>2</sup>=32,9%) (B) l'ensemble des variables correspondant aux aldéhydes(R<sup>2</sup>=30,7%). Chaque point représente un individu \_\_\_\_\_ 176
- Figure 36 Fold change des mercapturates retrouvés significatifs (Wilcoxon + fdr) lors des comparaison (A) BOSA/PUSA (B) BOSAR/BOSA. Les valeurs sont exprimées en Log<sub>2</sub>(Fold change). Une « \* » signifie une annotation putative \_\_\_\_\_ 177
- Figure 37 Dosage de (A) TBARS dans les eaux fécales (B) DHN-MA urinaire. La significativité est déterminée à l'aide d'une ANOVA à deux facteurs. \_\_\_\_\_ 178
- Figure 38 Expression relative des gènes impliqués dans la détoxification des aldéhydes (A et B : Aldo keto reductase family 1 member ; B et D : Aldehyde Dehydrogenase 1 Family Member A2) du catabolisme de l'hème (E et F : Heme oxygenase 1) et dans la synthèse de glutathion (G et H : Glutathione synthetase). Les valeurs correspondent à des moyennes ± écart-type. La significativité des différents facteurs a été déterminée par ANOVA à deux voies pour l'expérimentation de carcinogénèse et ANOVA à 3 voies pour l'expérimentation nutritionnelle. Seules les différences significatives sont indiquées \_ 179
- Figure 39 ACP réalisée sur l'ensemble des variables détectées dans les urines des rats issues de (A) l'expérimentation nutritionnelle (B) l'expérimentation de carcinogénèse. Chaque point représente un individu. \_\_\_\_\_ 181
- Figure 40 Corrélation entre les mercapturates annotés au niveau 1 et les métabolites discriminants entre les groupes bœuf et poulet (A) expérimentation nutritionnelle (B) expérimentation de carcinogénèse \_ 182
- Figure 41 Intensité du signal du glutathion [M+K-2H]- (A) dans les colons du groupe BOSA (B) dans les colons du groupe PUSA) \_\_\_\_\_ 183
- Figure 42 Viabilité cellulaire des cellules Apc +/+ exposées aux eaux fécales (A) issues de l'expérimentation de carcinogénèse (B et C) issues de l'expérimentation nutritionnelle. Des lettres identiques signifient aucune différence significative entre les groupes \_\_\_\_\_ 184
- Figure 43 Dosage urinaire du 8-OH-deoxyguanosine (A) dans l'expérimentation de carcinogénèse (B) dans l'expérimentation nutritionnelle \_\_\_\_\_ 185
- Figure 44 Taux de phosphorylation de l'histone H2AX en fonction du type de régime dans l'expérimentation nutritionnelle. \_\_\_\_\_ 186

<b>Figure 45 Lésions prénéoplasiques observées au microscope optique (A) ACF présentant une multiplicité de 5 cryptes (B) MDF présentant une multiplicité de 13 cryptes.</b>	<b>186</b>
<b>Figure 46 Dénombrement des lésions prénéoplasiques dans les colons des rats chimio induits (A) Nombre d'ACF au niveau du colon proximal (B) Nombre d'ACF au niveau du colon distal (C) Comparaison du nombre total d'ACF entre les groupes BOSAR et BOSA (D) Nombre total de MDF (E) Comparaison du nombre total de MDF entre les groupes BOSAR et BOSA. La significativité des comparaisons est déterminée par ANOVA à 2 facteurs pour les comparaisons A, B et D et par t-test pour les comparaisons C et E</b>	<b>187</b>
<b>Figure 47 Mécanisme proposé quant à l'impact de la consommation de viande rouge sur le métabolisme du glutathion</b>	<b>190</b>
<b>Figure 48 Représentation schématique du fonctionnement de (A) une cellule de mobilité classique et (B) une cellule TIMS. E correspond au champ électrique appliqué et <math>v</math> à la vitesse du gaz.</b>	<b>215</b>
<b>Figure 49 Schéma de l'architecture du spectromètre de masse TIMS ToF</b>	<b>216</b>
<b>Figure 50 Schéma général d'une acquisition en mode PASEF (Bruker©)</b>	<b>216</b>

## **Liste des tableaux**

<b>Tableau 1 Facteurs associés à une augmentation ou diminution du risque de CCR. Adapté du rapport du WCRF de 2018 (World Cancer Research Fund International., 2018)</b>	<b>23</b>
<b>Tableau 2 Paramètres du logiciel UNIFI utilisés pour la détection des mercapturates dans les données MS<sup>E</sup></b>	<b>91</b>
<b>Tableau 3 Aire du pic mesurée par le logiciel UNIFI pour deux standards de mercapturates à différentes concentrations : DHN-MA et Nac-Phe-Cys</b>	<b>96</b>
<b>Tableau 4 Composition des régimes expérimentaux. * Toutes les quantités sont exprimées en g pour 100 g eq sec</b>	<b>166</b>

## **Liste des abréviations**

4-HNE *4-hydroxynonéal*  
9-AA *9-aminoacridine*  
AAH *amines aromatiques hétérocycliques*  
ACF *aberrant crypt foci*  
ACP *analyse en composante principale*  
AICR *American Institute for Cancer Research*  
AIF *all ion fragmentation*  
AOM *azoxyméthane*  
APC *adenomatous polyposis coli*  
APCI *atmospheric pressure chemical ionisation*  
APPI *atmospheric pressure photo-ionisation*  
ATNC *Apparent Total Nitroso Compounds*  
AUC *Area under the curve*  
BaP *benzo[a]pyrène*  
BBHA *1-((ammoniooxy)methyl)-2-bromobenzene chloride*  
BHT *2,6-Di-tert-butyl-4-méthyl-phénol*  
BO *boeuf*  
CCR *cancer colorectal*  
CCS *cross collision section*  
CE *électrophorèse capillaire*  
CI *chemical ionisation*  
CIN *instabilité chromosomique*  
CIRC *Centre International de Recherche sur le Cancer*  
CO *monoxyde de carbone*  
CON *contrôle*  
CV *coefficient de variation, coefficient de variation*  
DCC *depleted in colon carcinoma*  
DDA *data driven analysis*  
DHB *acide 2,5-dihydroxy benzoïque*  
DHN-MA *di-hydroxynonane mercapturic acid*  
DIA *data independant analysis*  
DiMeIQx *2-amino-3,4,9-trimethylimidazo[4,5-f]quinoxaline*  
DMEM *Dulbecco/Vogt modified Eagle's minimal essential medium*  
DSS *sodium dextran sulfate*  
EFSA *European Food Safety Authority*  
EGF *facteur de croissance*  
EI *electron ionisation*  
ESI *electrospray ionisation*  
FO *huile de poisson*

## Liste des abréviations

GC *chromatographie gazeuse*  
GLOBOCAN *Global Cancer Observatory*  
Gss *Glutathione synthetase*  
HAP hydrocarbures aromatiques polycycliques  
HCP1 *heme carrier protein 1*  
HID-AB *high iron diamine/alcian blue*  
HILIC *hydrophilic interaction liquid chromatography*  
HMDB *Human Metabolome Database*  
HNPCC *Hereditary Non Polyposis Colorectal Cancer*  
HO-1 *hème-oxygénase 1*  
HO-2 *hème-oxygénase 2*  
IC *ion exchange chromatography*  
IDH *indice de développement humain*  
IMC *indice de masse corporel*  
INF- $\gamma$  *interféron gamma*  
IQ *imidazoquinoline ou imidazoquinoxaline*  
ITO *indium tin oxide*  
KO *knockout*  
LC *chromatographie liquide*  
LD<sub>50</sub> *dose létale médiane*  
LOO\* *radical peroxy*  
MALDI *matrix assisted laser desorption ionisation*  
MCV *maladies cardio-vasculaire*  
MDA *malondialdéhyde, malondialdéhyde*  
MDF *mucin depleted foci*  
MeIQ *MeIQ*  
MeIQx *2-amino-3,8-dimethylimidazo[4,5-f]quinoxaline*  
MeOH) *méthanol*  
Min *Multiple Intestinal Neoplasia*  
MMR *mismatch repair*  
MRM *multiple reaction monitoring*  
MS *spectrométrie de masse*  
MSI *instabilité des microsatellites*  
MSTUS *MS total useful signal*  
MTBE *methyl-tertiobutyl-ether*  
NAC-Phe-Cys *N-acétyl-S-phenylcystéine, N-acetyl-S-phenylcystéine*  
NCE *énergie de collision normalisée*  
Neu5Gc *acide N-glycolylneuraminique*  
NO *oxyde nitrique*  
NOC *composés N-nitrosés*  
OPLS-DA *orthogonal PLS-DA*  
PAF *polypose adénomateuse familiale*



## Liste des abréviations

PASEF *Parallel Accumulation Serial Fragmentation*

PhiP *-amino-1-methyl-6-phenylimidazo [4,5-b]pyridine*

Pirc *Polyposis in the rat colon*

PLS-DA *partial least square-discriminant analysis*

PPCA *Prévention et Promotion de la Carcinogénèse par l'Alimentation*

PU *poulet*

PUFA *acides gras poly-insaturés*

Q *analyseurs quadripolaires*

QC *contrôle qualité*

R *antioxydants*

RMN *résonance magnétique nucléaire*

ROC *receiving operator curve*

ROS *espèces réactives de l'oxygène*

RP *phase inverse ou reverse phase*

SA *huile de carthame*

SMILES *Simplified Molecular Input Line Entry Specification*

SPE *extraction sur phase solide*

SRM *single reaction monitoring*

SVF *sérum de veau fœtal*

SWATH *Sequential window acquisition of all theoretical mass*

TBA *acide thiobarbiturique*

TBARS *acide thiobarbiturique*

TIC *total ion current*

ToF *time of flight*

VIP *variable importance on projection*

WCRF *World Cancer Research Fund*

## Chapitre I : Introduction générale

### A. Généralités sur le cancer du colon

#### A.1. Incidence et mortalité

Selon l'étude la plus récente du Global Cancer Observatory (GLOBOCAN) le cancer colorectal (CCR) représentait le 3<sup>ème</sup> cancer le plus diagnostiqué dans le monde en 2018, soit 1,8 millions de nouveaux cas. (Bray *et al.*, 2018; Ferlay *et al.*, 2019). Tous sexes confondus, le CCR était en 2018 le deuxième cancer le plus mortel et représentait 9,2% des morts par cancer, soit 880 000 morts. L'incidence, qui traduit le nombre de nouveaux cas pour 100 000 habitants sur une période donnée et qui est exprimée en Age-Standardized Rate (ASR), était de 19,7 tous sexes confondus (23,6 pour les hommes contre 16,3 pour les femmes) comme le montre la Figure 1 (Rawla *et al.*, 2019).

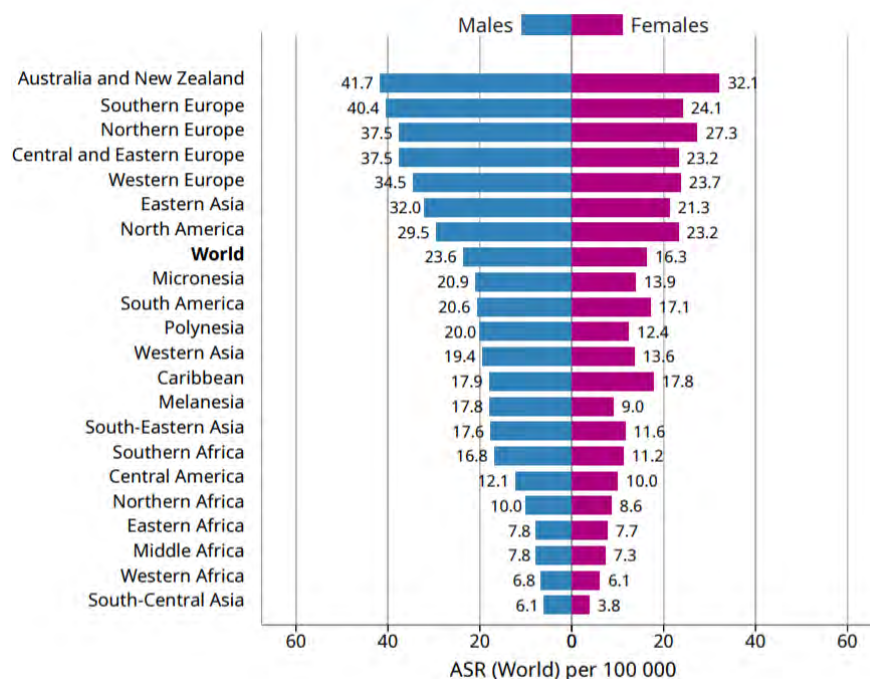


Figure 1 Incidence du cancer colorectal en 2018 (Bray *et al.*, 2018)

A l'instar de nombreux cancers, l'incidence du CCR est positivement corrélée à l'indice de développement humain (IDH). En effet, l'incidence est 5 fois plus élevée dans les pays à IDH élevé comparativement aux pays dont l'IDH est le plus bas (Fidler *et al.*, 2016; Rafiemanesh *et al.*, 2016). Bien que dans les pays aux plus hauts IDH (France, Etats-Unis...) l'incidence et la mortalité aient tendance à baisser en raison des stratégies de prévention et de l'avancée des traitements, l'apparition précoce de CCR est en augmentation pour la tranche des 20-49 ans aux USA (Rawla *et al.*, 2019). L'occidentalisation des modes de vies, caractérisée notamment par une alimentation pauvre en fruits et légumes, une consommation élevée de viande rouge et de charcuterie ou encore peu d'activité

physique est désignée comme le facteur principal participant au développement du CCR (Arnold *et al.*, 2017; Keum and Giovannucci, 2017).

## A.2. Etiologie des cancers colorectaux

Il est possible de distinguer deux grandes classes de CCR, à savoir les formes héréditaires et les formes sporadiques.

Les formes héréditaires représentent 2 à 8% des CCR mais 20% des cancers diagnostiqués avant l'âge de 50 ans (Valle *et al.*, 2019). Deux mutations génétiques autosomiques dominantes (mutation dominante sur un chromosome non sexuel) principales sont à l'origine des CCR familiaux : la polypose adénomateuse familiale (PAF) et le syndrome de Lynch ou *Hereditary Non Polyposis Colorectal Cancer* (HNPCC).

### **A.2.1. La polypose adénomateuse familiale**

Cette maladie se caractérise par l'apparition de très nombreux polypes coliques observables par coloscopie vers la fin de l'adolescence. Ces adénomes multiples prolifèrent dans tout le colon jusqu'à une dégénérescence maligne dans la quasi-totalité des cas autour de 40-50 ans (Lynch and de la Chapelle, 2003). Le nombre de polypes étant déjà trop important dès le début de la maladie pour réaliser une polypectomie par coloscopie, le seul traitement viable repose sur une colectomie totale ou sub-totale (Lynch and de la Chapelle, 2003).

L'origine de cette maladie provient de la mutation d'un gène suppresseur de tumeur situé sur le chromosome 5q21 : **le gène APC** (*adenomatous polyposis coli*). La mutation APC est présente chez tous les sujets atteints de PAF et dans 60-80% des cas de CCR sporadiques (Powell *et al.*, 1992; Smith *et al.*, 1993). En conditions normales, la protéine APC fait partie d'un complexe protéique impliqué dans la voie de signalisation Wnt induisant la phosphorylation de la  $\beta$ -caténine et ainsi son adressage au protéasome. Lorsque le gène APC est muté, la phosphorylation de la  $\beta$ -caténine n'a plus lieu et elle s'accumule dans le cytoplasme. S'en suit une translocation dans le noyau et une liaison au facteur de transcription Tcf altérant l'expression de nombreux gènes impliqués dans la prolifération, la différenciation et la migration cellulaire (Galiatsatos and Foulkes, 2006; Näthke, 2004).

### **A.2.2. Le syndrome de Lynch**

Le syndrome de Lynch est également une mutation autosomale dominante qui représente selon les études, 2 à 7% des cas de CCR héréditaires (Heinen, 2010). Ce syndrome se caractérise par l'apparition précoce de lésions néoplasiques (avant 45 ans) et par une instabilité microsatellitaire. Les microsatellites sont des séquences répétées du génome qui sont altérées chez les patients atteints du syndrome de Lynch et sont la conséquence de mutations du système de réparation des erreurs de

réplication de l'ADN : le système **mismatch repair** (MMR). Bien que 7 gènes aient été identifiés chez l'homme comme participant au système MMR, la majeure partie des mutations rencontrées dans les patients HPNCC se retrouvent sur les gènes *MSH2*, *MLH1* et *MSH6* (Heinen, 2010). Les sujets porteurs de ces mutations ont, outre le risque accru de développer un cancer colorectal, un risque également augmenté de développer des tumeurs extra-coliques au niveau de l'endomètre, de l'estomac, de l'urètre... (Lynch and Smyrk, 1996), ce qui complique le diagnostic. C'est pourquoi en 1991, suite à l'initiative de « l'International Collaborative Group on HNPCC », des critères ont été établis pour aider l'identification des familles atteintes de HPNCC : les critères d'Amsterdam (Rodriguez-Bigas *et al.*, 1997; Umar *et al.*, 2004).

### A.2.2. Les cancers colorectaux sporadiques

Bien que les cas héréditaires ne représentent qu'une mince fraction des cas de CCR, l'étude des mécanismes génétiques impliqués dans le développement des formes héréditaires a permis une meilleure compréhension du développement des formes sporadiques. En effet, la mutation APC et les mutations du système MMR sont retrouvées dans respectivement 70 à 80 % et 15 % des cas de CCR sporadiques (Fearon, 2011).

Le développement d'un CCR sporadique est un **processus multi-étapes**, initié par le développement d'un adénome (tumeur bénigne) évoluant possiblement en carcinome (tumeur maligne) sous l'action de mutations séquentielles activatrices d'oncogènes et inactivatrices de gènes suppresseurs de tumeurs. En 1990, après avoir compilé plusieurs travaux analysant les mutations retrouvées à différents stades de progression des tumeurs coliques, Fearon et Vogelstein proposaient un premier modèle pour décrire la **filiation adénome-carcinome** comme le montre la Figure 2 (Fearon and Vogelstein, 1990).

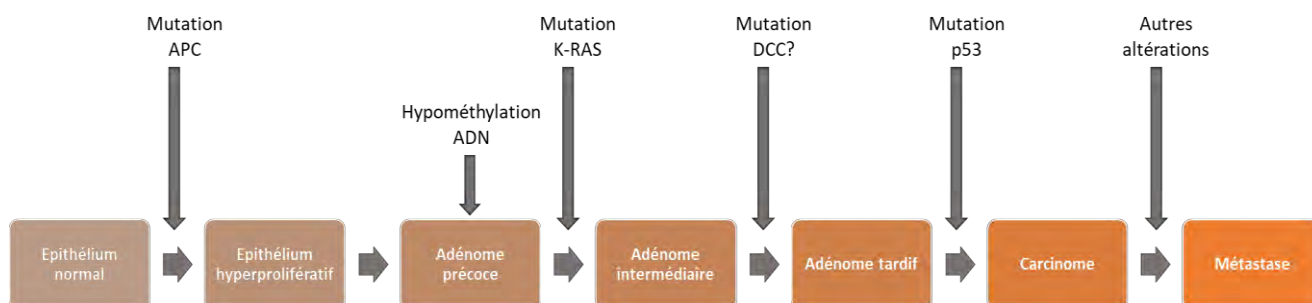


Figure 2 Premier modèle adénome-carcinome proposé par Fearon et Vogelstein (d'après Fearon and Vogelstein, 1990)

Comme décrit dans ce modèle, une succession d'altérations génétiques conférant un avantage sélectif aux cellules porteuses de ces mutations est à l'origine du développement adénome-cancer. Des

mutations au niveau du gène APC constituent la première étape de ce processus séquentiel en induisant une hyper prolifération au niveau de l'épithélium colique. Cependant, la seule mutation d'APC ne suffit pas pour induire le développement d'un adénome. Des mutations au niveau de l'oncogène KRAS sont fréquemment retrouvées (environ 40% des tumeurs colorectales) dans les adénomes précoces et sont fortement impliquées dans la transition de l'épithélium hyperprolifératif en adénome (Nguyen *et al.*, 2020). De plus, il a été montré qu'en l'absence de la combinaison de mutation APC et KRAS, la mutation KRAS seule n'induisait pas d'hyper prolifération, supportant ainsi l'idée d'un processus multifactoriel (Agnantis and Goussia, 2012). A ce stade, la majorité des adénomes formés n'évoluent pas vers un cancer. Les plus petits (< 1 cm) ont 5% de chance d'évoluer vers un cancer avec un risque croissant en fonction de la taille de la tumeur (Agnantis and Goussia, 2012). Dans le modèle de Fearon et Vogelstein, une mutation sur le chromosome 18q entraînant une perte d'expression du gène suppresseur de tumeur DCC (*depleted in colon carcinoma*) constitue la troisième étape de la carcinogenèse, suivie par une mutation du gène TP53 qui constitue la dernière étape de ce processus. TP53 est un régulateur majeur de la transcription des gènes impliqués dans la réparation de l'ADN retrouvé plus fréquemment altéré dans les carcinomes que dans les adénomes et indiquant le caractère tardif de ces mutations dans le processus de tumorigénèse (Russo *et al.*, 2005).

Depuis la première présentation de ce modèle en 1990, d'autres modèles de séquence adénome-carcinome ont été proposés. Parallèlement au modèle d'**instabilité chromosomique** (CIN) présenté par Fearon et Vogelstein, un autre modèle basé sur l'**instabilité des microsatellites** (MSI) a vu le jour permettant ainsi de catégoriser les tumeurs colorectales en deux types : non-hyper-mutées (modèle CIN) et hyper-mutées (modèle MSI) (Cancer Genome Atlas Network, 2012). La Figure 3 illustre les différences entre les modèles CIN et MSI. Comme pour le syndrome de Lynch, le modèle MSI est caractérisé par une instabilité des microsatellites de l'ADN suite à une altération somatique du système MMR et en particulier suite à l'hyperméthylation des îlots CpG du promoteur de MLH1. Bien que ce phénotype soit retrouvé chez tous les patients atteints du syndrome de Lynch, la majorité des tumeurs MSI sont sporadiques et représenteraient 15% des cas de CRC. L'évènement initiateur de la tumorigénèse dans le modèle MSI peut varier et être commun avec le modèle CIN, la mutation APC étant retrouvée dans 35-50% des tumeurs MSI (Nguyen *et al.*, 2020). Par la suite, le développement du CCR *via* la voie MSI fait intervenir des altérations génétiques différentes mais analogues à celle de la voie CIN, en particulier la mutation V600E (remplacement d'un résidu valine par un résidu glutamate au 600<sup>ème</sup> acide aminé de la protéine) de l'oncogène BRAF est retrouvée plus fréquemment que la mutation KRAS (Rajagopalan *et al.*, 2002). Les tumeurs MSI sont en général plus fréquemment

retrouvées chez les femmes âgées et associées à un meilleur pronostic (Samadder *et al.*, 2013; Walther *et al.*, 2009).

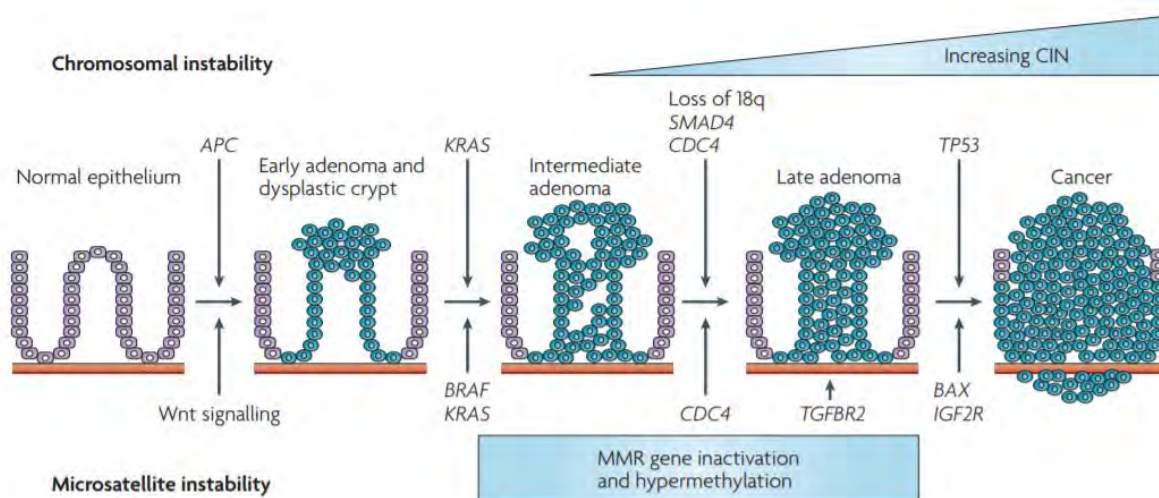


Figure 3 Séquence adénome-carcinome suivant la modèle d'instabilité chromosomique ou d'instabilité des microsatellites (Walther *et al.*, 2009)

## **B. Viande rouge, charcuteries, fer hémérique et cancer**

Pour la suite de ce manuscrit la définition de la viande rouge et de la charcuterie telle que définie par Bouvard *et al.* sera utilisée (Bouvard *et al.*, 2015). A savoir, la viande rouge peut-être définie comme de la viande issue des muscles de mammifères n'ayant pas subi de procédé de transformation. En revanche la charcuterie fait référence à la viande qui a subi des procédés de transformation tels que la salaison ou encore le fumage afin de renforcer sa conservation ou ses propriétés organoleptiques. A l'inverse de la viande rouge, les charcuteries ne se limitent pas à la viande issue de mammifères et peuvent contenir de la volaille ou des sous-produits tels que des abats ou du sang.

### **B.1. Bases épidémiologiques**

Comme vu dans la partie précédente la grande majorité des cancers colorectaux sont sporadiques. En fait, leur développement est étroitement corrélé au mode de vie et à l'exposition à divers facteurs de risques. Selon un consensus maintenant bien établi, le CCR fait partie des cancers les plus évitables en limitant les facteurs de risques. En 1981, Doll et Peto (Doll and Peto, 1981) estimaient que 10 à 70 % des cancers pouvaient être imputés à l'alimentation et particulièrement 90% des cancers colorectaux. Malgré une incertitude élevée, ce rapport pointait l'importance de la relation alimentation-cancer et tout particulièrement du lien étroit avec le CCR. Depuis lors, les estimations ont été revues à la baisse, mais le mode de vie reste un facteur prépondérant dans le développement du CRC. En 2000, Platz *et al.* estimaient que 30 à 50% des cas de CCR étaient directement liés au mode de vie et aux principaux

facteurs de risque : consommation d'alcool, tabagisme, obésité, inactivité physique et consommation de viande rouge et de charcuterie (Platz *et al.*, 2000). Les études plus récentes sont en accord avec ces chiffres et démontrent l'importance du mode de vie par rapport à la prédisposition génétique (Carr *et al.*, 2018; Erdrich *et al.*, 2015). Parmi les facteurs de risque précédemment cités, la consommation de viande rouge et de charcuterie sera tout particulièrement détaillée dans ce manuscrit.

En 1997, un premier rapport d'experts mandatés par le *World Cancer Research Fund* (WCRF) et l'*American Institute for Cancer Research* (AICR) ayant pour objectif de compiler l'ensemble des travaux effectués sur les liens entre mode de vie, nutrition et cancer, a été publié (Glade, 1999). Outre les recommandations sur la réduction de la consommation d'alcool, du tabagisme et du maintien d'une activité sportive régulière, les experts concluent également que la consommation de viande ne doit pas représenter plus de 10% de l'apport énergétique des régimes et recommandent de privilégier la viande blanche ou le poisson à la viande rouge. Depuis sa première parution, ce rapport a été régulièrement mis à jour, notamment en 2007 (Wiseman, 2008) et en 2018, et présente ses recommandations pour limiter les risques de cancer. Le dernier rapport datant de 2018 fait état de dix recommandations visant à réduire le développement de l'ensemble des cancers (World Cancer Research Fund International., 2018) :

- Avoir un poids optimal ( $18,5 < \text{IMC (indice de masse corporelle)} < 25$ )
- Pratiquer une activité physique régulière (45-60 minutes par jour)
- Avoir une alimentation riche en céréales complètes, légumes, fruits et légumineuses
- Limiter la consommation de « *fast-foods* », d'aliments transformés riches en gras ou en sucre
- **Limiter la consommation de viande rouge et de charcuterie**
- Limiter la consommation de boissons sucrées et sodas
- Limiter la consommation d'alcool
- Ne pas utiliser de compléments alimentaires pour atteindre les recommandations nutritionnelles
- Pour les mères : allaiter si possible
- Après un cancer : suivre ces recommandations

La limitation de la consommation de viande rouge et de charcuterie est tout particulièrement recommandée dans un objectif de prévention du CCR. En effet, après une méta-analyse de 99 études à travers le monde, les experts de ce comité ont établi une liste d'aliments et habitudes alimentaires ayant un impact sur le risque du développement du CRC, en les classant par niveau de preuve d'association (Tableau 1).

**Tableau 1 Facteurs associés à une augmentation ou diminution du risque de CCR. Adapté du rapport du WCRF de 2018 (World Cancer Research Fund International., 2018)**

		Réduction du risque	Augmentation du risque
<b>Preuves solides</b>	Convaincant	Activité physique	<b>Charcuteries</b> Alcool Surcharge pondérale Taille et âge adulte
	Probable	Céréales complètes Fibres Produits laitiers Suppléments en calcium	<b>Viande rouge</b>
<b>Preuves limitées</b>	Limité – suggestif	Aliments contenant de la vitamine C Poisson Vitamine D Suppléments multivitaminés	Faible apport en légumes non amylicés Faible apport en fruits Nourriture contenant du fer hémique
	Limité - sans conclusion	Produits céréaliers, pomme de terre, graisse animale, volaille, produits de la mer, composition en acide gras, cholestérol, acides gras $\omega$ -3 alimentaires <i>via</i> le poisson, légumineuses, ail, sources de calcium non quotidiennes, produits avec sucres ajoutés, saccharose, café, thé, caféine, glucides, graisse totale, amidon, charge et index glycémique, folate, vitamine A, E et B6, sélénium, peu de graisse, méthionine, $\beta$ -carotène, $\alpha$ -carotène, lycopène, rétinol, apport énergétique, fréquence des repas	

Des travaux similaires ont été réalisés en 2015 par un comité d’experts réunis par le Centre International de Recherche sur le Cancer (CIRC) (Bouvard *et al.*, 2015). Dans ce travail, plus de 800 études épidémiologiques et expérimentales portant sur l’association viande rouge/charcuteries et cancer ont été considérées pour déterminer le risque de développement du CCR et aussi d’autres cancers associés à la consommation de ces produits. Ainsi, les experts ont pu démontrer une relation dose-réponse significative avec une augmentation du risque de 17% (95% IC : 1.05-1.31) pour 100 g quotidiens de viande rouge et 18% (95 % IC : 1.10-1.28) pour 50 g quotidiens de charcuterie. Suite à cette étude la **charcuterie** a été classée dans le Groupe 1 selon la classification du CIRC qui regroupe les agents (composés chimiques, facteurs physiques, aliments, mélanges, etc...) qui sont **cancérogènes pour l’homme**. La **viande rouge** quant à elle a été classée dans le Groupe 2A qui regroupe les agents **probablement cancérogènes pour l’homme**.

En dehors de ces études menées par des organismes internationaux, de nombreuses études épidémiologiques viennent appuyer ces associations. Une étude publiée récemment basée sur une cohorte anglaise a estimé que les sujets mangeant plus de 76 g / jour de viande rouge et 21 g / jour de charcuteries avaient un risque de développement de CCR augmenté de 20% (Bradbury *et al.*, 2020). Une méta-analyse récente basée sur 148 articles a déterminé que les plus gros consommateurs de charcuteries avaient un risque de cancer colorectal accru de 18% en comparativement aux plus faibles consommateurs, 17% quand l’apport total de viande rouge (viande rouge + charcuteries) était



considéré, et 10% quand il s'agissait seulement de viande rouge (Farvid *et al.*, 2021). La multiplication de ce type d'études épidémiologiques et la concordance des conclusions ont contribué à renforcer l'association entre viande rouge et transformée et cancer colorectal mais ont également conduit à étudier l'impact de cette consommation sur d'autres types de pathologies.

### **B.2. Autres pathologies**

Compte-tenu de l'association étroite entre viande rouge et charcuteries et CCR, d'autres travaux visant à évaluer l'impact de cette consommation sur d'autres types de cancer ou encore sur diverses maladies chroniques ont été menés. En particulier, une augmentation de l'incidence de maladies cardio-vasculaire (MCV) associée à une consommation élevée de viande rouge et de charcuterie a été rapportée dans plusieurs études (Abete *et al.*, 2014b; Zhong *et al.*, 2020). En 2014, une méta-analyse a déterminé qu'en comparaison au groupe des plus faibles mangeurs, les gros consommateurs présentaient un risque accru de MCV de 18% pour la charcuterie et de 16% quand il s'agissait de viande rouge (Abete *et al.*, 2014a). Aucune association entre viande blanche et MCV n'a été déterminée dans cette étude. A l'inverse, une étude récente a déterminé qu'une consommation élevée de viande rouge, charcuteries mais aussi volailles était associée à un risque accru de MCV (Zhong *et al.*, 2020). Au-delà des fréquentes associations avec les MCV, une consommation élevée de viande rouge et de charcuterie a également été associée à d'autres maladies chroniques telles que le diabète de type II (Pan *et al.*, 2011) ou encore d'autres localisations de cancers telles que le cancer du sein (Diallo *et al.*, 2018) du pancréas ou de la vessie (Les cancers en France, Les Données, INCa, édition 2015). Tous ces travaux ont conduit à l'établissement d'associations entre mortalité toutes causes confondues et la consommation élevée de viande rouge et de charcuterie. Ces résultats sur la mortalité globale sont particulièrement convaincants pour les charcuteries tandis que le niveau de preuve pour la viande rouge semble moins élevé (Abete *et al.*, 2014a; Larsson and Orsini, 2014; Wang *et al.*, 2016).

### **B.3. Modèles d'étude de la carcinogénèse colorectale**

Les études épidémiologiques utilisées pour déterminer l'existence d'associations entre l'exposition à un ou plusieurs facteurs et le développement d'une pathologie ne suffisent pas à conclure quant à l'existence d'une causalité entre les deux phénomènes. Pour cela, les associations observées en épidémiologie doivent être complétées de données expérimentales. Pour des raisons éthiques évidentes, les études portant sur des agents cancéreux ou possiblement cancéreux sur l'Homme sont impossibles et il est donc nécessaire d'avoir recours à des modèles animaux ou cellulaires permettant de confirmer ou non les associations mises en évidence par les études épidémiologiques et d'étudier les mécanismes impliqués dans cette relation alimentation/développement du cancer. Concernant le

CCR, de nombreux modèles sont disponibles présentant tous des avantages et inconvénients comme détaillé dans le paragraphe suivant.

### ***B.3.1. Modèles in vivo***

La force des modèles *in vivo*, pour l'étude de pathologies complexes telles que le CCR, réside dans le fait que de tels modèles permettent de mesurer les effets d'un composé cancérigène à l'échelle d'un organisme entier. De par leur proximité génétique avec l'homme et surtout en raison d'un processus de cancérisation proche de celui de l'homme, les rongeurs sont considérés comme de bons modèles pour étudier la carcinogénèse colorectale (Corpet and Pierre, 2005). On peut distinguer deux grands types de modèle : les modèles chimio-induits et les modèles génétiques.

Le modèle chimio-induit repose sur l'administration d'un composé inducteur de la carcinogénèse tel que l'azoxyméthane (AOM) à des rats ou des souris. Ce modèle, qui est celui utilisé au cours de cette thèse, est largement employé pour étudier l'effet promoteur ou protecteur de divers composés sur le développement de la carcinogénèse colorectale (Silva *et al.*, 2015; Suzui *et al.*, 2005). Ce modèle développe spontanément des lésions prénéoplasiques 2 à 3 mois après injection, et des tumeurs se développant 7 à 9 mois après injection (Femia and Caderni, 2008). Le rat présente l'avantage d'avoir une LD<sub>50</sub> (dose létale médiane) plus élevée que la souris ce qui permet de ne réaliser qu'une seule injection, contrairement à la souris, permettant ainsi de limiter les risques pour l'expérimentateur. Outre l'aspect économique, l'avantage du modèle chimio-induit, par rapport aux animaux transgéniques, repose sur une séquence de développement des tumeurs relativement similaire à celle observée chez l'Homme. De plus, des altérations génétiques présentes chez l'Homme, telles que des mutations sur le gène K-RAS, sont également retrouvées chez les animaux chimio-induits. En revanche, malgré de nombreux avantages, ce modèle comporte quelques limitations en dehors du fait que l'Homme n'est jamais exposé à ce composé carcinogène. La principale limitation réside dans le fait que la mutation APC qui est retrouvée dans environ 80% des tumeurs colorectales humaines n'est retrouvée dans ce modèle qu'à un taux plus faible d'environ 33% (Perše and Cerar, 2011).

Par opposition aux modèles chimio-induits, les animaux transgéniques portent une mutation, souvent sur le gène Apc, et développent ainsi spontanément des tumeurs. La souris Min (*Multiple Intestinal Neoplasia*) et le rat Pirc (*Polyposis in the rat colon*) sont les deux modèles transgéniques les plus utilisés pour l'étude de la carcinogénèse colorectale. La principale limitation du modèle Min est due à la localisation des tumeurs qui se forment essentiellement au niveau intestinal (40 tumeurs intestinales pour 1 tumeur colique environ) contrairement au rat Pirc (Amos-Landgraf *et al.*, 2007). De plus, il est admis que ces modèles constituent principalement des modèles de PAF, qui ne représentent qu'une faible proportion des cas de CCR (Nascimento-Gonçalves *et al.*, 2021). Des alternatives aux modèles murins ont également été proposés avec le développement d'un modèle porcin de PAF (Flisikowska

*et al.*, 2012).

Plus récemment des modèles basés sur des xénogreffes ont été décrits afin d'étudier des stades plus tardifs de la pathologie (Nascimento-Gonçalves *et al.*, 2021). La génération de ces modèles repose sur l'implantation de cellules tumorales ou de tumeurs au niveau de leur site de développement d'origine (modèle orthotopique) ou sur un site différent (modèle hétérotopique). Ces modèles permettent notamment d'étudier l'invasion tumorale, les processus de vascularisation des tumeurs ainsi que les métastases mais nécessitent cependant des compétences techniques importantes (Oliveira *et al.*, 2020).

### ***B.3.2. Marqueurs précoces de la carcinogénèse colorectale***

Afin de déterminer l'effet de certains composés ou aliments sur la carcinogénèse colorectale, de nombreuses études s'appuient sur l'étude des lésions pré-néoplasique. Idéalement, l'utilisation de ces marqueurs permet de conclure quant à l'effet carcinogène des composés étudiés sans aller jusqu'au stade tumoral ce qui permet de réduire à la fois la durée des expérimentations mais surtout la souffrance animale induite par les études de carcinogénèse. De plus, l'impact de l'alimentation étant plus important aux stades précoces du développement du cancer, de nombreuses études visent à évaluer l'effet de certains aliments sur le développement de marqueurs précoces du cancer. En particulier, dans le cas de la relation viande rouge/charcuterie et CCR, l'effet promoteur de la consommation est étudié *via* le dénombrement de lésions préneoplasiques, les ACF (*aberrant crypt foci*) et les MDF (*mucin depleted foci*).

### B.3.2.1. ACF

Les ACF sont des lésions généralement caractérisées par un épithélium épaissi et des cryptes présentant un orifice déformé et élargi comme observable sur la Figure 4.

Aberrant crypt foci (ACF)

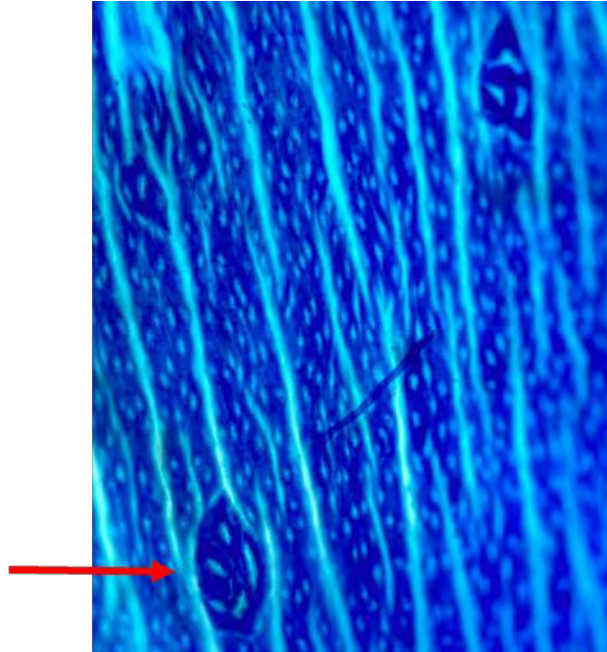


Figure 4 Foyer de cryptes aberrantes (ACF) de 5 cryptes observé au microscope après coloration au bleu de méthylène

Ces changements morphologiques sont liés à une hyper prolifération des cellules des cryptes. Depuis leur première description en 1987 (Bird, 1987), les ACF sont largement utilisés comme marqueurs précoces de la carcinogénèse colorectale en comparaison aux autres lésions (Femia and Caderni, 2008). Compte-tenu de l'importance d'un marqueur précoce prédictif d'une évolution vers le stade cancer, beaucoup d'études se sont intéressées au lien ACF-tumeur. Il est à noter que la majorité des ACF ne vont pas jusqu'au stade tumoral et régressent souvent d'eux-mêmes (Clapper *et al.*, 2020). Cependant, la corrélation entre nombre d'ACF et le degré de dysplasie a été fréquemment décrite (Kowalczyk *et al.*, 2020; Sakai *et al.*, 2011; Siu *et al.*, 1997) ce qui justifie son utilisation comme marqueur précoce de carcinogénèse colorectale. Il est en revanche à noter que l'utilisation de ce marqueur reste controversée car toutes les études ne concluent pas quant à une corrélation entre ACF et risque de CCR (Quintanilla *et al.*, 2019). Outre l'absence de corrélation, les principales critiques émises vis-à-vis des ACF concernent les altérations génétiques retrouvées dans ces lésions. En effet, la mutation du gène APC, qui constitue l'altération génétique la plus fréquemment retrouvée chez les patients développant un CCR sporadique, n'est quasiment jamais retrouvée dans les ACF de ces mêmes patients (Clapper *et al.*, 2020). En revanche la mutation K-Ras, fréquemment présente dans la séquence adénome-carcinome, est retrouvée dans plus de 50% des ACF. Cependant, cette mutation est autant retrouvée dans les ACF dysplasiques, c'est-à-dire présentant des modifications épithéliales

similaires à celles observées dans les adénomes (par exemple une perte de polarité de la cellule, un rapport noyau/cytoplasme important...), que non-dysplasiques ce qui limite son utilisation pour prédire l'évolution de la dysplasie (Gupta *et al.*, 2007).

#### B.3.2.2. MDF

Les MDF, décrits plus récemment par le groupe de G. Caderni (Caderni *et al.*, 2003), sont des lésions préneoplasiques caractérisées par (1) une absence ou une très faible production de mucine (2) une distorsion de la crypte et (3) une surépaisseur de la région par rapport au reste de l'épithélium colique. Contrairement aux ACF qui sont caractérisés par une coloration plus intense, les MDF présentent, après une coloration spécifique des mucines (coloration HIDAB), une absence de coloration due à l'absence de sécrétion de mucine (Figure 5).

Mucin depleted foci (MDF)

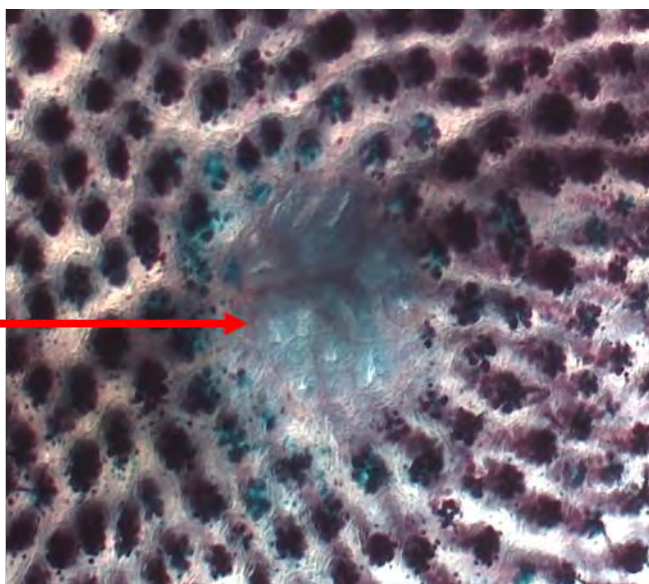


Figure 5 Foyer déplété en mucine observé au microscope après coloration HIDAB

Plusieurs études chez le rat et l'homme laissent à penser que les MDF pourraient être des marqueurs prédictifs de carcinogénèse. Chez l'Homme, Sakai *et al.* ont dénombré les ACF et MDF dans 53 échantillons issus de patients présentant un CCR sporadique et ont observé que les ACF ne présentaient pas de dysplasie contrairement aux MDF qui présentaient une dysplasie de bas-grade (Sakai *et al.*, 2011). Concernant les altérations génétiques retrouvées dans les MDF, il a été montré que, chez le rat chimio-induit au DMH, 25% des MDF, 30% des tumeurs et 0% des ACF présentaient une mutation APC (Femia *et al.*, 2007). Cependant, ACF et MDF ne sont pas à opposer car plusieurs travaux montrent une forte corrélation entre les deux marqueurs et en particulier pour les « flat-ACF » (Femia *et al.*, 2008; Sødning *et al.*, 2016).

## **B.4. Hypothèses sur la relation viande rouge et cancer colorectal**

Afin d'expliquer d'un point de vue mécanistique les associations observées entre viande rouge/charcuteries et développement du CCR, de nombreuses hypothèses sont actuellement débattues dans la littérature. Bien qu'il n'y ait pas à ce jour de consensus sur les mécanismes, plusieurs composés, soit présents de manière endogène, soit formés lors de la cuisson, ajoutés pour la conservation ou encore formés au cours de la digestion sont suspectés pour leur effets pro-carcinogènes. Parmi les hypothèses proposées dans la littérature, l'implication du fer héminique contenu dans la viande rouge et charcuterie semble être un élément central dans la carcinogénèse colorectale induite par la consommation de ces produits. Dans cette introduction une partie spécifique sera donc dédiée à la présentation de cette hypothèse

### ***B.4.1. Les amines aromatiques hétérocycliques (AAH)***

#### *B.4.1.1. Formation des AAH*

Parmi les composés suspectés, les AAH sont depuis longtemps étudiées pour leurs potentiels effets carcinogènes. Ces composés sont présents dans les gaz d'échappement, la fumée de cigarette mais aussi dans les viandes cuites à haute température (viande frite ou cuisson de type barbecue). On peut distinguer deux types d'AAH, les AAH formées entre 100 et 300°C, appelées AAH thermiques ou de type IQ (imidazoquinoline ou imidazoquinoxaline) et celles formées au-delà de 300 °C, appelées AAH pyrolytiques (Gibis, 2016). Les AAH thermiques sont présentes dans la majorité des produits animaux cuits à haute température puisque les précurseurs de leur formation, à savoir la créatine, certains acides aminés et hexoses sont présents dans ces matrices. Leur formation résulte d'une réaction de Maillard entre un acide aminé et un hexose suivie par une dégradation de Strecker, puis une réaction entre ce produit et la créatinine formée par action de la chaleur sur la créatine comme le montre la Figure 6 (Jägerstad *et al.*, 1998). Ce schéma réactionnel est retrouvé dans la formation de la quasi-totalité des AAH à l'exception du 2-amino-1-méthyl-6-phenylimidazo [4,5-b]pyridine (PhiP). Les AAH les plus fréquemment retrouvées dans les produits carnés sont : la 2-amino-3,8-diméthylimidazo[4,5-f]quinoxaline (MeIQx), la 2-amino-3,4,9-triméthylimidazo[4,5-f]quinoxaline (DiMeIQx), la 2-Amino-3-méthylimidazo[4,5-f]quinoline (IQ), la 2-Amino-3,4-diméthylimidazo[4,5-f]quinoline (MeIQ) et le PhiP (Barzegar *et al.*, 2019; Sugimura *et al.*, 2004) et tous appartiennent à la classe des AAH thermiques (ou polaires). Leur formation est influencée par de nombreux facteurs tels que le type et le temps de

cuisson (Khan *et al.*, 2008), le type de viande ou poisson (Oz and Kotan, 2016) ou encore le taux de graisse (Jägerstad *et al.*, 1998).

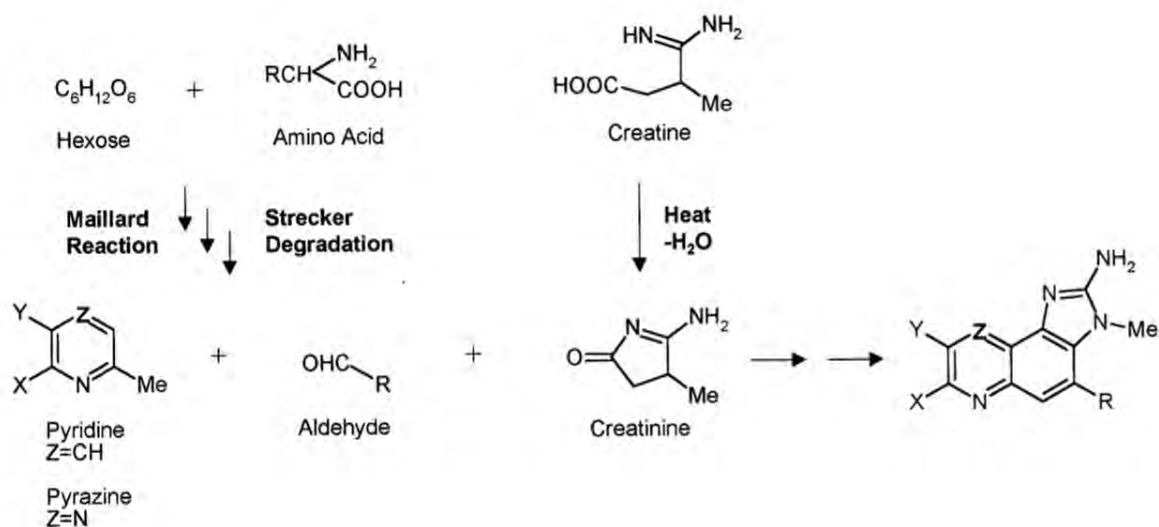


Figure 6 Formation des amines aromatiques hétérocycliques d'après Jägerstad *et al.*

#### B.4.1.2. Carcinogénicité des AAH

Bien que certaines AAH soient classées dans le groupe 2B (carcinogène possible) tels que : MeIQx, MeIQ et PhIP ou encore dans le groupe 2A (carcinogène probable) comme l'IQ, ces molécules ne sont pas intrinsèquement mutagènes. En effet, comme un grand nombre de molécules cancérigènes, leur action mutagène n'intervient qu'après bioactivation métabolique de ces molécules par l'organisme conduisant à leur bioactivation. Les AAH sont absorbées dans l'intestin grêle puis transportées dans le foie où elles sont activées, par le cytochrome P450 (principalement les cytochromes 1A1, 1A2 et 2B1). Une fois activées ces molécules peuvent s'adduire à l'ADN et induire des mutations (Turesky and Le Marchand, 2011). Dans les années 80-90, de nombreuses études ont été réalisées sur des rongeurs afin d'étudier le potentiel carcinogène d'une exposition alimentaire à ces composés (Kato *et al.*, 1989; Ohgaki *et al.*, 1987, 1984; Takayama *et al.*, 1984). Ainsi, une augmentation significative du nombre de tumeurs dans diverses localisations telles que le foie, le colon, l'intestin, les glandes mammaires a été observée. Dès lors, des études épidémiologiques ont été menées dans l'objectif d'évaluer l'association entre exposition aux AAH et développement de cancer (en particulier du CCR) et ont montré une augmentation du risque lié aux AAH (Fu *et al.*, 2011; Helmus *et al.*, 2013). Cependant, d'autres études n'ont pas montré de telles associations (Ollberding *et al.*, 2012). Les disparités observées dans les conclusions de ces différentes études s'expliquent notamment par la difficulté de dé-corréler l'effet spécifique des AAH et celui des autres carcinogènes de la viande. La viande blanche cuite contient également des AAH et sa consommation n'est pas corrélée à une augmentation du risque de cancer.

De plus, il a été montré que pour les mêmes AAH dérivées de la viande blanche ou rouge l'effet promoteur du cancer n'était observé qu'avec les AAH dérivées de la viande rouge (Helmus *et al.*, 2013), ce qui suggère qu'un autre facteur favorisant entre en jeu. Enfin, il semblerait que le **polymorphisme génétique** de certains gènes (en particulier NAT1, NAT2 et GST1) soit impliqué dans une sensibilité différente aux AAH (Eichholzer *et al.*, 2012).

#### ***B.4.2. Les hydrocarbures aromatiques polycycliques (HAP)***

De la même manière que les AAH, les HAP sont des composés également formés lors de la cuisson des viandes et proviennent de la combustion incomplète de composés organiques. Présents de manière quasi-ubiquitaire, les hommes sont exposés aux HAP *via* de nombreuses voies telles que l'alimentation, l'air ambiant, le tabagisme, les émissions industrielles... Dans un rapport publié en 2008, l'*European Food Safety Authority* (EFSA) annonçait qu'environ 500 HAP avaient été détectés dans l'air ambiant ("Polycyclic Aromatic Hydrocarbons in Food - Scientific Opinion of the Panel on Contaminants in the Food Chain," 2008). Cependant, pour des non-fumeurs, la majeure partie de l'exposition à ces composés intervient *via* l'alimentation, et notamment par la viande cuite. Les HAP sont produits lors de la cuisson de la viande, en particulier lors des cuissons au grill ou au barbecue et retrouvés à des niveaux de l'ordre de 130 µg / kg contre 0,01-1 µg/kg dans les viandes crues (Jeyakumar *et al.*, 2017) (Oz and Yuzer, 2016). Du fait de leur présence ubiquitaire et de leur action mutagène, les HAP ont été largement étudiés. Dès 1987, le benzo[a]pyrene (BaP) qui est le composé le plus étudié, a été reconnu comme cancérogène probable par le CIRC (groupe 2A) avant d'être reclassé dans le groupe I (cancérogène pour l'homme) en 2012 (Demeyer *et al.*, 2016). De même que pour les AAH, les HAP, une fois métabolisés, ont la capacité de former des liaisons covalentes avec les bases de l'ADN et induire ainsi un environnement mutagène propice au développement de cancers (Xiao *et al.*, 2016).

Malgré une génotoxicité démontrée *in vitro* et *in vivo* (Einaudi *et al.*, 2014; Muthusamy *et al.*, 2018), toutes les études épidémiologiques ne font pas état d'associations claires entre exposition alimentaire aux HAP *via* l'ingestion de viande rouge et développement du CRC. En effet, bien que certaines études épidémiologiques aient montré une association positive entre les HAP (en particulier le BaP) et un risque accru de CRC (Gunter *et al.*, 2005; Sinha *et al.*, 2005), d'autres études ne montrent pas d'association significative avec les HAP, mais plutôt avec les autres composés de la viande rouge (Cross *et al.*, 2010; Shin *et al.*, 2007). En revanche, de nombreuses études récentes montrent une association positive entre l'exposition aux HAP (toutes sources confondues) et le développement du cancer du sein (White *et al.*, 2016). De plus, il est notoire que certains aliments non associés au développement du CRC tels que les produits céréaliers, la volaille ou le poisson contiennent des niveaux similaires de



HAP renforçant ainsi l'hypothèse que les HAP seuls ne seraient pas l'élément principal de la promotion du CRC par la viande rouge (Demeyer *et al.*, 2016).

#### ***B.4.3. L'hypothèse Neu-5Gc***

Plus récemment, une nouvelle hypothèse faisant intervenir un composé exogène a été proposée pour expliquer le lien entre consommation de viande rouge et développement du CRC : l'hypothèse de l'acide N-glycolylneuraminique (Neu5Gc). Cet acide sialique, qui dérive de l'hydroxylation de l'acide sialique humain N-acétylneuraminique, n'est plus produit de manière endogène par les humains en raison d'une mutation évolutive sur le gène CMAH (Tangvoranuntakul *et al.*, 2003). En revanche, comme ce gène est actif chez la quasi-totalité des autres mammifères, des teneurs plus ou moins importantes de Neu5Gc sont retrouvées dans les produits issus de ces mammifères et notamment dans la viande rouge qui en contient des niveaux importants (Bashir *et al.*, 2020; Samraj *et al.*, 2015). D'un point de vue mécanistique, lors de la consommation de produits à haute teneur en Neu5Gc, ce composé s'incorpore dans les tissus humains avant d'être reconnu par des anticorps circulants anti-Neu5Gc. La formation de ce complexe antigène-anticorps promouvrait ou génèrerait une inflammation chronique favorisant ainsi le développement du cancer ou d'autres pathologies (Samraj *et al.*, 2015). En effet, dans une étude sur des souris KO (*knockout*) pour le gène Cmah et immunisées contre ce composé, il a été montré que, par rapport au groupe contrôle, les souris nourries avec un régime riche en Neu5Gc présentaient des niveaux élevés de marqueur d'inflammation dans le plasma ainsi qu'une augmentation de l'incidence de carcinomes hépatiques (Samraj *et al.*, 2015). Plus récemment, il a été montré qu'un régime enrichi en Neu-5Gc (à l'aide d'une supplémentation en mucine submaxillaire porcine) induisait une dysbiose intestinale, chez la souris KO, conduisant à un microbiote moins diversifié et qu'une sous-représentation des espèces microbiennes capables de métaboliser ce composé participerait au phénotype inflammatoire (Zaramela *et al.*, 2019). Cependant, dans une étude prospective humaine visant à déterminer la présence d'une association entre le taux d'anticorps anti-Neu5Gc et le risque de développement du cancer colorectal, aucune association n'a été rapportée (Samraj *et al.*, 2018).

Ainsi, bien que certaines études tendent à montrer l'implication du Neu5Gc dans la carcinogénèse colorectale, le nombre restreint d'études sur le sujet ne permet pas de conclure avec certitude quant au lien de causalité entre Neu5Gc et carcinogénèse. De plus la présence de ce composé dans d'autres produits animaux tels que le lait ou certains fromages qui ne sont pas associés à un risque accru de cancer reste un élément à considérer pour des études futures.

#### ***B.4.4. Les composés N-nitrosés***

Les nitrates et nitrites sont des additifs hautement utilisés dans l'industrie des viandes transformées, en majeure partie pour des raisons sanitaires. En effet, leur ajout lors de la fabrication de ces produits permet d'inhiber le développement de *Clostridium botulinum* (bactérie responsable du botulisme) et la croissance bactérienne de manière générale *via* la production d'oxyde nitrique (NO) (Majou and Christieans, 2018). Outre l'amélioration de la qualité sanitaire des aliments, les nitrites sont également utilisés pour renforcer les propriétés organoleptiques des viandes transformées et sont notamment responsables de la couleur rosée des charcuteries appréciée des consommateurs (Morales *et al.*, 2013). Cependant, leur usage est aujourd'hui remis en question en raison de leur capacité à induire la formation de composés N-nitrosés (NOC), dont certains sont connus pour leur activité mutagène. La formation de ces composés passe par une **réaction de nitrosation** entre un agent nitrosant tel que les nitrites et un substrat nitrosable, les plus fréquents étant les amines secondaires et les N-alkylamides, conduisant ainsi à la formation de nitrosamines et de nitrosamides respectivement (Mirvish, 1995). Il est à noter que seules les nitrosamines formées à partir d'amines secondaires donnent lieu à des produits stables (de La Pomélie *et al.*, 2017). Les nitrites peuvent également réagir avec le fer contenu dans la myoglobine (protéine majoritaire et caractéristique de la viande rouge) *via* une **réaction de nitrosylation** pour former de l'hème nitrosylé. L'exposition aux NOC peut être divisée en deux catégories : les NOC apportés par l'alimentation et les NOC formés de manière endogène, notamment *via* la capacité de la flore buccale à réduire les nitrates en nitrites.

##### B.4.4.1. Formation des NOC dans les charcuteries

Bien qu'il soit estimé que l'exposition aux NOC ait lieu surtout de manière endogène (de 50 à plus de 90% selon les études et selon le type de NOC) (de La Pomélie *et al.*, 2017; Tricker, 1997), l'Homme est également exposé aux NOC de manière exogène et en particulier *via* l'alimentation. Les charcuteries et viandes fumées sont une source importante à la fois de NOC mais aussi de précurseurs de ces produits. En effet, ces aliments sont riches en protéines, qui peuvent libérer des nitrosamines lors de la conservation ou de la digestion, et enrichis en nitrite (de La Pomélie *et al.*, 2017). Au pH de la viande, environ 5,5, la réaction de nitrosation entre les nitrites et les amines secondaires est très peu présente. En effet, cette réaction nécessite un pH acide afin que les nitrites soient convertis en ions  $\text{NO}^+$  (Honikel, 2008). Afin de réduire la formation de NOC lors des fabrications industrielles, l'ajout d'acide ascorbique est quasi systématique. Cependant, le pouvoir antioxydant de l'acide ascorbique est contrebalancé par sa capacité à catalyser la formation de nitrosyl-hème (Majou and Christieans, 2018).

#### B.4.4.2. Formation des NOC endogènes

Dans les années 2000-2010, les travaux de l'équipe de Sheila Anne Bingham ont permis une meilleure compréhension de la formation des NOC. Comme vu précédemment, le pH est un élément important dans les réactions de formation des NOC. Ainsi, tout au long de la digestion, les réactions conduisant à la formation de ces composés vont varier en fonction de l'environnement considéré. Dans les conditions acides de l'estomac, les réactions de nitrosation sont favorisées conduisant à la formation de nitrosamines, nitrosamides et S-nitrosothiols dans ce compartiment (Kuhnle *et al.*, 2007). Une fois que les S-nitrosothiols atteignent le milieu plus alcalin (pH compris entre 6 et 7,4) de l'intestin grêle, ces composés deviennent instables et sont dégradés en oxyde nitrique, qui peut réagir avec l'hème et ainsi participer à la formation d'hème nitrosylé (Steinberg, 2019). Outre le pH, d'autres facteurs jouent également un rôle dans la formation des NOC, notamment l'intervention du microbiote qui a été démontrée par Massey *et al.* comme étant impliquée dans la formation des NOC en comparant les niveaux de NOC totaux (*ATNC Apparent Total Nitroso Compounds*) chez des rats axéniques et conventionnels (Massey *et al.*, 1988). La formation de ces composés par la flore intestinale interviendrait plutôt dans le colon distal où le pH moins acide (pH compris entre 6,1 et 7,5) que le colon proximal (pH compris entre 5,5 et 7,5) serait plus optimal pour l'action des bactéries responsables de la formation des NOC (Jeyakumar *et al.*, 2017). Le groupe de Bingham a également étudié le rôle de l'hème et sa capacité à catalyser la formation des NOC en étudiant l'effet d'une supplémentation en hème sur le taux d'ATNC. Les auteurs ont ainsi pu montrer une augmentation significative des ATNC avec un régime supplémenté en hème par rapport aux groupes contrôles (Cross *et al.*, 2003), (Lunn *et al.*, 2007).

#### B.4.4.3. Carcinogénéité des NOC

Comme la plupart des composés carcinogènes, les NOC, notamment les nitrosamines, subissent une activation par le cytochrome P450 2E1 qui est exprimé dans le tractus gastro-intestinal (Turesky, 2018). Une fois activés ces NOC peuvent s'adduire à l'ADN, en particulier sur les résidus déoxyguanosine, pour former divers adduits, les plus connus étant la N7-méthyl-2'-déoxyguanosine (adduit de la N-nitrosodiméthylamine) et la O<sup>6</sup>-méthyl-2'-déoxyguanosine (O<sup>6</sup>-MeG). Ces deux adduits induisent des modifications de l'ADN et en particulier des transitions G → A et des transversions G → T qui contribuent à la formation d'un environnement mutagène propice au développement du CRC (Turesky, 2018). Ces adduits caractéristiques de l'exposition aux NOC ont été utilisés dans de nombreuses études évaluant l'impact de la consommation de viande rouge et de charcuterie sur la formation des NOC et démontrant l'augmentation des adduits à l'ADN dans le tractus gastro-intestinal après une consommation de viande (Hemeryck *et al.*, 2018; Leu *et al.*, 2015; Lewin *et al.*, 2006). Des études épidémiologiques récentes ont montré une association positive entre les taux de nitrate et

nitrite consommés et la fréquence d'apparition des CCR dans des conditions particulières. Dans une première étude l'association positive entre nitrates / nitrites et CRC n'a été observée que lors d'un apport en nitrate supérieur, et en vitamine C inférieur aux normes recommandées (DellaValle *et al.*, 2014), quand dans la seconde, seule une association positive avec le cancer du rectum a pu être mise en évidence suite à une exposition aux nitrates d'origine animale (issus de la consommation de viande rouge, viande blanche, charcuteries, produits laitiers...) (Espejo-Herrera *et al.*, 2016). Des expérimentations animales ont également donné des résultats contrastés. En effet, en 2016, Zhou *et al.* ne montraient pas d'augmentation significative du nombre de lésions précancéreuses (ACF) en comparant deux lots de souris traitées à l'AOM et nourries ou non avec des hot-dogs (Zhou *et al.*, 2016). En revanche Davis *et al.* ont montré en 2012 une augmentation significative de nombre de cryptes aberrantes après avoir administré des NOC extraits de hot-dogs à des souris induites à l'AOM (Davis *et al.*, 2012). Enfin une étude *in vivo* sur des rats visant à comparer l'impact de plusieurs types de charcuteries a montré une augmentation des MDF seulement pour les produits à haute teneur en nitrites (Santarelli *et al.*, 2010).

## **B.5. Le fer héminique**

### **B.5.1. Généralités-Introduction**

La viande rouge, par opposition à la viande blanche, contient deux protéines en grande quantité et responsables de sa couleur rose/rouge : la myoglobine (majoritairement) et l'hémoglobine qui sont respectivement impliquées dans le stockage et le transport de l'oxygène. En leur sein, ces protéines contiennent un groupement prosthétique (molécule non-protéique contenue dans une structure protéique) : **l'hème**. L'atome de fer contenu dans l'hème fixe l'oxygène et permet ainsi son stockage et son transport *via* la myoglobine et l'hémoglobine. L'hémoglobine est une métalloprotéine tétramérique de 64 kDa, chaque sous-unité contenant un noyau hème, tandis que la myoglobine est une protéine monomérique de 17 kDa contenant une seule molécule d'hème. Outre son activité de transport d'oxygène dans ces protéines, l'hème est également un élément essentiel pour de nombreux processus tels que le transfert d'électrons et la génération d'énergie dans les cytochromes, la dégradation de H<sub>2</sub>O<sub>2</sub> dans les catalases... (Kumar and Bandyopadhyay, 2005). Ainsi, comme pour de nombreux éléments essentiels, une dérégulation entraînant une augmentation ou une diminution anormale de la quantité d'hème biodisponible est associée à de multiples processus pathologiques.

### **B.5.2. Hème dans l'alimentation : métabolisme et absorption**

Le fer est un élément essentiel pour l'organisme de par son implication quasi-ubiquitaire dans les processus biologiques. Dans l'alimentation il est présent sous deux formes, le fer héminique (fer

organique) qui provient essentiellement de produits animaux comme la viande rouge ou encore certains fruits de mer, et le fer non-héminique (fer inorganique) présent en quantité importante notamment dans les légumineuses. Les besoins alimentaires en fer se situent aux environs d'1 mg par jour, majoritairement pour pallier la perte due à l'exfoliation des cellules épithéliales du tractus intestinal (Carpenter and Mahoney, 1992). Bien qu'il ne représente qu'un tiers du fer ingéré, le fer héminique peut constituer jusqu'à deux tiers du fer absorbé (West and Oates, 2008), ce qui peut expliquer une plus forte prévalence de l'anémie chez les végétariens (Pawlak *et al.*, 2016).

Avant d'être absorbé par les entérocytes, l'hème doit être libéré des protéines auxquelles il est complexé. Cette étape fait intervenir les enzymes protéolytiques du système digestif et en particulier la trypsine (Carpenter and Mahoney, 1992). Une fois sous forme libre, l'hème peut alors être absorbé au niveau de l'intestin grêle *via* deux mécanismes : *via* son transporteur HCP1 (*heme carrier protein 1*) exprimé au niveau du pôle apical des entérocytes ou par endocytose à récepteur suite à la fixation de l'hème sur son récepteur (Fuqua *et al.*, 2012). Une fois internalisé, l'hème est ensuite métabolisé et dégradé par les enzymes hème-oxygénase 1 et 2 (HO-1, HO-2). L'enzyme HO-1, exprimée à la surface du réticulum endoplasmique, est une enzyme inductible par de nombreux facteurs tels que les espèces réactives de l'oxygène (ROS), l'hème lui-même, les radiations UV, les cytokines pro-inflammatoires... (West and Oates, 2008). A l'inverse, l'enzyme HO-2, présente dans les vésicules, est une enzyme exprimée constitutivement et également impliquée dans le catabolisme de l'hème. Cette réaction de catabolisme produit du monoxyde de carbone (CO), du  $Fe^{2+}$  et de la biliverdine convertie ensuite en bilirubine par la biliverdine réductase. Le fer ainsi libéré participe au pool de fer intracellulaire, à l'instar du fer inorganique, et peut-être pris en charge par la ferroportine puis excrété dans la circulation comme le montre la Figure 7. Le fer peut également être stocké au sein de l'entérocyte *via* sa conjugaison à la ferritine, une protéine impliquée dans le stockage du fer (Fuqua *et al.*, 2012). Bien que n'ayant jamais été démontrée, l'hypothèse d'une réutilisation directe de l'hème et son relargage dans la circulation sanguine *via* le transporteur FLVCR1 est présente dans la littérature (Hoda *et al.*, 2014).

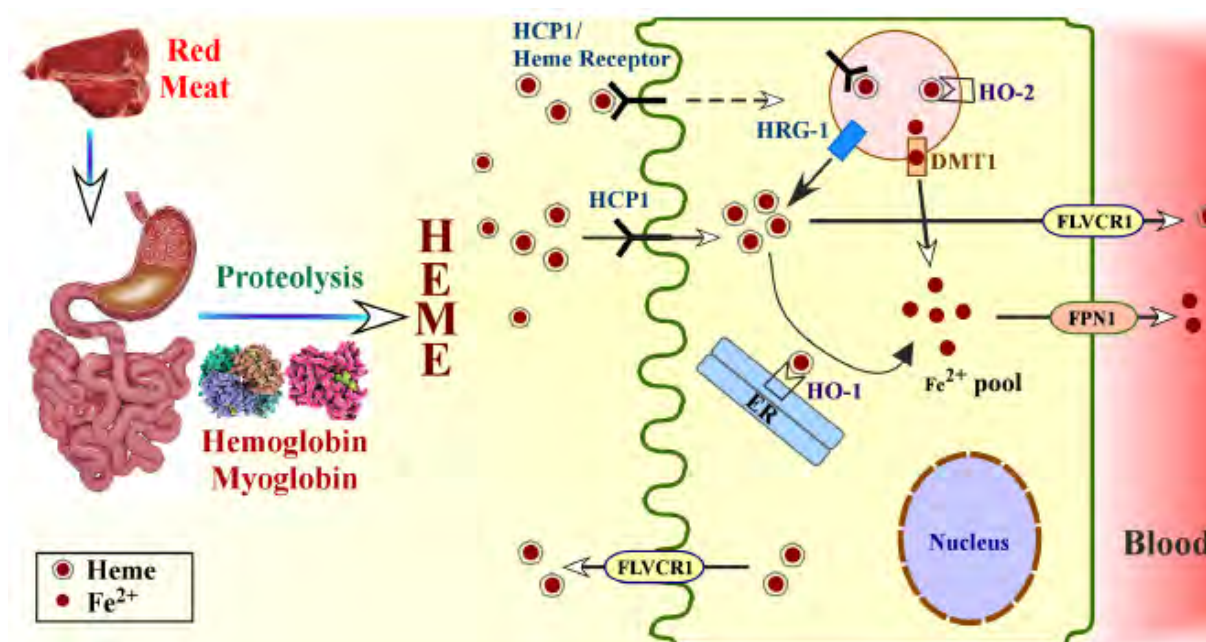


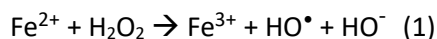
Figure 7 Absorption et métabolisme de l'hème dans le tractus intestinal (Hooda *et al.*, 2014)

La meilleure absorption du fer héminique par rapport au fer inorganique s'explique notamment par le fait qu'au pH de l'intestin grêle (pH neutre) le fer inorganique se retrouve sous forme oxydée (Fe<sup>3+</sup>) et précipite, tandis que le fer héminique est soluble et directement absorbable (West and Oates, 2008). Tout comme le fer inorganique, l'absorption du fer héminique peut être modulée par d'autres composés présents dans le bol alimentaire. En effet, des travaux sur des modèles *in vitro* de cellules épithéliales coliques humaines Caco2 ont montré une diminution de l'absorption du fer héminique en présence de polyphénols (Ma *et al.*, 2011). D'autres études chez l'Homme ont montré que le calcium, l'élément le plus connu pour limiter l'absorption du fer, inhibait l'absorption du fer de manière dose-dépendante (Gaitán *et al.*, 2011; Hallberg *et al.*, 1991). En revanche, d'autres composés tels que l'ascorbate ou encore des peptides issus de la dégradation des protéines peuvent augmenter la biodisponibilité de l'hème (Carpenter and Mahoney, 1992; Ma *et al.*, 2011). Cependant, d'une manière générale, il est admis que l'absorption de l'hème est moins modulable par les autres composés du régime alimentaire que le fer inorganique (Lynch *et al.*, 1985; West and Oates, 2008). En raison de la stabilité relative de l'absorption du fer héminique, plusieurs études ont déterminé des taux d'absorption moyen du fer héminique. D'après les données disponibles dans la littérature, 10 à 20% de l'hème ingéré serait absorbé par les entérocytes (Cediél *et al.*, 2012; Pizarro *et al.*, 2003; Raffin *et al.*, 1974). Ces données suggèrent que 80 à 90% de l'hème est retrouvé au niveau de la lumière intestinale, puis colique ce qui, compte-tenu de la capacité de l'hème à catalyser la formation de radicaux libres, constitue l'hypothèse principale de la carcinogénicité de l'hème.

### B.5.3. Hème catalyseur de la peroxydation lipidique.

Mécanismes de peroxydation par l'hème :

Bien qu'essentiel, le fer (héminique ou non), est connu pour ses propriétés pro-oxydantes et notamment sa capacité à catalyser la formation de radicaux libres *via* la réaction de Fenton. De par sa capacité à alterner entre plusieurs états d'oxydation, il est impliqué dans de nombreuses réactions d'oxydation faisant intervenir le transfert d'électrons comme le montre par exemple la réaction suivante :



Le radical hydroxyle  $\text{HO}^\bullet$  formé par cette réaction est hautement réactif et à l'origine de nombreux phénomènes d'oxydation. Sous forme oxydée ( $\text{Fe}^{3+}$ ) le fer peut être à nouveau réduit, en général par l'ascorbate, sous forme  $\text{Fe}^{2+}$  et ainsi intervenir à nouveau dans la réaction de Fenton (Carlsen *et al.*, 2005). Le mécanisme général de cette réaction est proposé en Figure 8A Le fer héminique peut également suivre un mécanisme Fenton-like comme proposé sur la Figure 8B Cependant, compte-tenu des vitesses de réaction, il apparait que le recyclage du  $\text{Fe}^{3+}$  en  $\text{Fe}^{2+}$  semble trop lent pour être le seul mécanisme de formation de radicaux libres par le fer héminique.

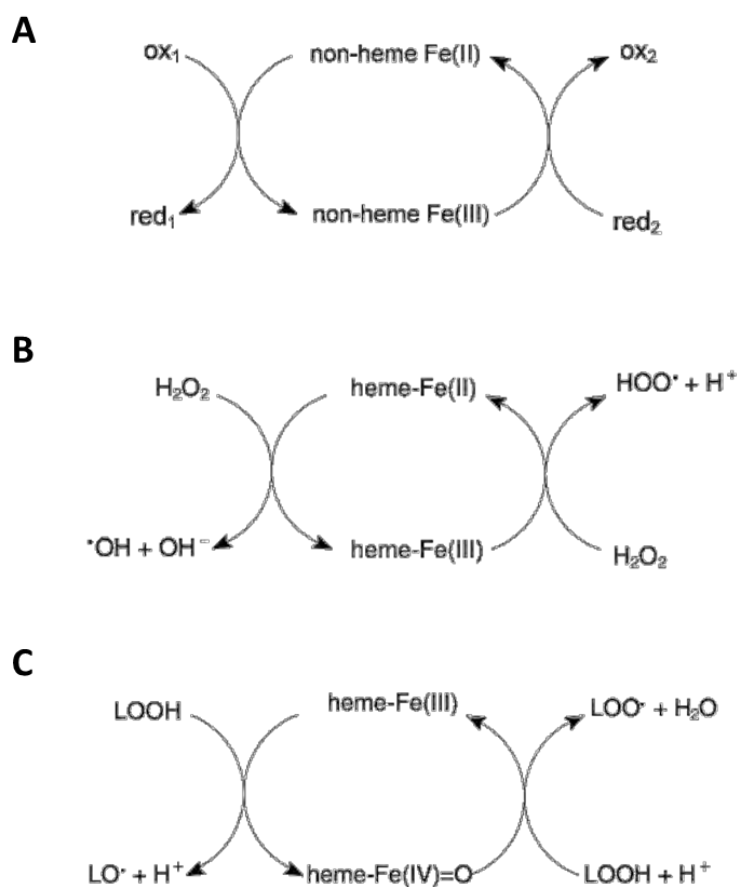


Figure 8 Mécanisme général de la réaction de Fenton (A) Mécanisme proposé de la réaction Fenton-like par le fer héminique (B) Mécanisme proposé de l'oxydation par le fer héminique hypervalent. D'après Carlsen *et al.*

Lors de la prise alimentaire, le fer, les lipides mais aussi d'autres composés interagissent étroitement dans le tractus digestif, favorisant ainsi l'oxydation des lipides (Lorrain *et al.*, 2012).

La peroxydation lipidique est un phénomène séquentiel en 3 étapes initié par l'abstraction d'un atome d'hydrogène, par un radical libre, sur un carbone situé entre deux doubles liaisons d'un lipide : cette étape correspond à l'**initiation** comme le montre la Figure 9 (Cillard and Cillard, 2006). Ainsi les radicaux libres, en particulier le radical hydroxyle HO<sup>\*</sup>, produit par la réaction de Fenton, mais aussi d'autres espèces radicalaires, vont pouvoir initier la séquence de peroxydation lipidique par la formation d'un radical alkyle instable. Ce radical va réagir rapidement avec l'oxygène et conduire à la formation d'un radical peroxyde (LOO<sup>\*</sup>) qui va alors se stabiliser par l'abstraction d'un atome d'hydrogène sur un lipide voisin conduisant ainsi à la formation d'un hydroperoxyde (LOOH) et d'un nouveau radical alkyle : c'est la **propagation**. La réaction en chaîne se propage jusqu'à atteindre l'étape de **terminaison**, où les nombreux radicaux peroxydes formés vont se dimériser pour former des espèces stables non radicalaires.

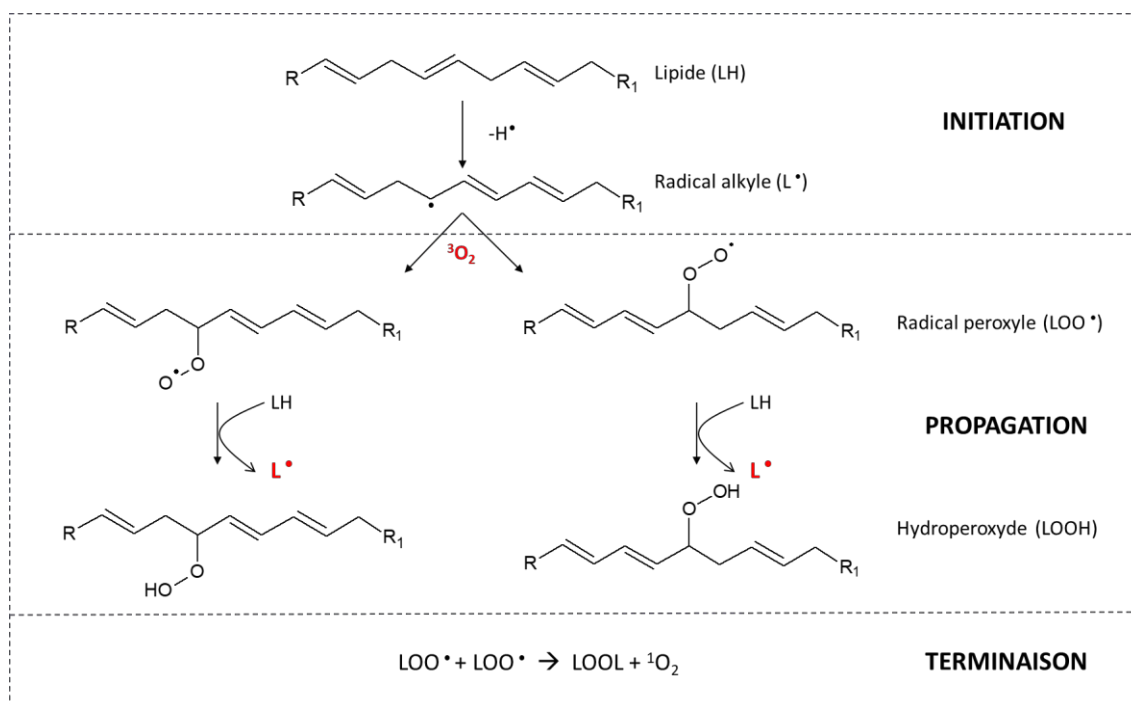


Figure 9 Mécanisme de la peroxydation lipidique

Comme proposé sur la Figure 8 C et également par Lorrain *et al.* (Lorrain *et al.*, 2012) le fer hémunique peut atteindre un état d'oxydation supérieur en réagissant directement avec un hydroperoxyde pour former une espèce radicalaire LO<sup>\*</sup> qui peut alors arracher un atome d'hydrogène d'un autre hydroperoxyde pour former un radical peroxyde, participant ainsi à la **propagation** de la peroxydation lipidique. En raison d'une haute teneur en oxygène, en peroxyde d'hydrogène et d'un faible pH, l'estomac est un milieu particulièrement favorable à la peroxydation des lipides alimentaires tandis que l'augmentation du pH dans l'intestin ralentirait la genèse de ces radicaux libres (Bechaux *et al.*,



2018; Lorrain *et al.*, 2012).

Les hydroperoxydes formés lors de la peroxydation lipidique sont des espèces instables ayant tendance à se décomposer pour donner naissance à une grande variété de composés en fonction de la position du groupement –OOH et des doubles liaisons sur la chaîne lipidique. Parmi ces composés de nombreux aldéhydes sont formés, les plus décrits dans la littérature étant le **4-hydroxynonanal (4-HNE)**, dérivant des acides gras en  $\omega$ -6 et le **malondialdéhyde (MDA)**, dérivant d'acides gras présentant plus de deux doubles liaisons (Esterbauer and Zollern, 1989; Schneider *et al.*, 2001). Ces composés issus de la peroxydation lipidique, en particulier les hydroxy-alcénals tels que le 4-HNE, sont hautement réactifs et sont considérés comme des seconds messagers toxiques (Catalá, 2009). En raison de leur relative stabilité par rapport aux radicaux libres, les hydroxy-alcénals peuvent exercer des effets en dehors de leur lieu de formation et ainsi s'adduire aux protéines, en particulier aux résidus cystéine, histidine et lysine, ou encore former des adduits au niveau des bases de l'ADN (Guéraud, 2017), et présentent donc des propriétés cytotoxiques et génotoxiques. Cette variété des effets induits par ces produits de peroxydation lipidique ont conduit certaines équipes, en particulier l'équipe PPCA (Prévention et Promotion de la Carcinogénèse par l'Alimentation) de l'Unité Toxalim au sein d'INRAE Toulouse, à les placer au centre de l'hypothèse de la carcinogénèse induite par l'hème (Bastide *et al.*, 2015).

## B.6. Carcinogénicité de l'hème

### **B.6.1. Associations épidémiologiques**

Plusieurs études épidémiologiques ont étudié les associations entre consommation d'hème et CCR, dont notamment deux études menées par l'équipe PPCA, au sein de laquelle les travaux présentés dans cette thèse ont été effectués. Dans la première, il a été déterminé que les sujets du quintile supérieur de consommation d'hème présentaient un risque accru de développement du CCR de 18% en comparaison aux personnes du quintile de plus faible consommation (Bastide *et al.*, 2011). Dans leur seconde étude basée sur les participantes de la cohorte française E3N, une association entre la quantité d'hème ingérée et le développement du CRC a été reportée (Bastide *et al.*, 2016). Des résultats basés sur une cohorte britannique ont également montré une association entre la consommation d'hème et le CCR (Ferrucci *et al.*, 2012). Plus récemment, Luo *et al.* ont montré une association positive entre le CCR et la consommation d'hème et une association négative entre le fer provenant des plantes et le CCR (Luo *et al.*, 2019). Cependant, il est à noter que toutes les études épidémiologiques sur le lien entre cancer colorectal et hème ne sont pas parvenues aux mêmes associations. Kabat *et al.* n'ont pas observé d'association dans une étude sur 50 000 femmes canadiennes tandis que Balder *et al.* ont observé une association seulement lorsque le ratio hème /

chlorophylle augmentait (Balder *et al.*, 2006; Kabat *et al.*, 2007).

Malgré une certaine disparité des conclusions provenant des études épidémiologiques, le fer héminique apparaît comme l'un des éléments clé de la promotion de la CCR par la viande rouge. Ainsi, de nombreux travaux ont été menés pour lier ces associations épidémiologiques à des données expérimentales.

### ***B.6.2. Carcinogénicité induite par les produits de peroxydation lipidique***

Comme détaillé précédemment, le fer héminique possède une capacité de catalyse de la peroxydation lipidique *via* la formation de radicaux libres ou en réagissant directement avec des hydroperoxydes. Compte-tenu de la réactivité de certains produits finaux de cette séquence réactionnelle tels que des alcénals, de nombreuses études se sont intéressées aux liens entre hème, peroxydation lipidique et promotion de la CCR. Dès 1998, il a été montré que des rats nourris avec un régime riche en hème (3% du régime sous forme d'hémoglobine) et en acides gras poly-insaturés (PUFA), développaient plus de tumeurs coliques en comparaison à un régime contrôle sans hème (Sawa *et al.*, 1998). Dans cette étude, les auteurs lient cette augmentation à la capacité de l'hème à former des radicaux LOO<sup>\*</sup>. Deux études *in-vivo* menées au sein de l'équipe PPCA sur des rats F344 ont également apporté des conclusions similaires. En effet, dans une première étude une augmentation du nombre d'ACF avec des régimes riches en hémine ou hémoglobine en comparaison à des régimes contrôles a été observée. En lien avec cette augmentation, la cytotoxicité des eaux fécales ainsi que la concentration en substances réactives à l'acide thiobarbiturique (TBARS) dans ces dernières étaient augmentées dans les régimes riches en hème, renforçant ainsi l'hypothèse d'un effet médié par la peroxydation lipidique (Pierre *et al.*, 2003). Dans la seconde étude les auteurs ont obtenu des résultats similaires montrant également un effet dose-réponse entre la quantité d'hème et le nombre d'ACF et de MDF (Pierre *et al.*, 2004a). Une étude plus récente de l'équipe a permis de renforcer l'hypothèse du fer héminique comme élément principal de la relation viande rouge – cancer. En effet, en 2015 Bastide *et al.* ont mené une étude sur des rats nourris avec de l'hémoglobine (1% du régime), des amines hétérocycliques et des nitrates seuls ou en mélange. Seuls les groupes nourris avec de l'hème présentaient une augmentation du nombre de lésions prénéoplasiques indépendamment de l'ajout d'autres facteurs. Comme dans les études précédentes, cette augmentation était associée à une augmentation des TBARS, de la cytotoxicité et de l'excrétion du di-hydroxynonane mercapturic acid (DHN-MA, métabolite urinaire du 4-HNE). Une absence de cytotoxicité et de réponse positive au test TBARS lors du « piégeage » des aldéhydes dans les eaux fécales à l'aide d'une résine spécifique a également été démontrée. Enfin, à l'aide d'une lignée cellulaire murine mutée ou non pour le gène APC, un différentiel de cytotoxicité lors de l'exposition des cellules aux eaux fécales a été mis en évidence, posant ainsi l'hypothèse d'une sélection darwinienne des cellules mutées en faveur des

cellules normales (Bastide *et al.*, 2015). Cette hypothèse a notamment été appuyée par d'autres travaux de l'équipe PPCA (Baradat *et al.*, 2011; Surya *et al.*, 2016). Cependant, il est intéressant de noter que la plupart des travaux cités précédemment mettent en avant un lien entre hème et lésions préneoplasiques, qui n'a pas toujours été confirmé lorsque l'expérimentation a été menée jusqu'au stade tumoral. En effet, si Constante *et al.* ont reporté une augmentation du nombre de tumeurs chez des souris nourries avec des régimes riches en hème, une autre étude n'a reporté aucune augmentation du développement tumoral chez des rats nourris avec des régimes riches en hème (Constante *et al.*, 2017a; de Moura *et al.*, 2019a).

L'ensemble de ces travaux mettent en exergue le possible rôle de la peroxydation lipidique dans la promotion de la carcinogénèse. De nombreux travaux ont également été menés visant à caractériser l'effet direct de ces produits de peroxydation. Notamment vers la fin des années 80 le groupe de H. Esterbauer a démontré le caractère génotoxique du 4-HNE dans plusieurs études *in vitro*. Une activité génotoxique, *via* une induction de la fragmentation d'ADN, de plusieurs 4-hydroxy-alcénals dont le 4-HNE a été reportée (Brambilla *et al.*, 1986). Dans une autre étude portant sur l'effet de différents alcénals, il a été démontré la capacité de ces composés à induire des échanges de chromatides-sœurs à de faibles concentration (0.1  $\mu\text{M}$ ), ainsi que des aberrations chromosomiques à de plus hautes concentration (10  $\mu\text{M}$ ) (Eckl *et al.*, 1993).

### ***B.6.3. Autres effets***

Outre les effets médiés par les produits de peroxydation lipidique, l'hème exerce également de nombreux autres effets délétères en lien plus ou moins direct avec l'action pro-oxydante de ce composé. L'équipe de R. Van der Meer s'est particulièrement intéressée à l'induction d'une hyperprolifération des cellules épithéliales coliques sous l'action du fer héminique. Dans un modèle *in vivo*, elle a notamment démontré que la cytotoxicité induite par l'hème entraînait une hyperprolifération des cellules épithéliales coliques (Sesink *et al.*, 1999). Des travaux plus récents ont permis de poser l'hypothèse que cette hyperprolifération serait due à une perturbation des signaux de régulation de la prolifération transmis des cellules du haut de la crypte aux cellules du fond de la crypte (Ijssennagger *et al.*, 2012b).

Comme introduit précédemment, l'hème catalyse également la formation de NOC, qui peuvent donner naissance à des adduits à l'ADN particulièrement mutagènes tels que l'O<sup>6</sup>-MeG. (Seiwert *et al.*, 2020). Une augmentation dose-dépendante des ATNC dans les fèces de volontaires sains nourris avec de la viande rouge a été observée (Hughes *et al.*, 2001). Plus récemment, des taux plus importants d'O<sup>6</sup>-MeG ont été retrouvés dans des colonocytes de souris exposées à des régimes riches en hème (Winter *et al.*, 2011; J. Winter *et al.*, 2014).

Enfin, depuis quelques années, des travaux mettent en avant le lien entre la prolifération épithéliale

hème-induite et une perturbation du microbiote intestinal également causée par ce composé. En effet, l'administration d'antibiotiques à large spectre à des souris nourries avec de l'hème a montré une réduction drastique de l'hyperprolifération des cellules épithéliales coliques en comparaison avec des souris non traitées (Ijssennagger *et al.*, 2015). Dans cette étude, il a également été montré une réduction de la peroxydation lipidique colique, *via* le dosage des TBARS, chez les souris traitées aux antibiotiques. De la même manière, dans une autre étude, un changement profond du microbiote fortement corrélé à une augmentation de la peroxydation lipidique hème induite a été observé *in vivo* chez le rat (Martin *et al.*, 2019a). L'altération du microbiote observée dans cette étude, à savoir une diminution des Lactobacilles et une augmentation des Entérobactéries, a également été retrouvée dans d'autres expérimentations animales (Ijssennagger *et al.*, 2012a; Schepens *et al.*, 2011). Des perturbations similaires, ainsi qu'une chute de l' $\alpha$ -diversité ont également été observées chez des souris supplémentées en hème (Constante *et al.*, 2017a). Il est intéressant de noter que dans cette étude, les souris soumises à un régime contrôle + sodium dextran sulfate (DSS), un agent induisant une inflammation colique, présentaient les mêmes modifications que les souris nourries avec de l'hème sans DSS.

En somme, bien que les mécanismes sous-jacents restent flous, l'hème semble provoquer une altération de la composition du microbiote intestinal en défaveur de bactéries « bénéfiques » telles que les Lactobacilles. Des dysbioses similaires ont été observées lors d'analyses d'ARN 16S à partir de fèces de patients atteints de cancer colorectaux (Gagnière *et al.*, 2016).

#### ***B.6.4. Prévention des effets délétères de l'hème***

Compte-tenu des nombreux effets toxiques de l'hème, des stratégies de prévention ont émergé pour limiter ces phénomènes. En particulier, de nombreuses stratégies visant notamment à réduire la capacité de l'hème à former des radicaux libres, *via* l'utilisation de molécules aux propriétés anti oxydantes ou encore de calcium, ont été proposées dans la littérature.

L'utilisation de calcium pour prévenir la cytotoxicité, la peroxydation lipidique et l'hyperprolifération hème-induite a largement été étudiée. Des données du groupe de Van der Meer suggèrent que la capacité du calcium à réduire les effets de l'hème réside dans la formation d'un précipité diminuant sa solubilité et sa biodisponibilité (Sesink *et al.*, 2001a). En effet, 53% de l'hème ingéré était retrouvé dans les fèces des rats nourris avec un régime pauvre en calcium contre 83% avec un régime riche en calcium. Dans une expérimentation sur des rats soumis à un régime riche en hème, la capacité du calcium, mais aussi d'autres composés comme des antioxydants, à inhiber la carcinogénèse hème-induite ainsi que la formation de TBARS a été démontrée (Pierre *et al.*, 2003). Cependant, dans une expérience similaire en remplaçant l'hémine par de la viande rouge seul le calcium a permis d'inhiber la carcinogénèse (Pierre *et al.*, 2008a). A l'instar du calcium, d'autres molécules, telles que la

chlorophylle, semblent agir directement sur l'hème en réduisant sa solubilité (de Vogel et al., 2005a). Hormis la réduction des effets de l'hème par sa chélation, d'autres stratégies visant à inhiber ses effets au niveau de la peroxydation lipidique et de la genèse de radicaux libres ont été proposées. Il a notamment été montré que des composés antioxydants comme l' $\alpha$ -tocophérol (vitamine E) et la quercétine (polyphénol) réduisaient la peroxydation de l'acide linoléique induite par l'hème, avec un effet synergique lorsque les deux étaient utilisés en parallèle (Goupy *et al.*, 2007). L'utilisation d' $\alpha$ -tocophérol ou d'extraits de polyphénols dans un modèle animal a montré une inhibition de la promotion de la carcinogénèse hème-induite (Bastide et al., 2017a). L'action conjointe des vitamines C (ascorbate) et E (tocophérol) semble également inhiber la formation de composés *N*-nitrosés par l'hème (Sasso and Latella, 2018). Enfin, d'autres composés alimentaires tels que les fibres, connus pour leur rôle protecteur dans le développement du CCR, pourraient réduire l'effet toxique de l'hème en accélérant le transit intestinal et donc limiter le temps d'exposition à l'hème (Sasso and Latella, 2018). Certains composés comme le cinnamaldéhyde (composé contenu entre autres dans la cannelle) ou encore les polyphénols contenus dans l'huile d'olive peuvent également agir en inducteur de défense antioxydante *via* l'activation du gène Nrf2 qui contrôle l'expression de HO-1, l'enzyme impliquée dans le catabolisme de l'hème (Serreli and Deiana, 2020; Wondrak *et al.*, 2010).

## **C. L'analyse métabolomique pour la toxicologie**

### **C.1. Préambule sur les approches omiques**

Jusqu'au milieu du XX<sup>ème</sup> siècle, la plupart des approches visant à comprendre les systèmes biologiques étaient basées sur des approches réductionnistes (Edwards, 2017). Le réductionnisme est une approche visant à expliquer un phénomène ou un système en étudiant ses constituants de manière indépendante. Par la suite, une vision plus holistique a été proposée afin de considérer les systèmes biologiques non pas comme un ensemble d'éléments indépendants mais bien comme un ensemble d'éléments interconnectés, ce qui a donné naissance à la biologie des systèmes. Bien qu'à ce jour il n'existe pas de consensus clair quant à la définition de la biologie des systèmes, cette dernière pourrait être définie comme « l'étude des interactions entre les constituants d'un système biologique » (Carvunis *et al.*, 2009). Ainsi, face à cette volonté de considérer les systèmes dans leur ensemble, les méthodes d'analyses ont dû évoluer en conséquence afin de pouvoir observer et mesurer ces derniers dans leur entièreté. Grâce aux progrès technologiques réalisés en biotechnologie, chimie analytique ou encore bio-informatique et afin de répondre à ce besoin, les approches dites « omiques » se sont développées dans le but de fournir une vue d'ensemble des organismes biologiques à plusieurs niveaux (du génome jusqu'au métabolome). L'objectif final de la biologie des systèmes étant de

considérer l'interconnexion entre les gènes, les protéines et les métabolites afin de mieux comprendre le fonctionnement de ces systèmes.

### C.2. Introduction générale et définitions

La métabolomique constitue le dernier élément de la séquence largement décrite des approches appelées « omiques », génomique → transcriptomique → protéomique → métabolomique illustrée en Figure 10. La première description de la métabolomique date de 1999 par le professeur J. Nicholson qui décrit la métabolomique comme étant « *la mesure quantitative de la réponse métabolique dynamique et multiparamétrique de systèmes vivant à des stimuli pathophysiologiques ou des modifications génétiques* » (Nicholson *et al.*, 1999). Le développement d'une telle approche était motivé par le fait que la génomique et la protéomique ne permettent pas d'obtenir une vue d'ensemble de la réponse métabolique d'un organisme à un stimulus (Nicholson *et al.*, 1999). Ainsi, comme la métabolomique constitue le dernier maillon de la chaîne présentée en Figure 10, il est communément admis que cette approche constitue le **lien le plus direct avec le phénotype** (Fiehn, 2002). En effet la génomique pourrait être décrite comme l'étude de « ce qu'il peut se produire », la transcriptomique de « ce qu'il va probablement se produire », la protéomique de « ce qu'il fait que cela se produit » et enfin la métabolomique de « ce qu'il se produit ».

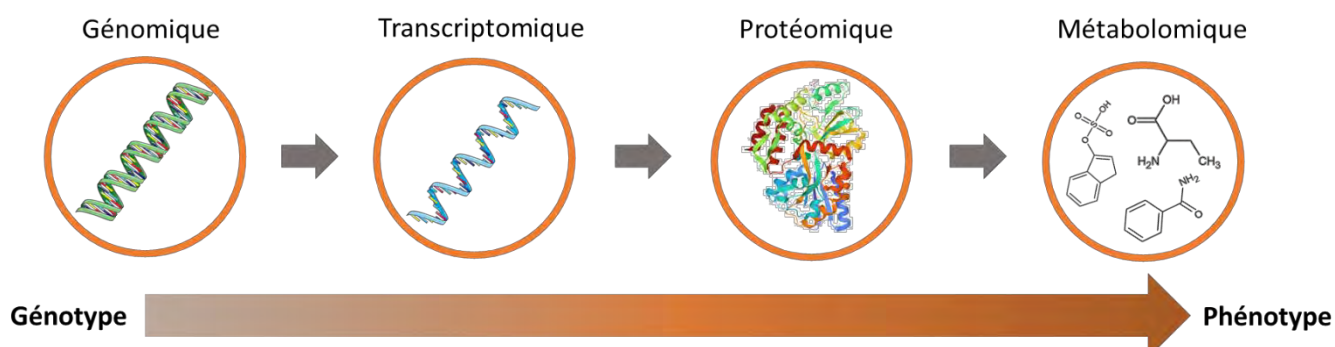


Figure 10 Schéma représentatif de la cascade des approches omiques

Comme les autres sciences omiques, la métabolomique repose sur l'étude d'un ensemble, le **métabolome**, qui correspond à l'ensemble des métabolites. Cet ensemble est constitué d'un mélange de molécules très diversifié tant au niveau de leurs propriétés physico-chimiques que de leurs concentration (Roca *et al.*, 2021; Yanes *et al.*, 2011). En effet, les métabolites couvrent une gamme dynamique allant jusqu'à 11 ordres de grandeur, compris entre la millimole et la picomole (Psychogios *et al.*, 2011; Dunn and Ellis, 2005). Les métabolites sont définis comme étant des composés de faible poids moléculaire (< 1000-1500 Da) qui participent aux réactions du métabolisme en tant que produit, substrat ou encore effecteur. Le métabolome comprend à la fois des métabolites spécifiques de l'organisme considéré qui est divisé en **endo-métabolome**, qui correspond aux métabolites intra-

cellulaires, et en **exo métabolome**, correspondant aux métabolites extracellulaires mais aussi aux métabolites d'autres organismes (par exemple chez l'Homme les métabolites du microbiote). Enfin, le métabolome comprend aussi des composés exogènes issus de l'alimentation, de l'environnement ou autre, constituant le **xeno-métabolome**. Le nombre de métabolites décrits est en constante évolution. En effet, si l'on prend l'exemple de la *Human Metabolome Database* (HMDB) qui est une base de données de métabolites créée par le groupe de D. Wishart en 2009, celle-ci contenait 883 métabolites détectés et quantifiés à sa création (Wishart *et al.*, 2009). Dans la dernière version parue en 2018 (HMDB 4.0), ce nombre s'élève à 18 557 métabolites détectés et plus de 90 000 métabolites prédits (Wishart *et al.*, 2018).

Le domaine de la métabolomique peut-être subdivisé en deux grands axes, à savoir la métabolomique ciblée et la **métabolomique non-ciblée**. L'approche ciblée repose sur la quantification d'un ensemble métabolites connus appartenant à une ou plusieurs voies biochimiques. De manière générale, le choix des métabolites quantifiés est guidé par une hypothèse biologique *a priori* sur les voies biochimiques impactées par le facteur étudié. A l'inverse, les approches non-ciblées tendent à mesurer sans *a priori* le plus grand nombre de métabolites possible dans un échantillon biologique donné et sont plutôt employées dans une optique de génération d'hypothèses (Patti *et al.*, 2012). Les champs d'application de la métabolomique sont vastes, avec des applications allant de la clinique jusqu'à l'agriculture. En effet, l'utilisation de la métabolomique a, par exemple, permis de mettre en évidence l'importance du métabolisme des hydrates de carbone dans la protection des plantes contre la sécheresse (Gechev *et al.*, 2013) ou encore de proposer de nouveaux biomarqueurs pour l'étude d'erreurs innées du métabolisme telles que la phénylcétonurie (Václavík *et al.*, 2018). L'utilisation de la métabolomique s'inscrit également dans des problématiques de médecine de précision avec une volonté de caractériser différents stades de pathologies, comme par exemple les maladies rénales chroniques (Gagnebin *et al.*, 2020). Enfin, il est également important de noter l'apport de la métabolomique dans le domaine de la toxicologie, pour la détermination de nouveaux biomarqueurs d'effets par exemple (van Ravenzwaay *et al.*, 2007; Xia *et al.*, 2020). Dans ce manuscrit, nous traiterons essentiellement de la partie non-ciblée bien que de nombreuses similarités existent entre les deux approches.

### C.3. Les outils analytiques en métabolomique

En raison de la diversité des propriétés physico-chimiques des composés, la gamme dynamique importante des métabolites et la variété des matrices analysées, l'analyse métabolomique constitue un réel challenge analytique. A ce jour, deux grandes techniques analytiques complémentaires sont utilisées pour répondre à ce défi, la résonance magnétique nucléaire (RMN) et la **spectrométrie de masse** (MS). Les avancées technologiques de ces 20 dernières années ont permis d'augmenter le

nombre de métabolites détectés mais aussi la sensibilité de détection de manière considérable. En effet, les spectromètres de masse actuels permettent de détecter des quantités infimes de l'ordre de la femtomole ( $10^{-15}$  mol) (Roca *et al.*, 2021).

La RMN est une technique d'analyse spectroscopique rapide, robuste et répétable basée sur la mesure de l'énergie émise par un noyau atomique lors de sa relaxation après son excitation dans un champ électromagnétique externe. En fonction de l'environnement du noyau et du champ magnétique, la relaxation se produit à des fréquences bien précises qui fournissent des informations sur la structure de la molécule considérée. L'intensité du signal produit fournit quant à elle des informations sur son abondance. Bien qu'elle présente certains avantages, la limitation majeure de l'utilisation de la RMN pour la métabolomique réside dans sa faible sensibilité qui permet seulement la détection des composés les plus abondants ( $> 1 \mu\text{M}$ ) (Emwas, 2015).

Bien que moins reproductible que la RMN, sa **sensibilité supérieure** ainsi que la possibilité de détecter des **milliers de signaux** ont fait de la MS la technique analytique la plus couramment utilisée en métabolomique (Emwas *et al.*, 2019). La Figure 11 témoigne à la fois de l'essor de la métabolomique sur les 20 dernières années, mais aussi de la prépondérance de l'utilisation de la MS dans les laboratoires en comparaison à la RMN, principalement en raison du coût plus faible et des contraintes d'installation moins importantes des équipements de MS.

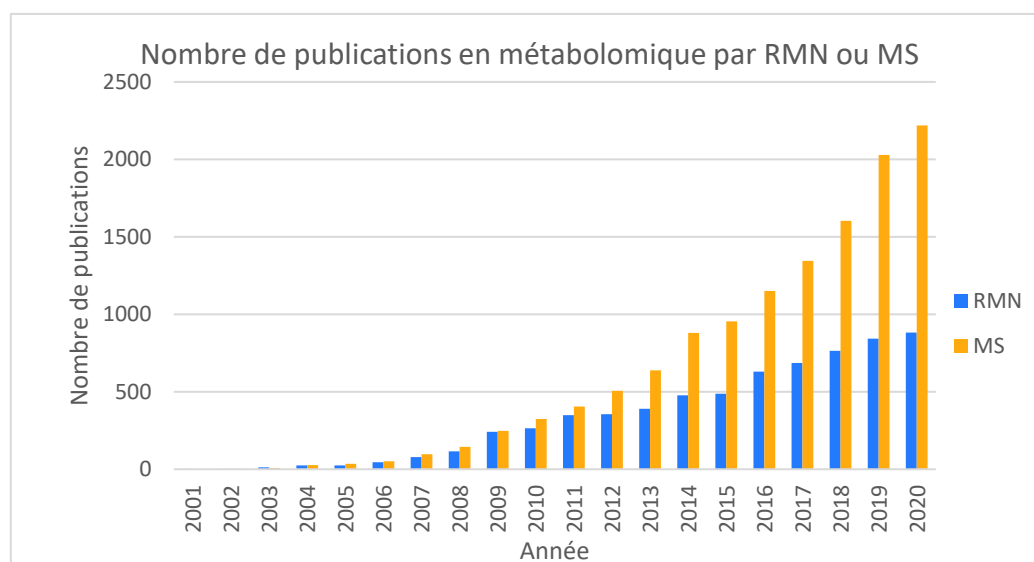


Figure 11 Nombre de publications dans PubMed pour les termes « NMR + Metabolomics » et « MS + metabolomics » entre 2001 et 2020

L'autre avantage majeur qui a contribué à l'envolée de l'utilisation de la MS réside dans sa capacité de couplage avec des méthodes séparatives telles que la chromatographie liquide (LC), la chromatographie gazeuse (GC) ou encore l'électrophorèse capillaire (CE), ce qui renforce les capacités de la MS en ajoutant une dimension analytique supplémentaire. Au cours de cette introduction, seule l'utilisation de la LC-MS pour l'analyse métabolomique sera détaillée. L'utilisation de la RMN ou de la



MS avec d'autres techniques de couplage, telle que la GC, sont détaillées dans de nombreux articles de revue récents (Emwas *et al.*, 2019; Markley *et al.*, 2017; Papadimitropoulos *et al.*, 2018).

#### C.4. La spectrométrie de masse couplée à la chromatographie liquide.

##### **C.4.1. Principe général**

La MS est une technique analytique qui permet d'analyser en phase gazeuse des molécules en **mesurant leur rapport masse sur charge ( $m/z$ )**. Afin d'être détectables, il est donc nécessaire que les molécules soient ionisées et en phase gazeuse. Cette opération intervient dans le premier élément d'un spectromètre de masse, **la source d'ionisation**. Une fois ionisés et en phase gazeuse, les ions formés vont être séparés au niveau d'un ou plusieurs **analyseurs** en fonction de leur rapport  $m/z$  avant d'arriver au niveau du **détecteur** qui va générer un signal électronique à partir des différents ions. Enfin ce signal va être amplifié, puis enregistré et traité par un logiciel dédié qui va permettre d'obtenir le spectre de masse (intensité des ions détectés en fonction du rapport  $m/z$ ) (Figure 12).

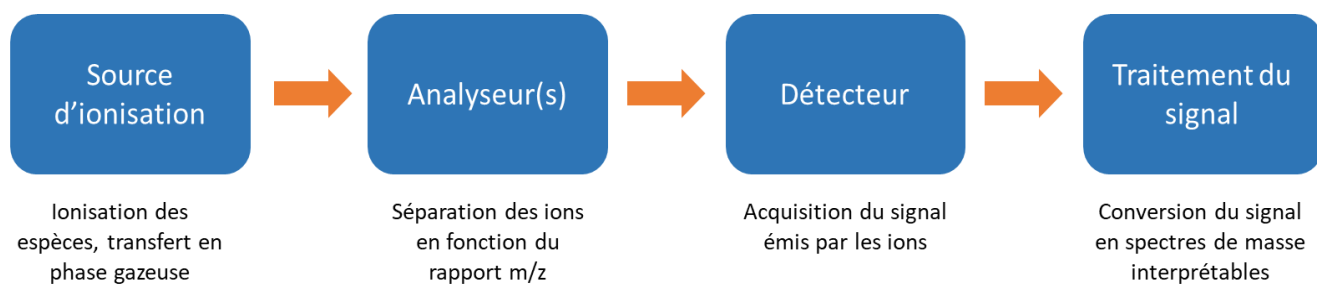


Figure 12 Représentation schématique de la structure d'un spectromètre de masse

Les sources et méthodes d'ionisation sont nombreuses et leur utilisation dépend à la fois du type de matrice, de la nature des composés à étudier et du type de méthode séparative employée. Parmi les principales sources d'ionisation, il est possible de citer l'EI (*electron ionisation*) et la CI (*chemical ionisation*) généralement utilisés en couplage GC-MS, le MALDI (*matrix assisted laser desorption ionisation*) généralement utilisé pour l'analyse de macro-molécules sans technique séparative en amont, et enfin l'ESI (*electrospray ionisation*), l'APCI (*atmospheric pressure chemical ionisation*) et l'APPI (*atmospheric pressure photo-ionisation*), utilisées en couplage LC-MS. Parmi ces techniques, l'ESI constitue la plus répandue dans les laboratoires et est plus particulièrement utilisée en métabolomique, et nous nous limiterons dans ce manuscrit à la description de cette technique. Le principe de l'ESI repose sur l'application d'une différence de potentiel élevée entre l'extrémité d'un capillaire au travers duquel passent les analytes en phase liquide et un système de contre électrodes. Sous l'effet du champ électrique intense créé par cette tension, un spray constitué de gouttelettes chargées est généré. L'application d'un gaz permet la désolvatation et donc la réduction du volume des gouttelettes. Lorsque les forces de répulsion des charges présentes dans la gouttelette deviennent

supérieures aux forces de cohésion des gouttelettes (phénomène appelé limite de Rayleigh), ces dernières explosent et conduisent à la formation de gouttelettes plus petites, la cascade d'explosions conduisant à la libération des composés ionisés dans la phase gazeuse. Les ions sont alors transférés et guidés de la zone d'ionisation à pression atmosphérique jusqu'à l'analyseur qui opère sous vide poussé. De la même manière que pour les sources d'ions, plusieurs types d'analyseurs existent. Les plus fréquemment rencontrés sont les analyseurs quadripolaires (Q), à temps de vol (ToF pour *time of flight*) ou les Orbitrap, chaque type d'analyseur possédant des caractéristiques particulières en termes de gamme de masse, de vitesse d'acquisition ou de pouvoir de résolution (capacité à distinguer deux rapports  $m/z$  proches). En fonction du type d'appareil, il est possible de retrouver des combinaisons d'analyseurs comme par exemple Q-q-ToF, Q-q-Q, Q-q-Orbitrap... Enfin, en fonction de l'analyseur plusieurs types de détecteurs peuvent être rencontrés, basés par exemple sur la multiplication d'électrons, qui sont les plus courants, ou encore sur la détection de courants induits (instruments à transformée de Fourier de type Orbitrap).

L'un des avantages de la spectrométrie de masse pour l'analyse métabolomique repose sur la possibilité de couplage avec une **technique séparative** en ligne telle que la LC. Bien que le temps d'analyse soit fortement rallongé par rapport à une analyse en introduction directe, l'utilisation d'une technique séparative permet de séparer les molécules d'un échantillon donné et ainsi étaler, dans le temps, leur arrivée au niveau de la source d'ionisation. L'ajout d'une dimension séparative avant analyse MS permet de limiter certains phénomènes tels que la suppression d'ionisation (caractérisée par une suppression du signal due à un nombre trop élevé d'espèces présentes simultanément dans la source) fréquemment observable en introduction directe (Chekmeneva *et al.*, 2018; Pöhö *et al.*, 2019).

La chromatographie liquide permet de séparer des molécules en solution en faisant intervenir plusieurs mécanismes de rétention. La chromatographie de partage, qui est le type de rétention le plus fréquemment rencontré, est basée sur une différence d'affinité entre une **phase stationnaire**, contenue dans la colonne chromatographique, et une **phase mobile** entraînant les analytes à travers la colonne jusqu'à la source d'ionisation. Il existe un grand nombre de phases stationnaires, principalement basées sur des particules de silice greffée, donnant accès à des colonnes permettant de générer des interactions hydrophobes (phase inverse, RP), hydrophiles (HILIC pour *hydrophilic interaction liquid chromatography*) ou ioniques (IC pour *ion exchange chromatography*). Chaque type de remplissage fournit des interactions différentes permettant la rétention de composés de caractéristiques physico-chimiques diverses. Plusieurs travaux, notamment en métabolomique ont pu mettre en évidence la spécificité et la complémentarité de différentes phases stationnaires. Une étude de 2012 visant à étudier l'apport de l'utilisation de plusieurs plateformes analytiques pour

l'annotation du métabolome de *Fasciola hepatica* a révélé que seulement 14 métabolites sur 142 étaient communs aux 5 techniques utilisées à savoir les chromatographies liquides en mode HILIC et RP (en ionisation positive et négative) et l'électrophorèse capillaire (Saric *et al.*, 2012). Plus récemment, Boudah *et al.* ont démontré l'intérêt de la combinaison de différents types de colonnes pour l'étude du métabolome du sérum humain avec 128/266 métabolites détectés avec les 3 colonnes testées. Sur ces 128 métabolites communs aux 3 colonnes, seuls 17 présentaient un facteur de rétention acceptable ( $k > 1$ ) sur l'ensemble des colonnes. (Boudah *et al.*, 2014). Ces travaux mettent en exergue à la fois l'intérêt de l'utilisation d'une dimension séparative avant analyse, mais aussi le fait qu'une meilleure couverture du métabolome peut être obtenue par l'utilisation de plusieurs méthodes complémentaires.

#### ***C.4.2. L'utilisation de la fragmentation en spectrométrie de masse***

Nous venons de voir que l'ESI représentait la technique d'ionisation la plus utilisée en métabolomique par LC-MS. Si cette technique d'ionisation dite « douce » permet d'accéder à l'information sur la masse de la molécule (par l'intermédiaire du rapport  $m/z$  de l'ion correspondant mesuré en haute résolution), elle ne génère que rarement des ions fragments. Pour pallier cette limitation, à savoir le manque d'information structurale, il est possible de réaliser des **acquisitions MS/MS** permettant de fragmenter les ions (moléculaires) afin d'obtenir des informations sur la structure de ces derniers. Les spectres MS/MS générés par ces acquisitions permettent par exemple de confirmer l'identité d'un métabolite en comparant son spectre MS/MS à celui d'un standard chimique, de tenter l'élucidation structurale de composés inconnus par l'interprétation des mécanismes de fragmentation ou encore de faciliter l'annotation par l'utilisation de réseaux moléculaires basés sur les similarités de fragmentation.

Plusieurs modes d'acquisition MS/MS sont disponibles en fonction de l'information recherchée et du type d'instrument utilisé : le mode ciblé, le mode donnée-dépendant (DDA pour *data driven analysis*) et le mode donnée-indépendant (DIA pour *data independant analysis*).

Le **mode ciblé** correspond à la fragmentation d'un ou plusieurs composés préalablement sélectionnés. Dans ce mode, il est nécessaire de connaître la masse du composé que l'on veut fragmenter, ce qui implique, dans la majorité des cas, d'avoir au préalable analysé l'échantillon dans lequel se trouve le composé d'intérêt et donc le réinjecter pour acquérir le spectre MS/MS voulu après sélection du précurseur. Au sein des configurations en tandem évoquées précédemment (de type Q-q-ToF, Q-q-Orbitrap...), la fragmentation se déroule en 3 étapes : dans un premier temps, l'ion parent est sélectionné dans le premier quadripôle, fragmenté dans la cellule de collision, puis les ions fils sont séparés dans le dernier analyseur afin de générer un spectre de fragmentation. En fonction de l'objectif, plusieurs modes de fragmentation sont disponibles, le *product ion scan* permet la détection

de l'ensemble des ions fragments produits par un parent donné, le SRM (*single reaction monitoring*) et le MRM (*multiple reaction monitoring*) permettent de suivre les transitions ions parents / ions fragments à des fins quantitatives. A l'inverse, des modes tels que le *parent ion scan* ou encore le *neutral loss scan* permettent de déterminer tous les ions parents ayant produit un ion fils donné ou généré une perte de neutre commune.

Grace aux progrès technologiques réalisés, en particulier sur l'amélioration des vitesses de scan, de nouveaux modes d'acquisition ont vu le jour permettant d'acquérir un grand nombre de spectres MS/MS en une seule analyse.

Le **mode DDA** correspond à l'acquisition successive et alternée de spectres MS et MS/MS en une seule analyse. Dans ce mode un scan MS est effectué puis un certain nombre de scans MS/MS sont réalisés dans un ordre établi, en général en fonction de l'intensité des ions détectés dans le scan MS. La dénomination « top  $n$  » est utilisée pour signifier que les  $n$  espèces les plus intenses sont fragmentées en MS/MS avant de réaliser un nouveau scan MS. Il s'agit donc d'une acquisition MS/MS ciblée dans laquelle le choix de l'ion précurseur n'est pas uniquement fixé par l'opérateur, mais aussi par l'instrument. En fonction des vitesses de scan des instruments et des méthodes chromatographiques utilisées, le nombre d'évènements MS/MS par cycle peut grandement varier (entre 3 et 20 acquisitions de spectres MS/MS par cycle) (Defossez *et al.*, 2021). Le principal avantage de ce mode d'acquisition repose sur le fait qu'il permet l'acquisition d'un **grand nombre de spectres MS/MS ciblés** de composés connus et inconnus en une seule analyse. Par exemple, dans une étude de 2016 combinant l'utilisation du mode DDA et des réseaux moléculaires pour la détection de composés issus de médicaments anti-hypertension, environ 5000 spectres MS/MS étaient acquis à chaque injection (van der Hooft *et al.*, 2016). En revanche, cette approche comporte quelques désavantages, notamment le fait que certaines espèces d'intérêt peuvent ne pas être fragmentés car trop peu intenses ou en dehors de critères de sélection (Barbier Saint Hilaire *et al.*, 2020). De plus, le paramétrage de l'acquisition constitue un élément crucial pour acquérir des spectres MS/MS de qualité. Ainsi, Defossez *et al.* ont récemment proposé un ensemble de directives afin d'assurer la bonne qualité des acquisitions DDA (Defossez *et al.*, 2021).

A l'inverse du mode DDA, le **mode DIA** consiste en la **fragmentation simultanée de tous les ions** sans sélection du précurseur. Ceci permet, en théorie, d'acquérir des données de fragmentation sur tous les composés présents dans l'échantillon, y compris les moins abondants (à condition que le composé se fragmente à l'énergie de collision utilisée). En fonction du type d'instrument, deux types d'acquisition DIA sont disponibles. Le premier, dit AIF (Geiger *et al.*, 2010) (pour *all ion fragmentation*) ou MS<sup>E</sup> (Silva *et al.*, 2005), correspond à l'alternance de deux scans MS/MS, respectivement à basse et à haute énergie de collision. Le scan basse énergie a pour but de *screener* l'ensemble des ions sans

fragmentation (accès aux masses moléculaires), et le scan haute énergie de les fragmenter (accès à l'information structurale). Le désavantage de cette approche est lié à la complexité des spectres générés qui représente un mélange de tous les ions parents et leurs fragments sur le même spectre, ce qui nécessite l'utilisation de logiciels de déconvolution performants pour réaliser l'association parent-fragment.

Afin de générer des spectres DIA moins complexes, une autre approche dite SWATH (Gillet *et al.*, 2012) (pour *Sequential window acquisition of all theoretical mass*) a été développée. Cette fois, la gamme de masse est découpée en fenêtres de masse, dont la taille est définie par l'utilisateur, et seuls les ions de la fenêtre considérée sont fragmentés. Ce mode a l'avantage de simplifier les spectres MS/MS résultants et ainsi de faciliter l'attribution parent/fragment. Le désavantage réside dans le temps nécessaire pour fragmenter l'ensemble de la gamme de masse puisque plusieurs étapes seront nécessaires pour reconstituer des spectres MS/MS complets.

#### **D. Le processus d'analyse métabolomique par LC-MS**

Le déroulement d'une analyse métabolomique se déroule en plusieurs étapes séquentielles communément appelées « *workflow* ». Ce processus, illustré en Figure 13, débute avec la mise en place d'un schéma expérimental adapté à la question biologique posée et se termine par l'identification de métabolites modulés par le facteur étudié, permettant la génération d'hypothèses sur les potentiels mécanismes sous-jacents. Chaque étape constitue un élément clé de l'analyse métabolomique et comporte des points critiques à prendre en considération, mis en avant dans plusieurs publications (Paglia and Astarita, 2017; Zheng *et al.*, 2020) et qui vont influencer la qualité des données produites. Les différentes étapes de ce processus seront détaillées dans cette section.

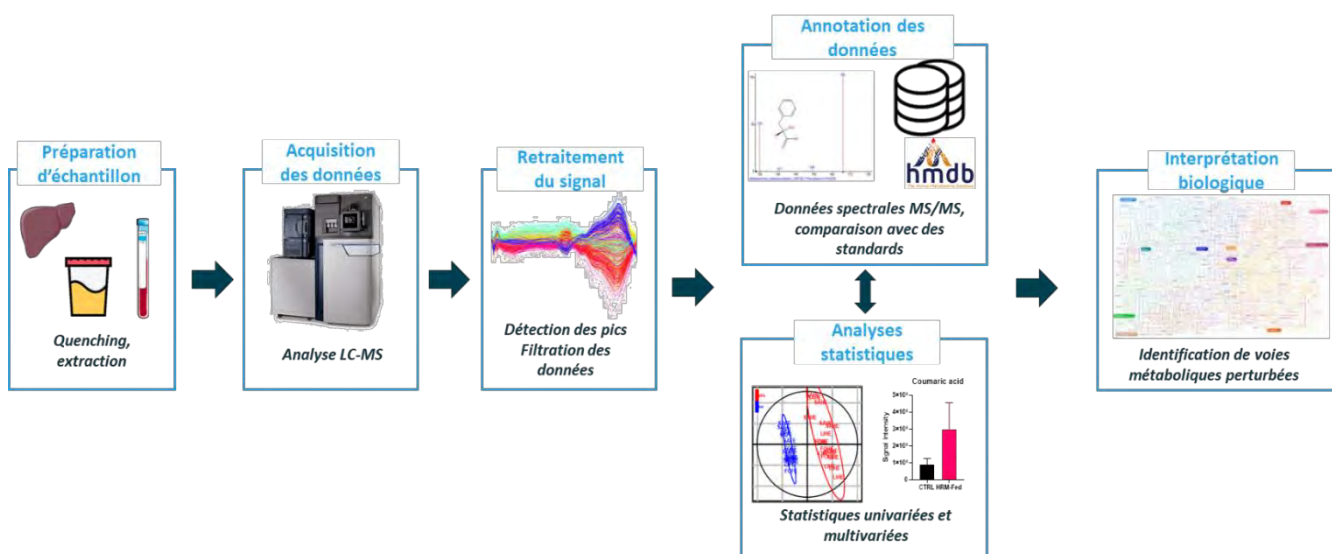


Figure 13 Représentation schématique du déroulement d'une analyse métabolomique non-ciblée par LC-MS

### D.1. Planification du schéma expérimental

L'une des utilisations principales de la métabolomique non-ciblée repose sur la comparaison d'au minimum deux populations dans le but de caractériser les altérations métaboliques induites par le facteur étudié. Il est par exemple possible de comparer des populations « traité VS contrôle », « malade VS sain », mais aussi de comparer des populations différant uniquement par leurs habitudes alimentaires, leur mode de vie, etc.. Ainsi, afin de répondre au mieux à la question biologique, il est crucial de définir les contours de l'expérience au préalable. En effet, outre le choix des échantillons et de leur méthode de prélèvement, de contrôles adéquats, de la méthode analytique adaptée à la question, il est également important de considérer toutes les sources de variabilité potentielles pouvant survenir lors de l'expérimentation (*i.e* variabilité technique, effet série d'analyses, variabilité interindividuelle...) (Martins *et al.*, 2018). Les différentes stratégies et considérations permettant de réduire la variabilité à chaque étape du processus seront détaillées dans les paragraphes suivants dédiés à chacune des étapes.

Afin d'établir des conclusions robustes, le nombre de sujets par groupe est également un élément crucial. De la même manière, il est communément admis que plus la force de l'effet étudié est faible plus le nombre de sujets nécessaires augmente (Dunn *et al.*, 2012). Une bonne caractérisation des échantillons est également recommandée. En effet, la collecte de données sur les échantillons analysés permet d'identifier d'éventuels facteurs confondants qui peuvent être pris en compte lors de la planification de l'analyse ou lors du traitement statistique des données. Il est cependant à noter que l'utilisation de méthodes statistiques visant à corriger les facteurs confondants nécessite un nombre d'individus par groupe plus important (Dunn *et al.*, 2012).

L'étape de planification de l'expérience est donc essentielle dans la mesure où elle détermine l'interprétation et les conclusions amenées par l'étude.

### D.2. Préparation des échantillons

L'objectif de cette étape est, dans le cas d'une analyse non-ciblée, d'extraire de manière reproductible le plus de métabolites possibles d'un échantillon donné. En fonction de la matrice étudiée les méthodes de préparation d'échantillon peuvent varier, mais tendent à se rapprocher de la méthode idéale qui doit être (1) **non-sélective**, (2) **rapide** et comportant un nombre minimal d'étapes, (3) **reproductible** et (4) comprenant si nécessaire une étape de *quenching* destinée à stopper le métabolisme (Fernández-Peralbo and Luque de Castro, 2012). Les modalités de prélèvement et de stockage des échantillons avant extraction constituent également un des paramètres à considérer. En effet, la période (par exemple matin ou soir) de prélèvement peut avoir un impact important sur

l'abondance de certains métabolites et introduire de la variabilité indésirable si cette dernière n'est pas standardisée. Dans une étude visant à caractériser les sources de variation lors d'analyses métabolomiques, Kim *et al.* ont démontré qu'une part importante de la variabilité des empreintes métaboliques dans le cas de l'urine était liée à l'heure du prélèvement (Kim *et al.*, 2014). Il est à noter, dans cette même étude, que le métabolome sanguin était moins impacté par ce facteur. Ce résultat peut notamment s'expliquer par le fait que de nombreux métabolites impactés par le cycle circadien sont excrétés dans l'urine (Giskeødegård *et al.*, 2015). Dans le cas de l'urine, il est possible de limiter la variabilité liée au cycle circadien par le prélèvement de l'urine sur une période de 24 h (Bi *et al.*, 2020). Outre la période de prélèvement, l'étape de *quenching* peut également être une étape cruciale de la préparation d'échantillon en fonction de la matrice analysée. Le but du *quenching* est de stopper le métabolisme le plus rapidement possible après le prélèvement de l'échantillon afin d'éviter la dégradation d'espèces labiles ou des modifications enzymatiques pouvant conduire à une faible corrélation entre le métabolome observé et le métabolome réel. Bien qu'essentiellement appliqué aux études sur cellules, bactéries et plantes (Kapoor and Vaidyanathan, 2018; Smith *et al.*, 2020) son usage pour les biofluides tels que l'urine ou le sérum est débattu. Certains métabolites sensibles à l'oxydation tels que l'ascorbate ou la dopamine ont été retrouvés altérés dans l'urine en l'absence de *quenching* (van de Merbel, 2008). A l'inverse, une étude de Dunn *et al.* n'a pas retrouvé de différences statistiques significatives dans une expérience de métabolomique non-ciblée entre des urines directement placées à -80 °C après prélèvement et des urines stockées à 4 °C pendant 24 h (Dunn *et al.*, 2008).

Dans les biofluides, la méthode la plus utilisée pour l'extraction des métabolites polaires repose sur l'ajout d'un solvant organique type méthanol (MeOH), ou d'un mélange méthanol/éthanol dans le but de faire précipiter les protéines (Sitnikov *et al.*, 2016). La précipitation est en général suivie d'une centrifugation afin de culotter les protéines et de les éliminer. Le but étant de limiter le nombre de signaux provenant des protéines et d'éviter d'éventuels bouchages de colonne LC dus à une précipitation de ces composés au sein du système chromatographique (Rico *et al.*, 2014). Il existe de nombreuses études comparatives visant à déterminer quelle méthode ou combinaison de méthodes permet d'obtenir la meilleure couverture métabolique (Rico *et al.*, 2014; Sitnikov *et al.*, 2016; Yang *et al.*, 2013). Dans leur étude, Yang *et al.* ont comparé l'extraction simple au MeOH avec une extraction au methyl-tertiobutyl-ether (MTBE) suivie d'un fractionnement par extraction sur phase solide (SPE) conduisant à la formation de 5 fractions distinctes (Yang *et al.*, 2013). Bien que la couverture métabolique ait été largement augmentée par la méthode MTBE-SPE (3806 métabolites putatifs VS 1851), cette dernière rallonge considérablement le temps d'analyse. De la même manière, Stikinov *et al.* ont comparé 7 méthodes d'extraction avec des taux de recouvrement entre les méthodes allant de

34 à 80% (Sitnikov *et al.*, 2016). Ce genre d'étude met en exergue l'impossibilité d'obtenir une couverture métabolique parfaite *via* l'utilisation d'une seule méthode d'extraction et que le meilleur compromis doit être adopté en fonction de la question biologique d'intérêt.

Bien que nécessaire pour la majorité des biofluides, l'extraction des métabolites et la précipitation des protéines n'est pas une étape obligatoire lors de l'analyse d'urine en raison d'une teneur en protéine bien plus faible comparativement au plasma ou sérum (Khamis *et al.*, 2017). Ainsi, la stratégie la plus simple consiste à simplement diluer l'urine dans un solvant compatible, et à l'injecter directement (stratégie « *Dilute and Shoot* »). L'étape de dilution reste néanmoins cruciale afin de limiter les effets de suppression d'ionisation ainsi que l'encrassement de l'appareil.

### D.3. Acquisition des données

Comme vu précédemment, l'acquisition des données se fait *via* l'utilisation d'outils analytiques tels que la MS ou la RMN. Outre les considérations analytiques telles que l'évaluation de la stabilité du système, la calibration, la répétabilité des analyses valables pour n'importe quel type d'analyse, quelques recommandations spécifiques à l'analyse métabolomique existent. En effet, afin de limiter les biais possibles lors de l'analyse, il est conseillé d'analyser les échantillons dans un ordre aléatoire, et, dans le cas de plusieurs séries d'injections, de répartir les différents groupes expérimentaux de façon équitable au travers des séries (Begou *et al.*, 2018). L'utilisation d'échantillons de contrôle qualité (QC) est également préconisée (Theodoridis *et al.*, 2011). Ces échantillons correspondent en général à un mélange d'un volume identique de tous les échantillons. Begou *et al.* recommandent de placer ces échantillons régulièrement au cours de la séquence analytique à raison d'un QC tous les 5 échantillons pour les séquences <100 échantillons. Pour les séquences plus longues, les auteurs préconisent que les QC doivent représenter 10% des injections (Begou *et al.*, 2018). Ces échantillons peuvent également être utilisés dans le but de conditionner le système avant l'analyse des échantillons (Theodoridis *et al.*, 2011). Enfin, pour déterminer la présence de *carry-over* ou identifier des composés issus du système analytique (contaminants) et non des échantillons analysés, il est possible d'injecter des « blancs analytiques » (Want, 2018). Cependant, il est préférable de placer ces échantillons en fin de séquence car l'injection d'échantillons « non-matrice » peut perturber l'équilibre du système analytique (Begou *et al.*, 2018; Gika *et al.*, 2016).

### D.4. Traitement du signal

Cette étape a pour but de transformer les signaux acquis lors de l'analyse LC-MS en une matrice de données interprétable. L'objectif est de regrouper les informations qui caractérisent chaque ion détecté, à savoir le rapport  $m/z$ , le temps de rétention et l'intensité (regroupé sous le terme *features*),



en une table de données. Etant donné que les approches non-ciblées visent à analyser le plus grand nombre d'espèces, plusieurs milliers de signaux sont collectés à chaque injection et il est donc nécessaire d'avoir recours à des outils de détection automatiques des signaux. De très nombreux logiciels ont été développés pour la détection et le traitement des données parmi lesquels il est possible de citer les logiciels libres d'accès : XCMS (Smith *et al.*, 2006), MS-DIAL (Tsugawa *et al.*, 2015), mzMine (Pluskal *et al.*, 2010), MetaboAnalyst (Xia and Wishart, 2016) ou certains logiciels commerciaux tels que Compound Discoverer™, Metaboscape ou encore Progenesis Q1. Spicer *et al.* ont recensé, en 2017, les outils libres d'accès disponibles pour le traitement des données de la détection à l'annotation (Spicer *et al.*, 2017). Compte-tenu de la diversité des outils disponibles, plusieurs études se sont intéressées à la comparaison de ces différents logiciels avec cependant peu de convergence des conclusions quant au meilleur outil dédié au traitement de données (Hao *et al.*, 2018; Hemmer *et al.*, 2020; Li *et al.*, 2018). A l'heure actuelle, la solution de traitement de données métabolomique qui semble la plus utilisée est la solution libre d'accès XCMS.

Bien que tous différents, la plupart des outils de détection et d'intégration automatique des pics suivent les mêmes étapes, parfois dans des ordres différents.

- La première étape consiste en la détection de tous les pics présents dans chaque échantillon indépendamment des autres. A cette étape, un certain nombre de filtres, tels que des seuils d'intensité, de rapport signal/bruit, sont applicables afin de limiter la détection de pics mal résolus ou trop peu intenses.
- L'étape suivante est due au fait que, lors de l'analyse LC-MS de multiples échantillons, de légères déviations sont observables, principalement au niveau du temps de rétention mais aussi du rapport  $m/z$  mesuré (Burton *et al.*, 2008). Il est donc important d'aligner les spectres obtenus afin de regrouper les pics communs à tous les échantillons. Lors de cette étape, le paramétrage est également important afin de ne pas regrouper des pics différents ou à l'inverse considérer une même espèce comme deux ions distincts. Là encore, certains filtres basés sur des fréquences de détection (par exemple ne conserver que les pics détectés dans > 70% des échantillons d'un même groupe) sont applicables. Dans le cas d'XCMS, une première étape de « *grouping* » des pics est réalisée avant d'appliquer une correction de la dérive en temps de rétention. Enfin, suite à cette correction, une dernière étape de regroupement des pics entre les échantillons est réalisée.
- La dernière étape permet de pallier d'éventuels signaux non extraits lors de la détection automatique des pics, en identifiant les échantillons dans lesquels certaines espèces n'auraient pas été intégrées. Pour XCMS, cette stratégie de complétion des données est basée sur l'intégration de la zone où le pic est attendu (Smith *et al.*, 2006).

Malgré l'application de seuils lors de la détection automatique des pics, les données peuvent encore contenir de nombreux signaux artéfactuels à l'issue de cette étape. De nombreuses stratégies ont été proposées afin d'éliminer ces signaux. Des échantillons « blancs » peuvent être utilisés afin d'éliminer des ions qui ne seraient pas liés à la matrice analysée, en conservant par exemple les *features* dont le signal excède 10 fois le bruit de fond ou dont le ratio « moyenne échantillon » / « moyenne blanc » >3 (Broadhurst *et al.*, 2018; Schiffman *et al.*, 2019; Want, 2018). Il n'existe toutefois pas de directives précises quant aux seuils à fixer qui restent, à ce jour, à l'appréciation de l'expérimentateur. Les QC sont également fréquemment utilisés pour s'assurer de la qualité des données et éliminer les *features* indésirables. L'élimination des variables dont le coefficient de variation (CV) excède 30% dans les QC est une recommandation fréquemment rencontrée dans la littérature (Broadhurst *et al.*, 2018; Theodoridis *et al.*, 2011). Les QC sont aussi utilisés afin d'évaluer et de corriger les éventuelles dérives du signal en fonction de l'ordre d'injection. En effet, puisqu'il s'agit du même échantillon injecté régulièrement tout au long de la séquence, il est possible de « lisser » les variations observées pour chaque métabolite à l'aide des QC. Pour cela, de nombreux outils ont été développés ou adaptés tels que les régression LOESS (Dunn *et al.*, 2011) ou l'utilisation de fonctions splines cubiques (Kirwan *et al.*, 2013). Enfin, il a été proposé plus récemment de réaliser des séries de dilution de QC et d'éliminer les variables qui ne corrélaient pas à la dilution (Lewis *et al.*, 2016). Il est estimé que, bien que redondante, la combinaison de ces filtres conduit à l'élimination de 40% du jeu de données initial (Broadhurst *et al.*, 2018).

#### D.5. Stratégies de normalisation

Afin d'étudier les variations du métabolome dues à un facteur d'intérêt, il est préférable d'éliminer les autres sources de variabilité indésirable pouvant masquer la variabilité biologique d'intérêt. Pour cela, il existe de nombreuses méthodes de normalisation à différentes étapes du *workflow*. Il est possible de distinguer les méthodes de normalisation pré-analytiques des méthodes post-analytiques. La normalisation pré-analytique repose la plupart du temps sur la dilution des échantillons afin de normaliser un paramètre biologique donné. Il peut s'agir de la quantité de protéine ou du poids de l'échantillon pour les tissus (Andres *et al.*, 2020), du nombre de cellules lors de l'analyse d'extraits cellulaires (Artati *et al.*, 2019), ou encore de divers paramètres tels que la teneur en créatinine (Vogl *et al.*, 2016), l'osmolalité (Gagnebin *et al.*, 2017) ou le volume de 24 h dans le cas de l'urine (Godzien *et al.*, 2011). Dans le cas de l'urine, de nombreuses études existent visant à déterminer une méthode optimale de normalisation car la concentration de ce biofluide peut varier de manière importante (jusqu'à un facteur 14) en raison de paramètres non-corrélés au facteur étudié tels que l'hydratation (Warrack *et al.*, 2009). Bien que de nombreuses études aient été menées, la disparité des conclusions

n'a pas permis d'établir un consensus quant à un paramètre à privilégier pour la normalisation d'échantillons urinaires.

Outre les méthodes de normalisation pré-acquisition, il existe également un grand nombre de méthodes développées pour éliminer la variabilité non désirée. Le *package* R Normalyzer illustre bien cette diversité de méthodes puisqu'il permet l'utilisation et la comparaison de 12 techniques de normalisation différentes (Chawade *et al.*, 2014). Les normalisations basées sur le signal telles les normalisations TIC (*total ion current*) ou MSTUS (*MS total useful signal*) restent parmi les plus utilisées (Warrack *et al.*, 2009). A l'instar des méthodes de normalisation pré-acquisition, les conclusions des études comparatives n'ont pas permis de déterminer une solution idéale pour la normalisation des signaux.

#### D.6. Annotation des données

Une fois les données filtrées et normalisées, l'étape suivante consiste à annoter les variables puis confirmer ces annotations afin d'identifier les métabolites qui constituent le jeu de données. De manière générale, comme les données de métabolomique contiennent des milliers de signaux, ces dernières sont annotées de manière automatique en faisant correspondre la matrice de données obtenue suite à l'analyse, à une base de données publique ou privée. Une variable présentant le même rapport  $m/z$  ( $\pm$  une tolérance fixée en fonction de la précision de mesure de masse de l'appareil utilisé) qu'une entrée dans la base de données est donc susceptible de correspondre au métabolite correspondant à cette entrée. L'utilisation de bases de données publiques comporte cependant quelques limitations. En effet, bien que de très nombreux métabolites y soient référencés, seule une correspondance sur la base du rapport  $m/z$  est possible, ce qui peut conduire à de nombreuses annotations erronées. A l'inverse, l'utilisation de bases de données constituées en interne au sein des laboratoires permet des annotations souvent plus robustes puisque les informations de  $m/z$  et de temps de rétention peuvent-être utilisées pour réaliser la correspondance. La contrepartie de ces bases de données internes réside dans le nombre de métabolites qui y sont contenus, souvent restreint par rapport aux bases publiques.

Lors de l'annotation, il est important de considérer qu'une *feature* dans le jeu de données ne correspond pas nécessairement à un composé unique. En effet, en considérant les ions adduits, les ions fragments, les artefacts et autres signaux, le nombre de métabolites uniques dans un jeu de données est bien inférieur au nombre de *features*, comme l'ont montré Mahieu et Patti (Mahieu and Patti, 2017). Ainsi, afin d'augmenter la couverture des signaux annotés, il est possible de rechercher les différentes espèces découlant d'un même composé (adduits, fragments connus etc...). Cependant,

la recherche de l'ensemble des adduits, fragments et autres signaux augmente les chances d'annotations erronées.

Une fois le jeu de données annoté, il est important de pouvoir confirmer ou infirmer les annotations proposées. Etant donné qu'il s'agit d'une étape cruciale dont dépend fortement l'interprétation de l'opérateur, des recommandations visant à standardiser les pratiques d'identification ont émergé. Le groupe d'analyse chimique de la « *Metabolomic Standard Initiative* » a ainsi proposé de classer l'identification des métabolites selon 4 niveaux de confiance (Sumner *et al.*, 2007).

- Niveau 1 : Deux critères orthogonaux communs avec un standard de référence, par exemple masse et temps de rétention
- Niveau 2 : Même conditions que le niveau 1 mais en comparaison avec une base de données externe ou la littérature
- Niveau 3 : Pas d'identification proposée mais une appartenance à une classe chimique
- Niveau 4 : Composés non-identifiés

Des améliorations de ces critères ont été proposées, comme notamment un système de score afin de donner plus d'importance aux identifications combinant spectres MS et MS/MS (Chaleckis *et al.*, 2019; Creek *et al.*, 2014).

Cependant, si les bases de données contiennent de plus en plus de métabolites, l'annotation reste un enjeu majeur des approches non-ciblées car un nombre important de *features* demeurent inconnues après l'étape d'annotation. Face à ce défi, le pouvoir de résolution et la précision de mesure croissants des appareils permettant de proposer un nombre de plus en plus restreint de formules brutes pour une masse donnée, ou l'utilisation de réseaux moléculaires basés sur des similarités de fragmentation sont autant d'outils permettant d'augmenter le nombre de *features* annotées.

### D.7. Analyses statistiques

L'utilisation principale de la métabolomique étant d'étudier des différences d'abondances de certains métabolites entre deux ou plusieurs groupes expérimentaux, il est nécessaire d'avoir recours à des analyses statistiques afin de conclure quant à la significativité des phénomènes observés. Parmi les outils classiquement utilisés en métabolomique, on peut distinguer deux grands types d'analyses statistiques, les analyses **univariées** et **multivariées**.

Les analyses univariées correspondent à des tests statistiques réalisés variable par variable. La significativité statistique de ces tests repose en général sur la « *p*-value » qui traduit la probabilité de rejeter une hypothèse alors que cette dernière est correcte, le seuil de cette valeur est par convention fixé à 0.05. Les tests de Student (pour la comparaison de deux groupes) ou les ANOVA (pour la

comparaison de plusieurs groupes) font partie des tests paramétriques les plus utilisés. Un test paramétrique est applicable uniquement si l'hypothèse de normalité est vérifiée. Dans le cas inverse, des tests non-paramétriques doivent être employés, les équivalents non paramétriques des tests de Student ou de l'ANOVA étant respectivement, les tests de Wilcoxon et de Kruskal-Wallis. Cependant, l'utilisation des tests univariés comporte plusieurs désavantages et limitations. L'analyse de chaque variable indépendamment ne permet pas de prendre en compte les possibles interactions entre les métabolites, par exemple les métabolites issus d'une même voie biochimique (Alonso *et al.*, 2015). De plus, compte-tenu de la structure des données issues des approches non-ciblées, l'utilisation de tests univariés est propice à l'apparition de faux-positifs, conduisant à conclure qu'une différence est significative alors qu'elle ne l'est pas. Ainsi, lors de la multiplication de tests univariés, il est nécessaire d'appliquer des méthodes de comparaisons multiples afin de corriger la  $p$ -value. Il existe de nombreuses méthodes de comparaisons multiples, les plus employées étant les méthodes de Bonferroni ou de Benjamini-Hochberg. La correction de Bonferroni n'est pas recommandée pour des données de métabolomique car le seuil de correction est trop proche de 0 lorsque le nombre de tests devient trop important (Xi *et al.*, 2014).

Contrairement aux méthodes univariées, les méthodes multivariées considèrent les métabolites (variables) dans leur ensemble et permettent ainsi la prise en compte des relations entre les variables. Au sein des méthodes multivariées, il est possible de distinguer deux types d'approches, les approches **supervisées** et les approches **non-supervisées**. Les approches non-supervisées sont des approches descriptives qui ne tiennent pas compte de l'appartenance des échantillons à un groupe expérimental dans l'élaboration du modèle. A l'inverse, les approches supervisées utilisent cette information pour déterminer quels sont les *features* les plus associés au facteur étudié.

**L'analyse en composante principale (ACP)** est la méthode non-supervisée la plus couramment utilisée en métabolomique car cette méthode est bien adaptée pour les données à haute dimensionnalité, c'est-à-dire lorsque le nombre de variables excède fortement le nombre d'échantillons (Alonso *et al.*, 2015). Le principe de l'ACP repose sur la combinaison linéaire de l'ensemble des *features* en un nombre réduit de variables dont la covariance est minimisée, ces variables étant appelées composantes principales. Cette approche permet de regrouper les variables dont la covariance est forte sous une même composante principale tout en minimisant la covariance entre les composantes. En métabolomique, l'ACP est fréquemment utilisée en première intention pour identifier des biais dans les données telles que des variations analytiques ou pour étudier la stabilité d'échantillons de façon globale (Gika *et al.*, 2008).

A l'inverse des méthodes non-supervisées qui sont à visée descriptive, les **approches supervisées** ont vocation à **classer les échantillons** selon leur appartenance à un groupe expérimental et à identifier

quels sont les métabolites permettant d'expliquer la séparation des groupes. L'analyse discriminante par régression des moindres carrés partiels (PLS-DA pour *partial least square-discriminant analysis*) est la méthode supervisée la plus employée dans les études métabolomiques non-ciblées (Gromski *et al.*, 2015). La PLS-DA est une méthode qui vise à optimiser la séparation des groupes en maximisant la covariance entre un ensemble de variables  $X$  et une variable explicative  $Y$ . De la même manière que l'ACP, la PLS-DA a pour objectif de regrouper l'ensemble des *features* en un nombre réduit de composantes. Les avantages de la PLS-DA découlent de sa capacité à traiter des données hautement colinéaires, ainsi qu'à fournir des informations sur les variables telles que l'importance de la variable dans la projection (VIP pour *variable importance on projection*). Cette valeur est souvent utilisée pour conclure quant à l'importance d'une variable dans le modèle étudié. Une variable dont le score VIP > 1 étant considérée par convention comme ayant un poids important dans la séparation des groupes (Bujak *et al.*, 2016). Cependant, la PLS-DA présente quelques limitations, notamment un risque « d'*over-fitting* », c'est-à-dire que le modèle classe trop bien les échantillons utilisés pour le construire mais se révèle inefficace pour discriminer d'autres échantillons (Gromski *et al.*, 2015). Un autre désavantage de cette méthode est lié au fait que certaines variables non-corrélées au facteur d'intérêt peuvent influencer les résultats. Pour pallier cette limitation, l'OPLS-DA (orthogonal PLS-DA) a été développée (Trygg and Wold, 2002) afin d'éliminer les variables  $X$  qui ne sont pas corrélées au facteur  $Y$ , ou autrement dit dont la variabilité est orthogonale au facteur d'intérêt. D'autres évolutions de la PLS-DA ont également été développées, notamment par Boccard *et al.* qui ont proposé un *workflow* combinant une OPLS-DA multi-blocs avec une ANOVA sur les différents blocs afin de pouvoir étudier simultanément l'effet de plusieurs facteurs et d'évaluer le poids de chaque facteur sur les variables (Boccard and Rudaz, 2016). Il existe également des méthodes multivariées supervisées non basées sur la PLS-DA, comme le détaillent Gromski *et al.* dans leur revue (Gromski *et al.*, 2015), mais celles-ci sont généralement moins utilisées en métabolomique.

#### D.8. Interprétation biologique

L'étape finale du *workflow* de l'analyse métabolomique consiste à déterminer la significativité biologique des modifications métaboliques observées. Autrement dit, cela consiste à générer des hypothèses sur les causes et les conséquences des modifications des niveaux des métabolites entre les différents groupes comparés. Cette étape reste, à l'heure actuelle, majoritairement dépendante de l'expertise de l'expérimentateur. Cependant, certains outils peuvent être utilisés pour faciliter l'interprétation des données. L'approche la plus communément employée consiste à déterminer à quelle(s) voie(s) métabolique(s) appartiennent les métabolites retrouvés significativement altérés par le facteur étudié afin d'émettre des hypothèses sur l'impact de ces modifications sur le métabolisme.

Cependant, ce type d'approche a tendance à minimiser l'interconnexion des voies métaboliques et le fait qu'une altération observée dans une voie métabolique peut avoir des répercussions sur d'autres parties du réseau métabolique (qui correspond à l'ensemble des voies métaboliques d'un organisme donné). Face à cette limitation, des outils tels que MetScape (Basu *et al.*, 2017) ou MetExplore (Cottret *et al.*, 2018), basés sur les réseaux métaboliques, sont développés pour aider l'expérimentateur à interpréter, de façon plus globale, la signature métabolique obtenue. Grâce à des fonctionnalités telles que l'extraction de sous-réseaux, ces outils permettent de connecter des métabolites non connectés dans les voies métaboliques traditionnelles (Perez De Souza *et al.*, 2020). Cette extraction de sous-réseaux permet, à partir du profil métabolique de l'étude, d'extraire seulement la partie du réseau permettant de relier les métabolites du profil obtenu de façon biologiquement pertinente (Frainay and Jourdan, 2017). Une preuve de concept de l'utilisation de sous-réseaux métaboliques a été réalisée en 2014 pour l'étude de la réponse de levures à une exposition au cadmium (Milreu *et al.*, 2014).

Cependant, malgré les récents développements, l'interprétation des données métabolomique reste un des points-clés du *workflow* (Barupal *et al.*, 2018).

## **E. La métabolomique au service de la toxicologie**

La toxicologie chimique peut être définie comme l'étude des effets toxiques ou indésirables de composés chimiques sur un organisme vivant (Eaton and Klaassen, 2008). Dès sa description en 1999, la métabolomique a été appliquée à des questions de toxicologie : en effet dans sa première utilisation, l'approche métabolomique consistait à mesurer des variations métaboliques dans l'urine de rats exposés à différentes toxines (Nicholson *et al.*, 1999). Depuis, dans ce domaine, la plupart des études visent à identifier des modifications de la concentration de métabolites suite à l'exposition à divers composés potentiellement toxiques, avec l'objectif de détecter ces changements avant l'apparition de signes cliniques (Roux *et al.*, 2011). Compte-tenu de la vue d'ensemble que fournissent les études métabolomiques non-ciblées, ces dernières ont été largement utilisées en toxicologie pour déterminer des biomarqueurs d'exposition, d'effet ou formuler des hypothèses sur les mécanismes d'action. Les biomarqueurs d'exposition correspondent à la présence dans l'échantillon, du composé toxique étudié, de son métabolite ou encore des produits issus de l'interaction entre le composé étudié et une molécule endogène. Les biomarqueurs d'effets correspondent aux molécules endogènes dont l'abondance est modifiée suite à l'exposition au composé étudié et qui traduisent un impact sur la physiologie (Mistretta and Charlier, 2013). L'intérêt de la métabolomique pour la toxicologie repose notamment sur le lien direct entre phénotype et métabolome mais aussi sur le fait que les informations puissent-être obtenues à partir d'échantillons périphériques (fèces, urines...) (Robertson *et al.*, 2011). De plus, en comparaison aux autres approches omiques, les données

généralisées par les expériences de métabolomique se rapprochent davantage des paramètres étudiés lors des études de toxicologie classiques (Ramirez *et al.*, 2013).

Que ce soit à travers l'environnement, le mode de vie, l'activité professionnelle ou encore l'alimentation, les Hommes, et d'une manière plus générale l'ensemble des êtres vivants, sont exposés tout au long de leur vie à de nombreuses molécules potentiellement toxiques dont il est important d'évaluer les effets sur la santé. Ainsi de nombreuses études métabolomiques ont été menées, parmi lesquelles nous pouvons citer par exemple l'étude de l'impact sur le métabolome de composés ubiquitaires tels que le benzo(a)pyrène (Wang *et al.*, 2015), dans laquelle les auteurs ont reporté une diminution dose-dépendante du taux de certains glycérophospholipides dans le plasma de rats exposés à ce composé. L'utilisation d'approche métabolomique par RMN a par exemple permis de générer des hypothèses sur d'éventuelles perturbations du métabolisme énergétique chez des femmes enceintes plus ou moins exposées à des pesticides en fonction de leur lieu d'habitation (Bonvallot *et al.*, 2013). Enfin, il est également possible de citer de multiples études menées *in vivo* et *in vitro* sur l'impact de l'exposition au bisphénol A, qui ont permis de révéler des altérations métaboliques induites par ce composé dès l'exposition à de faibles doses (Cabaton *et al.*, 2018; Tremblay-Franco *et al.*, 2015).

En dehors de l'exposition à des xénobiotiques tels que le bisphénol A ou le cadmium, la toxicologie s'intéresse également aux effets de l'exposition à des composés toxiques naturels tels que les mycotoxines (Rodríguez-Carrasco *et al.*, 2020). En effet, la caractérisation de l'exposition à ces composés ainsi que l'effet de celle-ci présente un intérêt majeur dans la mesure où, du fait de leur origine naturelle, la présence et la quantité de ces composés est moins « contrôlable ». Pour répondre à ce besoin de suivi de l'exposition, de nombreux travaux ont été menés afin de déterminer des biomarqueurs témoins de l'exposition à ces composés (*biomonitoring*). Parmi les composés formés de manière naturelle, il est possible de citer les AAH ou les HAP sur lesquels plusieurs études ont été menées afin de caractériser les métabolites formés lors de l'exposition à ces composés (Bessette *et al.*, 2009; Jacques *et al.*, 2010). A l'instar des composés exogènes, des approches métabolomiques sont également menées afin d'évaluer l'impact de ces composés naturels sur le métabolome. Dans une étude récente il a été montré que l'exposition au PhiP induisait une perturbation du métabolisme lipidique chez des rats, avec en particulier une diminution de certains précurseurs d'hormones stéroïdiennes (Zhao *et al.*, 2021).

Afin d'évaluer l'effet de notre exposition à divers composés toxiques, plusieurs stratégies analytiques coexistent, allant d'approches ciblées telles que la détection spécifique et quantification de certains métabolites connus tels que des AAH (Chevolleau *et al.*, 2020) ou certains adduits tels que des adduits



lysine ou glutathion (Karlstetter and Mally, 2020), jusqu'à des approches plus globales comme par exemple le screening non ciblé des adduits aux protéines ou à l'ADN (Hemeryck *et al.*, 2017). Enfin, la vue la plus globale de notre exposition est apportée par des approches dites « exposomique » qui visent à caractériser l'ensemble des composés auxquels nous sommes exposés, qu'ils proviennent de l'environnement ou de l'alimentation (Jamin *et al.*, 2014; Pourchet *et al.*, 2020). Ces dernières approches mettent en exergue la diversité et le nombre de composés auxquels nous sommes exposés, soulignant les possibles effets additifs ou synergiques de ces composés.

En raison de la multiplicité des études et de l'intérêt croissant pour les sciences omiques (en particulier la métabolomique) pour la toxicologie, des initiatives, à l'image de l'« *OECD omics reporting framework* », sont prises afin de faciliter l'intégration des résultats d'études métabolomiques dans la prise de décisions réglementaires (Harrill *et al.*, 2021). L'objectif de ce groupe de travail étant d'établir des directives afin de standardiser et harmoniser les pratiques de métabolomique (mais aussi d'autres omiques) afin que les données générées par ces approches soient considérées par les agences d'évaluation du risque.

### **E.1. La voie des acides mercapturiques comme marqueur d'exposition**

De nombreux xénobiotiques subissent des réactions de biotransformation au sein de l'organisme. Bien que certains composés soient activés par ces réactions, la majorité sont inactivés *via* des réactions d'oxydation, de réduction, d'hydrolyse ou encore de conjugaison (Iyanagi, 2007). On distingue deux voies de métabolisation : le métabolisme de phase I et le métabolisme de phase II. Les réactions de phase I peuvent être catalysées par de nombreuses enzymes (déshydrogénases, réductases, mono-oxygénases) et visent à ajouter ou modifier un groupement fonctionnel, afin de réduire son hydrophobicité et permettre soit son excrétion ou sa métabolisation par les enzymes de phase II. Les enzymes de phase II ont pour rôle de catalyser des réactions de conjugaison entre le composé à détoxifier et des groupements hydrosolubles pour former des conjugués sulfates, glucuronides ou encore glutathion (Grant, 1991). Les 3 familles d'enzymes majeures du système de phase II sont les glucuronyltransférases, les sulfotransférases et les glutathion-S-transférases, toutes les 3 étant, entre autres, sous le contrôle du gène Nrf2, régulateur de la réponse antioxydante (Itoh *et al.*, 1997).

Parmi ces différentes voies de détoxification, la conjugaison au glutathion et la voie métabolique qui en découle, à savoir la voie des acides mercapturiques, est particulièrement intéressante sur le plan toxicologique en raison de sa capacité à détoxifier principalement les composés électrophiles (Hanna and Anders, 2019). En effet, les composés électrophiles sont hautement réactifs vis-à-vis des macromolécules telles que l'ADN et les protéines et peuvent donc conduire à la formation d'adduits covalents aux effets délétères. Suite à un screening de plus de 300 composés, Ashby et Tennant

concluaient en 1991 que 80% des composés génotoxiques étaient électrophiles (Ashby and Tennant, 1991). Ainsi, compte-tenu de la prépondérance des composés électrophiles parmi les espèces génotoxiques et du rôle clé du glutathion dans la détoxification de ces composés, de nombreux travaux se sont intéressés à la voie des acides mercapturiques décrite en Figure 14. Le premier évènement consiste en la conjugaison d'une espèce électrophile avec le glutathion, dans le foie mais aussi dans d'autres organes, soit par réaction spontanée (réactivité chimique intrinsèque) soit sous l'action d'une glutathion-S-transférase (Mathias and B'Hymer, 2014). Deux réactions enzymatiques successives faisant intervenir la  $\gamma$ -glutamyltransférase et une dipeptidase conduisent à la formation d'un conjugué cystéine. Ce dernier est *N*-acétylé pour former *in fine* un conjugué à l'acide mercapturique. La première étape de cette séquence métabolique, à savoir la conjugaison au glutathion, a pour but d'inactiver le composé électrophile l'empêchant ainsi de réagir avec les autres composants cellulaires comme les protéines, les lipides ou encore les acides nucléiques. L'étape de *N*-acétylation consiste, de la même manière que les conjugaisons de phase II, à augmenter la polarité du composé et son hydrophilie afin de faciliter son excrétion dans les urines (Mathias and B'Hymer, 2014).

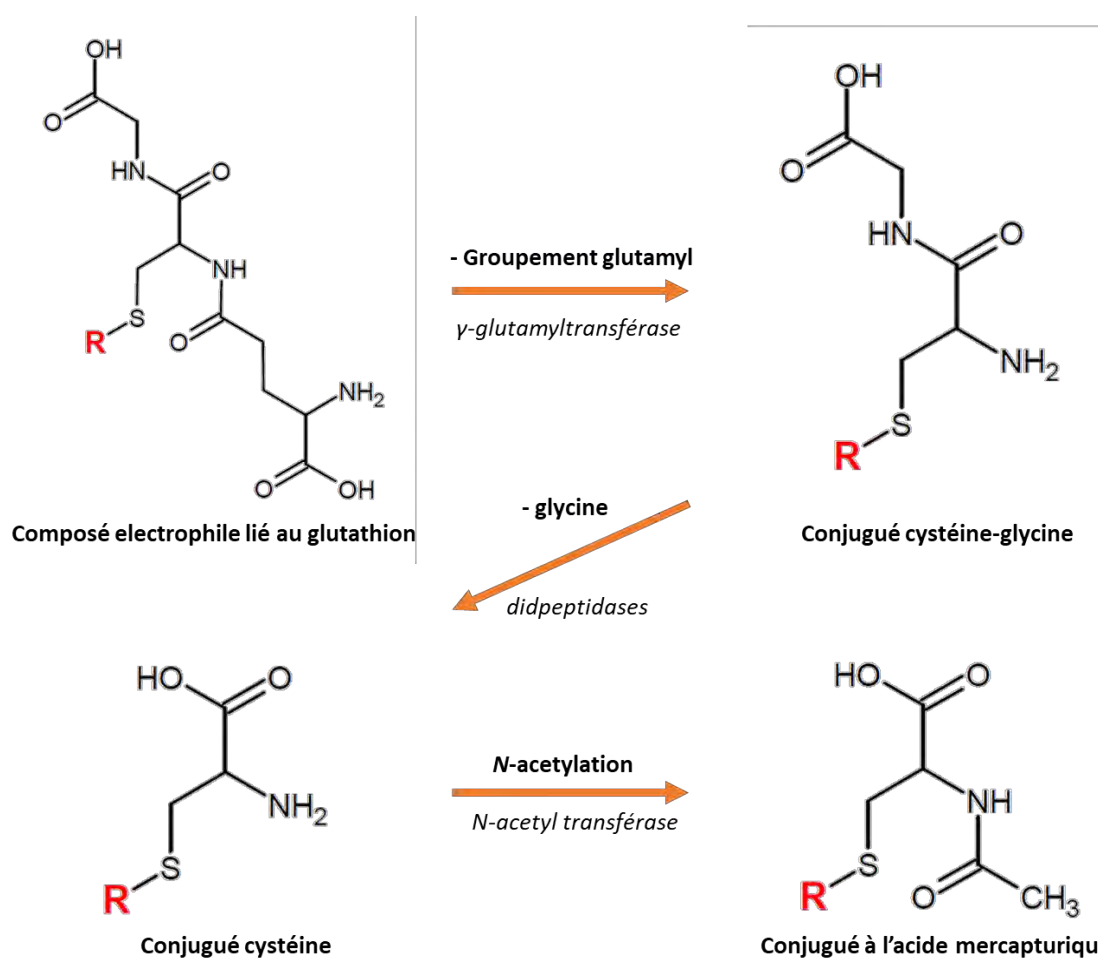


Figure 14 Voie de synthèse des acides mercapturiques. R correspond à un composé électrophile. Adapté de Mathias *et al.*

Etant donné la diversité des composés électrophiles, de nombreuses approches ont été développées pour doser des conjugués mercapturiques spécifiques. En revanche, seulement quelques groupes se sont intéressés à détecter de manière non-ciblée les conjugués afin d'évaluer une exposition globale aux dérivés électrophiles. Dans le cas des approches ciblées, l'objectif est de quantifier un nombre restreint de mercapturates pour mesurer précisément l'exposition au composé toxique de départ. Par exemple, Kellert *et al.* ont développé une méthode LC-MS/MS pour doser les conjugués mercapturiques de l'acrylamide et de son métabolite la glycidamide dans des urines de fumeurs et non-fumeurs (Kellert *et al.*, 2006). En effet, en raison de leur haute réactivité ces composés ne sont pas excrétés sous forme libre et leurs conjugués mercapturiques sont utilisés comme marqueurs d'exposition. L'acrylamide est un composé largement utilisé dans l'industrie, formé spontanément lors de la cuisson à haute température de certains aliments mais également hautement présent dans la fumée de cigarette et il est estimé que 60% de l'acrylamide ingéré est conjugué au glutathion pour être ensuite excrété dans les urines *via* la voie des acides mercapturiques (Dearfield *et al.*, 1988). D'autres méthodes ont été développées afin de doser les conjugués mercapturiques de polluants de l'air tels que le benzène, le toluène et le xylène (Sabatini *et al.*, 2008) ou encore les métabolites de l'atrazine, un herbicide largement utilisé entre les années 1960 et 2000 (Panuwet *et al.*, 2008).

Outre les composés exogènes, des composés électrophiles peuvent également être formés de manière endogène avant d'être détoxifiés et excrétés par la voie des acides mercapturiques. C'est notamment le cas des produits de peroxydation lipidique, qui, comme vu précédemment dans cette introduction, possèdent des propriétés réactives vis-à-vis de l'ADN, des protéines et d'autres composés. Compte tenu de sa réactivité largement décrite, le 4-HNE a fait l'objet de nombreuses études visant à caractériser sa métabolisation. Dès 1995, suite à des travaux menés au sein du laboratoire d'accueil, il a été montré que, suite à une administration par intraveineuse à des rats de 4-HNE radiomarqué, la majorité de la radioactivité était retrouvée dans les urines (67% après 48 h) (Alary *et al.*, 1995a) majoritairement sous forme de conjugués mercapturiques avec au moins 4 métabolites distincts. Plus récemment, une étude basée sur l'administration par voie orale de 4-HNE radiomarqué a donné des résultats similaires, à savoir qu'environ 50% de la radioactivité était retrouvée dans les urines après 24 h (Keller *et al.*, 2015). De la même manière, le HNE administré était retrouvé sous forme de plusieurs conjugués mercapturiques dans les urines dont le DHN-MA. Considérant l'intérêt toxicologique de doser les acides mercapturiques issus de la métabolisation du 4-HNE, plusieurs méthodes de dosage par LC-MS/MS ont été proposées (Kuiper and Stevens, 2011; Rathahao *et al.*, 2005). En particulier grâce à une méthode développée au sein du laboratoire d'accueil par Rathahao *et al.*, les auteurs ont pu mettre en évidence l'intérêt de ce biomarqueur pour évaluer la peroxydation lipidique en appliquant la méthode à des urines de rats traités ou non au BrCCl<sub>3</sub>, un

composé connu pour induire la peroxydation lipidique hépatique. D'autres produits de peroxydation lipidique tels que le 4-HHE, un produit de l'oxydation des acides gras oméga-3, sont également détoxifiés *via* la voie des acides mercapturiques, mais il existe peu de travaux visant à détecter et quantifier ces composés dans les urines. En revanche, une méthode de quantification de ces composés dans les eaux fécales à l'aide d'une dérivation spécifique des aldéhydes a récemment été proposée (Chevolleau *et al.*, 2018). Afin d'étendre la gamme des mercapturates analysés, et d'évaluer de manière plus globale l'exposition aux composés électrophiles, des méthodes visant à la quantification de plusieurs mercapturates ont été développées. Alwis *et al.* ont ainsi développé une méthode de quantification de 28 métabolites, dont une majorité de mercapturates, issus de composés organiques volatils réactifs (Alwis *et al.*, 2012). Plus récemment, une étude dosant 17 mercapturates a été réalisée pour déterminer l'impact du tabagisme sur le taux de ces composés (Frigerio *et al.*, 2020a). Ces mêmes auteurs ont réalisé une analyse métabolomique non-ciblée sur ces mêmes échantillons dans le but d'évaluer la capacité d'une méthode non-ciblée à détecter et retrouver les différences observées avec la méthode ciblée. Dans cette seconde étude, seules les espèces les plus abondantes dans l'approche ciblée ont pu être détectées avec la méthode non-ciblée, témoignant du challenge analytique que représente l'analyse des mercapturates, qui sont souvent des composés minoritaires (Frigerio *et al.*, 2020b).

Enfin, afin d'obtenir une vue globale de l'exposition aux composés électrophiles, quelques études ont été menées afin de détecter de manière non-ciblée l'ensemble des mercapturates dans l'urine. Les travaux du groupe de Wolfgang Völkel ont notamment ouvert la voie à la caractérisation non-ciblée des mercapturates urinaires. En effet, en 2005, ce groupe a développé une méthode de détection des acides mercapturiques basée sur une perte de neutre caractéristique de ces espèces (à savoir une perte de 129 Da) correspondant à la fragmentation de la liaison thioéther (Scholz *et al.*, 2005). La stratégie reposait sur l'utilisation d'une méthode MS/MS ciblée sur un appareil de type triple quadrupole afin de fragmenter tous les ions précurseurs de rapport  $m/z$  compris entre 200 et 450 et d'identifier ceux produisant la perte de neutre recherchée. Dans cette étude de preuve de concept, les auteurs ont surchargé des urines de rats avec 16 standards de mercapturates et ont pu tous les retrouver grâce à l'approche non-ciblée basée sur la perte de neutre. De plus, les auteurs ont également identifié un autre mercapturate non présent dans la liste initiale de standard à savoir le DHN-MA, un métabolite du 4-HNE, mentionné plus haut dans cette introduction. Afin d'éprouver leur méthode, une seconde étude a été réalisée visant à séparer deux groupes de volontaires sains et traités ou non à l'acétaminophène, un composé connu pour être métabolisé et excrété dans les urines sous forme d'acide mercapturique (Wagner *et al.*, 2006). En utilisant leur approche basée sur la perte de neutre, les auteurs ont pu mettre en évidence que les composés responsables de la discrimination

statistique des groupes correspondaient aux métabolites urinaires de l'acétaminophène. Bien que prometteuse, cette approche souffrait d'une limitation principale qui était l'utilisation d'un instrument à basse résolution rendant l'identification de composés inconnus difficile. Depuis lors, malgré un intérêt toxicologique pointé dans plusieurs revues (Gonçalves-Dias *et al.*, 2019; Mathias and B'Hymer, 2014), peu d'études ont été menées pour l'analyse non-ciblée des acides mercapturiques. Récemment, la même stratégie combinant une détection non-ciblée des mercapturates par l'utilisation de la perte de neutre combinée à l'infusion d'un composé en post-colonne afin de corriger l'effet matrice a été développée (Bloch *et al.*, 2019). Dans cette étude, les auteurs ont pu mettre en avant les limites de la détection non-ciblée, car les mercapturates les moins abondants n'étaient plus détectés en comparaison à une approche ciblée, mais aussi les forces, en raison des nombreux mercapturates putatifs détectés dans les échantillons. Néanmoins, cette étude a également été menée en utilisant un instrument à basse résolution limitant ainsi les perspectives d'identification des composés inconnus. De récents travaux menés au sein du laboratoire d'accueil ont visé à pallier ces limitations en utilisant un instrument à haute résolution combiné à un mode de fragmentation non spécifique (le DIA). Cette étude menée par Jamin *et al.* sera présentée plus en détail dans le Chapitre II : Développement et optimisation de la méthode d'analyse, de ce manuscrit.

## **Objectifs de la thèse**

Si la consommation de viande rouge possède des avantages nutritionnels tels qu'une haute teneur en protéines, un apport en fer important et une sensation de satiété élevée (Wyness, 2016), les données de la littérature pointent aussi ses implications plus négatives. En effet, les associations entre une consommation élevée de viande rouge et de charcuterie et des effets négatifs sur la santé sont maintenant bien établies, en particulier pour le cancer colorectal mais aussi d'autres pathologies telles que les maladies cardio-vasculaires ou encore le diabète. En revanche, les mécanismes sous-jacents qui tendent à expliquer ces associations, sont eux, moins bien établis et toujours sujet à débat dans la littérature si bien que de nombreuses hypothèses, souvent complémentaires, sont proposées. Les travaux menés par l'équipe PPCA ont permis d'établir que l'implication du fer héminique semblait être un des éléments clés dans la promotion de la carcinogénèse colorectale induite par la consommation de viande rouge et de charcuterie. Ces effets délétères seraient notamment médiés par la capacité du fer héminique à catalyser la formation de ROS, qui agiraient sur les lipides polyinsaturés du régime alimentaire et conduiraient à la formation de produits de peroxydation lipidique. Certains de ces produits, tels que le 4-HNE, présentent une réactivité élevée vis-à-vis des macromolécules telles que l'ADN ou les protéines mais peuvent également altérer de nombreux processus physiologiques. Le 4-HNE, mais de manière plus générale l'ensemble des hydroxy-alcénals formés lors de la peroxydation lipidique, sont des composés électrophiles qui, comme de nombreux composés électrophiles sont (1)

génotoxiques et (2) détoxifiés par la voie des mercapturates. En revanche, contrairement aux effets du 4-HNE qui sont largement décrits dans la littérature, peu de travaux se sont intéressés à la caractérisation de l'ensemble des produits issus de la peroxydation lipidique ni à leurs potentiels effets.

La voie des acides mercapturiques est étudiée depuis de nombreuses années comme étant le reflet de notre exposition aux composés électrophiles. Ces derniers étant des composés réactifs peu stables, la mesure de leurs dérivés mercapturiques, leurs métabolites finaux stables, constitue une alternative adaptée pour l'évaluation de notre exposition. L'autre avantage réside dans le fait que la majorité des mercapturates sont excrétés dans l'urine qui est un biofluide facilement récoltable en grande quantité en raison du caractère non-invasif de son prélèvement. La majorité des études portant sur l'analyse des produits issus de la voie des acides mercapturiques sont des études ciblées visant à quantifier uniquement certains produits ne permettant pas d'estimer de manière globale notre exposition aux composés électrophiles. Bien que prometteuses, peu d'études se sont intéressées à la caractérisation globale de notre exposition *via* la mise en place d'une approche de screening non-ciblé des mercapturates.

Ainsi, en raison de l'essor de la métabolomique, qui est notamment lié aux progrès technologiques réalisés sur les instruments analytiques, le développement d'une approche non-ciblée permettant, dans la même analyse, la caractérisation du métabolome urinaire mais aussi du sub-métabolome des conjugués mercapturiques permettrait de répondre à un besoin jusque-là non satisfait.

**L'objectif de cette thèse est de développer une approche de métabolomique non-ciblée visant à caractériser le métabolome urinaire ainsi que l'ensemble des acides mercapturiques présents dans cette matrice. La stratégie de détection des mercapturates repose sur l'utilisation du mode DIA associé au suivi de la perte de neutre caractéristique de 129,0426 Da correspondant à la fragmentation de la liaison thio-éther. Le second objectif est d'éprouver l'approche en l'appliquant à la thématique du lien entre consommation de viande rouge et cancer du côlon. En effet, compte-tenu de l'implication probable des produits de peroxydation lipidique dans la pathogénèse et du manque de connaissance vis-à-vis de la diversité des produits formés, il apparaît pertinent de s'intéresser aux acides mercapturiques formés lors d'un régime riche en viande rouge. Un schéma général du devenir des mercapturates dans l'organisme est présenté en Figure 15.**

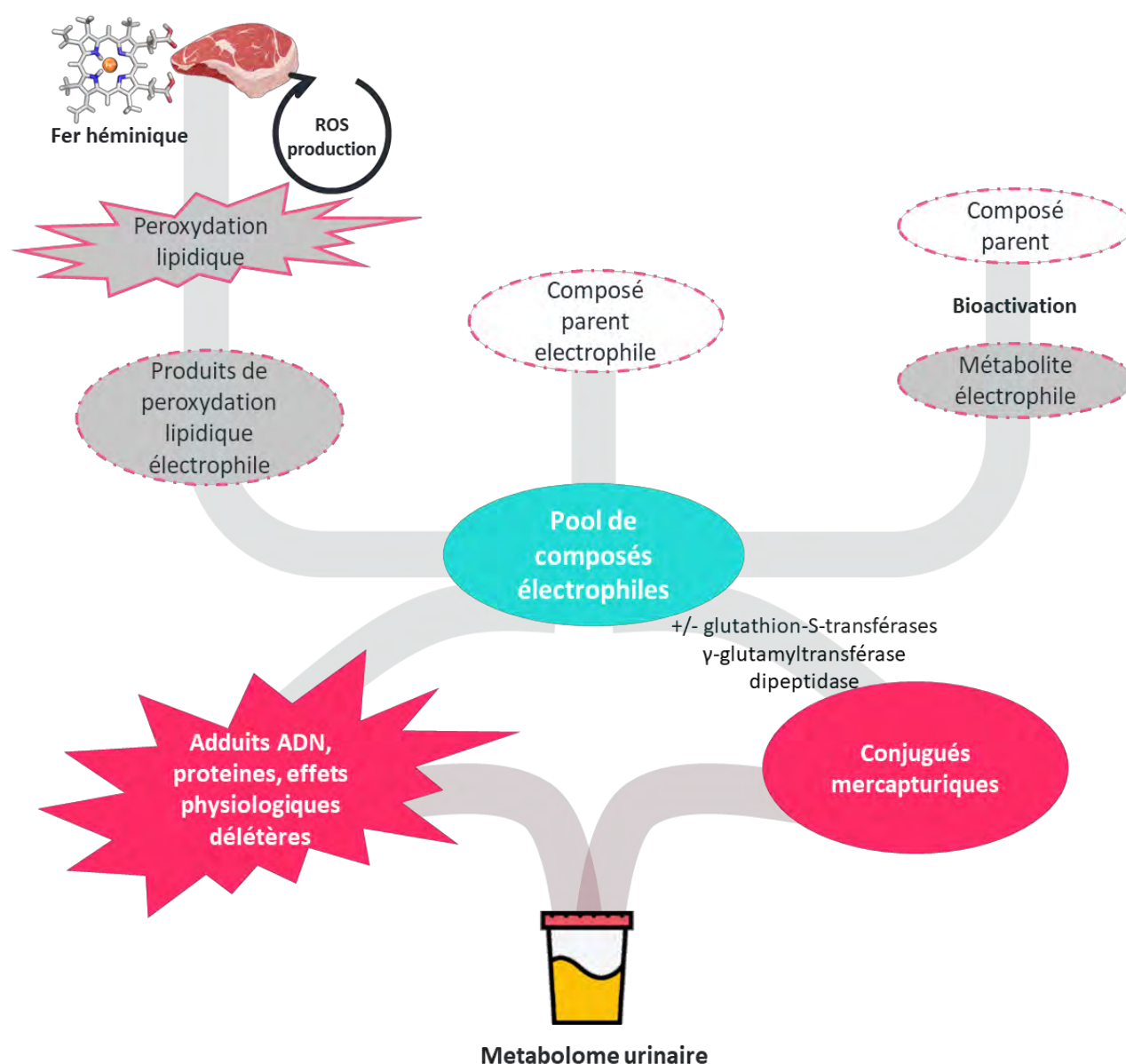


Figure 15 Schéma général du devenir des composés électrophiles au sein de l'organisme

Dans un premier temps, les travaux de cette thèse ont porté sur l'optimisation d'une méthode de caractérisation non ciblée des mercapturates par DIA déjà développée au sein du laboratoire. Ces optimisations ont permis une meilleure détection des conjugués mercapturiques, ainsi qu'une amélioration du traitement des données DIA pour assurer une meilleure fiabilité des données. Un autre pan du développement méthodologique consistait à optimiser les paramètres de l'analyse d'urine en métabolomique en menant notamment une étude sur la normalisation de ce biofluide. L'ensemble de ces développements constituent le chapitre II de ce manuscrit.

Dans un second temps cette approche a été appliquée à la thématique de la consommation de viande rouge en lien avec le cancer du côlon. Ces applications sont décrites dans les chapitres III et IV de ce manuscrit.

Le chapitre III se concentre sur l'application de cette méthodologie aux échantillons de la cohorte nutritionnelle française Nutrinet-Santé. Dans cette partie, 240 échantillons de la cohorte classés en 3 groupes de consommation de viande rouge ont été analysés à l'aide la méthode mise en place. Si les résultats obtenus n'ont pas permis de mettre en évidence de modulation profonde du mercapturome, l'analyse métabolomique non-ciblée a quant à elle permis d'identifier des potentiels biomarqueurs de consommation de viande rouge et de charcuterie liés à une potentielle modification du microbiote intestinal.

Enfin, le chapitre IV de ce manuscrit sera consacré aux expérimentations *in vivo* menées durant la thèse. Deux expérimentations *in vivo*, l'une nutritionnelle et l'autre de carcinogénèse, ont été menées sur des rats afin d'étudier l'impact de consommation de viande rouge sur la promotion du cancer colorectal. Dans cette partie, l'impact du régime lipidique sur la diversité des produits formés ainsi qu'une stratégie de prévention basée sur une supplémentation en antioxydant ont été testés. Ces deux expérimentations ont permis de confirmer l'hypothèse d'une peroxydation lipidique hème-induite ainsi qu'une promotion de la carcinogénèse avec un régime riche en viande rouge en comparaison à un régime riche en viande blanche.



## Chapitre II : Développement et optimisation de la méthode d'analyse

### A. Objectifs

L'objectif des travaux réalisés dans cette partie est de contribuer à l'optimisation de la méthode combinée d'analyse non ciblée du métabolome urinaire et du screening des conjugués mercapturiques par l'utilisation du mode DIA (désigné par MS<sup>E</sup> sur l'instrument Waters utilisé au laboratoire). Le développement initial de la méthode et une preuve de concept ont fait l'objet d'une publication disponible ci-après et dont la stratégie analytique est présentée en Figure 16. Ce chapitre regroupe l'ensemble des optimisations complémentaires réalisées afin d'évaluer et améliorer les performances de cette méthode décrite dans la publication. Les développements présentés dans cette partie concernent la plupart des étapes du *workflow* présenté en Figure 16 et ont pour but (i) d'optimiser les paramètres de transfert des ions de la source d'ionisation vers le spectromètre de masse afin de maximiser la détection des mercapturates, (ii) de déterminer le facteur de dilution et la méthode de normalisation à appliquer aux échantillons urinaires, et (iii) d'évaluer les performances de l'approche MS<sup>E</sup> en terme de reproductibilité et de justesse et d'optimiser les paramètres de traitement des données MS<sup>E</sup> pour le screening des conjugués mercapturiques.

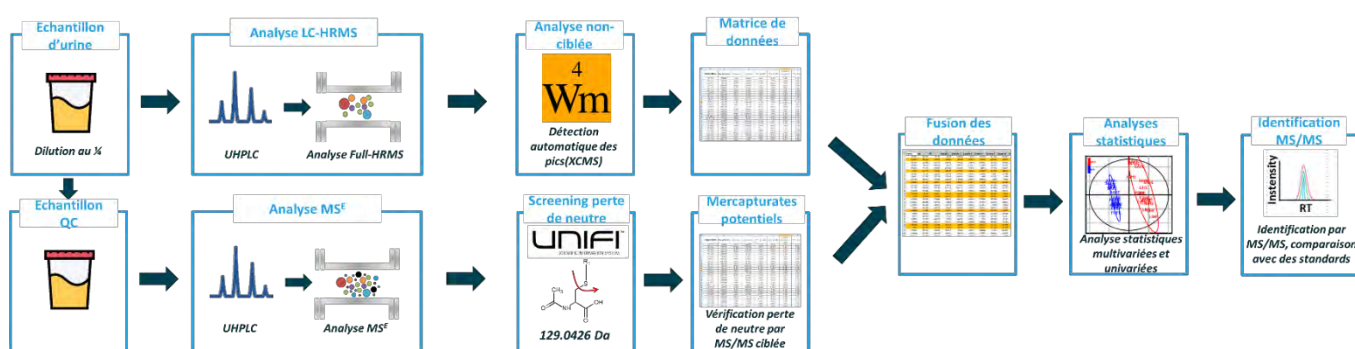


Figure 16 Schéma général de la stratégie analytique utilisée pour l'analyse du métabolome urinaire et du screening des mercapturates par MS<sup>E</sup>

### A.1. Introduction à l'article

#### A.1.1. Contexte

Au travers de leur mode de vie ou de leur alimentation, les Hommes sont exposés à un grand nombre de composés en quantités variables et aux propriétés diverses. Bien que la plupart de ces composés ne montrent pas d'effets délétères pour la santé, l'exposition à certains composés tels que des pesticides ou autres polluants est largement étudiée afin d'évaluer leurs possibles effets sur la santé. Parmi les composés présentant des propriétés toxiques, certains sont réactifs (sous forme native ou après bio-activation) vis-à-vis des macromolécules telles que les protéines ou l'ADN en raison de leur caractère électrophile. Comme mentionné dans l'introduction de ce manuscrit, la majorité des

composés génotoxiques sont des composés électrophiles. Ainsi, la caractérisation globale de l'exposition aux composés électrophiles *via* la mise en place d'approches non-ciblées s'avère particulièrement pertinente d'un point de vue toxicologique. Ainsi, dans le cadre d'une démarche de caractérisation de l'exposome dans laquelle est engagé le laboratoire, cette approche vise à se rapprocher de la notion de toxome en s'intéressant plus particulièrement à la partie réactive du xéno-métabolome.

Comme détaillé dans l'introduction de ce manuscrit, l'association entre la consommation de viande rouge et de charcuterie et le développement de pathologies telles que le CCR, serait médiée en particulier par les produits de peroxydation lipidique hème-induite. Cependant, l'exposition à ces produits est généralement limitée au dosage de marqueurs connus et bien identifiés comme le 4-HNE ou le DHN-MA, sous-estimant probablement la quantité et la diversité des composés formés. Le développement d'une méthode non-ciblée, orientée vers ces produits, permettant d'évaluer l'exposition aux composés électrophiles permettrait (i) de répondre à un besoin de biosurveillance et (ii) dans le cas de l'association viande rouge / CCR de mieux caractériser les effets du fer héminique et les mécanismes sous-jacents.

#### ***A.1.2. Objectif de l'étude***

L'une des voies de détoxification majeure des composés électrophiles étant celle des acides mercapturiques (voir Chapitre I : Introduction générale), l'objectif principal de ces travaux est de développer une méthode d'analyse non-ciblée de ces composés étant, *a priori*, le reflet direct d'une exposition aux molécules réactives. Plus précisément, il s'agit de développer une méthode d'analyse par spectrométrie de masse basée sur la fragmentation caractéristique des conjugués acides mercapturiques conduisant à une perte de neutre de 129,0426 Da. Notre raisonnement est donc de mettre en place une approche DIA basée sur une détection de cette perte de neutre spécifique, qui permettrait la caractérisation globale des conjugués mercapturiques en tant que marqueurs d'exposition.

#### ***A.1.3. Méthodes***

Pour répondre à l'objectif de l'étude et apporter une preuve de concept de l'approche, des échantillons urinaires de rats soumis ou non à un régime favorisant la peroxydation lipidique ont été utilisés. Plus précisément, les urines de 3 groupes de 6 rats nourris avec des régimes riches en hème avec soit des acides gras polyinsaturés riches en  $\omega$ -3, soit des acides gras polyinsaturés riches en  $\omega$ -6, soit des acides gras saturés (donc pas sensibles à l'oxydation) ont été utilisées pour cette étude. Les échantillons ont été analysés selon le protocole présenté en Figure 16. L'ensemble des échantillons ont été analysés en mode Full-MS et les QC de groupe en mode MS<sup>E</sup> afin d'accéder à la fragmentation

de l'ensemble des composés présents dans les urines et filtrer les fragmentations spécifiques des conjugués acides mercapturiques.

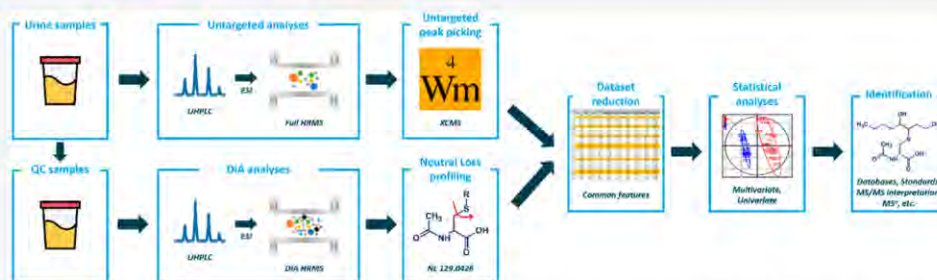
## Global Profiling of Toxicologically Relevant Metabolites in Urine: Case Study of Reactive Aldehydes

Emilien L. Jamin,<sup>\*,†,‡,§</sup> Robin Costantino,<sup>†,‡,§</sup> Loïc Mervant,<sup>†,‡</sup> Jean-François Martin,<sup>†,‡</sup> Isabelle Jouanin,<sup>†,‡</sup> Florence Blas-Y-Estrada,<sup>‡</sup> Françoise Guéraud,<sup>§,‡</sup> and Laurent Debrauwer<sup>§,†,‡</sup>

<sup>†</sup>Metatoul-AXIOM Platform, National Infrastructure for Metabolomics and Fluxomics: MetaboHUB, Toxalim, INRAE, 31300 Toulouse, France

<sup>‡</sup>Toxalim (Research Centre in Food Toxicology), Toulouse University, INRAE UMR 1331, ENVT, INP-Purpan, UPS, Toulouse, 31300 France

### Supporting Information



**ABSTRACT:** Among the numerous unknown metabolites representative of our exposure, focusing on toxic compounds should provide more relevant data to link exposure and health. For that purpose, we developed and applied a global method using data independent acquisition (DIA) in mass spectrometry to profile specifically electrophilic compounds originating metabolites. These compounds are most of the time toxic, due to their chemical reactivity toward nucleophilic sites present in biomacromolecules. The main line of cellular defense against these electrophilic molecules is conjugation to glutathione, then metabolization into mercapturic acid conjugates (MACs). Interestingly, MACs display a characteristic neutral loss in MS/MS experiments that makes it possible to detect all the metabolites displaying this characteristic loss, thanks to the DIA mode, and therefore to highlight the corresponding reactive metabolites. As a proof of concept, our workflow was applied to the toxicological issue of the oxidation of dietary polyunsaturated fatty acids, leading in particular to the formation of toxic alkenals, which lead to MACs upon glutathione conjugation and metabolization. By this way, dozens of MACs were detected and identified. Interestingly, multivariate statistical analyses carried out only on extracted HRMS signals of MACs yield a better characterization of the studied groups compared to results obtained from a classic untargeted metabolomics approach.

Throughout their lives, humans are exposed to many substances, natural or not, in varying amounts depending on their environment, lifestyle, or diet. Through diet, people are exposed to a wide variety of compounds, mostly essential (e.g., nutrients, vitamins, etc.), but also, to a much lesser extent, to compounds with possible harmful effects such as pesticides,<sup>1</sup> mycotoxins,<sup>2</sup> or neo-formed compounds produced during the processing, cooking, and digestion of food.<sup>3</sup> Since the concept of exposomics has been introduced,<sup>4</sup> new approaches aim at assessing this exposure in the most comprehensive way by the implementation of untargeted methods.<sup>5</sup> However, in front of the number and complexity of these contaminants representing our chemical exposome, focusing on toxic compounds appears essential to estimate the impact of this exposure on health. Many of these toxicologically relevant compounds are toxic due to their electrophilicity, which makes them reactive toward nucleophilic sites of biomolecules such as DNA, RNA,

or proteins, and lead to the formation of various adducts with potential harmful effects. Structure–activity relationship analyses carried out by Ashby and Tennant,<sup>6</sup> on the basis of carcinogenesis data from US-NTP, showed that more than 80% of genotoxic agents are electrophilic compounds or compounds that are able to be bioactivated into electrophilic metabolites. In particular, meat represents a complex matrix that may contain a variety of contaminants. Some of them are involved in colon cancer development according to a commonly accepted hypothesis involving the formation of DNA adducts.<sup>7,8</sup>

In terms of food safety, the cooking of meat at high temperature induces the formation of numerous electrophilic

Received: July 11, 2019

Accepted: December 19, 2019

Published: December 19, 2019

compounds from heterocyclic amines (HCAs) and polycyclic aromatic hydrocarbons (PAHs), such as 2-amino-3,8-dimethylimidazo-[4,5-f]-quinoxaline (MeIQx) or benzo[a]pyrene (BaP), respectively, which are known to be pro-mutagenic by giving rise to DNA adducts after the formation of reactive metabolites.<sup>9</sup> Alternatively, curing processes lead to the formation of *N*-nitroso compounds (NOCs).<sup>10</sup> Some of these NOCs are known to be mutagenic because of their ability to form DNA adducts.<sup>11</sup> Finally, several studies showed convincing evidence that heme iron contained in red meat could be one of the main factors in carcinogenesis promotion.<sup>12–14</sup> The main hypothesis to explain heme iron toxicity relies on its ability to promote the formation of radical species by Fenton-like reactions, which leads to the peroxidation of dietary polyunsaturated fatty acids into secondary lipid peroxidation electrophilic products such as 4-hydroxynonenal (4-HNE).<sup>15</sup> As reported by Sesink et al.<sup>16</sup> and further confirmed by Bastide et al.,<sup>17</sup> heme iron as such does not participate in carcinogenesis promotion, which is mainly mediated through lipid peroxidation product formation in the colon lumen.

Targeted methods are commonly used to monitor these reactive metabolites, able to covalently bind to biomolecules.<sup>18</sup> Concerning lipid peroxidation, 4-HNE has been one of the most studied lipid peroxidation secondary products.<sup>19–21</sup> However, depending on the fatty acid precursor, lipid peroxidation may induce the formation of many other reactive compounds (potentially unknown) with various abundances and reactivities. Developing untargeted methods is then necessary to get a more comprehensive view of their formation.

One of the main detoxification routes employed to detoxify such reactive/electrophilic compounds is enzymatic conjugation with glutathione, a tripeptide known as an efficient scavenger of electrophilic species. Those glutathione conjugates are then cleaved by  $\gamma$ -glutamyltransferase and cysteinylglycine dipeptidase and finally *N*-acetylated to form mercapturic acid conjugates (MACs) that are mostly urine-excreted.<sup>22</sup> Thus, MACs are the result of detoxification processes of reactive electrophilic species and can be considered as biomarkers of exposure to these harmful compounds. Monitoring MACs could therefore be useful for a more accurate characterization of exposure to reactive compounds and a better evaluation of the role of reactive metabolites in toxicology in general.<sup>23</sup>

Despite the interest this approach, only a few studies have focused on the global analysis of MACs.<sup>24–26</sup> These studies aimed at evaluating the electrophilic burden of the organism by screening mercapturates in an untargeted way by liquid chromatography coupled with mass spectrometry (LC-MS) using constant neutral loss-like acquisition mode. Other available studies of MACs were based on MS targeted analysis to study specific exposure situations like atrazine<sup>27</sup> or neofenol products in food such as acrylamide or acrolein.<sup>28</sup> All these works use the characteristic loss of 129 Da observed from MACs in urine samples by the cleavage of the thioether bond and rely on low-resolution mass spectrometry.

In the present study, we present an original method for the global screening of MACs. A high-resolution mass spectrometry (HRMS) based workflow was developed using a data-independent acquisition mode (DIA) for monitoring the specific neutral loss of mercapturates through gas phase fragmentation in mass spectrometry. As a proof of concept, this paper presents results from a study based on an animal

experiment conducted with different diets containing various oils (fish oil, safflower oil, and hydrogenated coconut oil) with different polyunsaturated fatty acid profiles, in combination with heme iron to catalyze lipid peroxidation. Thus, monitoring urinary MACs in these animals provided an accurate overview of their exposure to lipid peroxidation products conjugated with glutathione.

## EXPERIMENTAL SECTION

**Chemicals.** All solvents were purchased from Fisher Scientific (Thermo Fisher Scientific, Illkirch, France) and were LC-MS grade. Standard 4-hydroxynonan-1-ol mercapturic acid (DHN-MA), 4-hydroxyoctan-1-ol-MA, 4-hydroxyheptan-1-ol-MA, 4-hydroxyhexan-1-ol-MA (DHH-MA), 4-hydroxynonan-1-ol-MA, 4-hydroxyoctan-1-ol-MA, and 4-hydroxyheptan-1-ol-MA were synthesized in the laboratory as previously described.<sup>29</sup> Standard *N*-acetyl-S-phenylcysteine was purchased from Sigma-Aldrich (Saint Quentin Fallavier, France), and the MS calibration mix solution was purchased from Waters (Manchester, UK). Heme iron (Hemin from bovine) was purchased from Sigma-Aldrich, and fish oil (Refined Menhaden oil), safflower oil, and hydrogenated coconut oil came from MP Biomedicals (Illkirch-Graffenstaden, France).

**Animals and Diets.** Young adult Fischer male rats (six rats/group) were housed individually in plastic metabolic cages. They were allowed 3 days of acclimatization to their cage. The room was kept at a temperature of 22 °C on a 12 h light–dark cycle. Animals had free access to tap water and to their respective diet for 15 days. Diets were given each day at the end of the afternoon in order to limit oxidation, and 24 h urine samples were collected at day 12 and frozen (–20 °C) until analysis. A first set of diets contained 5% fish oil (containing long chain  $\omega$ -3 fatty acids). A second set of diets contained 5% safflower oil (containing  $\omega$ -6 fatty acids). A final set of control diets contained 5% hydrogenated coconut oil (containing almost only saturated fatty acids). All diets contained heme iron (0.094% w/w) to catalyze lipid peroxidation. All animal experiments were performed in accordance with EU directive 2010/63/EU and approved by the local Animal Care and Use Committee of Toulouse Midi-Pyrénées (agreement TOXCOM/0006/FG).

**Sample Preparation.** Urine samples were allowed to thaw at room temperature. Aliquots of 500  $\mu$ L were prepared for each urine sample and diluted in 500  $\mu$ L of mobile phase A (95% H<sub>2</sub>O, 5% CH<sub>3</sub>OH, 0.1% CH<sub>3</sub>CO<sub>2</sub>H). After dilution, the aliquots were vortex-mixed and centrifuged at 10 000 rpm for 5 min. The supernatants were transferred into vials to be directly used for analyses. This sample preparation protocol was previously developed for mass spectrometric detection of urinary metabolites of xenobiotics, including conjugated metabolites.<sup>30</sup>

**Liquid Chromatography and Mass Spectrometry Analyses.** Analyses were performed using ultra-performance liquid chromatography (UPLC; ACQUITY, Waters Manchester, UK) coupled to a quadrupole time-of-flight mass spectrometer (Q-ToF Synapt G2-Si, Waters, Manchester, UK). A volume of 10  $\mu$ L of samples was injected into a Hypersil Gold C18 (1.9  $\mu$ m, 100  $\times$  2.1 mm) analytical column (Thermo Fisher Scientific, Illkirch, France) at a flow rate of 0.3 mL/min and maintained at 40 °C. According to a previously developed chromatographic protocol,<sup>30</sup> a linear gradient program was set up with mobile phase A, 95% H<sub>2</sub>O/5% methanol/0.1% acetic acid, and mobile phase B, 100%

Table 1. UHPLC-ESI-HRMS/MS of MAC Standards, Achieved at a Collision Energy of 15 eV

MAC	<i>m/z</i>	Rt (min)	neutral formula	fragment ions (relative abundance)				
4-hydroxyheptanal-MA	290.1067	7.74	C <sub>12</sub> H <sub>21</sub> NO <sub>3</sub> S	290.1061 (20)	161.0637 (53)	143.0532 (100)	125.0426 (3)	115.0579 (25)
4-hydroxyoctanal-MA	304.1223	10.67	C <sub>13</sub> H <sub>23</sub> NO <sub>3</sub> S	304.1219 (23)	175.0792 (68)	157.069 (100)	129.0739 (27)	
4-hydroxynonanal-MA (HNE-MA)	318.1380	13.18	C <sub>14</sub> H <sub>25</sub> NO <sub>3</sub> S	318.1383 (16)	189.0946 (25)	171.0839 (100)	143.0880 (17)	
4-hydroxyheptanol-MA	292.1223	7.35	C <sub>12</sub> H <sub>23</sub> NO <sub>3</sub> S	292.1219 (30)	163.0790 (100)	145.0490 (34)	130.963 (10)	115.0583 (39)
4-hydroxyoctanol-MA	306.1380	10.27	C <sub>13</sub> H <sub>25</sub> NO <sub>3</sub> S	306.1373 (27)	177.0949 (100)	159.0847 (23)	129.0741 (32)	
4-hydroxyhexanol-MA (DHH-MA)	278.1068	4.47	C <sub>11</sub> H <sub>21</sub> NO <sub>3</sub> S	278.1062 (20)	149.0637 (100)	131.0530 (38)	128.0346 (3)	101.0424 (42)
4-hydroxynonanol-MA (DHN-MA)	320.15374	13.21	C <sub>14</sub> H <sub>27</sub> NO <sub>3</sub> S	320.1536 (4)	191.1106 (100)	173.1000 (16)	143.0898 (44)	128.0352 (2)
<i>N</i> -acetylphenylcysteine	238.0543	9.92	C <sub>11</sub> H <sub>13</sub> NO <sub>3</sub> S	109.0112 (100)				

methanol/0.1% acetic acid. Initial conditions were 100% of A, followed by a linear gradient from 0 to 100% of B in 30 min. These conditions were held for 4 min prior to switching in 1 min to the starting conditions and held for 5 min to equilibrate the column. MS<sup>E</sup> acquisitions (*i.e.*, DIA mode) were achieved with electrospray ionization (ESI) in the negative mode. Parameters were set up using two standard compounds (namely DHN-MA and *N*-acetyl-S-phenylcysteine) as detailed in the Experimental Section: capillary voltage 0.5 kV, sampling cone voltage 30 V, source temperature 120 °C, desolvation temperature 550 °C, cone gas flow rate 30 L/h, desolvation gas flow rate 600 L/h, and a trap collision energy of 15 eV for the high energy spectra. MS<sup>E</sup> spectra were acquired with a range of mass-to-charge ratio (*m/z*) set to 50–800. MS<sup>3</sup> experiments were achieved on a LTQ Orbitrap XL mass spectrometer (Thermo Scientific, Les Ulis, France).

A quality control (QC) sample corresponding to the pool of all samples was analyzed repetitively (*n* = 7) along the injection sequence to ensure that no instrumental deviation was observed during the analysis. The quality of analyses was checked by monitoring the variation of the chromatographic pressure (lower than 10 bar between the first and the last injection), the variation of MS signal (Relative Standard Deviation (RSD) of signal area of 8.6%), retention time (RSD of 0.15%), and mass measurement accuracy (lower than 5 ppm) based on the tryptophan signal detected at 3.2 min.

**Data Processing.** Data from MS<sup>E</sup> experiments were processed with the UNIFI software (Waters, Manchester, UK). Extraction of data from full scan mass spectra was carried out using XCMS offline software<sup>31</sup> with the centwave algorithm<sup>32</sup> (ppm = 10, peakwidth = (10,70), snthresh = 10, noise = 2000, bw = 5, mzwid = 0.01). To avoid urine dilution variation among samples, data were normalized using probabilistic quotient normalization.<sup>33</sup> All data analysis steps were performed using workflow4metabolomics, a collaborative portal dedicated to metabolomics data processing.<sup>34</sup>

Statistical analyses were carried out using SIMCA software v15 (Umetrics, Umeå, Sweden). Results were examined using unsupervised principal component analysis (PCA) in order to check the validity of acquisition and to detect potential outliers. Projection to latent structure discriminant analysis modeling (PLS-DA) was then carried out to find diet discriminant features. The PLS-DA *Q*<sup>2</sup> parameter (fraction of samples correctly predicted by the model) was used to assess the quality of models. Models with *Q*<sup>2</sup> > 0.4 were considered valid. For valid models, a permutation test was performed to assess the robustness of models. For valid and robust models, significant features were selected by their variable importance on projection (VIP). A feature displaying a VIP > 1 was considered significant.

## RESULTS AND DISCUSSION

**Method Development.** To set up ionization parameters for the profiling of MACs, two standards were used. The first one corresponds to the reduced conjugate between *N*-acetylphenylcysteine and 4-HNE, namely, DHN-MA. Since this compound is generated from lipid peroxidation, it is representative of structures we would like to detect. The second standard corresponds to a smaller structure, which displays no heteroatom other than those of the *N*-acetylphenylcysteine moiety, namely, *N*-acetyl-S-phenylcysteine, the conjugate between a phenyl group and *N*-acetylphenylcysteine. This compound could be considered as representative of smaller and less ionizable compounds.

UPLC coupled with HRMS through ESI is an analytical approach particularly suited for the qualitative and quantitative analysis of urinary metabolites.<sup>5</sup> ESI in the negative mode gave the most intense signal, due to the carboxylic acid function of the mercapturate that easily undergoes deprotonation. Fragment ions were observed in the full MS spectrum of standards even with no collision energy applied into the collision cell of the mass spectrometer. The [M-H]<sup>-</sup> ion of DHN-MA was detected at *m/z* 320.1531, but another intense ion was observed at *m/z* 191.1111, corresponding to the loss of 129 Da, representative of the *N*-acetylphenylcysteine moiety. The same phenomenon was observed with *N*-acetyl-S-phenylcysteine. This fragmentation was caused by an excess of internal energy of ions coming from the desolvation conditions of the electrospray ion source and the conditions of transfer of the ions into the Triwave and StepWave cells of the Synapt mass spectrometer. Nevertheless, this fragmentation has to be limited to increase the signal of the deprotonated ions of MACs, and therefore to increase the sensitivity of the method. For that purpose, we decreased the excess of internal energy during the steps of ionization/desolvation and ion transfer by limiting the acceleration of ions into the different cells of the Synapt mass spectrometer where remaining gas was present, *i.e.*, into the StepWave cell for the last step of desolvation and into the Triwave cell for the transfer of ions.

We based our profiling method on the property of MACs to undergo a particular fragmentation during the collision induced dissociation (CID) process, leading to a neutral loss of C<sub>3</sub>H<sub>7</sub>NO<sub>3</sub> = 129.0426 Da, corresponding to the cleavage of the thioether bond as already mentioned in previous works.<sup>24–28</sup> In this study, this fragmentation pattern was investigated from seven MAC standard compounds originating from lipid peroxidation products conjugated with glutathione, namely, 4-hydroxynonan-1-ol mercapturic acid (DHN-MA), 4-hydroxyoctan-1-ol-MA, 4-hydroxyheptan-1-ol-MA, 4-hydroxyhexan-1-ol-MA (DHH-MA), 4-hydroxynonanal-MA, 4-hydroxyoctanal-MA, and 4-hydroxyheptanal-MA, as well as the standard of *N*-acetyl-S-phenylcysteine. Results summarized in

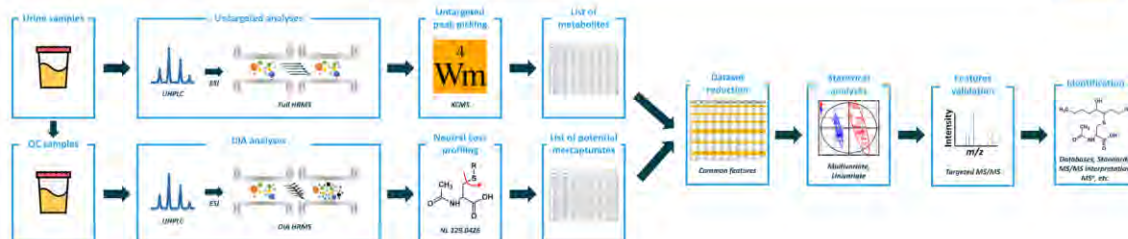


Figure 1. Workflow of the untargeted profiling of mercapturic acid conjugates.

Table 1 show that all MACs displayed the neutral loss of  $C_5H_7NO_3 = 129.0426$  Da. As an example, in the case of DHN-MA detected at a retention time (Rt) of 13.2 min, the neutral loss of 129.0426 Da was observed with a resulting fragment ion at  $m/z$  191.1106. Consecutive losses of  $H_2O$  and  $CH_3O$  led to the formation of  $m/z$  173.1000 and 143.0898 ions. The complementary fragment ion  $[C_5H_7NO_3-H]^-$  detected at  $m/z$  128.0352 was also previously observed for three metabolites of B(a)P<sup>35</sup> and used as a multiple reaction monitoring transition (MRM) for detecting some MACs.<sup>36</sup> However, this ion was observed only in the MS/MS spectra of DHH-MA and DHN-MA and with a weak relative abundance (<3%).

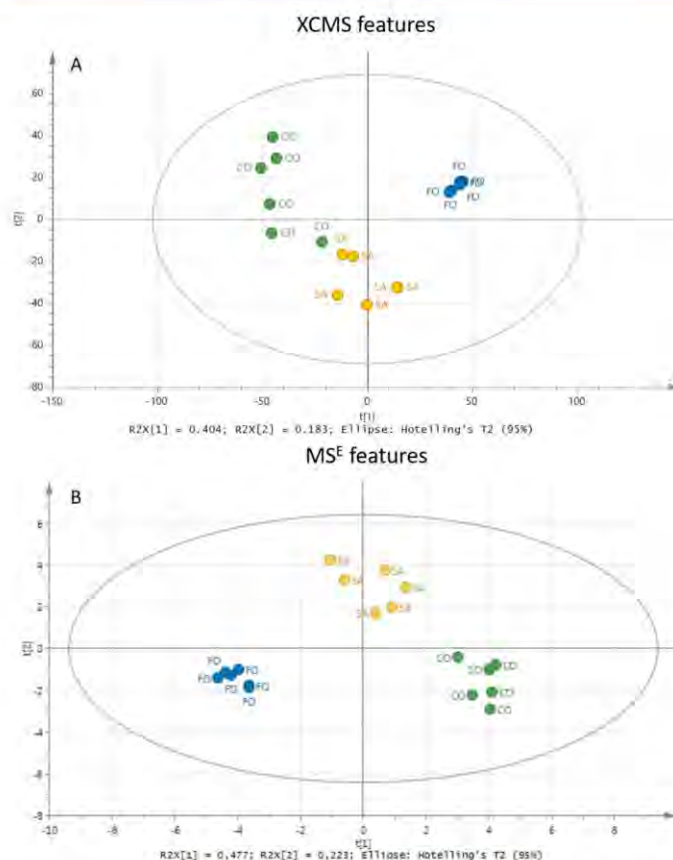
We applied a data-independent acquisition (DIA) process allowing the fragmentation of all ions. In the Waters Q-ToF instruments, the DIA mode is called MS<sup>E</sup> and is based on the continuous switching between a low-energy CID scan and a higher energy CID scan without any precursor ion selection. The first scan is used to detect the precursor ion, whereas the second scan is used to detect the fragment ions.<sup>37</sup> Data are then processed using the UNIFI software to associate the two scans and to build product ion spectra from the collected data. Within these data, the software also offers the possibility to screen every ion giving a particular neutral loss. The neutral loss tolerance was set at 5 mDa as it was the value which gave the most notable results without being too large. The signal threshold of the low energy scan was set at 2000 counts to minimize the signals mixed to the background noise, whereas the signal threshold of the high energy scan was set at 20 counts. At the end, only candidates displaying the neutral loss of 129.0426 Da detected with a mass accuracy below 10 ppm were kept for subsequent processing.

The limit of detection (LOD) of the method was evaluated by analyzing successive dilutions of DHN-MA and *N*-acetyl-S-phenylcysteine standard solutions prepared in blank urine samples. A signal ratio of 3 was observed at concentrations of 5 ng/mL and 10 ng/mL for DHN-MA and *N*-acetyl-S-phenylcysteine, respectively. Although the dynamic range of the detector is limited when working in the profile mode, it can be increased by using the dynamic range extension (DRE) in the centroid mode. Since the centroid mode is not compatible with UNIFI, a parallel untargeted acquisition of data in the centroid mode should be performed for quantitative purposes. Thus, according to the workflow displayed in Figure 1, all samples were analyzed by UPLC-ESI-HRMS, and raw data were processed using XCMS.<sup>31</sup> In parallel and using exactly the same chromatographic and ionization parameters, QC samples were also analyzed by UPLC-ESI-HRMS<sup>E</sup> to detect MACs. Features of MACs were then annotated into the XCMS data set to allow their statistical analyses using unsupervised or supervised statistical tools.

In DIA mode, no precursor ion is selected, and fragmentation patterns are built *in silico*. Thus, false positive MACs might be detected by the profiling of the considered neutral loss. A first possible origin of artifactual detection could be chromatography. However, since features detected by DIA are matched to features detected by XCMS, only features displaying chromatographic characteristics in agreement with parameters used for XCMS (*i.e.*, signal-to-noise ratio, peak width, peak resolution) are kept. The second origin of potential false positive MACs could be the nonspecificity of the fragmentation (*i.e.*, neutral loss of  $C_5H_7NO_3$  129.0426 Da not specific of MACs), or if the deconvolution algorithm of DIA is not robust enough to align the right fragment ion with the right precursor ion. However, since next steps of the method aim at highlighting discriminant features with statistical tools, and then identifying them with targeted MS/MS, false positive MACs will be eliminated at this last step.

**Application.** This method was validated on a proof-of-concept application on urine samples coming from rats fed diets containing representative lipid peroxidation prone fatty acids expected to yield various glutathione conjugates, further metabolized into MACs. Thus, the global profiling method of MACs should be more adapted to detect MACs by comparison with an untargeted metabolomics approach.

Since lipid peroxidation products depend on their precursor unsaturated fatty acid, three different diets were selected. The first one was based on hydrogenated coconut oil known to contain almost only saturated fatty acids and which should produce no MACs (except those possibly observed coming from endogenous peroxidation). The second one was based on fish oil containing long chain omega-3 fatty acids, and finally the third one was based on safflower oil, which is rich in omega-6 fatty acids. Thus, these three contrasted diets were expected to lead to specific peroxidation products that are detoxified by conjugation with glutathione, and finally eliminated as specific MACs in urine. Urine samples coming from the different diets were randomly analyzed by UPLC-ESI-HRMS. Since the detection of standard MACs analyzed for the setup of the method was more sensitive in the negative ionization mode, samples were analyzed only with this mode. After data acquisition, features were detected using XCMS. Raw data were filtered according to a blank sample of mobile phase A, to keep only features whose signal intensity is at least three times higher in the QC samples than in the blank samples. A total of 9614 features were detected by this untargeted metabolomics approach (Supporting Information, Table S1). In parallel, the QC sample was also analyzed according to our designed MS<sup>E</sup> workflow. By MS<sup>E</sup>, eighty-two MACs candidates were detected. Twenty-six of them displayed poor chromatographic characteristics, including a low signal-



**Figure 2.** Principal component analyses of (A) all features detected with XCMS and (B) features detected by  $MS^E$ . Yellow dots (SA) represent samples from safflower oil diet. Blue dots (FO) represent samples from fish oil diet, and green dots (CO) represent coconut oil diet.

to-noise ratio, a large peak width or a very low chromatographic resolution. All these false positive MACs arising from chromatography were not present in the XCMS data matrix (Supporting Information, Table S2) illustrating the relevant data curation achieved by using a parallel processing of data using XCMS.

An unsupervised statistical analysis by PCA was achieved on all the features detected with XCMS in all samples and in the QC. No outlier sample was detected, and a good correlation of the injections of the QC sample demonstrated that no deviation has occurred along the analysis (Supporting Information, Figure S1). The resulting PCA shows in Figure 2A a discrimination of the fish oil diet (FO), but a weak separation of the coconut oil diet (CO) and safflower oil diet (SA). By comparison, the same PCA achieved on features filtered thanks to our  $MS^E$  approach obviously displays a better discrimination of diets (Figure 2B). Results match perfectly the expected effects of the lipid peroxidation induced by heme iron according to the composition of oil in polyunsaturated fatty acids.

To evaluate the contribution of the  $MS^E$  profiling approach to discriminate populations, a PLS-DA was achieved on both the data sets validated by permutation tests. Samples were discriminated according to a two component model ( $Q^2 = 0.94$ ) of 3944 features with a  $VIP > 1$  from the data set

generated by XCMS (Supporting Information, Figure S2). Conversely, a two component model ( $Q^2 = 0.95$ ) of 22 features with a  $VIP > 1$  was obtained from the data set generated by  $MS^E$  (Supporting Information, Figure S3). In the XCMS data set, discriminating MACs were ranked between the 21 and 4354 position of  $VIP$  in component 1 and between the 13 and 2929 position in component 2 (Supporting Information Table S3). This shows that the discriminating MACs were diluted into thousands of features representative of other metabolites using the whole metabolomics data set, making it very difficult to identify them. In contrast, the PLS-DA achieved on the  $MS^E$  data set allowed direct highlighting of discriminating MACs and detection of two more significant MACs. This better discrimination can be explained by the fact that the drastic data matrix reduction enables focusing the data set only on a class of metabolites of interest. The studied populations were expected to be discriminated according to the lipid peroxidation products, detoxified by conjugation with glutathione, and finally metabolized into MACs. Therefore, by looking only at MACs, which include lipid peroxidation product conjugates, the data matrix was filtered only on metabolites of interest, excluding other potential interfering signals.

Another interesting issue with the analysis of metabolites is the identification of discriminating compounds. Since a



Table 2. Annotated MACs of Lipid Peroxidation Products Detected by UPLC-ESI-HRMS<sup>E</sup> in Urine Samples of Rats

XCMS feature	<i>m/z</i>	Rt (min)	[M-H] <sup>-</sup> formula	ppm	annotation of MACs	annotation of reactive metabolite	identification level <sup>10</sup>
n248.0955T6.5	248.0955	6.40	C <sub>10</sub> H <sub>18</sub> NO <sub>4</sub> S	2.8	pentanol-MA	pentenol	3
n262.1113T9.5	262.1114	9.51	C <sub>11</sub> H <sub>20</sub> NO <sub>4</sub> S	1.6	hexanol-MA	hexenol	3
n266.0849T13.8	266.0847	13.78	C <sub>13</sub> H <sub>16</sub> NO <sub>3</sub> S	3.4	ethylbenzene-MA	styrene	3
n274.0746T6	274.0737	5.44	C <sub>11</sub> H <sub>16</sub> NO <sub>5</sub> S	6.6	ethylactone-MA	hydroxyhexenoic ac	3
n276.0905T10	276.0908	10.02	C <sub>11</sub> H <sub>16</sub> NO <sub>5</sub> S	1.0	4-hydroxyhexanal-MA (HHE-MA)	4-hydroxyhexenal (HHE)	2
n276.1271T12.5	276.1270	12.50	C <sub>12</sub> H <sub>22</sub> NO <sub>4</sub> S	1.8	heptanol-MA	heptenol	3
n278.1054T4.6	278.1059	4.52	C <sub>11</sub> H <sub>20</sub> NO <sub>5</sub> S	3.1	4-hydroxyhexanol-MA (DHH-MA)	4-hydroxyhexenol (DHH)	1
n278.1054T5.6	278.1062	4.64	C <sub>11</sub> H <sub>20</sub> NO <sub>5</sub> S	2.0	hydroxyhexanol-MA	hydroxyhexenol	3
n288.1269T13.2	288.1271	13.17	C <sub>13</sub> H <sub>22</sub> NO <sub>4</sub> S	1.5	octanal-MA	octenal	3
n288.1269T12.5	288.1269	12.81	C <sub>13</sub> H <sub>22</sub> NO <sub>4</sub> S	2.1	octanal-MA	octenal	3
n288.1271T11.5	288.1268	11.50	C <sub>13</sub> H <sub>22</sub> NO <sub>4</sub> S	2.3	octanal-MA	octenal	3
n290.106T6.2	290.1057	6.07	C <sub>12</sub> H <sub>20</sub> NO <sub>5</sub> S	3.9	hydroxyheptanal-MA	hydroxyheptenal	3
n290.1425T15.2	290.1425	15.22	C <sub>13</sub> H <sub>24</sub> NO <sub>4</sub> S	2.0	octanol-MA	octenol	3
n304.0855T4.2	304.0854	4.54	C <sub>12</sub> H <sub>18</sub> NO <sub>6</sub> S	2.0	hydroxypropylactone-MA	4,7-dihydroxyheptenoic ac	3
n306.1019T4.7	306.1014	5.12	C <sub>12</sub> H <sub>20</sub> NO <sub>6</sub> S	1.7	hydroxyheptanoic ac-MA	hydroxyheptenoic ac	3
n314.1433T14.4	314.1424	14.45	C <sub>15</sub> H <sub>24</sub> NO <sub>4</sub> S	2.1	decenal-MA	decadienal	3
n318.1374T10.6	318.1375	10.70	C <sub>14</sub> H <sub>24</sub> NO <sub>4</sub> S	1.9	hydroxynonenol-MA	hydroxynondienol	3
n320.1177T7.1	320.1170	7.20	C <sub>13</sub> H <sub>22</sub> NO <sub>6</sub> S	1.0	dihydroxyoctanal-MA	dihydroxyoctenal	3
n320.1529T12.9	320.1530	13.07	C <sub>14</sub> H <sub>26</sub> NO <sub>4</sub> S	2.3	4-hydroxynonanol-MA (DHN-MA)	4-hydroxynonenol (DHN)	1
n336.1115T3.6	336.1114	3.41	C <sub>13</sub> H <sub>22</sub> NO <sub>5</sub> S	2.3	dihydroxyoctanoic ac-MA	dihydroxyoctenoic ac	3
n362.1272T6.4	362.1274	6.34	C <sub>15</sub> H <sub>24</sub> NO <sub>7</sub> S	1.5	wCOOH-hydroxydecenal-MA	wCOOH-hydroxydecenal	3

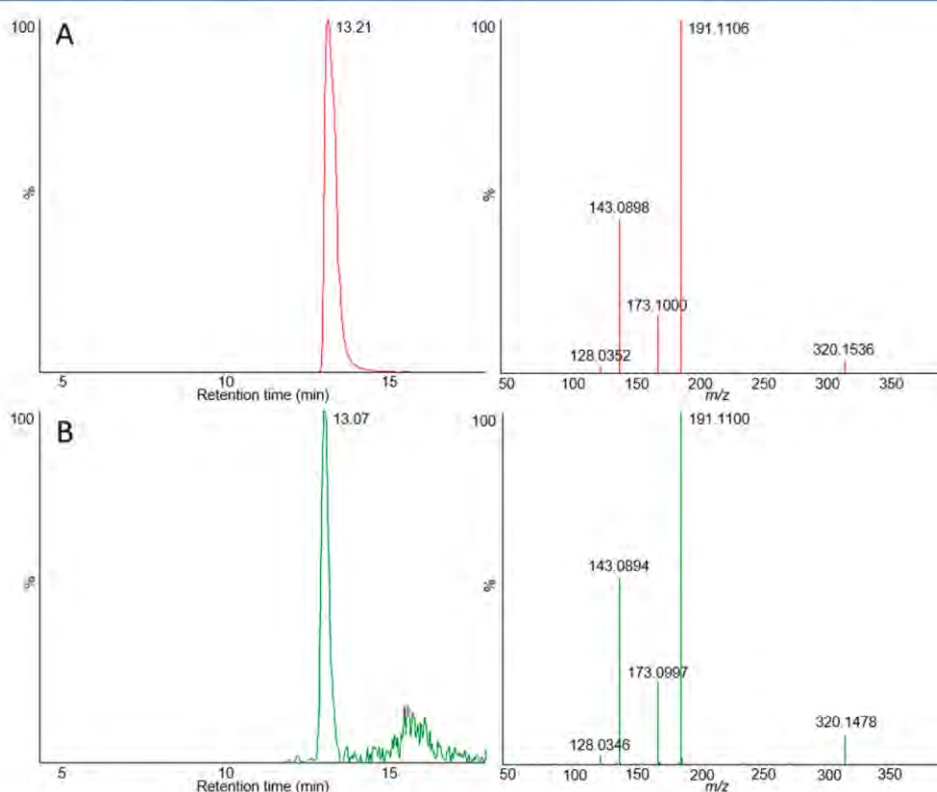


Figure 3. Extracted ion chromatogram and product ion spectrum at a collision energy of 15 eV of (A) the standard of DHN-MA and (B) DHN-MA detected in urine samples.

fragmentation pattern highly specific of MACs is monitored, this identification criterion can be proposed in a first attempt.

To validate the loss of 129.0425 as representative of MACs, targeted MS/MS analyses were attempted on the 56 MACs

candidates detected in the XCMS data matrix. Some MS/MS spectra could not be generated likely due to a loss of sensitivity related to the decrease of the scan time, which has to be set to allow numerous parallel MS/MS acquisitions, and/or the weak signal of corresponding precursor ions. In all other MS/MS spectra, the expected neutral loss was successfully observed, showing that no interference occurred during its detection by MS<sup>E</sup>. To assess the possible detection of a potential loss of 129.0425 from a metabolite that was not a MAC, the chemical formula of each candidate was generated. Only two candidates displayed a loss of 129.0425 from a chemical formula bearing no sulfur atom (n230.1032T4.9 and n261.0879T5.7). The list of features detected by MS<sup>E</sup> was also matched with the HMDB database. Some features matched with metabolites that are not a MAC. However, none of these other potential metabolites had MS/MS spectra in HMDB displaying a [(M - H) - 129]<sup>-</sup> fragment ion.

On the basis of this proposed identification of MACs, a further hypothesis of the reactive metabolite before its conjugation with glutathione can also be formulated (Table 2), with a particular focus on lipid peroxidation products in this study. A chemical formula of the precursor reactive metabolite can be hypothesized from the formula of the corresponding MAC, from which the *N*-acetylcysteine formula was subtracted. Then, putative lipid peroxidation products were matched with the LIPID MAPS database. In the case of unknowns or not well characterized lipid peroxidation products, annotation was performed from the bibliography of lipid peroxidation products.<sup>29,38,39</sup>

An interesting feature was detected at  $m/z$  320.1529 and  $R_t = 13.07$  min. This metabolite displayed an  $m/z$  ratio consistent with a [C<sub>14</sub>H<sub>27</sub>NO<sub>5</sub>S-H]<sup>-</sup> ion with a precision of 2.3 ppm. Four fragment ions were detected in its MS/MS spectrum (Figure 3B) at  $m/z$  191.1100, 173.0997, 143.0894, and 128.0346. Relative abundances of these ions were 100%, 23%, 52%, and 2%, respectively, which were highly similar to those of standard DHN-MA (Figure 3A, Table 1). These two independent and orthogonal data (i.e., retention time and MS/MS spectrum) comparable with a standard analyzed under identical experimental conditions allowed us to identify this metabolite as DHN-MA at level 1 of identification, in full concordance with the nomenclature of identification proposed by the metabolomics standards initiative (MSI).<sup>40</sup> Interestingly, the consecutive fragmentation of the [(M - H) - 129]<sup>-</sup> ion of DHN-MA produced the  $m/z$  143.0894 fragment ion by losses of H<sub>2</sub>O and CH<sub>2</sub>O. This loss of CH<sub>2</sub>O could be representative of the primary hydroxyl function of DHN-MA, which could give some information on the structure of other MACs.

Related to DHN-MA, the mercapturate conjugate coming from the glutathione conjugation of HNE (HNE-MA), before its metabolization into DHN-MA, could also be annotated (n318.1374T10.6 in Table 2). However, its targeted MS/MS spectrum (Supporting Information, Table S2) and its retention time did not match with synthesized standard of HNE-MA (Table 1). The consecutive fragmentation of the [(M - H) - 129]<sup>-</sup> ion of the standard produced the  $m/z$  143.0888 fragment ion by losses of H<sub>2</sub>O and CO. This loss of CO could be representative of the aldehyde function of HNE-MA. By comparison, the consecutive fragmentation of the [(M - H) - 129]<sup>-</sup> ion of the unknown n318.1374T10.6 produced the  $m/z$  141.0749 fragment ion by losses of H<sub>2</sub>O and CH<sub>2</sub>O. This observation, which is similar to DHN-MA, suggests that this

unknown MAC may include a primary hydroxyl function, another hydroxyl function, and an unsaturation, which can be a ring or a double bond.

Whereas DHN-MA is originating from  $\omega$ -6 polyunsaturated fatty acids,  $\omega$ -3 ones are known to lead to the formation of 4-hydroxyhexenal (HHE), which can be conjugated with glutathione to form HHE-MA.<sup>41</sup> HHE-MA can be metabolized similarly to HNE-MA to generate the 4-hydroxyhexanal-MA (DHH-MA), which was detected at  $m/z$  278.1054 and  $R_t = 4.52$  min. This latter identification was confirmed at level 1 by analysis of the corresponding standard (Supporting Information, Figure S4), which displayed the same retention time and fragmentation pattern. The mercapturate conjugate coming from the glutathione conjugation of HHE (HHE-MA), before its metabolization into DHH-MA, could also be annotated (n276.0905T10 in Table 2). On the basis of its MS/MS spectra (Supporting Information, Table S2) and the biological context, the HHE-MA was putatively identified (level 2, according to MSI<sup>40</sup>). The  $m/z$  290.1057 ion detected at  $R_t = 6.1$  min (Supporting Information, Table S2) was not attributed to 4-hydroxyheptanal-MA, based on the MS/MS spectrum and  $R_t$  of the corresponding standard (Table 1).

Some annotated MACs displayed a high degree of unsaturation, which can correspond to double bonds or rings. This latter possibility has already been described as lactone structures for lipid peroxidation products.<sup>39</sup> For example, the metabolite detected at  $m/z$  304.0854 and  $R_t = 4.5$  min corresponds to a conjugate formed between *N*-acetylcysteine and a structure displaying seven carbon atoms, three oxygen atoms, and two double bond equivalents. Since it is likely that this metabolite is a product of lipid peroxidation, one unsaturation may be involved in a carbonyl function at the carbon in position one. Then, the second unsaturation can be either a carbon-carbon double bond or another carbonyl function. It is also possible that a dihydroxylated acid alken, hydroxylated in particular at position 4, can be conjugated as a dihydroxylated heptanoic acid mercapturate, which could further undergo a condensation to generate a lactone compound, like the hydroxypropylactone, according to previously described metabolic processes for HNE.<sup>38,39</sup>

As observed on the MS/MS spectra of standards (Table 1), MACs display poor fragmentation rates, which limits available information for unknowns identification. Neutral loss profiling was used to give a first step of identification (i.e., MAC), and then statistics on particular populations highlighted some particular metabolites to have another hypothesis of identification (i.e., lipid peroxidation products). On the basis of these hypotheses and since standard MACs are not commercially available, seven chemical syntheses were achieved, among which five did not validate corresponding putative annotations. Some consecutive fragmentations could occur from the [(M - H) - 129]<sup>-</sup> fragment ion, but the weak signal of the precursor ion prevented their detection in many cases (Supporting Information, Table S3). The use of MS<sup>3</sup> experiments could be an alternative solution to get more structural information. However, the analysis of the samples on an LTQ Orbitrap XL mass spectrometer led to the detection of only DHH-MA, hydroxynonenol-MA, DHN-MA, dihydroxyoctanal-MA, and dihydroxyoctanoic ac-MA, and the corresponding MS<sup>3</sup> spectra displayed only a loss of water.

## CONCLUSIONS

Urinary MACs are originated from detoxification conjugates between glutathione and reactive electrophilic compounds that are able to bind to biomolecules, such as DNA and proteins, and to induce deleterious effects. These compounds can be electrophilic as such or be the result of the bioactivation of xenobiotics to which we are exposed. Therefore, by monitoring MACs in urine, it is possible to monitor their precursor electrophilic metabolites and, by extension, the molecules from which these metabolites are produced and which can be considered as toxicologically relevant compounds. The global profiling of MACs combined with the detection of known and unknown MACs by using MS<sup>E</sup>, and their relative quantification, pave the way for addressing new perspectives in toxicology. The application of this method to the study of lipid peroxidation catalyzed by heme iron allowed the highlighting of various unknown structures of reactive lipid peroxidation products, which may be implicated in the promotion of colorectal cancer, similarly to 4-HNE.<sup>42</sup> This detection of unknown potentially reactive metabolites was greatly facilitated by the simplification of statistical analyses due to data set reduction provided by the specific monitoring of metabolites of interest using the MS<sup>E</sup> profiling. Although our workflow suffers from some complexity because of the difficulty in obtaining a large amount of quantitative data with the software we used for DIA processing, this approach could be adapted to other mass spectrometers, DIA approaches, or fast data dependent acquisitions.

This case study was only focused on the study of MACs to demonstrate the proof of concept. However, the approach could be applied to monitor other classes of metabolites and markers, such as conjugates with glucuronic acid, sulfate, cysteine, and glutathione, as well as DNA and RNA adducts as recently published,<sup>43</sup> which all display characteristic neutral losses.

Furthermore, it is noteworthy that this method provides not only the profiling of MACs by MS<sup>E</sup> but also in parallel an untargeted analysis of the metabolome by HRMS, in the same samples. Indeed, an untargeted metabolomics analysis by reversed phase liquid chromatography coupled with HRMS using electrospray ionization is able to detect amino acids, organic acids, polyols, nucleosides, indoles, etc.,<sup>44</sup> which can give access to the metabolome modifications induced by the studied diets. Therefore, this method allows the characterization of a series of toxicologically relevant compounds, on one hand, as well as the study of their effects on the metabolome, on the other hand, in the same analysis of samples. This may represent a valuable way to link exposure to toxics and possible health adverse outcomes.

## ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.analchem.9b03146>.

Figure S1: Principal component analysis of all features detected with XCMS. Figure S2: Projection to latent structure discriminant analysis modeling (PLS-DA) of all features detected with XCMS. Figure S3: Projection to latent structure discriminant analysis modeling (PLS-DA) of features detected by MSE. Figure S4: Extracted ion chromatograms of DHH-MA obtained by UPLC-ESI-HRMS from (A) the standard of DHH-MA and (B)

urine samples. Corresponding product ion spectra obtained by electrospray in the negative mode from (C) the standard of DHH-MA and (D) urine samples (PDF)

Table S1: Data matrix obtained from UPLC-ESI-HRMS using XCMS (XLSX)

Table S2: Candidates of MACs detected in the XCMS matrix and validated by targeted MS/MS (XLSX)

Table S3: VIP of the PLS-DA of the XCMS data set and the MS<sup>E</sup> data set (XLSX)

## AUTHOR INFORMATION

### Corresponding Author

\*E-mail: [emilien.jamin@inra.fr](mailto:emilien.jamin@inra.fr).

### ORCID

Emilien L. Jamin: 0000-0002-4568-9177

### Author Contributions

<sup>§</sup>Equal contribution

### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

The Ph.D. doctoral fellowship of Loïc Mervant was cofunded by INRAE (Nutrition, Chemical Food Safety and Consumer Behaviour Scientific Division) and French region Occitanie. All MS experiments were performed on the instruments of the MetaToul-AXIOM platform, partner of the national infrastructure of metabolomics and fluxomics: MetaboHUB [MetaboHUB-ANR-11-INBS-0010, 2011].

## REFERENCES

- (1) Lu, C.; Barr, D. B.; Pearson, M. A.; Waller, L. A. *Environ. Health Perspect.* **2008**, *116* (4), 537–542.
- (2) Pestka, J. J. *Arch. Toxicol.* **2010**, *84* (9), 663–679.
- (3) Cascella, M.; Bimonte, S.; Barbieri, A.; Del Vecchio, V.; Caliendo, D.; Schiavone, V.; Fusco, R.; Granata, V.; Arra, C.; Cuomo, A. *Infect. Agents Cancer* **2018**, *13* (1), 3.
- (4) Wild, C. P. *Cancer Epidemiol., Biomarkers Prev.* **2005**, *14* (8), 1847–1850.
- (5) Jamin, E. L.; Bonvallet, N.; Tremblay-Franco, M.; Cravedi, J. P.; Chevrier, C.; Cordier, S.; Debrauwer, L. *Anal. Bioanal. Chem.* **2014**, *406* (4), 1149–1161.
- (6) Ashby, J.; Tennant, R. W. *Mutat. Res., Genet. Toxicol. Test.* **1988**, *204* (1), 17–115.
- (7) Turesky, R. J. *Chimia* **2018**, *72* (10), 718–724.
- (8) Hemeryck, L. Y.; Van Hecke, T.; Vossen, E.; De Smet, S.; Vanhaecke, L. *Food Chem.* **2017**, *230*, 378–387.
- (9) Sinha, R.; Knize, M. G.; Salmon, C. P.; Brown, E. D.; Rhodes, D.; Felton, J. S.; Levander, O. A.; Rothman, N. *Food Chem. Toxicol.* **1998**, *36* (4), 289–297.
- (10) Santarelli, R.; Pierre, F.; Corpet, D. *Nutr. Cancer* **2008**, *60* (2), 131–144.
- (11) de La Pomélie, D.; Santé-Lhoutellier, V.; Gatellier, P. *Food Chem.* **2018**, *239*, 86–93.
- (12) Guéraud, F.; Taché, S.; Steghens, J.-P.; Milkovic, L.; Borovic-Sunjic, S.; Zarkovic, N.; Gaultier, E.; Naud, N.; Héliers-Toussaint, C.; Pierre, F.; et al. *Free Radical Biol. Med.* **2015**, *83*, 192–200.
- (13) Surya, R.; Héliers-Toussaint, C.; Martin, O. C.; Gauthier, T.; Guéraud, F.; Taché, S.; Naud, N.; Jouanin, L.; Chantelauze, C.; Durand, D.; et al. *Carcinogenesis* **2016**, *37* (6), 635–645.
- (14) Ijssennagger, N.; Rijnierse, A.; de Wit, N.; Jonker-Termont, D.; Dekker, J.; Müller, M.; van der Meer, R. *Gut* **2012**, *61* (7), 1041–1049.
- (15) Carlsen, C. U.; Møller, J. K. S.; Skibsted, L. H. *Coord. Chem. Rev.* **2005**, *249* (3–4), 485–498.

- (16) Sesink, A. L. A.; Termont, D. S. M. L.; Kleibeuker, J. H.; Van der Meer, R. *Cancer Res.* **1999**, *59* (22), 5704–5709.
- (17) Bastide, N. M.; Pierre, F. H. F.; Corpet, D. E. *Cancer Prev. Res.* **2011**, *4* (2), 177–184.
- (18) Chevolleau, S.; Noguer-Meireles, M.-H.; Jouanin, I.; Naud, N.; Pierre, F.; Gueraud, F.; Debrauwer, L. *J. Chromatogr. B: Anal. Technol. Biomed. Life Sci.* **2018**, *1083*, 171–179.
- (19) Guéraud, F. *Free Radical Biol. Med.* **2017**, *111*, 196–208.
- (20) Kuiper, H. C.; Miranda, C. L.; Sowell, J. D.; Stevens, J. F. *J. Biol. Chem.* **2008**, *283* (25), 17131–17138.
- (21) Rossin, D.; Calfapietra, S.; Sottero, B.; Poli, G.; Biasi, F. *Free Radical Biol. Med.* **2017**, *111*, 186–195.
- (22) Hinchman, C. A.; Ballatori, N. *J. Toxicol. Environ. Health* **1994**, *41* (4), 387–409.
- (23) Mathias, P. L.; B'Hymer, C. *J. Chromatogr. B: Anal. Technol. Biomed. Life Sci.* **2014**, *964*, 136–145.
- (24) Scholz, K.; Dekant, W.; Völkel, W.; Pähler, A. *J. Am. Soc. Mass Spectrom.* **2005**, *16* (12), 1976–1984.
- (25) Wagner, S.; Scholz, K.; Donegan, M.; Burton, L.; Wingate, J.; Völkel, W. *Anal. Chem.* **2006**, *78* (4), 1296–1305.
- (26) Jian, W.; Yao, M.; Zhang, D.; Zhu, M. *Chem. Res. Toxicol.* **2009**, *22* (7), 1246–1255.
- (27) Panuwet, P.; Nguyen, J. V.; Kuklenyik, P.; Udunka, S. O.; Needham, L. L.; Barr, D. B. *Anal. Bioanal. Chem.* **2008**, *391* (5), 1931–1939.
- (28) Koszucka, A.; Nowak, A. *Crit. Rev. Food Sci. Nutr.* **2019**, *59*, 3579–3596.
- (29) Alary, J.; Debrauwer, L.; Fernandez, Y.; Cravedi, J.-P.; Rao, D.; Bories, G. *Chem. Res. Toxicol.* **1998**, *11* (2), 130–135.
- (30) Jamin, E. L.; Bonvallot, N.; Tremblay-Franco, M.; Cravedi, J.-P.; Chevrier, C.; Cordier, S.; Debrauwer, L. *Anal. Bioanal. Chem.* **2014**, *406* (4), 1149–1161.
- (31) Smith, C. A.; Want, E. J.; O'Maille, G.; Abagyan, R.; Siuzdak, G. *Anal. Chem.* **2006**, *78* (3), 779–787.
- (32) Tautenhahn, R.; Böttcher, C.; Neumann, S. *BMC Bioinf.* **2008**, *9* (1), 504.
- (33) Dieterle, F.; Ross, A.; Schlotterbeck, G.; Senn, H. *Anal. Chem.* **2006**, *78* (13), 4281–4290.
- (34) Giacomoni, F.; Le Corguille, G.; Monsoor, M.; Landi, M.; Pericard, P.; Petera, M.; Duperier, C.; Tremblay-Franco, M.; Martin, J.-F.; Jacob, D.; et al. *Bioinformatics* **2015**, *31* (9), 1493–1495.
- (35) Yang, Y.; Griffiths, W. J.; Sjövall, J.; Gustafsson, J.-Å.; Rafter, J. *J. Am. Soc. Mass Spectrom.* **1997**, *8* (1), 50–61.
- (36) Frigerio, G.; Mercadante, R.; Polledri, E.; Missineo, P.; Campo, L.; Fustinoni, S. *J. Chromatogr. B: Anal. Technol. Biomed. Life Sci.* **2019**, *1117*, 66–76.
- (37) Plumb, R. S.; Johnson, K. A.; Rainville, P.; Smith, B. W.; Wilson, I. D.; Castro-Perez, J. M.; Nicholson, J. K. *Rapid Commun. Mass Spectrom.* **2006**, *20* (13), 1989–1994.
- (38) Alary, J.; Fernandez, Y.; Debrauwer, L.; Perdu, E.; Guéraud, F. *Chem. Res. Toxicol.* **2003**, *16* (3), 320–327.
- (39) Alary, J.; Debrauwer, L.; Fernandez, Y.; Paris, A.; Cravedi, J.-P.; Dolo, L.; Rao, D.; Bories, G. *Chem. Res. Toxicol.* **1998**, *11* (11), 1368–1376.
- (40) Sumner, L. W.; Amberg, A.; Barrett, D.; Beale, M. H.; Beger, R.; Daykin, C. A.; Fan, T. W.-M.; Fiehn, O.; Goodacre, R.; Griffin, J. L.; et al. *Metabolomics* **2007**, *3* (3), 211–221.
- (41) Winter, C. K.; Segall, H. J.; Jones, A. D. *Drug Metab. Dispos.* **1987**, *15* (5), 608–612.
- (42) Bastide, N. M.; Chenni, F.; Audebert, M.; Santarelli, R. L.; Tache, S.; Naud, N.; Baradat, M.; Jouanin, I.; Surya, R.; Hobbs, D. A.; et al. *Cancer Res.* **2015**, *75* (5), 870–879.
- (43) Guo, J.; Villalta, P. W.; Turesky, R. *J. Anal. Chem.* **2017**, *89* (21), 11728–11736.
- (44) Contrepolis, K.; Jiang, L.; Snyder, M. *Mol. Cell. Proteomics* **2015**, *14* (6), 1684–1695.

## Global profiling of toxicologically relevant metabolites in urine: case study of reactive aldehydes

Emilien L. Jamin<sup>\*(1)(2)</sup>, Robin Costantino<sup>\*(1)(2)</sup>, Loïc Mervant<sup>(1)(2)</sup>, Jean-François Martin<sup>(1)(2)</sup>, Isabelle Jouanin<sup>(1)(2)</sup>, Florence Blas-y-Estrada<sup>(2)</sup>, Françoise Guéraud<sup>\*(2)</sup> and Laurent Debrauwer<sup>\*(1)(2)</sup>

(1) Metatoul-AXIOM platform, National Infrastructure for Metabolomics and Fluxomics: MetaboHUB, Toxalim, INRAE, 31300 Toulouse, France

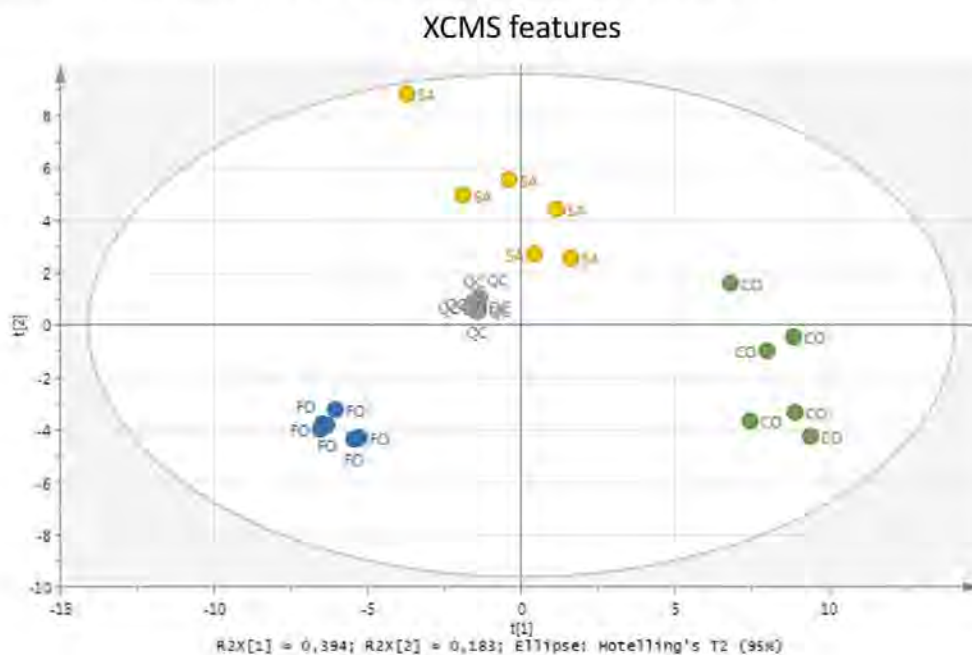
(2) Toxalim (Research Centre in Food Toxicology), Toulouse University, INRAE, ENVT, INP-Purpan, F-31300 Toulouse, France.

\* Equal contribution

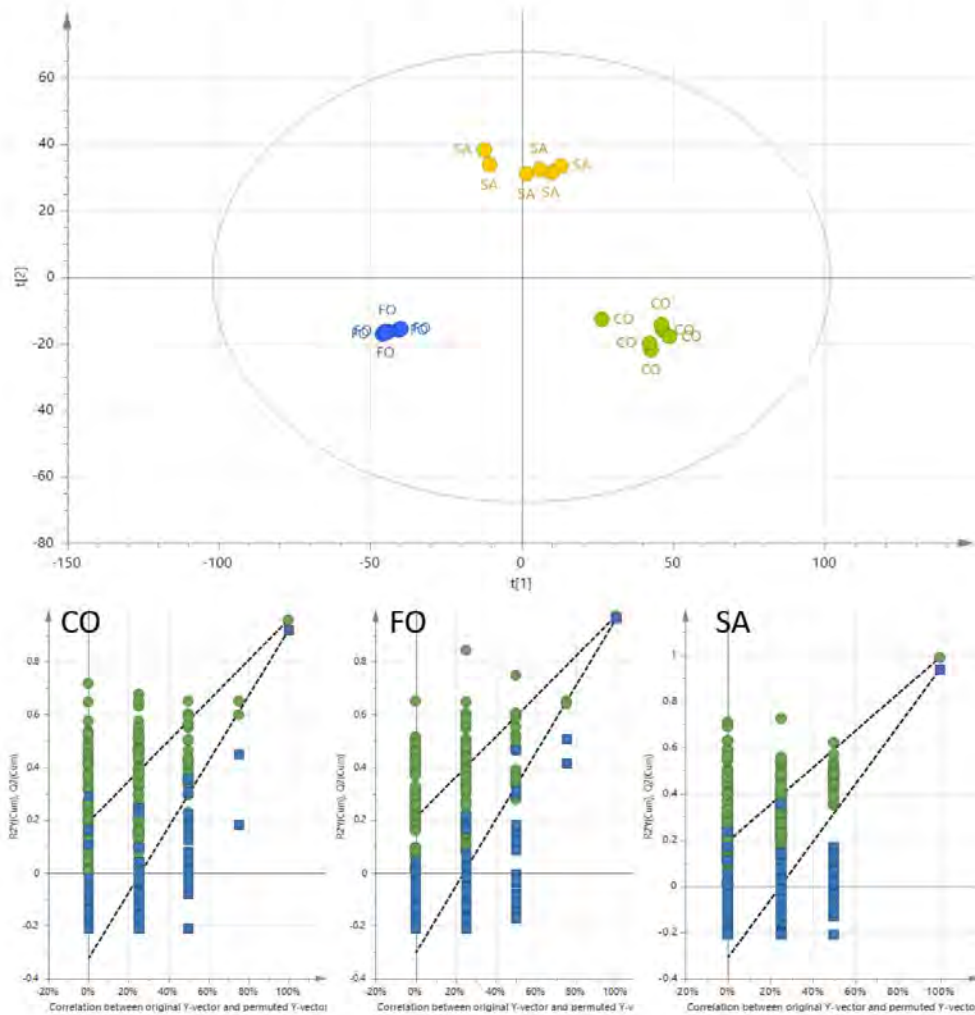
° Corresponding author: emilien.jamin@inra.fr

### SUPPORTING INFORMATION

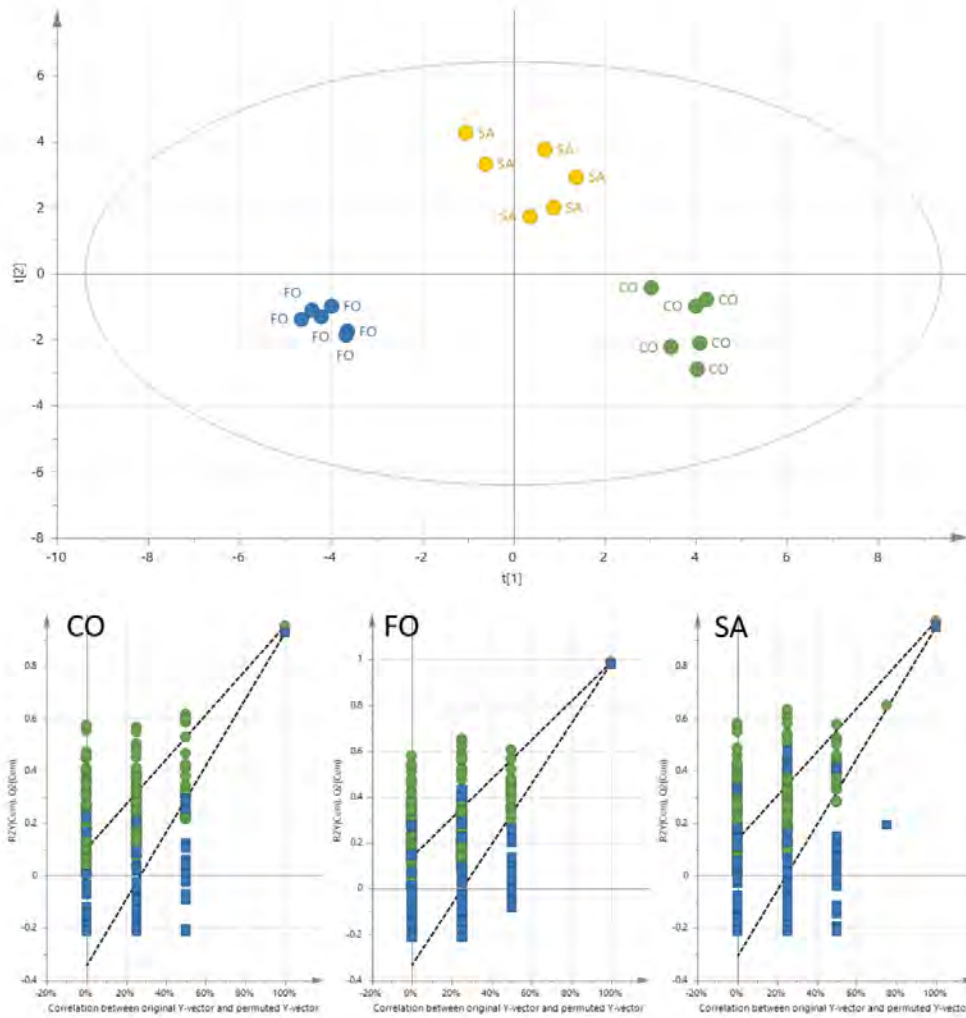
**Figure S1:** Principal Component Analysis of all features detected with XCMS. Yellow dots (SA) represent samples from safflower oil diet, blue dots (FO) represent samples from fish oil diet, green dots (CO) represent hydrogenated coconut oil diet, and grey dots represent QC samples



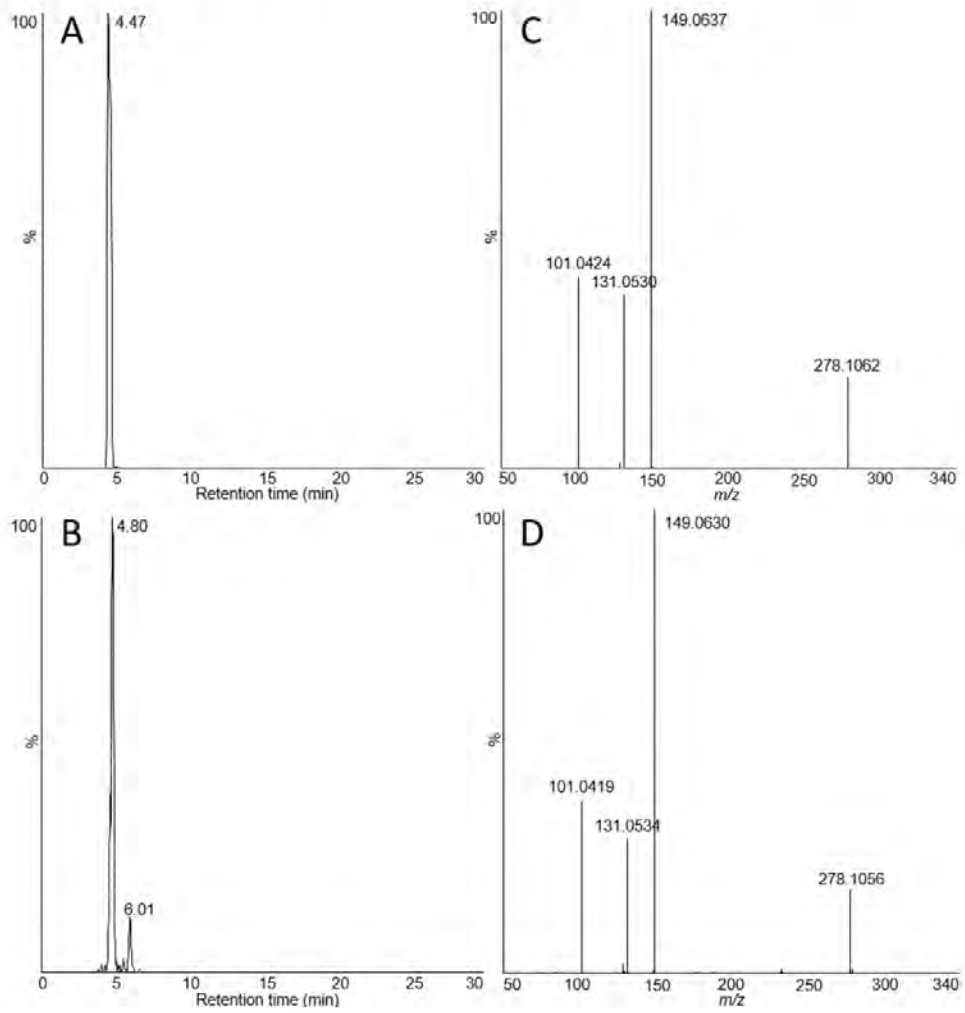
**Figure S2:** Projection to latent structure discriminant analysis modeling (PLS-DA) of all features detected with XCMS. Yellow dots (SA) represent samples from safflower oil diet, blue dots (FO) represent samples from fish oil diet, green dots (CO) represent hydrogenated coconut oil diet, Q2 = 0.94, permutation test of each diet.



**Figure S3:** Projection to latent structure discriminant analysis modeling (PLS-DA) of features detected by MS<sup>E</sup>. Yellow dots (SA) represent samples from safflower oil diet, blue dots (FO) represent samples from fish oil diet, green dots (CO) represent hydrogenated coconut oil diet, Q2 = 0.95, permutation test of each diet.



**Figure S4:** Extracted ion chromatograms of DHH-MA obtained by UPLC-ESI-HRMS from (A) the standard of DHH-MA and (B) urine samples. Corresponding product ion spectra obtained by electrospray in the negative mode from (C) the standard of DHH-MA and (D) urine samples.





## **A.2. Principaux résultats et conclusion de l'article**

Ce travail a permis de démontrer l'intérêt de la mise en place d'une approche de *screening* des conjugués mercapturiques par DIA en raison de (i) la détection d'un nombre important de conjugués mercapturiques (56 espèces produisant la perte de neutre d'intérêt) (ii) l'amélioration de la discrimination statistique des groupes expérimentaux lorsque seuls les mercapturates étaient considérés. Ces travaux mettent également en avant le peu de connaissances spectrales disponibles vis-à-vis de ces composés. En effet, sur les 56 conjugués mercapturiques détectés, seulement 2 ont pu être annotés au niveau 1, 1 seul au niveau 2 et le reste étant au niveau 3.

Outre l'intérêt toxicologique de l'approche démontré par cette étude, ces résultats mettent également en évidence une forte peroxydation lipidique induite par des régimes riches en hème ainsi que la diversité de ces produits en fonction du régime lipidique, ce qui pose la question de l'effet biologique de ces divers produits. Cependant, lors du développement de cette méthode, une forte fragmentation en source avait été constatée, ainsi qu'un nombre important de faux positifs lors de la vérification de la perte de neutre par MS/MS ciblé. Les développements présentés ci-après s'inscrivent dans une optique d'optimisation de cette méthodologie afin de pallier, entre autres, ces deux limitations.

## **B. Matériel et méthodes**

### **B.1. Réactifs et tampons utilisés**

Tous les solvants utilisés pour les analyses LC-MS sont de qualité LC-MS et proviennent de Fisher Scientific (Thermo Fisher Scientific, Illkirch, France). Le DHN-MA a été synthétisé au laboratoire d'après une méthode décrite précédemment (Alary *et al.*, 1998). La *N*-acetyl-S-phenylcystéine (NAC-Phe-Cys) provient de Sigma-Aldrich (Saint Quentin Fallavier, France). L'urine humaine utilisée provient de volontaires sains du laboratoire. L'urine de synthèse utilisée a été préparée au laboratoire, avec la composition suivante : 500 mL d'eau Milli-Q, 3,8 g de chlorure de potassium, 8,5 g chlorure de sodium, 24,5 g urée, 1,03 g acide citrique, 0,34 g acide ascorbique, 1,18 g phosphate de potassium, 1,4 g créatinine, 0,64 g hydroxyde de sodium, 0,47 g bicarbonate de sodium, 0,28 mL acide sulfurique.

### **B.2. Préparation des échantillons urinaires**

Les échantillons d'urine sont dilués au ¼ dans la phase mobile A (95 % H<sub>2</sub>O/ 5 % méthanol + 0,1 % d'acide acétique) puis centrifugés 10 min à 10 000 g. Le surnageant est récupéré puis mis en vial pour analyse.

### B.3. Analyse LC-MS

Les analyses LC-MS ont été réalisées sur un spectromètre de masse Q-ToF Synapt G2-Si (Waters, Manchester, UK) couplé à un système chromatographique ACQUITY™ (Waters, Manchester, UK). La séparation LC a été réalisée à l'aide d'une colonne chromatographique Hypersil Gold C18 (1,9 µm, 100 x 2,1 mm, Thermo Fisher Scientific, Illkirch, France). La température de la colonne est fixée à 40 °C tout le long de l'analyse et le débit est fixé à 0.3 mL.min<sup>-1</sup>. Pour l'élution, un gradient linéaire est réalisé à l'aide des phases A : 95 % H<sub>2</sub>O/ 5 % méthanol + 0,1 % d'acide acétique et B : 100 % méthanol + 0,1 % d'acide acétique. Le gradient est constitué d'une étape linéaire de 0 % à 100 % de B en 30 min, puis d'un plateau à 100 % de B pendant 4 min avant de revenir aux conditions initiales en 1 min puis de conserver ces conditions 5 min afin de rééquilibrer le système.

Concernant les analyses MS, les spectres ont été acquis sur une gamme de masse allant de  $m/z$  50 à  $m/z$  800 en ionisation négative par electrospray en utilisant les paramètres suivants : voltage du capillaire 0,5 kV, voltage du sampling cone 30 V, température de source 120 °C, température de désolvatation 550 °C, débit du gaz de cône 30 et du gaz de désolvatation 600 L.h<sup>-1</sup>. Les mêmes paramètres sont utilisés pour le mode MS<sup>E</sup>. Dans ce mode l'énergie de collision est fixée à 15 eV pour le scan de haute énergie.

### B.4. Traitement des données

Les données issues des analyses LC-MS ont été traitées avec XCMS 3.0 disponible sous l'environnement collaboratif Workflow4Metabolomics (Giacomoni *et al.*, 2015). Le détail des paramètres utilisés est disponible en Annexe 1. Les données issues des acquisitions MS<sup>E</sup> ont été traitées à l'aide du logiciel UNIFI™ (Waters, Manchester, UK). Les données ciblées ont été analysées avec le logiciel QuanLynx™.

### B.5. Détermination de la dilution de travail

Trois dilutions d'urine ont été réalisées selon le protocole présenté ci-avant : 1:1, 1:3, 1:7 (v:v). Chaque dilution a été injectée en triplicat en mode MS<sup>E</sup>. Les données MS<sup>E</sup> sont analysées *via* le logiciel UNIFI et les paramètres utilisés sont reportés en **Erreur ! Source du renvoi introuvable.** Le logiciel UNIFI ne procédant pas à l'alignement des données en temps de rétention et en  $m/z$ , un *matching* est réalisé sur les 3 matrices de données provenant de chacune des dilutions afin de dénombrer les espèces qui sont retrouvées dans un seul, deux ou l'ensemble des réplicats pour une dilution donnée. Les paramètres utilisés pour le *matching* sont les suivants : tolérance en précision de mesure de masse de ± 10 ppm pour le  $m/z$  et ± 5 sec pour le temps de rétention.

Parallèlement, 3 séries de 60 injections successives de chaque dilution ont été réalisées en mode full-MS puis analysées *via* XCMS sous W4M afin d'analyser la dérive du signal en fonction de l'ordre d'injection et de la dilution utilisée. Suite au traitement des données, les *features* dont le rapport  $\frac{\text{Moyenne \acute{e}chantillon}}{\text{Moyenne blanc}} < 3$  et dont le CV < 30 % sont éliminés.

### **B.6. Evaluation de la justesse et sensibilité de l'acquisition MS<sup>E</sup>**

Trois réplicats d'urine (dilution 1:3 (v:v)) ont été analysés en mode MS et MS<sup>E</sup> et les données ont été traitées *via* XCMS et UNIFI comme décrit précédemment. Les mercapturates à la fois détectés dans chaque réplicat MS<sup>E</sup> et également présents dans la *peak list* issue de l'extraction XCMS ont été fragmentés en MS/MS ciblée en utilisant une énergie fixée à 15 eV (fenêtre d'isolation du précurseur : 2 Da) afin de confirmer la présence de la perte de neutre de 129,0426 Da caractéristique des conjugués mercapturiques.

L'évaluation de la sensibilité a été déterminée par l'injection en triplicat d'une gamme d'urine de synthèse surchargée avec deux standards de conjugués mercapturiques : DHN-MA et NAC-Phe-Cys. 100 µL d'urine de synthèse ont été surchargés avec 50 µL d'une solution de NAC-Phe-Cys à 1 µg.mL<sup>-1</sup> et 50 µL d'une solution de DHN-MA à 1 µg.mL<sup>-1</sup>. A partir de cette solution une série de 5 dilutions successives au ¼ a été réalisée pour atteindre une concentration de 0,24 ng.mL<sup>-1</sup>.

### **B.7. Optimisations supplémentaires du traitement des données MSE**

Un total de 240 échantillons d'urines issus de la cohorte NutriNet-Santé, classés en 3 groupes de consommation de viande rouge et de charcuterie, ont été utilisés pour réaliser cette étude. La sélection des échantillons et la description de la cohorte sont décrites de manière détaillée dans le Chapitre III : Etude de l'impact de la consommation de viande rouge et de charcuterie, de ce manuscrit. En résumé, 80 individus qualifiés de gros consommateurs de viande rouge et de charcuterie ont été sélectionnés parmi le quintile de consommation le plus haut, 80 non-consommateurs de viande rouge et de charcuterie et 80 autres participants sélectionnés au hasard dans l'ensemble des quintiles excepté celui des non-consommateurs. L'osmolalité des échantillons a été mesurée à l'aide d'un osmomètre à point de congélation (Loeser Messtechnik, Berlin, Germany) puis une dilution spécifique de chaque échantillon a été réalisée, en utilisant de l'eau MilliQ, afin de normaliser l'osmolalité. Trois pools de ces échantillons correspondant à chacun des 3 groupes ont été constitués par mélange d'un volume équivalent de chaque échantillon d'un groupe donné. Enfin, les pools ont été injectés en réplicat (n=6 par groupe) en LC-MS<sup>E</sup> en utilisant les paramètres décrits précédemment. Les données acquises ont été traitées à l'aide du logiciel UNIFI et 6 jeux de paramètres décrits dans le Tableau 2

ont été testés afin de maximiser la détection des mercapturates. Comme décrit précédemment, un *matching* est réalisé afin de déterminer la fréquence de détection des mercapturates potentiels dans chaque réplicat (paramètres utilisés :  $\pm 10$  ppm et  $\pm 5$  sec). Ces mercapturates potentiels détectés sont ensuite fragmentés en MS/MS ciblée en utilisant une énergie de 15 eV et une fenêtre d'isolation de 2 Da afin de confirmer la présence de la perte de neutre de 129,0426 Da.

**Tableau 2 Paramètres du logiciel UNIFI utilisés pour la détection des mercapturates dans les données MS<sup>E</sup>**

		Méthode 1	Méthode 2	Méthode 3	Méthode 4	Méthode 5	Méthode 6
	Rt range (min)	2.5-20	2.5-20	2.5-20	2.5-20	2.5-20	2.5-20
<b>High energy</b>	Intensity treshold	200	200	200	20	500	20
	<i>m/z</i> start	10	10	10	10	10	10
	<i>m/z</i> end	800	800	800	800	800	800
<b>Low energy</b>	Intensity treshold	2500	2500	2500	2500	5000	5000
	<i>m/z</i> start	230	230	230	230	230	230
	<i>m/z</i> end	800	800	800	800	800	800
	Background filter	High	High	High	High	High	High
<b>3D isotope clustering and high to low energy association</b>	Fraction of the chromatographic peak width to apply during cluster creation	0.25	0.25	0.33	0.33	0.33	0.33
	Fraction of the chromatographic peak width to apply during high to low energy association	0.25	0.25	0.33	0.33	0.33	0.33
<b>Common Neural Loss</b>	Tolerance (mDa)	2	4	5	5	5	5

### **B.8. Comparaison de l'approche MS<sup>E</sup> avec une méthode fast-DDA**

Afin de comparer la performance d'une analyse fast-DDA et la méthode MS<sup>E</sup> mise en place pour la détection des conjugués mercapturates, des échantillons QC issus de la cohorte NutriNet-Santé ont été analysés par LC-MS sur un spectromètre de masse timsTOF-fleX (Bruker, Brême, Allemagne) couplé à un système chromatographique ELUTE™ (Bruker, Brême, Allemagne). Cet instrument possède la capacité de réaliser des analyses en utilisant le mode PASEF™ (pour *Parallel Accumulation Serial Fragmentation*), permettant d'allier l'usage de la mobilité ionique et une approche DDA afin d'acquérir un grand nombre de spectres MS/MS ciblés. La théorie et le fonctionnement du mode

PASEF sont décrits en Annexe 15. Les analyses LC sont opérées à l'aide des mêmes paramètres que ceux décrits précédemment. Concernant les analyses en mode PASEF™, les spectres ont été acquis sur une gamme de  $m/z$  de 60 à 1100 en ionisation négative en utilisant les paramètres de source suivants : *capillary voltage* 3,6 kV, *end plate offset* 500 V, *nebulizer pressure* 2.5 Bar, *drying gas* 6 L.min<sup>-1</sup> et *drying temp* 220 °C. L'analyse PASEF nécessite l'activation de la cellule de mobilité ionique pour laquelle les paramètres suivants ont été utilisés : gamme de mobilité de 0,42 V.s.cm<sup>2</sup> (1/K0 Start) à 1,32 V.s.cm<sup>2</sup> (1/K0 End) avec un *ramp time* fixé à 100 ms. Enfin la partie DDA du mode PASEF est paramétrée comme suit : *collision energy* : 30 eV, *active exclusion* : 0.3 min, nombre de scans MS/MS PASEF : 10, re-fragmentation du précurseur pour une intensité > 5 fois son intensité précédemment mesurée.

Les données acquises en mode PASEF ont été traitées à l'aide du logiciel MetaboScape® qui permet la recherche automatique de la perte de neutre d'intérêt de 129,0426 Da caractéristique des conjugués mercapturiques.

## C. Résultats

### C.1. Optimisation et évaluation des performances de la méthode

#### C.1.1. Détection et fragmentation en source des mercapturates

Afin d'assurer une détection optimale des conjugués mercapturiques, les paramètres de transfert des ions au sein du spectromètre de masse ont été optimisés. En effet, en raison de la présence d'une liaison thioéther fragile, ces composés sont propices à se fragmenter en source, ce qui peut limiter la capacité de détection des ions parents. Ainsi, l'optimisation des paramètres a été réalisée dans le but

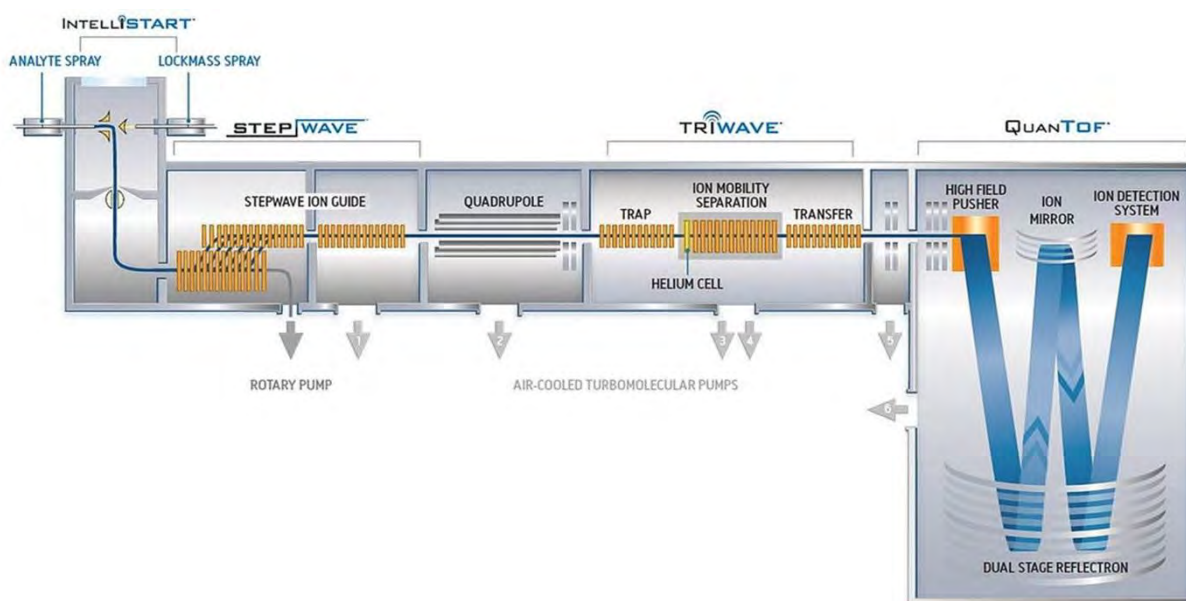


Figure 17 Représentation schématique du Synapt-G2-Si (Waters)

d'assurer une meilleure détection possible des ions parents en limitant la fragmentation en source. Pour cela, l'aire des composés parents mais aussi la valeur du rapport des aires entre ion fragment/ion parent ont été utilisés. La fragmentation en source correspond à une dissociation non-désirée des molécules, dans les parties du spectromètre de masse où règne encore une pression élevée, due à la collision entre les ions et des molécules de gaz (air résiduel). Ainsi, ce phénomène a essentiellement lieu, dans le cas du Synapt G2-Si, au niveau des compartiments StepWave et TriWave (voir Figure 17).

Afin d'évaluer l'influence des paramètres de transfert des ions au niveau du StepWave et du TriWave sur la fragmentation des mercapturates, 3 jeux de paramètres de transferts disponibles en Annexe 3 ont été testés et comparés. Pour cette étude, deux standards de conjugués mercapturiques ont été utilisés, le DHN-MA et NAC-Phe-Cys. Les résultats sont présentés en Figure 18. En abaissant les différences de potentiel entre l'entrée et la sortie du TriWave et du StepWave, une augmentation du signal correspondant aux ions monoisotopiques  $[M-H]^+$  des deux standards considérés a été constatée, comme présenté en Figure 18A et C. Une diminution du rapport fragment/parent témoignant d'une diminution de la fragmentation en source a également été observée (Figure 18B et D). Les résultats des optimisations présentés ont été obtenus à l'aide de solutions à  $1 \mu\text{g}\cdot\text{mL}^{-1}$  mais des résultats

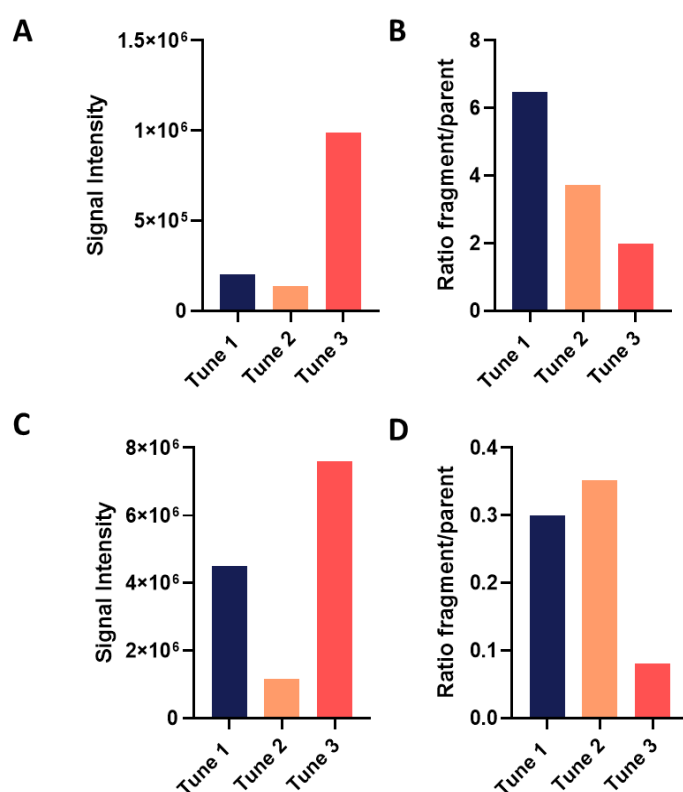


Figure 18 Comparaison de l'intensité du signal de deux standards de mercapturates en fonction des paramètres de tune. (A) Intensité du signal de l'ion parent du NAC-Phe-Cys (B) Ratio entre l'aire du signal de l'ion fragment et de l'ion parent (C) Intensité du signal de l'ion parent du DHN-MA (D) Ratio entre l'aire du signal de l'ion fragment et de l'ion parent

similaires ont été retrouvés pour d'autres concentrations (100 ng.mL<sup>-1</sup> pour le DHN-MA et 10 ng.mL<sup>-1</sup> pour le NAC-Phe-Cys).

### C.1.2. Détermination de la dilution de travail

La détermination de la dilution optimale de l'urine pour l'analyse des conjugués mercapturiques a été réalisée en se basant sur deux critères principaux, la dérive analytique lors d'injections répétées et le nombre de mercapturates détectés pour chaque dilution testée. L'analyse d'urine en métabolomique par LC-MS est souvent réalisée à l'état pur ou après dilution au ½ ou au ¼ afin de limiter l'encrassement de la source et les effets matrice (Chetwynd *et al.*, 2015). Pour des raisons de sensibilité, les dilutions au ½ ou au ¼ sont préférées aux dilutions plus importantes afin d'assurer la détection des espèces les moins abondantes.

Pour cette étude, chaque échantillon dilué au ½ au ¼ ou au ⅛, a été injecté en triplicat et le nombre de mercapturates détectés ainsi que leur fréquence de détection (i.e. détectés dans un, deux ou les trois réplicats) ont été utilisés pour déterminer le choix de la dilution. Les résultats présentés en Figure 19A témoignent de l'impact de la dilution sur le nombre total de mercapturates potentiels détectés. En effet, si des résultats similaires sont observés pour les dilutions ½ (1314 espèces) et ¼ (1354 espèces) il apparaît que le nombre de mercapturates potentiels détectés est bien inférieur avec la dilution 1/8 (637 espèces). Bien que moins marquée, cette différence persiste lorsque l'on considère uniquement les mercapturates détectés dans chaque triplicat pour une dilution donnée (Figure 19B).

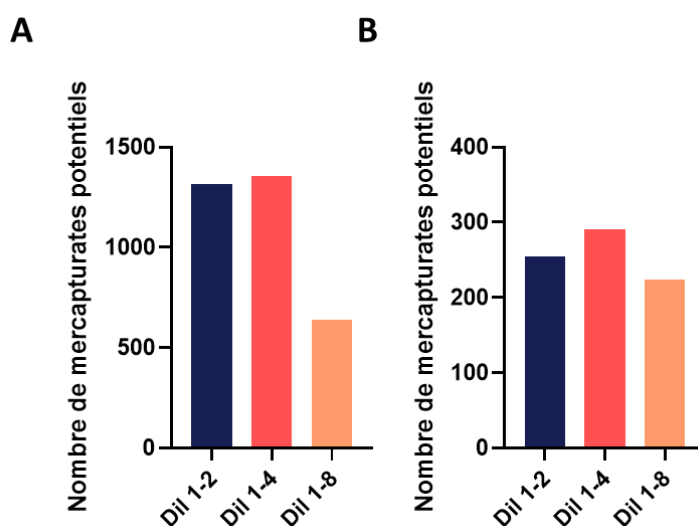


Figure 19 Nombre de mercapturates potentiels détectés (A) Somme des candidats détectés (B) Candidats retrouvés dans chaque triplicat

En effet, 291 mercapturates potentiels sont détectés dans les urines « Dil 1-4 » contre 254 et 223 dans les urines « Dil 1-2 » et « Dil 1-8 », respectivement. Ces résultats mettent en avant à la fois une faible reproductibilité de la méthode MS<sup>E</sup> (seulement 22% des espèces détectées sont retrouvées dans

chaque réplicat pour la dilution  $\frac{1}{4}$ ), mais aussi la performance de la dilution au  $\frac{1}{4}$  qui permet à la fois la détection du plus grand nombre de mercapturates potentiels, et surtout du plus grand nombre d'espèces dans chaque réplicat. Compte-tenu de ces résultats, il apparaît que la dilution de l'urine  $\frac{1}{4}$  semble optimale pour la détection des acides mercapturiques pour la suite de nos expériences. Afin d'étudier la décroissance du signal due à un possible encrassement de la source d'ionisation, 60 injections consécutives de chaque dilution ont été réalisées. Un suivi de l'intensité du signal global détecté en fonction de l'ordre d'injection a été réalisé afin d'évaluer cette décroissance. Les résultats présentés en Figure 20A montrent que, comme attendu, la dérive analytique est d'autant plus faible que l'urine est diluée. En effet au bout de 60 injections, le signal global diminue d'environ 15 % pour l'urine « Dil 1-8 », tandis que la diminution observée est de 35 et 45 % pour les « Dil 1-4 » et « Dil 1-8 », respectivement.

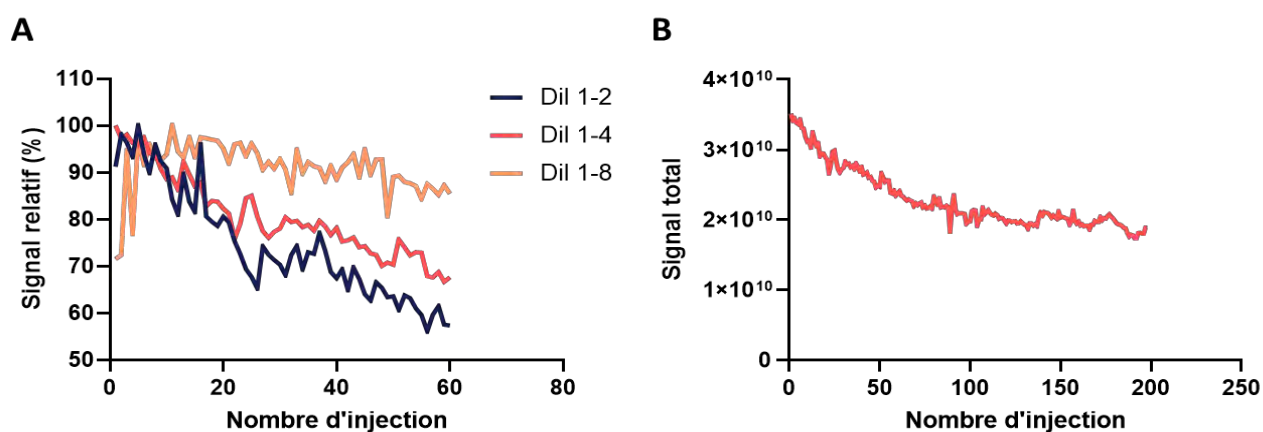


Figure 20 Evolution du signal en fonction de la dilution et du nombre d'injection (A) Signal relatif en % par rapport à la valeur maximale détectée pour chaque dilution (B) Signal observé pour 200 injections d'urine diluée au  $\frac{1}{4}$

L'objectif de notre développement de méthode étant de réaliser à la fois une analyse non-ciblée du métabolome urinaire tout en maximisant la détection des mercapturates, le nombre de variables détectées a également été étudié. La dilution « Dil 1-2 » a permis de conserver 8970 variables après l'application des filtres tandis que le nombre de variables détectées pour les dilutions « Dil-4 » et « Dil-8 » est de 6350 et 3077 respectivement.

L'ensemble de ces résultats montrent que la dilution « Dil 1-4 », malgré un nombre de variables détectées inférieur à la dilution « Dil 1-2 », apparaît comme un compromis satisfaisant permettant une détection optimale des mercapturates en limitant l'encrassement de la source et donc la décroissance du signal ce qui, dans le cas de longues séquences analytiques, permet de limiter le nombre de séries d'injections et donc de limiter les biais. Enfin, afin d'évaluer la décroissance du signal sur une série d'injections plus longue, 200 injections d'urine diluée au  $\frac{1}{4}$  ont été réalisées et une décroissance du signal d'environ 50% au bout de 200 injections a été observée, comme présenté en Figure 20B.



### C.1.3. Evaluation de la justesse et sensibilité de l'acquisition MS<sup>E</sup>

Une fois la dilution optimale de l'urine déterminée et compte-tenu de la faible répétabilité du mode MS<sup>E</sup>, le nombre de faux positifs (mercapturates potentiels ne présentant pas la perte de 129,0426 Da en MS/MS ciblée) a été déterminé. Après injection en triplicat d'échantillons d'urine dilués au ¼ puis réconciliation des données XCMS et UNIFI, 42 mercapturates potentiels ont été fragmentés en MS/MS. Parmi ces 42 composés, 24 présentaient la perte de neutre caractéristique des mercapturates, cette dernière n'ayant pas été observée pour les 18 autres candidats. Ces résultats permettent de mettre en avant les limites du traitement de données automatique et de l'attribution parent/fragment réalisée par le logiciel UNIFI, car environ 40 % des mercapturates potentiels détectés par UNIFI s'avèrent être de faux positifs. L'étape de confirmation par MS/MS ciblée est donc nécessaire pour s'assurer de la qualité des données.

Enfin, l'injection en triplicat d'urine de synthèse surchargée à l'aide de standards de mercapturates à différentes concentrations, a permis une évaluation de la sensibilité de la méthode MS<sup>E</sup>. Les résultats présentés dans le Tableau 3 permettent d'établir que la limite de détection pour le DHN-MA se situe autour de 1 ng.mL<sup>-1</sup> tandis que celle du NAC-Phe-Cys est comprise entre 1 et 4 ng.mL<sup>-1</sup>. Ces résultats mettent en avant que, bien qu'appartenant à la même classe chimique, la réponse et l'efficacité d'ionisation diffère en fonction des composés et que l'établissement d'une limite de détection claire pour l'ensemble des mercapturates n'est pas possible.

**Tableau 3 Aire du pic mesurée par le logiciel UNIFI pour deux standards de mercapturates à différentes concentrations : DHN-MA et Nac-Phe-Cys**

[C] ng.mL	DHN-MA			NAC-Phe-Cys		
	Gamme 1	Gamme 2	Gamme 3	Gamme 1	Gamme 2	Gamme 3
<b>0.244141</b>	0	0	0	0	0	0
<b>0.976563</b>	6955	7982	8848	0	0	0
<b>3.90625</b>	29026	36466	34850	4352	6069	6076
<b>15.625</b>	104611	160354	146365	14119	24926	23739
<b>62.5</b>	370066	666443	605384	50534	99059	91518
<b>250</b>	1109263	1925097	1841763	144108	322615	303329

## C.2. Annotation des données

Un des enjeux majeurs en métabolomique réside dans l'annotation des composés détectés, qui reste l'une des limites principales dans les approches métabolomiques non ciblées. En particulier pour certaines classes de composés, comme les conjugués issus de métabolisation tels que les acides mercapturiques, peu de composés sont reportés dans les bases de données publiques telles que HMDB (3 entrées pour le terme « *mercapturic* »). Afin d'augmenter la couverture des mercapturates annotés, 2 stratégies ont été mises en place visant à la création d'une base de données de mercapturates connus et suspectés. La première, basée sur l'expertise du laboratoire et sur la littérature, vise à inclure les mercapturates reportés dans la littérature ainsi que certains suspectés à partir de métabolisations également décrites dans la littérature. Par exemple, l'hypothèse est faite que les voies de métabolisation qui permettent la formation des différents dérivés mercapturiques du 4-HNE, du 4-HHE ou du 4-oxo-nonanal, décrites dans la littérature (Alary et al., 1995b; Kuiper et al., 2008; Winter et al., 1987) pourraient être similaires pour d'autres hydroxy- ou oxo-alcénals et que l'ensemble de ces formes peuvent être potentiellement retrouvées sous forme de conjugués mercapturiques dans les urines.

La deuxième stratégie, basée sur la structure des molécules, vise à extraire des bases de données publiques tous les aldéhydes présentant un système conjugué et donc une possible réactivité vis-à-vis des nucléophiles. Pour cela, une fonctionnalité de l'outil Chemistry Development Kit (<https://cdk.github.io/>) est utilisée « *Unique SMARTS matches against a SMILES string* ». Cette fonctionnalité est basée sur l'utilisation des identifiants SMILES (*Simplified Molecular Input Line Entry Specification*) qui permettent de décrire la structure d'une molécule à l'aide d'une notation symbolique. L'utilisation des identifiants SMARTS, qui correspondent à des SMILES avec un degré de liberté sur certains atomes (voir Figure 21), permet de rechercher au sein d'une chaîne SMILES la sous-structure d'intérêt. A l'aide de cette stratégie, présentée en Figure 21, l'ensemble des aldéhydes conjugués reportés dans les bases de données PubChem et HMDB ont été ajoutés à la liste des

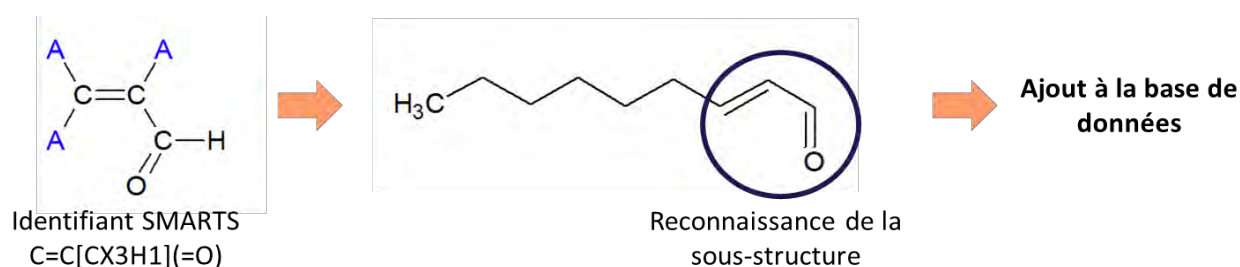


Figure 21 Stratégie d'extraction automatique des systèmes aldéhydes conjugués à l'aide d'un identifiant SMARTS (les « A » dans la structure correspondent à Any atom)

potentiels précurseurs de conjugués mercapturiques. Ainsi, grâce à l'utilisation de ces deux approches, environ 1000 composés ont été ajoutés à notre base de données des mercapturates.

### C.3. Etude de méthodes de normalisation

#### **C.3.1. Introduction à l'article**

##### C.3.1.1. Contexte

L'urine est un biofluide largement utilisé en métabolomique, notamment en raison du caractère non-invasif de son prélèvement, de la possibilité de récolter de grandes quantités d'échantillons à différent temps, permettant d'étudier des effets sur la durée et d'une préparation d'échantillons simple. En revanche, l'un des inconvénients majeurs de l'analyse d'urine, en comparaison aux autres biofluides, réside dans la grande variabilité de concentration entre les individus et qui, chez des individus sains, est majoritairement due au statut hydrique. Etant donné que la plupart des études métabolomiques reposent sur la comparaison de concentration (absolue ou relative) de métabolites entre deux ou plusieurs groupes d'individus, de telles variations peuvent conduire à des résultats biaisés et donc à de mauvaises interprétations. Malgré ce besoin, une étude de la littérature réalisée sur les publications de métabolomique visant à détecter des biomarqueurs urinaires entre 2012 et 2014 a montré que près de 50 % des études ne précisaient pas si les données avaient été normalisées (Zhang and Watson, 2015). De plus, il existe une grande disparité des techniques de normalisation, basées sur des paramètres biologiques ou physico-chimiques tels que le taux de créatinine, le volume d'urine sur 24 h, la gravité spécifique, l'osmolalité ou sur des paramètres liés à l'analyse tels que la somme du signal global détecté dans chaque échantillon. Bien que plusieurs études aient déjà été menées pour comparer l'efficacité de ces différents paramètres (Gagnebin *et al.*, 2017; Rosen Vollmar *et al.*, 2019; Vogl *et al.*, 2016) aucun consensus n'a, à ce jour, été établi.

##### C.3.1.2. Objectifs de l'étude

Face à la diversité des techniques de normalisation et au manque de standardisation des pratiques, l'objectif principal de cette étude a été de comparer les méthodes les plus couramment utilisées pour la normalisation d'échantillons urinaires en métabolomique. L'objectif a également été de comparer l'impact d'une normalisation pré-acquisition, à savoir la dilution des échantillons pour uniformiser la créatinine ou l'osmolalité, à celui d'une normalisation post acquisition basée soit sur ces mêmes paramètres, soit sur l'application de méthodes de normalisation du signal telles que PQN ou MSTUS (Dieterle *et al.*, 2006; Warrack *et al.*, 2009). Enfin, la pertinence de la combinaison de plusieurs méthodes de normalisation a également été évaluée.

#### C.3.1.3. Méthodes

Pour répondre à l'objectif de l'étude, 40 échantillons d'urine issus de 20 femmes et 20 hommes sains ont été sélectionnés au sein d'échantillons de la cohorte BioNutriNet analysés au laboratoire sur un autre projet. Les échantillons sont séparés en 3 puis soit dilués au 1/4, soit dilués pour normaliser les taux de créatinine ou l'osmolalité. L'ensemble de ces 120 échantillons sont randomisés puis analysés en 3 séries analytiques. Afin d'évaluer la performance de ces différentes approches de normalisation, les données issues des échantillons non-normalisés sont utilisées comme référence. Les résultats sont évalués selon plusieurs critères tels que la capacité des différentes méthodes à (i) réduire la variabilité globale du jeu de données, (ii) réduire la variabilité résiduelle et (iii) augmenter la part de variabilité biologique d'intérêt expliquée par le modèle, le tout conduisant à une amélioration de la prédictivité du modèle OPLS-DA utilisé pour discriminer les deux populations.

## **Osmolality-based normalization enhances statistical discrimination of untargeted metabolomic urine analysis: results from a comparative study**

Loïc Mervant<sup>1,2</sup>, Marie Tremblay-Franco<sup>1,2</sup>, Emilien L Jamin<sup>1,2</sup>, Emmanuelle Kesse-Guyot<sup>3</sup>, Pilar Galan<sup>3</sup>, Jean-François Martin<sup>1,2</sup>, Françoise Guéraud<sup>2</sup>, Laurent Debrauwer<sup>1,2</sup>

1. Metatoul-AXIOM platform, MetaboHUB, Toxalim, INRAE, Toulouse, France
2. Toxalim, Toulouse University, INRAE, ENVT, INP-Purpan, UPS, Toulouse, France
3. Sorbonne Paris Nord University, Inserm, INRAE, Cnam, Nutritional Epidemiology Research Team (EREN), Epidemiology and Statistics Research Center – University of Paris (CRESS), 93017 Bobigny, France

Corresponding author: marie.tremblay-franco@inrae.fr

### **Abstract**

**Introduction:** Because of its ease of collection, urine is one of the most commonly used matrices for metabolomics studies. However, unlike other biofluids, urine exhibits tremendous variability that can introduce confounding inconsistency during result interpretation. Despite many existing techniques to normalize urine samples, there is still no consensus on either which method is most appropriate or how to evaluate these methods.

**Objectives:** To investigate the impact of several methods and combinations of methods conventionally used in urine metabolomics on the statistical discrimination of two groups in a simple metabolomics study.

**Methods:** We applied 14 different strategies of normalization to forty urine samples analysed by liquid chromatography coupled to high-resolution mass spectrometry (LC-HRMS). To evaluate the impact of these different strategies, we relied on the ability of each method to reduce confounding variability while retaining variability of interest, as well as the predictability of statistical models.

**Results:** Among all tested normalization methods, osmolality-based normalization gave the best results. Moreover, we demonstrated that normalization using a specific dilution prior to the analysis outperformed post-acquisition normalization. We also demonstrated that the combination of various normalization methods does not necessarily improve statistical discrimination.

**Conclusions:** This study re-emphasized the importance of normalizing urine samples for metabolomics studies. In addition, it appeared that the choice of method had a significant impact on result quality. Consequently, we suggest osmolality-based normalization as the best method for normalizing urine samples.

Trial registration number: NCT03335644

**Keywords:** Untargeted metabolomics, Urine analysis, Normalization, Mass spectrometry, Osmolality

## Declarations

**Funding** PhD doctoral fellowship of LM was co-funded by INRAE (Nutrition, Chemical Food Safety and Consumer Behaviour Scientific Division) and French region Occitanie. The BioNutriNet project was supported by the French National Research Agency (Agence Nationale de la Recherche) in the context of the 2013 Programme de Recherche Systèmes Alimentaires Durables (ANR-13-ALID-0001).

**Conflict of interest** All authors declare that they have no potential conflict of interest.

**Ethics approval** NutriNet-Santé study is conducted according to the Declaration of Helsinki guidelines and was approved by the Institutional Review Board of the French Institute for Health and Medical Research (IRB Inserm n°0000388FWA00005831) and the “Commission Nationale de l’Informatique et des Libertés” (CNIL n°908450/n°909216). Clinicaltrials.gov number: NCT03335644. In addition, in 2011–2014, the NutriNet-Santé participants were invited, on a voluntary basis, to attend a visit for biological sampling and clinical examination in one of the local centres throughout France. All procedures were approved by the ‘Consultation Committee for the Protection of Participants in Biomedical Research’ (C09-42 on May 5th 2010) and the CNIL (no. 1460707).

**Consent to participate** The authors declare that electronic and paper written informed consent was obtained from all participant included in this study.

**Consent for publication** The authors declare that informed consent was obtained from all participant included in this study.

**Availability of data and material** Metabolomics data have been deposited to the EMBL-EBI MetaboLights database (DOI: 10.1093/nar/gkz1019, PMID: 31691833) with the identifier MTBLS1765. The complete dataset can be accessed here <https://www.ebi.ac.uk/metabolights/MTBLS1765>

**Code availability** Not applicable

**Author contributions** LM, LD, ELJ and FGU designed the study. LM has conducted MS analysis. LM MTF and JFM analyzed the data and performed statistical analysis. EK has provided urine samples. All authors co-wrote and approved submission of the final version of the manuscript.

**Acknowledgements** PhD doctoral fellowship of LM was co-funded by INRAE (Nutrition, Chemical Food Safety and Consumer Behaviour Scientific Division) and French region Occitanie. All MS experiments were performed on the instruments of the MetaToul-AXIOM platform, partner of the national infrastructure of metabolomics and fluxomics: MetaboHUB [MetaboHUB-ANR-11-INBS-0010, 2011]. In the context of NutriNet-Santé study, we especially thank Younes Esseddik, Nathalie Druesne-Pecollo, Nathalie Arnault, Fabien Szabo, Julien Allegre and Laurent Bourhis

## 1. Introduction

Along with other omics approaches, such as transcriptomics and proteomics, metabolomics methods are now widely used to probe qualitative and quantitative variations in endogenous metabolites to assess the impact of various diseases or environmental conditions on metabolic homeostasis (Jansson et al., 2009; Wikoff et al., 2009). The study of metabolome modulation constitutes a powerful approach since the metabolome represents the most direct link to the phenotype compared to other “omics” (Fiehn, 2002). Nuclear magnetic resonance (NMR) and mass spectrometry (MS) are now established as the reference techniques used in metabolomics. NMR is known to be more robust than MS but also to give access to a smaller set of metabolites due to a lower sensitivity. Considering this lower sensitivity, MS is generally more widely used than NMR for metabolomics studies and will be the focus of this article. Among the various matrices analysed in metabolomics, urine is one of the most commonly used because of the non-invasive nature of its collection and the quantity of available samples. However, in contrast to other biofluids, such as plasma and cerebrospinal fluid, urine is far less regulated and can exhibit up to 14-fold variation in total metabolite concentrations (Warrack et al., 2009). Focusing on urine, numerous normalization strategies have emerged to mitigate the variations that are mainly attributable to non-relevant parameters such as hydration level, which can bias conclusions based on changes in metabolite concentration and introduces confounding factors.

Several normalization techniques based on various factors have been proposed for normalizing metabolomics data. Among them, as proposed by Edmands et al. (Edmands et al., 2014) and more recently by Gagnebin et al. (Gagnebin et al., 2017), two different approaches can be distinguished, namely, pre-analysis normalization often based on a biological parameter such as creatinine, specific gravity, and osmolality (Chen et al., 2013; Vogl et al., 2016; Warrack et al., 2009), and post-acquisition curative normalization. The creatinine concentration is one of the most commonly used reference factors because the urinary creatinine level is supposed to reflect the overall concentration of metabolites (Ryan et al., 2011). Indeed, creatinine excretion levels are assumed to remain constant within and across individuals since they are linked to muscle degradation and glomerular filtration (Waikar et al., 2010). However, its use as a normalization factor is currently under debate since factors such as diet, muscle mass, age, physical activity and menstrual cycle can affect creatinine levels (Cross et al., 2011; Davison and Noble, 1981; Décombaz et al., 1979; James et al., 1988; Skinner et al., 1996). The urine volume excreted during 24 h is also used as a normalization factor for metabolomics studies (Godzien et al., 2011) (Shih et al., 2019), but this parameter is not always accessible, especially in the case of human samples. Osmolality, which is considered one of the most reliable methods for evaluating overall urine metabolite concentration (Chadha et al., 2001), can also be used as a normalization factor in metabolomics (Yamamoto et al., 2019). The advantage of measuring osmolality is that it is only influenced by the number of particles and chemical entities dissolved in urine, thus better reflecting the overall urine metabolite concentration than creatinine level or total urine volume. However, osmolality measurement requires specific equipment. Directly linked to osmolality and more easily routinely measurable, specific gravity is also used to estimate urine metabolite concentrations in metabolomics studies (Edmands et al., 2014). However, this parameter, representing the ratio between the mass

of a given volume of urine and the mass of the same volume of water, less faithfully reflects the solute concentration, is more prone to measurement errors due to the presence of large molecules (Chadha et al., 2001; Cook et al., 2020; Voinescu et al., 2002) and requires larger sample volumes. Other data, such as the total NMR signal, can also be used to estimate the overall urine metabolite concentration. It should be noted that all these parameters can be used pre-analysis by applying specific dilutions or post acquisition by dividing the signal of detected metabolites. As reviewed by Wu and Li (Wu and Li, 2016), sample normalization methods in metabolomics are diverse and vary from one study to another, as well as the conclusion of studies evaluating the performance of normalization methods. In their study, Vogl et al. concluded that creatinine was the most efficient normalization method, even over osmolality (Vogl et al., 2016), while more recently, Rosen Vollmar et al. and Burton et al. determined that specific gravity was more reliable than creatinine (Burton et al., 2014; Rosen Vollmar et al., 2019), and others found better results with osmolality normalization (Chetwynd et al., 2016; Gagnebin et al., 2017).

In line with the diversity of pre-analytical normalization techniques, the post-analytical panel of approaches is also very diverse. Some of the most employed post-acquisition normalization methods, such as total ion current (TIC) normalization, MS total useful signal (MSTUS) (Warrack et al., 2009) or probabilistic quotient normalization (PQN) (Dieterle et al., 2006), are based on a ratio between each feature and a given spectral reference. With TIC normalization, each feature intensity is divided by the sum of all feature intensities in the sample of interest. As TIC normalization is sensitive to orphan features and xenobiotics, Warrack et al. (Warrack et al., 2009) proposed MSTUS normalization based on a normalization factor calculated from features common to all samples. This technique assumes that the number of decreasing and increasing metabolites is relatively equivalent across the samples, although this condition is not always met in all studies (Jacob et al., 2014). Initially, proposed to normalize NMR metabolomics data, the PQN method was also used to normalize MS metabolomics data. This method is based on the calculation of a dilution coefficient in relation to a reference spectrum, such as the median of quality control (QC) samples. In addition to these normalization methods, which aim to correct concentration differences, other algorithms such as LOESS (Dunn et al., 2011) have been used to correct analytical drift, which is of particular importance in MS-based metabolomics. In the face of this diversity, some tools, such as NormalizeMets (De Livera et al., 2018) or Normalyzer (Chawade et al., 2014), have been developed to evaluate the performance of various normalization methods for metabolomics data. Despite these tools, there is at present no consensus on the use of normalization methods in metabolomics or on how to evaluate their performances.

Here, we propose to evaluate the performance of several of the most employed metabolomics data normalization methods on the basis of a simple metabolomics study carried out by liquid chromatography mass spectrometry (LC-MS) on urine samples. As the conclusions of the literature remain unclear, it is still necessary to conduct multiple studies evaluating normalization methods to draw general conclusions. A cohort of forty individuals free of pathological conditions (that could bias normalization based on biological factors such as creatinine or osmolality) was chosen to perform this study. Considering this particular design relying on subjects



from the general population, we suggest that extrapolation of our study results could be more generally applicable than those obtained from studies comparing a control population to a diseased population that may bias normalization. The impact of specific dilutions of samples to normalize either creatinine or osmolality was studied and compared with post-acquisition normalization using these parameters, as well as the PQN and MSTUS mathematical normalization methods. A total of 14 normalization strategies were tested and compared. The comparison was based on several complementary criteria, including intra-group variance and residual variability reduction, percentage of explained biological variability of interest, and finally predictive capacity of statistical models.

## 2. Materials and methods

### 2.1 Chemicals

All solvents were purchased from Fisher Scientific (Thermo Fisher Scientific, Illkirch, France) and were of LC-MS grade.

### 2.2 Sample Preparation

Urine samples were collected from forty healthy individuals (20 women and 20 men) from the NutriNet-Santé study, an ongoing web-based observational prospective study launched in France in 2009 (Herberg et al., 2010) and including adult volunteers. The NutriNet-Santé study was conducted according to the Declaration of Helsinki guidelines and was approved by the Institutional Review Board of the French Institute for Health and Medical Research (IRB Inserm n°0000388FWA00005831) and the "Commission Nationale de l'Informatique et des Libertés" (CNIL n°908450/n°909216). Clinicaltrials.gov number: NCT03335644. At the clinical visit, urine sample collection was performed using vessels allowing for closed-circuit urine transfer from the vessel to a Vacutainer® tube. Urine samples were collected either in autumn/winter (n=23) or spring/summer (n=17) from 2011 to 2014. The Vacutainer® tubes containing the spot urine samples were kept at +4 °C before and during transportation to the central laboratory. After splitting into aliquots, urine samples were stored at -80 °C for further analyses. Samples were divided into three parts and were prepared as follows: the first part of the samples was diluted 1:3 (v/v) in LC mobile phase A (95:5 water/methanol + 0.1% acetic acid). The second part of the samples was diluted to standardize creatinine concentration based on the smallest value, determined by <sup>1</sup>H NMR according to a method adapted from Bouatra et al. (Bouatra et al., 2013), and then diluted 1:3 (v/v) in mobile phase A. The third part was specifically diluted to standardize osmolality levels based on the smallest value, determined using a freezing point osmometer (Loeser Messtechnik, Berlin, Germany) and then diluted 1:3 (v/v) in mobile phase A. After dilution, all samples were vortexed and centrifuged at 9500 g for 5 min. The supernatant was collected

and transferred into vials to be directly analysed by UHPLC-MS. QC samples were prepared by pooling equivalent volumes of all samples. QCs were diluted to obtain dQC samples.

### 2.3 Instrumentation

LC-MS analyses were performed using an ultra-performance liquid chromatography (UPLC) instrument (ACQUITY, Waters, Manchester, UK) coupled to a quadrupole time-of-flight mass spectrometer (Q-ToF Synapt G2-Si, Waters, Manchester, UK) as previously described (Jamin et al., 2020). A volume of 10  $\mu$ L was injected into a Hypersil Gold C18 (1.9  $\mu$ m, 100 mm  $\times$  2.1 mm) analytical column (Thermo Fisher Scientific, Illkirch, France) at a flow rate of 0.3 mL/min and maintained at 40 °C. A linear gradient programme was set up with mobile phase A: 95% H<sub>2</sub>O/5% methanol/0.1% acetic acid, and mobile phase B: 100% methanol/0.1% acetic acid. The initial conditions were 100% A, followed by a linear gradient from 0 to 100% B in 30 min. These conditions were held for 4 min prior to switching in 1 min to the starting conditions that were held for 5 min to equilibrate the column. Samples were analysed in three analytical batches, and QC samples were injected at the beginning of each batch and then every six samples to evaluate analytical drift.

Full MS acquisitions were achieved with electrospray ionization in negative mode as follows: capillary voltage, 0.5 kV; sampling cone voltage, 30 V; source temperature, 120 °C; desolvation temperature, 550 °C; cone gas (N<sub>2</sub>) flow rate, 30 L/h; and desolvation gas (N<sub>2</sub>) flow rate, 600 L/h. Data were acquired from mass-to-charge ratios (*m/z*) of 50 to 800. Internal mass calibration was performed by injecting leucine enkephalin at 1 ng/mL at a flow rate of 10  $\mu$ L/min. Data were acquired by operating the Synapt spectrometer in sensitivity mode and using dynamic range extension (DRE) to enhance both the sensitivity and dynamic range.

### 2.4 Data processing

Metabolomics data are complex, and there are multiple sources of variability, so it is sometimes difficult to isolate variability due to biological factors of interest. For this reason, it is necessary to define and identify the different sources of variability to apply appropriate filters and decrease some unwanted variabilities. It is therefore mandatory to distinguish between variability linked to the biological factor studied, which is called biological variability of interest, and variability not linked to the biological factor. The latter can be divided into several factors, such as technical variability, confounding biological variability and residual variability. Technical variability can, for instance, originate from sample preparation or analytical drift. Confounding biological variability is related to a known biological factor other than the factor of interest (e.g., biofluid concentration, BMI, etc.). Finally, residual variability corresponds to the remaining variability not explained by these previous factors. Full-HRMS data were processed using XCMS 3.0 with the collaborative environment Workflow4Metabolomics (W4M) dedicated to the processing of metabolomics data (Giacomini et al., 2015). Extraction of data from full scan spectra was carried out using the centwave algorithm (ppm = 10, peak width = (5,40), sntresh = 10, noise = 10000, bw = 10, mzwid = 0.01). Data were then filtered using a ratio of the average

signal in QCs over the average signal in blanks with a threshold of 3. We also eliminated all features displaying a correlation coefficient with dQC samples lower than 0.6. LOESS regression was applied based on QC samples to correct intra- and inter-batch drift. At the end, only common features displaying an RSD lower than 30% in QC samples were retained. PQN and MSTUS normalization were performed using the W4M environment. For PQN normalization, the median of the QCs was chosen as a reference.

## 2.5 Tested normalization strategies

The combination of pre- and post-acquisition normalizations led to the test of fourteen strategies of normalization presented in Table 1.

## 2.6 Statistical analyses

Statistical analyses were performed using Simca-P 14.0 software (Umetrics, Umea, Sweden) and MATLAB 2017b (The MathWorks, Inc.). All data were log transformed and Pareto scaled prior to principal component analysis (PCA) and orthogonal partial least square discriminant analysis (OPLS-DA). Log transformation was applied to reduce heteroscedasticity (Li et al., 2016) and Pareto to reduce the differences in intensity between features without increasing the analytical noise too much (Yang et al., 2015). PCA models were assessed using the explained variance parameter ( $R^2$ ), and the number of principal components was chosen based on a scree plot (Cattell, 1966). Predictive capacity ( $Q^2$ ) and permutation tests were used to check the validity and robustness of the O-PLS-DA models.  $Q^2$ , whose values are between 0 and 1, corresponds to the fraction of the total variation of the biological factor of interest that can be predicted by the O-PLS-DA model. The permutation test assesses whether the specific classification of individuals in the two designated groups is significantly better than any other random classification in two arbitrary groups (Westerhuis et al., 2008). A model is considered robust when it displays a  $Q^2$  intercept  $< 0.05$  (Lapins et al., 2008). When the optimal number of orthogonal components was higher than 1, the orthogonal component was manually added as long as the percentage of explained variance by this component was higher than 5%. The numbers of predictive and orthogonal components included in the OPLSDA models are displayed in Supplementary Material S.2. Seven-fold cross-validation was used to estimate the predictive ability of the adjusted model ( $Q^2$ ). However, O-PLS-DA cannot separate variability corresponding to the different factors of the experimental design (sex and normalization in this study). To reach this, analysis of variance multiblock orthogonal partial least squares (AMOPLS) (Boccard and Rudaz, 2016) was applied to assess the percentage of explained variance associated with each factor of the experimental design. This method works in two steps. In the first step, total variability is split according to the factors (normalization, sex and interaction, in this study). The significance of each effect is tested based on the permutation test (Vis et al., 2007). Then, multiblock OPLS was used to fit a multivariate model based on submatrices obtained from ANOVA decomposition.

### 3. Results and discussion

The objective of this work was to evaluate the performance of several of the most employed normalization methods for urine metabolomics. Urine samples from forty male and female individuals were analysed by UPLC-MS with or without normalization by specific dilution prior to acquisition. Data obtained from samples without pre-acquisition dilution were used to evaluate the impact of pre- and post-acquisition normalization. The evaluation of the performance of these different strategies was carried out based on the predictive ability of the generated statistical models, as well as the percentage of variance that could be attributed to the biological factor of interest, i.e., sex, in this study. Since metabolomics aims to highlight quantitative or qualitative differences in metabolite levels as a function of a biological factor, an ideal normalization method would reduce the variability not linked to the factor of interest while retaining the maximum amount of information related to the studied condition.

#### 3.1 Urine exhibits a large concentration distribution

Creatinine concentrations and osmolality were measured to estimate the overall sample concentration, allowing for the calculation of individual dilution factors to normalize either creatinine concentrations or osmolality levels. A large distribution of concentration levels was observed, ranging from 1.67 to 20.27 mol.L<sup>-1</sup> for creatinine and from 217 to 1236 mOsmol for osmolality (Supplementary S.1). Surprisingly, a significant difference was observed (Wilcoxon paired test, significance threshold 0.05) between calculated dilution factors from creatinine and osmolality, as shown in Fig. 22A. Dilution factors ranged from 1 to 12 for creatinine and from 1 to 5.7 for osmolality. Despite this disparity, a good correlation ( $r = 0.87$ ) was observed between the dilution factors calculated from the two parameters, as presented in Fig. 22B, indicating that both of these estimators are suitable to evaluate the overall metabolite concentration in urine. Because creatinine normalized samples are more diluted, the number of values missing before the FillPeaks step in the XCMS process was studied. No major differences were found between the three conditions, even a slight decrease in the number of missing values was observed with creatinine-normalized samples (data not shown).

To ensure that urine normalization based on these two parameters would not introduce bias in the discrimination of the two population groups, univariate statistical analyses (Mann-Whitney test, significance threshold 0.05) were performed on creatinine and osmolality levels, and no significant differences between men and women were observed for either of these parameters (data not shown). These results underlined the large concentration distribution in urine sample metabolite concentrations.

#### 3.2 Normalization decreases overall variability

As shown in part 3.1, osmolality levels and creatinine concentrations exhibit large variations that were possibly responsible for a large part of the sample set variability. Dilution according to creatinine and osmolality levels

aimed to normalize the dataset and reduce the variability linked to overall urinary concentration.

To assess the impact of normalization methods on the overall variability, which corresponds to both biological variability and variability not linked to the factor of interest, PCA was first performed on log-transformed data. This allowed for the comparison of the non-normalized dataset to the pre- and post-acquisition normalized datasets using both creatinine and osmolality-based normalization. A total of 4758 out of the 17366 initially detected features were kept after the filtration process and subjected to PCA. The variance in the coordinates of individuals in each dataset on the first two components was calculated to assess the impact of normalization on data dispersion. As displayed in Fig. 23 A, samples diluted according to osmolality levels (red triangles) exhibit less dispersion on the PCA plot than samples adjusted according to creatinine concentration (green circles) or non-normalized samples (blue squares). This observation was confirmed by calculating the variance of individuals in each dataset where osmolality normalization showed a strong decrease in overall variance (var = 257.9) compared to creatinine normalization (var = 375.2) or the non-normalized dataset (var = 422.32).

Post-acquisition normalization achieved by dividing the signal intensity of each feature for a given sample by the calculated dilution factor was also assessed. The latter approach is less constrained since post-acquisition normalization avoids pre-acquisition sample dilution by a pre-determined factor, which can be tedious, especially with large sample cohorts. For the pre-acquisition normalized samples, PCA was performed from the non-normalized dataset and the dataset normalized by dilution factors calculated either from creatinine or osmolality levels. As presented in Fig. 23 B, when compared to the non-normalized dataset (var = 524.5), variances in the coordinates of individuals showed a decrease in overall variability when normalization was performed on osmolality (var = 329.4) levels but not with creatinine-based normalization (var = 582.6).

These results suggest that regardless of the strategy employed, osmolality-based normalization leads to a lower overall variability than the other normalization methods. PCA combining all datasets was performed to determine which of the pre-acquisition or post-acquisition strategies led to the greatest reduction in overall variability. The variance in coordinates was 239.2 for diluted samples and 329.4 for post-acquisition normalized samples (data not shown), suggesting that post-acquisition normalization is less efficient than pre-acquisition normalization for decreasing global variability. This can be tentatively explained by the fact that major differences in urine concentration induce strong differences in matrix effects, which can lead to ion suppression in MS analyses and thus to an increase in signal variability. Therefore, using a specific dilution to standardize urine concentration prior to mass spectrometric analysis could decrease the differences in matrix effects linked to urine concentration.

### 3.3 Normalization improves statistical discrimination

Most metabolomics studies are based on the comparison of relative metabolite abundances between two or more groups. As a consequence, decreasing variability not linked to the biological factor by applying normalization techniques would result in an increase in the statistical discrimination of groups. In this study, to assess the effect of various normalization methods, we relied on the statistical power of supervised OPLS-DA

models to discriminate two groups composed of 20 men or 20 women.

Fourteen normalization strategies were tested and compared with the non-normalized dataset chosen as the reference. The results displayed in Fig. 24 show that statistical models built with the pre-acquisition normalized dataset led to better results than those built with the non-normalized dataset (grey line), especially for osmolality normalization in terms of the predictive ability of the model ( $Q^2 = 0.52$ ).  $Q^2$  values for all models are displayed in Supplementary S.2

These results are in accordance with the decrease in overall variability observed in Fig. 23, suggesting that pre-acquisition normalization reduced variability not linked to the biological factor of interest while retaining biological information, according to Fig. 24. On the other hand, pre-acquisition normalization according to creatinine concentration only led to a slight improvement in statistical segregation ( $Q^2 = 0.42$ ) compared to the non-normalized dataset ( $Q^2 = 0.35$ ). Interestingly, when normalization was performed on osmolality or creatinine levels after data acquisition, no improvement was observed for fitted OPLS-DA models, although a reduction in overall variability was observed, especially for osmolality normalization.

We further tested whether post-acquisition methods such as PQN (Dieterle et al., 2006) or MSTUS (Warrack et al., 2009) could supplant pre-acquisition normalization for improving statistical discrimination. We observed that PQN led to similar results ( $Q^2 = 0.33$ ) as MSTUS ( $Q^2 = 0.33$ ), explained by the fact that we kept only features that were common to all samples during the data filtration. Indeed, the first step of the PQN normalization consisted of a TIC normalization before performing a normalization with respect to a reference sample (i.e., the median of QC samples in this case). As all the features of the dataset were shared between all the samples, MSTUS normalization corresponded here to TIC normalization. Thus, in this particular case, the only difference between PQN and MSTUS normalization was the division by a constant determined by the median of the QC samples and thus led to the same results. By comparing the results obtained with mathematical post-acquisition methods to those obtained with pre-acquisition normalization, post-acquisition methods were found to be less efficient than specific dilutions. Finally, the combination of both biological- and mathematical-based normalization was investigated. No statistical improvement was observed when applying this combination ( $Q^2 = 0.33$ ). Surprisingly, when combining two post-acquisition methods, the results were identical. Indeed, this combination consists of dividing a feature by its calculated dilution factor and then by the sum of features itself divided by the same dilution constant. In total, this nullifies the post-acquisition normalization based on the dilution factor. However, when pre- and post-acquisition methods were combined, no statistical improvement was observed, suggesting that part of the biological variability associated with sex was removed by this sequential process. Indeed, the statistical model performed with the Pre-acq osmo dataset led to better predictivity ( $Q^2 = 0.52$ ) than the models resulting from the combination of Pre-acq osmo with MSTUS ( $Q^2 = 0.46$ ) or with PQN ( $Q^2 = 0.46$ ). The same results were obtained for the creatinine normalized dataset.

Taken together, these results showed that the applied normalization method can act on the predictability of a statistical model. In total, 6 out of the 14 tested normalizations led to an increase in the discrimination power of the adjusted models, and notable differences were observed between normalization strategies. Normalization by specific dilution according to osmolality levels outperformed other methods. Interestingly, no improvement

(and even a decrease in predictability) was obtained when applying this method in combination with PQN or MSTUS.

### 3.4 Pre-acquisition normalization reduces residual variability

To further investigate the link between the overall decrease in variability and the statistical power improvement, the ability of these various normalization strategies to reduce residual variability while keeping biological information was investigated by splitting the total variance into parts corresponding to the different factors based on AMOPLS.

As presented in Fig. 25 A, the percentage of variability associated with sex ranged from 2.9 to 3.6% depending on the applied normalization method. In line with previous findings, normalization by specific dilution based on osmolality levels corresponded to the higher percentage of variability by explained sex with 3.6% variability. Post-acquisition normalization based on osmolality or creatinine levels did not lead to such an improvement, although the percentage of variability explained by sex was higher than that in the reference dataset. Interestingly, no difference was observed when applying post-acquisition methods such as PQN or MSTUS compared to the non-normalized dataset. This is also in accordance with our previous results and suggests that over-normalization can reduce variability with a negative impact on the discrimination power. Despite the slight improvement in the percentage of variability linked to sex, it is noteworthy that sex was only a significant factor ( $p=0.04$ , 500 permutations) when pre-acquisition osmolality normalization is performed, in contrast to other tested normalizations. Residual variability, which corresponds to unexplained variance, can mask variability linked to sex. An ideal method would reduce residual variability while increasing the percentage of variance explained by the biological factor of interest. As illustrated in Fig. 25 B, residual variability was also affected by normalization. The lowest residual variability was associated with the pre-acquisition normalization based on osmolality levels with 75% residual variability, while it was between 75 and 76.5% for other normalizations. Here, again, the combination of methods with MSTUS or PQN or the use of PQN or MSTUS alone did not lead to a greater reduction in residual variability.

These findings are in accordance with previous observations, showing that the statistical improvement observed in section 3.3 for pre-acquisition osmolality-based dilution is linked to a decrease in residual variability and an increase in the percentage of variance associated with the biological effect.

## 4. Conclusion

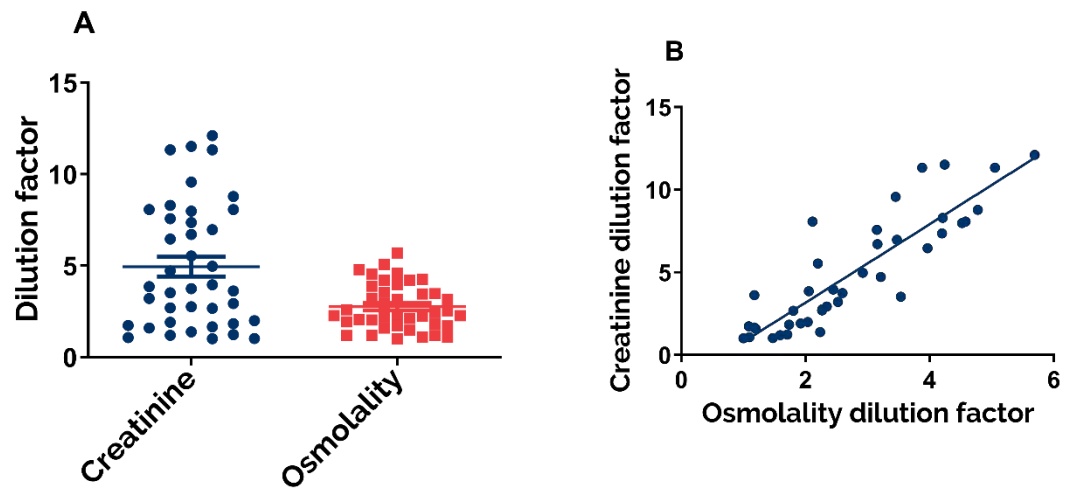
Normalization of data is often a pre-requisite for obtaining unbiased results in statistical analyses. When applied to metabolomics data, normalization is especially important for matrices such as urine, which are highly variable. Fourteen normalization strategies often used in metabolomics were compared to determine the most suitable normalization for urine metabolomics, considering their performance for overall variability reduction,

enhancement of statistical power and unwanted variability reduction. Our results showed that the normalization method using urine-specific dilution according to osmolality levels gave the best results. Surprisingly, the combination of this method with PQN or MSTUS did not improve our assessment criteria and even led to a decrease in statistical discrimination, likely due to over-normalization, which could partially suppress the variability of interest. This shows that the combination of normalization methods could decrease statistical power and should not be systematically applied. Moreover, not all fourteen tested methods led to improved performances compared to the non-normalized dataset, reinforcing the recommendation to apply a suitable normalization strategy for untargeted urine metabolomics studies.

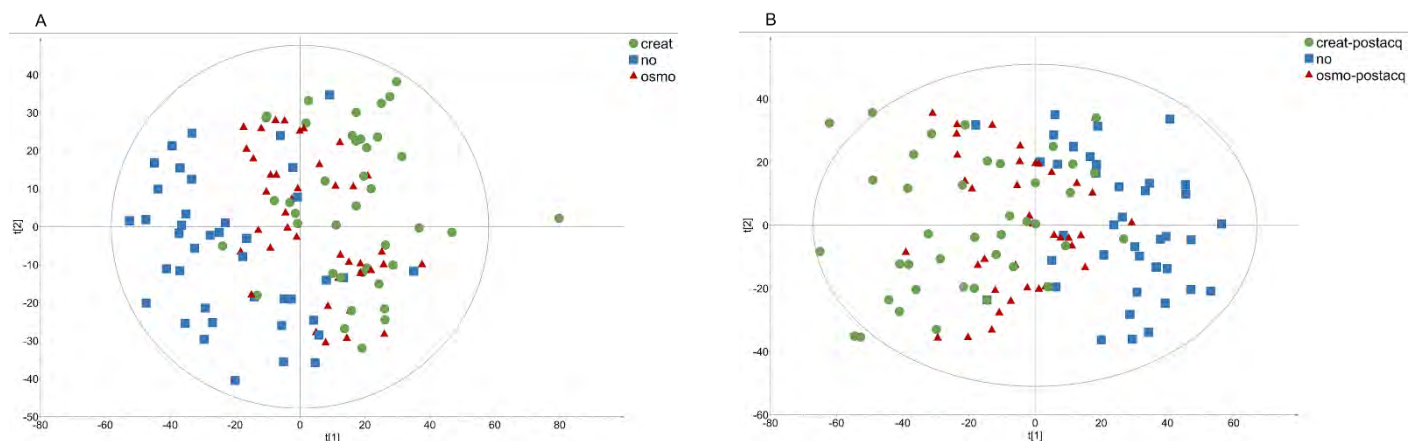


Table.1 Description of the fourteen normalization techniques assessed in this work

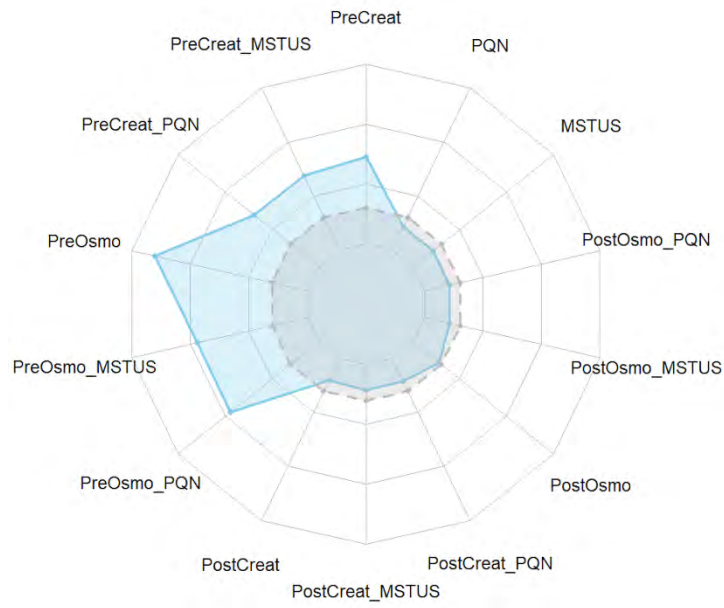
Name	Definition
Pre-acq osmo	Specific dilution applied prior to the analysis according to osmolality levels
Pre-acq creat	Specific dilution applied prior to the analysis according to creatinine cocentration
Post-acq osmo	Signal of each feature of a given sample divided by its calculated dilution factor according to osmolality level
Post-acq creat	Signal of each feature of a given sample divided by its calculated dilution factor according to creatinine concentration
PQN	PQN normalization
MSTUS	MSTUS normalization
Pre-acq osmo + PQN	Combination of pre-acq osmo and PQN
Pre-acq osmo + MSTUS	Combination of pre-acq osmo and MSTUS
Post-acq osmo + PQN	Combination of post-acq osmo and PQN
Post-acq osmo + MSTUS	Combination of post-acq osmo and MSTUS
Pre-acq creat + PQN	Combination of pre-acq creat and PQN
Pre-acq creat + MSTUS	Combination of pre-acq creat and MSTUS
Post-acq creat + PQN	Combination of post-acq creat and PQN



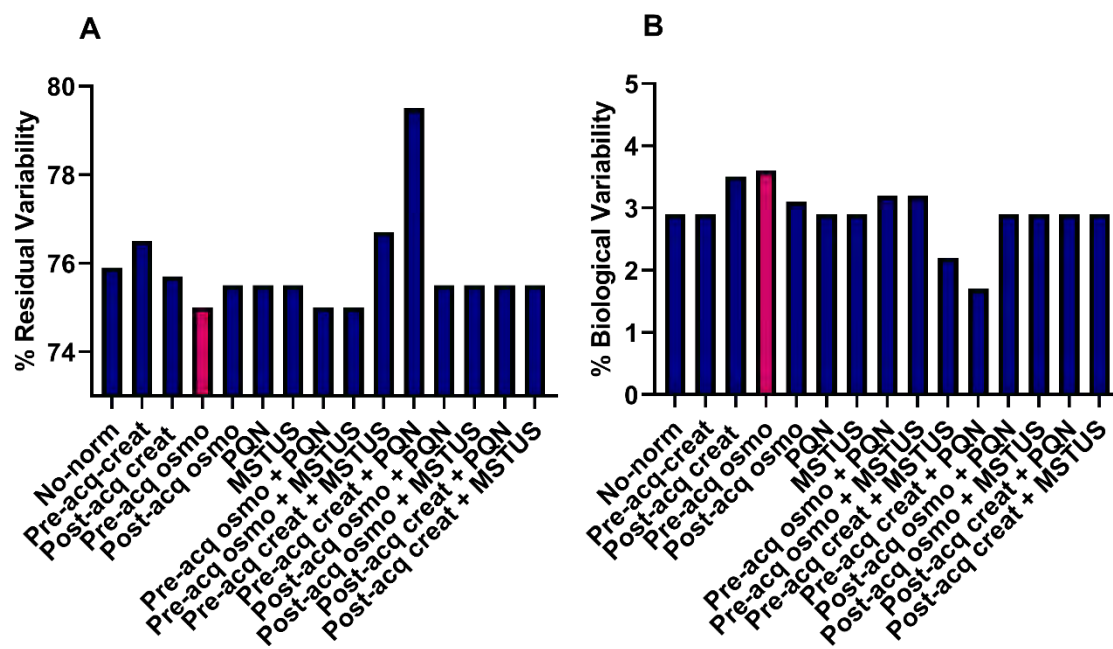
**Fig. 22** Concentration variations in urine samples. (A) Dilution factors calculated for both creatinine and osmolality. (B) Correlation of dilution factors calculated according to creatinine levels and osmolality levels



**Fig. 23** Two-dimensional PCA score plot of HR-MS spectra of urine samples. Each symbol represents an individual data point projected onto the first (horizontal axis) and second (vertical axis) PCA components. Normalization methods are shown in different colours: creatinine in green dots (n=40), none in blue squares (n=40), and osmolality in red triangles (n=40). The black ellipse determines the 95% confidence interval, which was drawn using Hotelling's T2 statistic. (A) Preacquisition normalization:  $A = 2$ ,  $R^2 = 20\%$ . (B) Post-acquisition normalization:  $A = 2$ ,  $R^2 = 24\%$



**Fig. 24** Radar chart of  $Q^2$  of OPLS-DA models generated for all tested normalization strategies (blue line) compared to  $Q^2$  of OPLS-DA generated from non-normalized model (grey line). For detailed values of  $Q^2$ , see Supplementary S.2



**Fig. 25** Percentage of (A) variability explained by sex. (B) Residual variability determined by AMOPLS for all tested normalizations. The red bar represents the results from data normalized prior to the acquisition based on osmolality levels

## 5. References

- Boccard, J., Rudaz, S., 2016. Exploring Omics data from designed experiments using analysis of variance multiblock Orthogonal Partial Least Squares. *Anal. Chim. Acta* 920, 18–28. <https://doi.org/10.1016/j.aca.2016.03.042>
- Bouatra, S., Aziat, F., Mandal, R., Guo, A.C., Wilson, M.R., Knox, C., Bjorndahl, T.C., Krishnamurthy, R., Saleem, F., Liu, P., Dame, Z.T., Poelzer, J., Huynh, J., Yallou, F.S., Psychogios, N., Dong, E., Bogumil, R., Roehring, C., Wishart, D.S., 2013. The Human Urine Metabolome. *PLOS ONE* 8, e73076. <https://doi.org/10.1371/journal.pone.0073076>
- Burton, C., Shi, H., Ma, Y., 2014. Normalization of urinary pteridines by urine specific gravity for early cancer detection. *Clin. Chim. Acta* 435, 42–47. <https://doi.org/10.1016/j.cca.2014.04.022>
- Cattell, R.B., 1966. The Scree Test For The Number Of Factors. *Multivar. Behav. Res.* 1, 245–276. [https://doi.org/10.1207/s15327906mbr0102\\_10](https://doi.org/10.1207/s15327906mbr0102_10)
- Chadha, V., Garg, U., Alon, U.S., 2001. Measurement of urinary concentration: a critical appraisal of methodologies. *Pediatr. Nephrol.* 16, 374–382. <https://doi.org/10.1007/s004670000551>
- Chawade, A., Alexandersson, E., Levander, F., 2014. Normalyzer: A Tool for Rapid Evaluation of Normalization Methods for Omics Data Sets. *J. Proteome Res.* 13, 3114–3120. <https://doi.org/10.1021/pr401264n>
- Chen, Y., Shen, G., Zhang, R., He, J., Zhang, Y., Xu, J., Yang, W., Chen, X., Song, Y., Abliz, Z., 2013. Combination of Injection Volume Calibration by Creatinine and MS Signals' Normalization to Overcome Urine Variability in LC-MS-Based Metabolomics Studies. *Anal. Chem.* 85, 7659–7665. <https://doi.org/10.1021/ac401400b>
- Chetwynd, A.J., Abdul-Sada, A., Holt, S.G., Hill, E.M., 2016. Use of a pre-analysis osmolality normalisation method to correct for variable urine concentrations and for improved metabolomic analyses. *J. Chromatogr. A* 1431, 103–110. <https://doi.org/10.1016/j.chroma.2015.12.056>
- Cook, T., Ma, Y., Gamagedara, S., 2020. Evaluation of statistical techniques to normalize mass spectrometry-based urinary metabolomics data. *J. Pharm. Biomed. Anal.* 177, 112854. <https://doi.org/10.1016/j.jpba.2019.112854>
- Cross, A.J., Major, J.M., Sinha, R., 2011. Urinary biomarkers of meat consumption. *Cancer Epidemiol. Biomark. Prev. Publ. Am. Assoc. Cancer Res. Cosponsored Am. Soc. Prev. Oncol.* 20, 1107–1111. <https://doi.org/10.1158/1055-9965.EPI-11-0048>
- Davison, J.M., Noble, M.C.B., 1981. Serial Changes in 24 Hour Creatinine Clearance During Normal Menstrual Cycles and the First Trimester of Pregnancy. *BJOG Int. J. Obstet. Gynaecol.* 88, 10–17. <https://doi.org/10.1111/j.1471-0528.1981.tb00930.x>

De Livera, A.M., Olshansky, G., Simpson, J.A., Creek, D.J., 2018. NormalizeMets: assessing, selecting and implementing statistical methods for normalizing metabolomics data. *Metabolomics* 14, 54. <https://doi.org/10.1007/s11306-018-1347-7>

Décombaz, J., Reinhardt, P., Anantharaman, K., von Glutz, G., Poortmans, J.R., 1979. Biochemical changes in a 100 km run: Free amino acids, urea, and creatinine. *Eur. J. Appl. Physiol.* 41, 61–72. <https://doi.org/10.1007/BF00424469>

Dieterle, F., Ross, A., Schlotterbeck, G., Senn, H., 2006. Probabilistic quotient normalization as robust method to account for dilution of complex biological mixtures. Application in <sup>1</sup>H NMR metabonomics. *Anal. Chem.* 78, 4281–4290. <https://doi.org/10.1021/ac051632c>

Dunn, W.B., Broadhurst, D., Begley, P., Zelena, E., Francis-McIntyre, S., Anderson, N., Brown, M., Knowles, J.D., Halsall, A., Haselden, J.N., Nicholls, A.W., Wilson, I.D., Kell, D.B., Goodacre, R., Human Serum Metabolome (HUSERMET) Consortium, 2011. Procedures for large-scale metabolic profiling of serum and plasma using gas chromatography and liquid chromatography coupled to mass spectrometry. *Nat. Protoc.* 6, 1060–1083. <https://doi.org/10.1038/nprot.2011.335>

Edmands, W.M.B., Ferrari, P., Scalbert, A., 2014. Normalization to Specific Gravity Prior to Analysis Improves Information Recovery from High Resolution Mass Spectrometry Metabolomic Profiles of Human Urine. *Anal. Chem.* 86, 10925–10931. <https://doi.org/10.1021/ac503190m>

Fiehn, O., 2002. Metabolomics--the link between genotypes and phenotypes. *Plant Mol. Biol.* 48, 155–171.

Gagnebin, Y., Tonoli, D., Lescuyer, P., Ponte, B., de Seigneux, S., Martin, P.-Y., Schappler, J., Boccard, J., Rudaz, S., 2017. Metabolomic analysis of urine samples by UHPLC-QTOF-MS: Impact of normalization strategies. *Anal. Chim. Acta* 955, 27–35. <https://doi.org/10.1016/j.aca.2016.12.029>

Giacomoni, F., Le Corguillé, G., Monsoor, M., Landi, M., Pericard, P., Pétéra, M., Duperier, C., Tremblay-Franco, M., Martin, J.-F., Jacob, D., Goulitquer, S., Thévenot, E.A., Caron, C., 2015. Workflow4Metabolomics: a collaborative research infrastructure for computational metabolomics. *Bioinformatics* 31, 1493–1495. <https://doi.org/10.1093/bioinformatics/btu813>

Godzien, J., Ciborowski, M., Angulo, S., Ruperez, F.J., Paz Martínez, M., Señorans, F.J., Cifuentes, A., Ibañez, E., Barbas, C., 2011. Metabolomic Approach with LC-QTOF to Study the Effect of a Nutraceutical Treatment on Urine of Diabetic Rats. *J. Proteome Res.* 10, 837–844. <https://doi.org/10.1021/pr100993x>

Hercberg, S., Castetbon, K., Czernichow, S., Malon, A., Mejean, C., Kesse, E., Touvier, M., Galan, P., 2010. The Nutrinet-Santé Study: a web-based prospective study on the relationship between nutrition and health and determinants of dietary patterns and nutritional status. *BMC Public Health* 10, 242. <https://doi.org/10.1186/1471-2458-10-242>

Jacob, C.C., Dervilly-Pinel, G., Biancotto, G., Le Bizec, B., 2014. Evaluation of specific gravity as normalization strategy for cattle urinary metabolome analysis. *Metabolomics* 10, 627–637. <https://doi.org/10.1007/s11306-013-0604-z>

James, G.D., Sealey, J.E., Alderman, M., Ljungman, S., Mueller, F.B., Pecker, M.S., Laragh, J.H., 1988. A Longitudinal Study of Urinary Creatinine and Creatinine Clearance in Normal Subjects Race, Sex, and Age Differences. *Am. J. Hypertens.* 1, 124–131. <https://doi.org/10.1093/ajh/1.2.124>

Jamin, E.L., Costantino, R., Mervant, L., Martin, J.-F., Jouanin, I., Blas-Y-Estrada, F., Guéraud, F., Debrauwer, L., 2020. Global Profiling of Toxicologically Relevant Metabolites in Urine: Case Study of Reactive Aldehydes. *Anal. Chem.* 92, 1746–1754. <https://doi.org/10.1021/acs.analchem.9b03146>

Jansson, J., Willing, B., Lucio, M., Fekete, A., Dicksved, J., Halfvarson, J., Tysk, C., Schmitt-Kopplin, P., 2009. Metabolomics Reveals Metabolic Biomarkers of Crohn's Disease. *PLOS ONE* 4, e6386. <https://doi.org/10.1371/journal.pone.0006386>

Lapins, M., Eklund, M., Spjuth, O., Prusis, P., Wikberg, J.E., 2008. Proteochemometric modeling of HIV protease susceptibility. *BMC Bioinformatics* 9, 181. <https://doi.org/10.1186/1471-2105-9-181>

Li, B., Tang, J., Yang, Q., Cui, X., Li, S., Chen, S., Cao, Q., Xue, W., Chen, N., Zhu, F., 2016. Performance Evaluation and Online Realization of Data-driven Normalization Methods Used in LC/MS based Untargeted Metabolomics Analysis. *Sci. Rep.* 6, 38881. <https://doi.org/10.1038/srep38881>

Rosen Vollmar, A.K., Rattray, N.J.W., Cai, Y., Santos-Neto, Á.J., Deziel, N.C., Jukic, A.M.Z., Johnson, C.H., 2019. Normalizing Untargeted Periconceptional Urinary Metabolomics Data: A Comparison of Approaches. *Metabolites* 9. <https://doi.org/10.3390/metabo9100198>

Ryan, D., Robards, K., Prenzler, P.D., Kendall, M., 2011. Recent and potential developments in the analysis of urine: A review. *Anal. Chim. Acta* 684, 17–29. <https://doi.org/10.1016/j.aca.2010.10.035>

Shih, C.-L., Wu, H.-Y., Liao, P.-M., Hsu, J.-Y., Tsao, C.-Y., Zgoda, V.G., Liao, P.-C., 2019. Profiling and comparison of toxicant metabolites in hair and urine using a mass spectrometry-based metabolomic data processing method. *Anal. Chim. Acta* 1052, 84–95. <https://doi.org/10.1016/j.aca.2018.11.009>

Skinner, A.M., Addison, G.M., Price, D.A., 1996. Changes in the urinary excretion of creatinine, albumin and N-acetyl- $\beta$ -D-glucosaminidase with increasing age and maturity in healthy schoolchildren. *Eur. J. Pediatr.* 155, 596–602. <https://doi.org/10.1007/BF01957912>

Vis, D.J., Westerhuis, J.A., Smilde, A.K., van der Greef, J., 2007. Statistical validation of megavariate effects in ASCA. *BMC Bioinformatics* 8, 322. <https://doi.org/10.1186/1471-2105-8-322>

Vogl, F.C., Mehrl, S., Heizinger, L., Schlecht, I., Zacharias, H.U., Ellmann, L., Nürnberger, N., Gronwald, W., Leitzmann, M.F., Rossert, J., Eckardt, K.-U., Dettmer, K., Oefner, P.J., GCKD Study Investigators, 2016. Evaluation



of dilution and normalization strategies to correct for urinary output in HPLC-HRTOFMS metabolomics. *Anal. Bioanal. Chem.* 408, 8483–8493. <https://doi.org/10.1007/s00216-016-9974-1>

Voinescu, G.C., Shoemaker, M., Moore, H., Khanna, R., Nolph, K.D., 2002. The Relationship between Urine Osmolality and Specific Gravity. *Am. J. Med. Sci.* 323, 39–42. <https://doi.org/10.1097/00000441-200201000-00007>

Waikar, S.S., Sabbiseti, V.S., Bonventre, J.V., 2010. Normalization of urinary biomarkers to creatinine during changes in glomerular filtration rate. *Kidney Int.* 78, 486–494. <https://doi.org/10.1038/ki.2010.165>

Warrack, B.M., Hnatyshyn, S., Ott, K.-H., Reily, M.D., Sanders, M., Zhang, H., Drexler, D.M., 2009. Normalization strategies for metabonomic analysis of urine samples. *J. Chromatogr. B* 877, 547–552. <https://doi.org/10.1016/j.jchromb.2009.01.007>

Westerhuis, J.A., Hoefsloot, H.C.J., Smit, S., Vis, D.J., Smilde, A.K., van Velzen, E.J.J., van Duijnhoven, J.P.M., van Dorsten, F.A., 2008. Assessment of PLS-DA cross validation. *Metabolomics* 4, 81–89. <https://doi.org/10.1007/s11306-007-0099-6>

Wikoff, W.R., Anfora, A.T., Liu, J., Schultz, P.G., Lesley, S.A., Peters, E.C., Siuzdak, G., 2009. Metabolomics analysis reveals large effects of gut microflora on mammalian blood metabolites. *Proc. Natl. Acad. Sci.* 106, 3698–3703. <https://doi.org/10.1073/pnas.0812874106>

Wu, Y., Li, L., 2016. Sample normalization methods in quantitative metabolomics. *J. Chromatogr. A, Editors' Choice X* 1430, 80–95. <https://doi.org/10.1016/j.chroma.2015.12.007>

Yamamoto, M., Pinto-Sanchez, M.I., Bercik, P., Britz-McKibbin, P., 2019. Metabolomics reveals elevated urinary excretion of collagen degradation and epithelial cell turnover products in irritable bowel syndrome patients. *Metabolomics Off. J. Metabolomic Soc.* 15, 82. <https://doi.org/10.1007/s11306-019-1543-0>

Yang, J., Zhao, X., Lu, X., Lin, X., Xu, G., 2015. A data preprocessing strategy for metabolomics to reduce the mask effect in data analysis. *Front. Mol. Biosci.* 2. <https://doi.org/10.3389/fmolb.2015.00004>

**Supplementary Material.1** Samples characteristics

Sample Name	sexe	age	BMI	Osmolality (mosmol)	Creatinine (mmol.L-1)
BOX2-27	H	47	27,72	238	1,79
BOX2-91	H	50	21,06	371	2,06
BOX2-8	H	71	28,96	236	2,89
BOX1-10	H	67	28,94	255	2,77
BOX2-85	H	49	25,19	256	6,06
BOX2-34	H	56	22,30	1236	20,27
BOX2-55	H	62	22,66	982	13,36
BOX2-81	H	59	23,83	914	13,88
BOX2-77	H	56	24,66	861	10,82
BOX2-51	H	68	27,07	446	6,45
BOX2-2	H	45	25,19	564	6,26
BOX2-21	H	65	22,31	531	6,62
BOX2-18	H	37	25,01	548	5,37
BOX2-10	H	56	26,91	1097	18,97
BOX2-60	H	71	27,75	754	11,67
BOX2-36	H	71	33,58	260	2,66
BOX2-43	H	70	23,15	842	18,97
BOX2-56	H	60	24,41	635	8,32
BOX2-25	H	61	24,34	994	13,50
BOX2-35	H	53	19,95	698	7,90
BOX2-38	F	53	17,82	217	1,67
BOX2-52	F	51	20,89	320	1,70
BOX2-4	F	68	26,62	346	2,00
BOX2-40	F	54	22,76	442	3,33
BOX2-53	F	63	24,06	493	4,62
BOX2-44	F	54	23,53	459	13,51
BOX2-42	F	70	20,43	478	9,26
BOX2-5	F	57	20,08	492	4,50
BOX2-32	F	32	21,11	418	3,20
BOX2-31	F	64	21,99	392	4,47
BOX1-4	F	63	24,23	912	12,32
BOX2-39	F	41	23,30	509	4,90
BOX2-19	F	73	24,22	377	3,07
BOX2-33	F	66	22,44	1037	14,70
BOX1-6	F	74	27,14	768	5,90
BOX2-14	F	47	19,54	921	19,27
BOX1-1	F	54	21,63	684	12,66
BOX2-50	F	48	32,39	750	16,02
BOX2-69	F	53	24,39	486	2,31
BOX2-59	F	61	20,60	686	11,23

**Supplementary Material.2** Q<sup>2</sup> values of OPLS-DA models generated depending on tested normalization

Tested Normalization	OPLS-DA			Permutation test (intercept value)	
	Number of latent variables	% of orthogonal variation removed	Cross validated Q <sup>2</sup> OPLS-DA	R <sup>2</sup>	Q <sup>2</sup>
Pre-acq osmo	1+2	15.4	0.518	0.983	-0.119
Pre-acq creat	1+2	13.1	0.421	0.987	-0.11
Post-acq osmo	1+1	8.0	0.347	0.929	-0.152
Post-acq creat	1+2	21.4	0.334	0.968	-0.137
PQN	1+2	14.9	0.336	0.985	-0.068
MSTUS	1+2	14.9	0.336	0.984	-0.082
Pre-acq osmo + PQN	1+2	15.1	0.457	0.984	-0.105
Pre-acq osmo + MSTUS	1+2	15.1	0.457	0.985	-0.107
Post-acq osmo + PQN	1+2	14.9	0.336	0.984	-0.076
Post-acq osmo + MSTUS	1+2	14.9	0.336	0.984	-0.085
Pre-acq creat + PQN	1+2	14.5	0.415	0.987	-0.088
Pre-acq creat + MSTUS	1+2	14.5	0.415	0.987	-0.119
Post-acq creat + PQN	1+2	14.9	0.336	0.984	-0.088
Post-acq creat + MSTUS	1+2	14.9	0.336	0.984	-0.958

### ***C.3.2. Principaux résultats et conclusions de l'article***

Cette étude a permis de souligner la grande variabilité interindividuelle en matière de concentration urinaire. En effet, en considérant la créatinine comme marqueur de concentration, un facteur 12 a été retrouvé entre l'échantillon le moins concentré et celui le plus concentré, tandis que ce facteur était de 6 en considérant l'osmolalité. Les données sur la variabilité globale, déterminée à partir de la dispersion des coordonnées des individus en ACP, ont montré que la normalisation pré-acquisition basée sur l'osmolalité permettait de réduire au maximum cette dispersion. Bien que, dans l'étude, la variance ait été calculée seulement sur les deux premières composantes, les mêmes tendances ont été retrouvées pour la variance des coordonnées sur l'ensemble des composantes. En plus de la réduction de la variabilité globale, la normalisation pré-acquisition sur l'osmolalité a montré la plus grande réduction de la variabilité résiduelle tout en maximisant la variabilité biologique d'intérêt. Finalement, l'ensemble de ces résultats a conduit à une meilleure discrimination statistique des deux groupes expérimentaux (hommes et femmes) lorsque les échantillons étaient normalisés sur l'osmolalité avant analyse. Il est également à noter que les combinaisons de plusieurs méthodes de normalisation n'ont pas conduit à de meilleurs résultats, probablement en raison de la suppression d'une partie de la variabilité d'intérêt. Cependant, même si les gains en matière de réduction de variabilité résiduelle et de pourcentage de variance expliquée par le facteur biologique d'intérêt semblent ténus, il est intéressant de constater que la normalisation sur l'osmolalité a conduit aux meilleurs résultats pour l'ensemble des critères d'évaluation. En conclusion, cette étude a de nouveau démontré l'importance de normaliser les échantillons urinaires et a également pointé l'importance du choix d'une méthode de normalisation appropriée. Les résultats obtenus suggèrent qu'une normalisation réalisée avant l'acquisition afin d'uniformiser l'osmolalité constitue la méthode de normalisation la plus performante pour des études de métabolomique non-ciblée.

#### C.4. Optimisations supplémentaires mode MS<sup>E</sup>

Lors de l'analyse des échantillons de la cohorte NutriNet-Santé, dont les résultats sont présentés dans le Chapitre III : Etude de l'impact de la consommation de viande rouge et de charcuterie, un nombre restreint de mercapturates ont été détectés avec les paramètres de traitement de données précédemment décrits. Ainsi, pour optimiser la détection des mercapturates par le logiciel UNIFI, un total de 6 méthodes (disponibles en Annexe 4) ont été comparées pour déterminer la combinaison de paramètres permettant la détection du plus grand nombre de mercapturates. La Figure 26 montre le nombre de mercapturates détectés en fonction de la méthode utilisée, ainsi que la fréquence de détection de ces espèces au sein de l'un des QC de groupe réalisés dans cette étude, le groupe des « gros-consommateurs ».

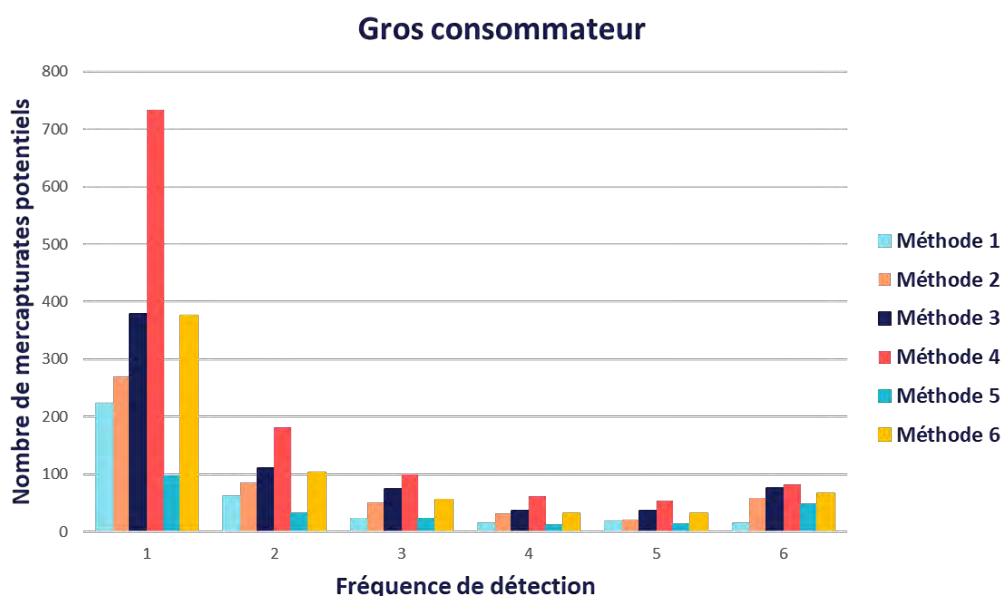


Figure 26 Fréquence de détection et nombre de mercapturates potentiels détectés au sein du groupe des gros consommateurs de viande rouge et de charcuterie en fonction de la méthode de traitement des données utilisée

Il apparaît que le jeu de paramètres correspondant à la Méthode 4 permet la détection du plus grand nombre d'espèces au total mais aussi le plus grand nombre d'espèces détectées dans chaque réplikat. Cette augmentation du nombre d'espèces détectées s'explique par l'abaissement du seuil d'intensité pour considérer les pics dans le scan « high-energy ». Bien que la diminution de ce seuil puisse conduire à un plus grand nombre de faux positifs, la méthode 4 sera conservée car permettant la détection du plus grand nombre de composés. Il est à noter que des résultats identiques sont observés pour les deux autres pools analysés. Afin de vérifier la justesse de la méthode, l'ensemble des candidats retrouvés dans chaque réplikat a été fragmenté en MS/MS ciblée. La Figure 27A illustre le recouvrement de la détection des mercapturates entre les différents pools. Au total, sur 147 espèces fragmentées, 56 mercapturates sont validés, ce qui correspond à un taux de faux positif de 62 %. Sur ces 56 mercapturates, 47 ont pu être retrouvés dans le jeu de données XCMS obtenu à partir de l'analyse en LC-HRMS de ces échantillons.

Afin de déterminer si les candidats détectés dans un seul des 6 répliquats sont effectivement des faux positifs, la même approche a été adoptée. La Figure 27B illustre le recouvrement entre les 3 groupes pour les espèces détectées une fois. Face au nombre important de mercapturates potentiels (1964 espèces uniques), seules les 33 espèces communes aux 3 groupes et 110 candidats sélectionnés aléatoirement ont été fragmentés en MS/MS. Finalement, 3 espèces sur les 143 présentaient la perte de neutre caractéristique des mercapturates, ce qui correspond à un taux de faux positif de 98 %, ce qui montre bien la nécessité de réaliser ce type d'expérience et de conserver un regard critique sur les données avant de valider les propositions des logiciels de traitement des données.

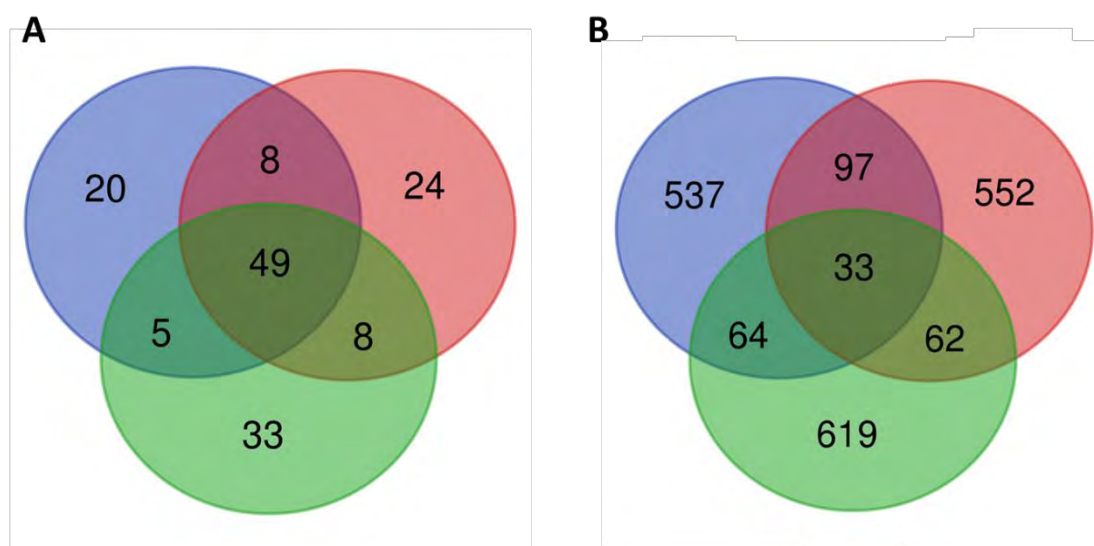


Figure 27 Recouvrement des mercapturates potentiels détectés entre chaque groupe pour (A) les mercapturates détectés au sein de chaque réplikat (B) les espèces détectées dans un seul des 6 répliquats. Le disque bleu correspond au groupe des gros consommateurs, le vert à celui des non-consommateurs et le rouge aux individus sélectionnés au hasard

### C.5. Comparaison MS<sup>E</sup> avec une méthode fast-DDA (PASEF)

L'objectif des travaux présentés dans cette partie est de comparer la méthode DIA utilisée dans ce manuscrit à une méthode DDA pour la détection des mercapturates. Pour cela, le mode d'acquisition PASEF (Parallel Accumulation Serial Fragmentation), disponible sur les instruments de configuration TIMS-ToF proposés par la société Bruker, a été utilisé. La comparaison a été effectuée en utilisant des échantillons de QC issus de la cohorte NutriNet-Santé et les résultats obtenus ont été comparés avec ceux du paragraphe ci-dessus.

L'utilisation du mode PASEF, combiné à la recherche automatique de perte de neutre *via* le logiciel MetaboScape<sup>®</sup>, a conduit à la détection de 81 mercapturates dont 31 parmi les 47 détectés lors de l'analyse MS<sup>E</sup>. En considérant le fait que l'analyse PASEF a été réalisée sur un échantillon de type QC global (mélange des 240 échantillons) tandis que les résultats pour le mode MS<sup>E</sup> sont issus de l'analyse des échantillons QC issus de chaque groupe expérimental, il est probable que le taux de recouvrement entre les deux techniques pourrait être augmenté.

Ainsi, le taux de recouvrement important entre les deux techniques ainsi que le nombre de mercapturates potentiels détectés *via* l'analyse PASEF, suggère que ce type d'acquisition fast-DDA permettrait un screening exhaustif de ces composés. De plus, le principal avantage du mode PASEF repose sur un accès direct à des spectres MS/MS ciblés facilitant ainsi l'identification des conjugués mercapturiques. Enfin, dans le cas de composés inconnus, l'accès à la mesure de CCS (*cross collision section*) peut-être un atout pour l'identification. Bien que les résultats de cette brève comparaison soient encourageants le mode PASEF n'a pas été utilisé pour les travaux proposés dans ce manuscrit, l'appareil TIMS-ToF n'ayant été installé dans le laboratoire qu'en fin de deuxième année de cette thèse.

### D. Discussion et perspectives

Les travaux proposés dans ce chapitre ont permis d'optimiser la plupart des étapes du *workflow* analytique présenté en Figure 16. Plusieurs optimisations ont conduit à une amélioration de la détection des composés mercapturiques, en particulier l'étude réalisée sur les paramètres de transfert des ions qui a grandement amélioré l'intensité du signal de l'ion pseudo-moléculaire des deux standards utilisés comme référence. L'optimisation des paramètres de transfert pour limiter la fragmentation en source et ainsi limiter la perte de sensibilité était un point clé du développement de cette méthode. Une récente étude menée sur des lipides a montré que l'optimisation des paramètres de transfert et de température permettait de réduire la fragmentation en source de 70 % et ainsi d'améliorer la capacité de détection de certaines espèces (Criscuolo *et al.*, 2020). Cette étude a également confirmé que les paramètres analytiques optimaux variaient en fonction de la classe de composés considérés sur un instrument donné, mais étaient également dépendants des architectures

des instruments utilisés (elles-mêmes dépendantes des différents brevets déposés par les constructeurs). Ainsi, bien que l'optimisation des paramètres de transfert sur l'instrument utilisé ait permis une nette amélioration de la détection de certaines espèces (*i.e.* les conjugués mercapturiques dans ce travail), il est possible que ces nouveaux paramètres limitent les capacités de détection de certaines autres classes de composés (donc potentiellement au détriment de l'acquisition métabolomique non ciblée incluse dans notre « *double-workflow* »). Le choix de la dilution de l'urine ainsi que l'optimisation des paramètres du logiciel UNIFI pour la détection des conjugués mercapturiques ont également contribué à l'amélioration de la détection du plus grand nombre de ces composés. En définitive, les résultats ont montré qu'une dilution au  $\frac{1}{4}$  semblait plus efficace pour la détection d'un maximum de composés en comparaison à des dilutions au  $\frac{1}{2}$  ou au  $\frac{1}{8}$ . Le fait d'améliorer la capacité de détection *via* l'utilisation d'un échantillon plus dilué peut s'expliquer notamment par une réduction de l'effet matrice permettant la détection de composés moins abondants. Des résultats similaires ont déjà été retrouvés dans d'autres études où l'intensité de certains métabolites était augmentée malgré une dilution plus importante (Stojiljkovic *et al.*, 2014). D'autres avantages tels qu'une réduction de l'effet matrice ou du *carry-over* (phénomène dans lequel un composé présent dans un échantillon est retrouvé dans les injections suivantes) ont également été reportés (Helfer *et al.*, 2017).

Bien que ces optimisations aient conduit à une couverture plus large du nombre de mercapturates potentiels détectés, les résultats présentés ci-avant suggèrent que cette méthodologie comporte encore quelques limitations notamment en matière de sélectivité (capacité à détecter les vrai / faux positifs) et de reproductibilité quant à l'identification des mercapturates potentiels. Les limitations principales reposent sur (i) un nombre élevé de faux positifs, c'est-à-dire des mercapturates potentiels repérés en première intention et qui ne présentent pas la perte de neutre d'intérêt lors de leur fragmentation en MS/MS ciblée, et (ii) un nombre important de mercapturates retrouvés dans une seule ou une partie des injections multiples réalisées sur un même échantillon. Ainsi, bien que plus consommatrices de temps, des stratégies palliatives ont été proposées (telles que ne garder que les espèces détectées dans chaque réplicat et une vérification des candidats en MS/MS ciblée) afin d'assurer une qualité optimale des données. Ce manque de sélectivité et de reproductibilité peut s'expliquer par la complexité des données issues du mode MS<sup>E</sup>. En effet, le mode MS<sup>E</sup> consiste à fragmenter l'ensemble des ions en un scan, sans sélection d'ion précurseur générant ainsi des données très complexes, ce qui pourrait compliquer l'attribution parent/fragment. D'autres constructeurs ont proposé des solutions alternatives telles que le SWATH<sup>®</sup> (proposé par Sciex), permettant de fragmenter des « tranches » de *m/z* définies par l'expérimentateur, ce qui pourrait atténuer la complexité des spectres de masse résultants. Une étude comparative entre une acquisition DIA en mode SWATH<sup>®</sup> ou DIA en mode MS<sup>E</sup> a montré que la qualité des spectres de fragmentation issus du



mode SWATH était supérieure au mode MS<sup>E</sup> (Zhu *et al.*, 2014). Cette étude pointe également l'impact la complexité de la matrice sur l'attribution parent/fragment. En effet, la qualité des spectres MS/MS issus des modes DIA s'est révélée être moins bonne en matrice urinaire qu'en tampon. L'utilisation du mode MS<sup>E</sup> a cependant montré de bonnes performances lors de son utilisation pour des approches ciblées. En effet, Chindarkar *et al.* ont reporté un taux de vrai positif de 92 % lors du développement d'une méthode de screening de composés d'intérêt toxicologique (Chindarkar *et al.*, 2014).

Il est cependant à noter que même si les approches DIA semblent adaptées pour le screening non-ciblé de composés d'intérêt, les avancées technologiques récentes permettant l'acquisition de nombreux spectres MS/MS ciblés en une seule analyse (type *fast-DDA*) ouvrent la voie à l'utilisation de ce type de technique pour des approches non-ciblées.

## **Chapitre III : Etude de l'impact de la consommation de viande rouge et de charcuterie**

### **A. Introduction à l'article**

Comme détaillé dans l'introduction de ce manuscrit (Chapitre I : Introduction générale) la consommation de viande rouge et de charcuterie en grande quantité est associée à de nombreux effets néfastes pour la santé. En effet, de nombreuses études épidémiologiques ont reporté des associations entre la consommation de ces produits et un risque accru de développement du CCR (Bradbury *et al.*, 2020; Farvid *et al.*, 2021; Yiannakou *et al.*, 2021). Outre cette étroite association, d'autres maladies chroniques, telles que le diabète ou les maladies cardiovasculaires, ont également été associées à la consommation de viande rouge et de charcuterie (Abete *et al.*, 2014a; Pan *et al.*, 2011). D'une manière plus générale, une mortalité globale plus élevée est reportée chez les gros consommateurs de ces produits, en particulier pour la charcuterie (Larsson and Orsini, 2014; Wang *et al.*, 2016).

Parmi les hypothèses avancées dans la littérature, le fer héminique, présent en quantité importante dans la viande rouge, est proposé comme étant un élément-clé dans le processus de promotion de la carcinogénèse induite (Bastide *et al.*, 2015). En effet, de par sa capacité de catalyse de production radicaux libres, le fer héminique induit l'oxydation d'acide gras polyinsaturés donnant ainsi naissance à des composés réactifs aux propriétés toxiques tels que les 2-alcénals. Plus récemment, certains travaux ont montré que le fer héminique pouvait avoir un impact sur le microbiote intestinal, mais aussi que l'effet hyper-prolifératif du fer héminique pourrait également être médié par le microbiote (Ijssennagger *et al.*, 2015; Martin *et al.*, 2019a).

#### **A.1. Objectifs de l'étude**

Dans ce travail, la méthode précédemment développée et optimisée présentée au chapitre précédent (Chapitre II : Développement et optimisation de la méthode d'analyse) a été utilisée dans l'objectif de caractériser les changements induits par une consommation élevée de viande rouge et de charcuterie, sur le profil des acides mercapturiques et le métabolome urinaire chez des individus sains. Ainsi, l'analyse des échantillons urinaires permet à la fois de suivre les effets de l'hème sur la peroxydation lipidique hème-induite, par l'analyse des conjugués mercapturiques, mais aussi l'effet de l'hème sur le microbiote dans la mesure où l'urine contient de nombreux métabolites qui en sont dérivés. Enfin, pour renforcer les résultats obtenus sur l'analyse des échantillons provenant des individus, ceux-ci ont été comparés avec des résultats obtenus dans une étude *in vivo* menée sur des rats nourris avec un régime riche en viande rouge ou un régime sans viande.

## A.2. Méthodes

Dans cette étude, 240 échantillons d'urines issus d'individus de la cohorte NutriNet-Santé ([etude-nutrinet-sante.fr](http://etude-nutrinet-sante.fr)) ont été analysés. Ces 240 échantillons ont été sélectionnés au sein de la bio banque de l'étude NutriNet-Santé et divisés en 3 groupes classés en fonction de leur consommation de viande rouge et de charcuterie : 80 individus représentant le plus haut quintile de consommation, 80 individus ne consommant pas de viande rouge et de charcuterie et 80 individus pris au hasard dans les différents quintiles (excepté le quintile des non-consommateurs). Afin de tenir compte de l'appariement des individus, une méthode de régression logistique conditionnelle pénalisée a été utilisée pour déterminer les métabolites et mercaptates significativement impactés par le régime alimentaire. Pour comparer les résultats obtenus sur la cohorte humaine, 20 échantillons urinaires obtenus à partir d'une expérimentation nutritionnelle *in vivo* sur des rats ont été analysés selon la même méthode.

## **Urinary metabolome analysis reveals potential microbiota alteration and electrophilic burden induced by high red meat diet: results from the French NutriNet-Santé cohort and an *in vivo* intervention study in rats**

L.Mervant<sup>1,2</sup>, M.Tremblay-Franco<sup>1,2</sup>, M.Olier<sup>1</sup>, E.Jamin<sup>1,2</sup>, J-F.Martin<sup>1,2</sup>, C.Buisson<sup>1</sup>, N.Naud<sup>1</sup>, C.Maslo<sup>1</sup>, C.Héliès-Toussaint<sup>1</sup>, E.Fouché<sup>1</sup>, B. Allès<sup>3</sup>, J. Baudry<sup>3</sup>, E. Kesse-Guyot<sup>3,4</sup>, S. Hercberg<sup>3,4</sup>, P. Galan<sup>3</sup>, M.Deschasaux-Tanguy<sup>3,4</sup>, M.Touvier<sup>3,4</sup>, F.Pierre<sup>1</sup>, F.Guéraud<sup>1</sup> & L.Debrauwer<sup>1,2</sup>

<sup>1</sup> Toxalim, Toulouse University, INRAE, ENVT, INP-Purpan, UPS, Toulouse, France.

<sup>2</sup> MetaboHUB-MetaToul, National Infrastructure of Metabolomics and Fluxomics, Toulouse, 31077, France

<sup>3</sup> Sorbonne Paris Nord University, INSERM U1153, INRAe U1125, CNAM, Nutritional Epidemiology Research Team (EREN), Epidemiology and Statistics Research Center – University of Paris (CRESS), 74 rue Marcel Cachin, 93017 Bobigny, France

<sup>4</sup> French Network for Nutrition and Cancer Research (NACRe Network), 78352 Jouy-en-Josas, France

### **Introduction**

Over the past decades, red and processed meat intake has been associated with the development of several chronic diseases. In particular, strong associations have been reported between red and processed meat consumption and the occurrence of colorectal cancer (CRC) (WCRF, 2018; Wolk, 2017). IARC classified processed meat as carcinogenic and red meat as a probably carcinogen to human (Bouvard *et al.*, 2015). The WCRF dose-response meta-analysis determined a 12% greater risk of developing colon cancer per 100 grams increase in red and processed meat consumed per day and a 16% greater risk per 50 grams increase in processed meat consumed per day (WCRF, 2018). Another recent meta-analysis reported similar conclusions (Farvid *et al.*, 2021). Associations have also been suggested with other cancer locations such as breast cancer but also with type 2 diabetes (Pan *et al.*, 2011) or cardiovascular diseases (CVD) mortality in a dose-response manner with a 16% higher risk of CVD mortality for red meat and 18% for processed meat (Abete *et al.*, 2014b). Strong associations were also observed between processed meat and all-cause mortality while evidence for unprocessed red meat remained sparse (Abete *et al.*, 2014a; Larsson and Orsini, 2014; Wang *et al.*, 2016).

To explain adverse health effects (notably carcinogenic) of red and processed meat, several hypotheses have been emitted involving various compounds among which heme iron (Bastide *et al.*, 2015; Sesink *et al.*, 1999), heterocyclic aromatic amines (Ohgaki *et al.*, 1987), polycyclic aromatic hydrocarbons (Sinha *et al.*, 2005), *N*-nitroso compounds (Hemeryck *et al.*, 2018) and *N*-

glycolylneuraminic acid (Neu-5Gc) (Samraj *et al.*, 2015) appear to be the most likely ones. In particular, some studies in animal and cellular models have emphasized the role of heme iron as a central element of the carcinogenesis induced by red and processed meat consumption (Bastide *et al.*, 2015). Increasing the level of evidence, a positive association between heme iron intake and risk of colorectal adenoma was found in the E3N prospective cohort study (Bastide CEBP 2016). Several mechanistic studies focused on the link between heme iron and CRC promotion. One of the most plausible mechanisms involves the catalytic effect of heme iron that potentialize oxidation of polyunsaturated fatty acids (PUFAs) *via* reactive oxygen species production (Carlsen *et al.*, 2005). PUFAs oxidation can lead to the formation of reactive compounds such as 2-alkenals, which have been identified as toxic compounds (Guéraud, 2017). 2-alkenals are electrophilic compounds which, as many electrophilic species, are mainly detoxified *via* glutathione conjugation and subsequent elimination in urine through the mercapturate pathway, as mercapturic acid derivatives (MACs) (Hinchman and Ballatori, 1994; Keller *et al.*, 2015). Considering the toxicological interest of MACs, several approaches using liquid chromatography coupled to mass spectrometry (LC-MS) have been developed to monitor the exposure to these compounds as the reflect of the electrophilic burden (Jamin *et al.*, 2020; Wagner *et al.*, 2006).

Besides lipid peroxidation, it has recently been demonstrated that red meat and heme iron can induce changes in the gut microbiome. Some studies have emphasized that dietary heme iron can induce profound and non-beneficial shifts in gut microbiota composition (Constante *et al.*, 2017b; Martin *et al.*, 2019a). It has been showed that in F344 rats that the consumption of a heme-enriched diet resulted in the alteration of gut microbiota composition with covariations between the production of 2-alkenals and a breakdown in microbiome homeostasis (Martin *et al.*, 2019a). Recent evidence suggested that gut microbiota alteration by the consumption of red and processed meat could be one of the contributors to CRC development (Song *et al.*, 2020). In addition to these findings, Ijssennagger *et al.* underlined the possible role of gut microbiota in heme-induced hyperproliferation of colonic cells (Ijssennagger *et al.*, 2015). Such changes in microbiota composition can have an impact on the urinary metabolome, since both have been shown to be closely linked (Zheng *et al.*, 2011).

Hence, by identifying biomarkers associated with red and processed meat consumption, this study would serve a dual purpose. Firstly, the determination of biomarkers of consumption would allow a better estimation of the quantity of meat consumed as a complement to questionnaires in cohort studies. Secondly, the discovery of biomarkers of effect as revealed in urine could help deciphering the mechanisms involved in the link between red/processed meat and health.

In this work, considering the ability of heme iron to both catalyse lipoperoxidation in the intestinal lumen and alter microbiota, we focused on the urinary metabolome as a resulting image of these deleterious effects, with a particular focus on MAC profile. We based our study on two complementary

and multidisciplinary approaches: urinary samples from participants to the French NutriNet-Santé cohort study and on an *in vivo* experiment conducted on rats fed with either a high red meat (HRM) or a control diet (CTL), with the objective to investigate the potential impact of red and processed meat consumption on health *via* the urinary metabolome. To this end, we used an LC-MS method to perform an in-depth characterization of the urinary metabolome, based on an original approach combining a hypothesis-driven and an untargeted analysis. This approach allowed a simultaneous collection of untargeted metabolomic data for assessing microbiota alteration as revealed by urinary metabolome.

## Methods

### Population and collection of data and urine samples

The NutriNet-Santé study is an ongoing web-based cohort launched in 2009 in France aiming to study the associations between nutrition and health. This cohort study has been previously described in details (Herberg *et al.*, 2010). Briefly, adults aged 18 years or older are continuously recruited among the general population and followed using a dedicated web interface ([etude-nutrinet-sante.fr](http://etude-nutrinet-sante.fr)). The NutriNet-Santé study is being performed in accordance with the tenets of the Declaration of Helsinki, and all procedures have been approved by the institutional review board at the French National Institute for Health and Medical Research (IRB INSERM #0000388FWA00005831) and by the National Commission on Informatics and Liberty (CNIL #908,450 and #909,216). All participants provided informed consent and an electronic signature. The study is registered at ClinicalTrials.gov (#NCT03335644).

Upon inclusion and then every year, participants are asked to complete validated questionnaires related to sociodemographic and lifestyle characteristics, health status and medication use, dietary intakes (Lassale *et al.*, 2016, 2015; Touvier *et al.*, 2011), physical activity (Craig *et al.*, 2003) and anthropometric characteristics (Lassale *et al.*, 2013; Touvier *et al.*, 2010). Dietary intakes are assessed every 6 months through 3 non-consecutive, validated 24-hour dietary records, randomly distributed over 2 weeks, including 2 weekdays and 1 weekend day. Portion sizes are estimated using validated photographs, standard containers, or directly in g/L. The food content in energy, alcohol, macro- and micro-nutrients are derived from the NutriNet-Santé food composition table which is continuously updated and currently comprise >3500 items. Amounts of food consumed from composite dishes are estimated using French recipes validated by food and nutrition professionals. Dietary energy under-reporters are detected *via* the method proposed by Black (Black, 2000). Between 2011 and 2014, 19,600 volunteers participated to a clinical examination including the collection of fasting blood and urine samples that were stored at -80°C ever since.

### **Selection of the study sample**

For this study, participants with at least six 24h dietary records in the 2 years preceding the urine sampling were classified according to their consumption of red and processed meat (mean daily intake in g/d calculated across all 24h dietary records filled in this 2y-time-frame). A total of 80 high consumers was selected among the fifth sex-specific quintile of red and processed meat consumption and matched to 80 non-consumers according to sex, age (5-year categories), season of urine sampling (spring-summer/fall-winter), smoking status (current smokers/non-smokers), energy intakes (quartiles) and menopausal status at urine sampling for women. Finally, 80 other participants were selected randomly and matched to the high consumer/non-consumer pairs using the same *matching* criteria.

### **Chemicals**

All solvents were purchased from Fisher Scientific (Thermo Fisher Scientific, Illkirch, France) and were LC-MS grade.

### **Animal and diets**

Four-week-old male Fischer 344 rats (10 rats/group) were purchased from Charles River. Animal experiment was authorized by the French Ministry for Higher Education, Research and Innovation (MESRI) in accordance with the local ethic committee evaluation (Project APAFiS #22200-2019093015339804v2). Rats were housed in conventional cages by groups of 3 or 4. They were allowed to 10 days of acclimatisation to their cage before the start of experiment. During this period, all rats consumed the control diet. Rats had unlimited access to tap water and diet. Diets were given every day at the end of afternoon to limit oxidation during the day, as rats eat during the night. After 16 days of experimental diet, rats were placed in plastic metabolic cages. Urine was collected and food consumption was recorded after 24h of housing. Once collected, urine was stored at -80°C prior analysis. Faeces were collected after 21 days of experiment and directly frozen in liquid nitrogen. The control diet was a modified AIN-76 prepared and formulated in a powdered form by UPAE (INRAE, Jouy-en-Josas, France). All diets contained 5% of safflower oil (MP Biomedicals 102888) and were low-calcium (3.5 g/kg). High red meat diet (HRM) contained 40% (on a dry weight basis) of 5% fat beef meat (Picard, France). Control diet was adjusted for protein and fat using casein and saturated animal fat.

### **Sample preparation**

Sample collection has already been described elsewhere (Mervant *et al.*, 2021). Once collected samples were specifically diluted in Ultrapure water according to their osmolality measured using a freezing point osmometer (Loeser Messtechnik, Berlin, Germany). After dilution samples were centrifuged at 9500g during 5 min. The supernatant was collected and transferred into vials to be

directly analyzed by UHPLC-MS. QC samples were prepared either by pooling equivalent volumes of all samples or by pooling all samples from one experimental group to obtain QC group. QC of all samples were diluted accordingly to obtain dQC samples.

### **DHN-MA Immunoassay**

Urinary DHN-MA concentration was determined by competitive enzyme immunoassay (EIA) as previously described (Guéraud *et al.*, 2006). Diluted and normalized urines were assayed in duplicates. Detection was recorded at 414 nm.

### **Bacterial community analysis**

Purified genomic DNA was obtained from snap frozen faecal samples and 16S rRNA gene was amplified as described elsewhere (Martin *et al.*, 2019b). The resulting purified amplicon libraries were loaded onto the Illumina Miseq cartridge. Valid raw sequences obtained after sequencing were processed using the FROGS pipeline (Escudié *et al.*, 2018) (Galaxy Version 3.2.3). Each pair-end valid denoised sequences were filtered, merged and clustered (swarm fastidious option using a maximum aggregation distance of 1) (Mahé *et al.*, 2014). After removal of putative chimera, 306 clusters satisfying the following criteria were kept: abundance representing at least 0,005 % of all sequences, presents in at least 3 times in a minimum of 5% of total samples. An average of 50 810 initial valid raw reads per sample was obtained and normalized to 27512 reads per sample after rarefaction. Affiliation of Operational Taxonomic Units (OTUs) was performed using the silva 138.1\_16S reference database and Blast+with equal multi-hits were manually verified and corrected if necessary. Community profiles (including  $\alpha$  and  $\beta$  diversity) were determined using Phyloseq R package (v1.34.0). Chao-1 index was used to determine the within sample community richness. Divergence of microbiota composition between samples was explored using the Unifrac distance matrices and model robustness was assessed by performing permutational multivariate analysis of variance (Adonis test with 9999 permutations). OTUs were agglomerated at the genus rank to examine differential abundances in response to the type of diet using the DESeq2 R package (v1.30.1) with Wald method. Univariate testings were corrected for false discovery rate (FDR) by Benjamini-Hochberg method, and adjusted p values less than 0.05 were considered statistically significant. LDA (Linear Discriminant Analysis) effect size (LEfSe, <https://doi.org/10.1186/gb-2011-12-6-r60>) was also estimated between the two diets (factorial Kruskal-Wallis sum-rank test) and discriminative *features* with  $p < 0.01$  and an absolute logarithmic LDA score  $> 2.5$  were represented using a cladogram.

### **LC-MS Analyses**

A Q-ToF Synapt G2-Si mass spectrometer (Waters, Manchester, UK) coupled to an ACQUITY™ liquid chromatographic system (Waters, Manchester, UK) was used to perform sample analyses. The analysis



method has already been described elsewhere (Jamin *et al.*, 2020). Briefly, chromatographic separation was performed by injecting 10  $\mu$ L of sample into a Hypersil Gold C18 (1.9  $\mu$ m, 100 x 2.1 mm) analytical column (Thermo Fisher Scientific, Illkirch, France). The column was kept at 40°C and the flow rate set at 0.3 mL/min. A linear gradient elution was used with mobile phase A: 95 % H<sub>2</sub>O / 5 % methanol / 0.1 % acetic acid, and mobile phase B: 100% methanol / 0.1% acetic acid. The mobile phase composition increased from 0 to 100 % of B in 30min, held for 4 min prior to switching back to initial condition in 1 min and held for 5 min to re-equilibrate the system. Samples were analysed in two pseudo-randomized analytical batches. dQC samples were injected at the beginning of the first batch then QC samples were regularly injected every 8 samples.

Mass spectra and Data Independent Acquisition (MS<sup>E</sup>) spectra were acquired from  $m/z$  50 to 800 using the negative ionization mode and the following set of parameters: capillary voltage 0.5 kV, sampling cone voltage 30 V, source temperature 120°C, desolvation temperature 550°C, cone gas (N<sub>2</sub>) and desolvation gas (N<sub>2</sub>) 30 and 600 L/H, respectively. The trap collision energy was set at 15eV for MS<sup>E</sup> acquisitions.

### Data Processing

Mass spectral data processing was achieved using XCMS 3.0 with the Workflow4Metabolomics (W4M) collaborative environment dedicated to the processing of metabolomics data (Giacomoni *et al.*, 2015). Centwave algorithm was used to perform automatic data extraction using the following set of parameters: ppm= 10, peakwidth = (5,40), sntresh = 10, noise = 10000, bw =10, mzwid = 0.01. Data filtering process was carried out using several filters as follows: first, all *features* displaying a correlation coefficient with dQC samples lower than 0.6 were eliminated. Then, *features* whose ratio of the average signal in QCs over average signal in blanks was lower than 3 were eliminated. Finally, after performing LOESS correction to avoid batch effects, *features* displaying a RSD lower than 30% in QC samples were kept for further processing.

Data from MS<sup>E</sup> experiments performed on group-QC samples were processed according to Jamin *et al.* (Jamin *et al.*, 2020) using the UNIFI software (Waters, Manchester, UK). Only *features* displaying the characteristic neutral loss of mercapturic acids (129.042) in every replicate (n=6) within a group QC-sample were considered. Then targeted MS/MS fragmentation experiments were carried out to confirm the detected neutral loss.

### Metabolite Identification

*Features* annotation was performed by matching in-house database to the obtained peak list. For every putative annotation, the experimental MS/MS spectrum was compared to our in-house database of

metabolite standards. *Features* that displayed similar  $m/z$ , retention time and fragmentation pattern were considered as level 1 identification according to Sumner et al (Sumner *et al.*, 2007).

### Statistical analyses

Statistical analyses were achieved using Simca-P 14.0 software (Umetrics, Umea, Sweden). All data were log transformed and Pareto scaled prior orthogonal partial least square discriminant analysis (OPLS-DA). Predictive capacity ( $Q^2$ ) and permutation test were used to check validity and robustness of O-PLS-DA models. OPLS-DA model was considered valid if displayed  $Q^2$  intercept  $< 0.05$  and  $Q^2_{cum} > 0.4$  (Lapins *et al.*, 2008).

However, OPLS-DA does not allow to adjust for confounding factors in the model and cannot account for the matching of individuals. Hence, penalized conditional logistic regression models were used to determine which metabolites were most associated to red and processed meat consumption as well as the direction of the associations (using Odds Ratios), while accounting for the matched design and adjusting for potential confounders. ElasticNet penalization (Zou and Hastie, 2005) available in the R `penalizedclr` package was used to allow variable selection by removing the less predictive variables. The first model included all annotated metabolites, identified MACs and: physical activity, BMI, level of education, alcohol consumption and energy intake. Once intensities of ions were scaled to unit-variance, the ElasticNet parameters  $\alpha$  and  $\lambda$  were optimized using 7-fold cross validation repeated 100 times.  $\alpha$  refers to the partition coefficient between RIDGE and LASSO regression ( $\alpha=1$  correspond to LASSO only) and  $\lambda$  correspond to the strength of penalization. Optimal values were determined as  $\alpha=0.3625$  and  $\lambda=2.44$ . The model was calculated with optimized parameters and its stability was evaluated by bootstrap resampling by selecting 80% of dataset and repeated 1000 times. Only variables selected in the initial model and over 40% of bootstrap iteration were considered. Dietary intakes of fibres and calcium were additionally included in the models using the same  $\alpha$  and  $\lambda$  parameters to investigate the potential influence of these variables in the discrimination of high consumers and non-consumers of red and processed meat. Further adjustments were tested by taking into account other nutritional parameters such as polyphenol, vitamin, fruits and heme iron intake. All tested combinations are displayed in the Supplementary Table1.

### Results

Table 1 describes the socio-demographic, lifestyle and dietary characteristics of the study population according to the three groups of subjects: high meat/processed meat consumers, non-consumers and randomly selected participants.

To investigate the potential deleterious effects of high red and processed meat consumption, untargeted metabolomics analysis was performed on urine with an LC-MS approach aiming at a dual purpose. In first instance, urinary metabolome was considered as a whole to investigate the global impact of red and processed meat consumption with respect to the three groups. In parallel, considering the toxicological interest of MACs as markers of exposure to reactive substances, special attention was paid to these compounds *via* a specific screening workflow enabling the extraction of mass spectrometric signals coming from MACs (Jamin *et al.*, 2020).

The LC-MS metabolomic analysis led to the extraction of 8834 *features*, yielding 4138 MS *features* after the filtration process of MS data. Among these *features*, 63 metabolites were annotated at level 1 according to MSI classification (Sumner *et al.*, 2007) and 30 MACs were formally identified as they displayed the characteristic neutral loss of the N-acetyl cysteine moiety (129.042) confirmed in targeted MS/MS experiments. However, among these 30 MACs only a few could be annotated. In first instance, an OPLS-DA analysis was performed using the whole dataset (i.e. the 4138 *features*). A valid model ( $Q^2= 0.409$ , intercept = -0.291) could be constructed for discriminating high red/processed meat consumers (HRPC) and non-red/processed meat consumers (NRPC). According to this model, 1189 *features* displayed a ViP score > 1. The related score plot and the permutation tests are available in Supplementary Data 1. This result suggested that many urinary metabolites are associated with red and processed meat consumption of participants.

In order to consider potential confounding factors and the matching of samples, a penalized logistic regression analysis was performed based on the selection of all the 63 annotated metabolites and the 30 confirmed MACs. In the first instance, only non-nutritional parameters were used as covariates to take into account other differences in dietary habits existing between HRPC and NRPC. Therefore, the first model was adjusted on BMI, physical activity, level of education, alcohol consumption and caloric intake.

Among the 93 ions constituting the dataset, 8 were found negatively associated to red and processed meat consumption and 1 was found positively associated, as shown in Figures 28A and 28B. When compared to the NRPC, HRPC exhibited lower level of: cinnammoylglycine (OR= 0.54, 95% CI = [0.33-0.84]), methoxysalicylic acid, (0.31, [0.13-0.65]), 2,2-dimethylglutaric acid (0.55, [0.33-0.87]), N-acetyl-leucine (0.58, [0.34-0.92]), p-coumaric acid (0.51, [0.28-0.85]), hydroxyphenyllactic acid (0.55, [0.3-0.93]), 10-hydroxydecanoic acid (0.3, [0.11-0.69]) and finally one unknown mercapturate, named as Merca14 (0.47, [0.28-0.74]). Interestingly, only one annotated metabolite and no mercapturate was found positively associated to a higher intake of red and processed meat. The only metabolite found to be positively associated with high intake of red and processed meat was indoxyl sulfate (1.59, [1.06-2.47]). Moreover, correlations between red and processed meat intake for all 240 participants and

metabolite levels are shown in Figure 28B. The strongest negative correlation was observed for cinnamoylglycine and the strongest positive correlation for indoxyl sulfate.

To further investigate the influence of other nutritional parameters and because of their significantly lower intake in high red and processed meat consumers and their association with health, notably CRC risk (Veettil *et al.*, 2021), dietary intakes of fibres and calcium were additionally included in the models. When the model was further adjusted for calcium intake, only a few changes occurred: indoxyl sulfate was no longer kept in over 40% of bootstrap iteration and two mercapturates, respectively named as Merca16 (0.48, [0.27-0.77]) and Merca53(0.56, [0.3-0.91]), were negatively associated with red and processed meat consumption. All other metabolites from the previous model were retained in the calcium adjusted model. Odds-ratio values for all models are available in Supplementary Table 1.

Interestingly, when the initial model was adjusted for dietary fibre intake, the resulting metabolic signature was drastically different. Only three compounds, namely *N*-acetyl-leucine (0.47, [0.25-0.82]), tyrosine (0.58, [0.33-0.96]) and one mercapturate named as Merca15 (0.42, [0.22-0.77]), were found negatively associated with red and processed meat intake and none was found positively associated.

Finally, no valid model could be generated when the initial model was adjusted for heme iron intake due to a high colinearity between the consumption of red and processed meat and the intake of heme iron.

Concerning the animal experiment, a valid OPLS-DA model ( $Q^2 = 0.899$ , intercept = -0.18), displayed in Supplementary Data 2, was generated to discriminate high red meat (HRM) diet and control (CTRL) diet. From MS data 3470 *features* were kept after data filtration and 920 exhibited a ViP score > 1. Among the nine metabolites found significantly different between compared groups in the human cohort, five were annotated in the animal experiment: cinnamoylglycine, indoxyl sulfate, *N*-acetyl-leucine, *p*-coumaric acid and 2,2-dimethylglutaric acid. Three of these five metabolites were significantly altered between the two experimental groups as displayed in Figure 28C to E: cinnamoylglycine (fold change= 0.19, *p*-value = < 0.0001, unpaired t-test), indoxyl sulfate (1.7, 0.043) and coumaric acid (3.39, 0.0014).

Among the 3470 *features*, 47 were formally identified as MACs and 16 were selected in the OPLS-DA model. As shown in Figure 29, an overall increase in the amount of MACs was observed between the HRM-fed and the CTRL groups. In addition, we were able to formally annotate three of the sixteen significant MACs as dihydroxynonene mercapturic acid (DHN-MA), 4-hydroxy-nonanal mercapturic acid (HNE-MA) and 4-hydroxy-heptenol mercapturic acid (HHEP-MA). However in the human study no change in the MACs profile was observed, especially for lipid peroxidation-derived MACs. This result

was reinforced by quantification of di-hydroxynonene-mercapturic acid (DHN-MA), a well-known mercapturic acid linked to lipid peroxidation, which exhibited no significant difference ( $p$ -value: 0.1137) between HRPC and NRPC.

Considering the microbial origin of some of the discriminant metabolites, faecal microbiota analysis was performed by sequencing V3-V4 region of the 16S rRNA gene. This analysis revealed that the composition of microbiota was significantly impacted by the diet. Even if Chao-1 index used to assess the bacterial richness was not affected by the diet (Supplementary Data 3A), the ordination based on UniFrac distances between samples demonstrated the occurrence of distinct profiles of bacterial community between high red meat and control diet (Supplementary Data 1B). Regarding the microbiota composition at the phylum level, significant changes in main phyla were observed, namely *Firmicutes*, *Actinobacteria*, *Bacteroidota*, *Proteobacteria* and *Deferribacterota* (Supplementary Figure 3C). The two first were increased with high red meat diet whereas the three others were decreased. Within each of them, significant changes at the genus level were observed, including notably significant decreases of *Bifidobacteria* and *Parasutterella*, and increases of *Lactobacilli* and an unclassified genus belonging to *Eggerthellaceae* with high red meat diet (see Supplementary Figures C-D).

## Discussion

In this work, the impact of high red and processed meat consumption on the urinary metabolome has been investigated in both a human cohort and a controlled *in vivo* experiment on rats. In human, modulation of this impact could be highlighted by considering other nutritional intakes differing between HRPC and NRPC, including e.g. calcium or fibres. MACs profile, which notably reflects the electrophilic burden of the diet, was slightly altered in the human study conversely to the *in vivo* experiment in which a tremendous increase of mercapturic species in urine of HRM-fed rats was observed.

In the human cohort, the urinary content of 8 metabolites and one unknown mercapturate were found modified in the initially adjusted model. This model reflects differences in eating habits between HRPC and NRPC which go beyond the consumption of red and processed meat alone. Interestingly, several of these 8 metabolites, including in particular cinnamoylglycine, hydroxyphenyllactic acid (HPLA) and indoxyl sulfate, were reported to have a microbial origin (Abdul Rahim *et al.*, 2019; Wikoff *et al.*, 2009). Although microbiota alteration by dietary habits is not surprising, the modulation of microbiota-derived compounds observed in this work is of particular interest because some of these metabolites had never been associated to red and processed meat consumption. Cinnamoylglycine, which abundance is lowered in HRPC, was proposed as a marker of a healthy microbiome, as the level of this

metabolite appears to be correlated with a higher diversity (Menni *et al.*, 2020; Shomorony *et al.*, 2020). Beyond its use as a marker of diversity, this metabolite has been proposed as a marker of resistance against the colonization by *Clostridium difficile* that is known to exhibit pathogenic effects (Obrenovich *et al.*, 2017). This lower level of cinnamoylglycine in HRPC could be associated to a lower microbiota diversity that could render more susceptible to colonization by pathogenic species (Obrenovich *et al.*, 2017). In the same way, HPLA was found to be excreted by a highly used probiotic strain of *Lactobacillus fermentum* (González *et al.*, 2017). *In vitro*, *Bifidobacteria* and *Lactobacilli* were found to produce high amounts of HPLA that has the ability to act as an antioxidant by scavenging ROS in mitochondria (Beloborodova *et al.*, 2012). On the contrary, urinary indoxyl sulfate levels were increased in HRPC subjects. Indoxyl sulfate is a well-studied uremic toxin which microbial origin has been demonstrated by comparing germ-free and conventional mice (Wikoff *et al.*, 2009). Indoxyl sulfate is originating from bacteria expressing tryptophanase that converts tryptophane into indole, which is then sulphated in the liver (Abdul Rahim *et al.*, 2019). Therefore, this end-metabolite can be used as a biomarker of indole production by the gut microbiota (Philippe *et al.*, 2021). Modulations of indoxyl sulfate by diet have already been described when increasing dietary protein (Poesen *et al.*, 2015) and particularly red meat (Mafra *et al.*, 2018). Patel *et al.* also showed a reduction of urinary indoxyl sulfate when comparing healthy volunteers following a vegetarian diet or unrestricted diet (Patel *et al.*, 2012). As it has been linked to an increased risk of cardiovascular diseases (Gao and Liu, 2017) some dietary strategies have emerged to lower the level of indoxyl sulfate such as lowering protein intake or probiotic supplementation with a *Lactobacillus* strain (Di Iorio *et al.*, 2019; Lee *et al.*, 2020). Considering those results, we hypothesize a modulation of the gut microbiota of HRPC in favour of both an increase of deleterious species and/or a decrease in beneficial ones such as *Bifidobacteria* and *Lactobacilli*. However, all these three metabolites were no longer found significant when the model was adjusted on dietary fibre intake, indicating that red and processed meat intake only is not sufficient to explain the observed associations. As fruits, vegetables, fibres, calcium intake and other nutritional intakes were significantly different between HRPC and NRPC in this cohort, reflecting different overall dietary patterns, we assume that all these factors have a varying impact on the observed microbiota alteration. Our results suggest that not all nutritional factors have the same weight on the observed metabolic perturbations. When the model was adjusted on calcium intake, only a few changes were observed compared to the original model, while when adjusting on fibre intake only two discriminant metabolites, *N*-acetyl-leucine and tyrosine, remained significantly different between NRPC and HRPC.

Despite the difficulty of deciphering the complexity of multifactor effects of diet in human based on questionnaires, a specific effect of red/processed meat has to be considered since similar differences in cinnamoylglycine and indoxyl sulfate levels were observed in the *in vivo* experiment conducted on

rats. As rat diets were adjusted for fat, protein and calories but not for iron/red meat, this argues for a specific effect of red meat consumption on cinnamoylglycine and indoxyl sulfate levels. Microbiota sequencing carried out on these animals strengthen the hypothesis of a dysbiosis induced by high red meat consumption. In fact, as revealed by  $\beta$ -diversity results and differential abundances analysis, alterations in bacterial composition were observed both at the OTUs, genus and phylum levels. In this study a great decrease of genus associated with health benefits such as *Bifidobacterium* and *Parasutterella* was observed (Ehrlich *et al.*, 2020; Ju *et al.*, 2019). This may explain the modification of the levels of some microbiota derived metabolites. In accordance with these findings, both a decrease of *Bifidobacteria* and an increase in *Lactobacilli* have already been described in other studies. In particular Zhu *et al.* have reported an increase in *Lactobacilli* with high red meat diet compared to a control casein diet (Zhu *et al.*, 2016, 2015) while Le Leu *et al.* reported a decrease of *Bifidobacteria* with high red meat diet (Le Leu *et al.*, 2013).

In the present work, we observed a tremendous increase of MACs in the rat experiment, and particularly of lipid peroxidation related species such as DHN-MA. In accordance to this finding, many other animal studies demonstrated an increase in lipid peroxidation products in heme rich diet (Guéraud *et al.*, 2015a; Ijssennagger *et al.*, 2013; Keller *et al.*, 2020a). Moreover, the modulation of mercapturate profile by heme iron was recently highlighted by Jamin *et al.* who observed an increase of both known and unknown MACs levels in urine of rats submitted to a diet combining heme and polyunsaturated fatty acids (Jamin *et al.*, 2020). In the present study, we observed the same phenomenon with a more complex food matrix. The high number of MACs altered by the HRM diet underlined the diversity and complexity of products generated by heme induced lipid peroxidation. Only three MACs could be formally annotated, highlighting the current lack of knowledge about these compounds and underlining the poor availability of analytical standards and spectral data. However, this mercapturome alteration has not been observed in the human study especially for lipid peroxidation derived mercapturates, probably because of a less contrasted situation due to a more diverse diet, containing, among other things, more antioxidants and lower proportion of red meat. However, it should be noted that mercapturates can also originate from other compounds such as polyphenols (Kotronoulas *et al.*, 2013). This could explain the observed decrease of some mercapturic species in HRPC. The absence of an increased level of urinary lipid peroxidation products in HRPC could be due to the calcium content of the diet which is known to reduce heme-induced lipid peroxidation (Pierre *et al.*, 2003; Sesink *et al.*, 2001b). In the rat study, the calcium content of diet is reduced while in the human cohort calcium consumption is in the range of the nutritional recommendations for NRPC and slightly below for HRPC. Finally, it is worth noting that a urinary mercapturic acid profile is transitory (Keller *et al.*, 2020a) and highlights only a recent exposure to electrophilic substances, therefore strongly depending on the sampling time slot. In the human study, red and processed meat

consumption was assessed as an average of repeated 24h dietary records (at least 6) over the two years preceding the urine sampling and therefore reflects long-term dietary habits rather than intakes occurring just before the urine sampling. Conversely to MACs profile, gut microbiota alterations may be less transitory and would be least sensitive to the sampling time.

The originality of this work lies in several points. First, the vast majority of LC-MS metabolomics studies aiming to highlight markers of red and processed meat consumption have been performed using positive ionisation, in particular for the detection of biomarkers such as acyl-carnitines or carnosine. Acyl-carnitines have been widely described as biomarkers of red and processed meat consumption (Cheung *et al.*, 2017; Khodorova *et al.*, 2019; Wedekind *et al.*, 2020). In a recent study, Wedekind *et al.* strengthened the link between increased acyl-carnitines in urine and red and processed meat consumption by a metabolomics analysis combining samples from both an intervention and an observational study (Wedekind *et al.*, 2020). Khodorova and colleagues confirmed the origin of acyl carnitine as well as other compounds such as phenylacetyl glutamine or 1-methylhistidine by the using of <sup>15</sup>N-labelled bovine meat in an intervention study (Khodorova *et al.*, 2019). In our work, the use of negative ionisation allowed highlighting other metabolites which may not be efficiently detected in the positive ionisation mode. This could increase the knowledge on biomarkers of red and processed meat consumption. Second, the human study compared three groups of participants differing by their usual consumption of red and processed meat and matched for several individual characteristics, which reinforced the capacity to highlight differences in urinary metabolome while reducing inter-individual variability. The extensive characterisation of participants from the large prospective NutriNet-Santé cohort, especially regarding dietary data, also allowed to take into account numerous nutritional and lifestyle parameters as potential confounding factors. In addition, participants in the NutriNet-Santé cohort display some particularities including: (i) they are predominantly women, and (ii) they overrepresent high socio-professional categories. Therefore, they are more likely to display a healthier nutritional profile, with consequently a lower consumption of red and processed meat, compared to the general French population (Andreeva *et al.*, 2016, 2015). In spite of this, we were still able to highlight biomarkers of consumption in this study, but this lower consumption of red and processed meat, resulting in a reduced contrast between high consumers and non-consumers, may have limited our ability to detect further significant differences in the urinary metabolome. Finally, our findings on the human cohort were supported by similar trends observed in an *in vivo* intervention experiment in rats, which allowed strengthening our hypothesis of a gut microbiota alteration induced by high red meat diet. However, considering the hypothesis of a red meat-induced dysbiosis, this type of study would, in the future, highly beneficiate from microbiome sequencing data in human in parallel with urinary metabolome analysis. Unfortunately, no sample is available to carry out a metagenomic study on the subjects investigated in this work. Further work is foreseen for assessing both



metabolomic and microbiota.

In conclusion, this work demonstrated a modification of the urinary metabolome associated with high red and processed meat consumption in both a human cohort and an *in vivo* experiment on rats. To the best of our knowledge, some metabolites such as cinnamoylglycine or HPLA have never been described as being associated with red meat consumption. This particular metabolic signature led us to assume a microbiota dysbiosis induced by red meat consumption, a hypothesis reinforced by microbiota sequencing data on rats. The ability of high red meat diet to induce the production of electrophilic and reactive compounds was demonstrated in rats but not in humans. Nevertheless, in view of the results of the *in vivo* experiment it would be worth studying mercapturic profile in human after an intervention study with a controlled diet. In conclusion, by providing new markers of consumption, this work may help to better understand the impact of red and processed meat consumption on health and in particular the potential deleterious effect on gut microbiota.

	High red/processed meat consumers (n=80)	Non red/processed meat consumers (n=80)	Randomly selected participants (n=80)
<b>Age</b>	50 ± 13	49.9 ± 12.8	50.2 ± 12.7
<b>Kcal</b>	1968.3 ± 453.7	1951.6 ± 414.9	1941.3 ± 415.3
<b>R.meat-Proc.meat<sup>a</sup></b>	123.4 ± 30.5	0 ± 0	65.9 ± 40.7
<b>BMI</b>	24.9 ± 4.7	21.5 ± 2.8	23.6 ± 4.6
<b>Alcohol (g/day)</b>	11.7 ± 12.6	4.6 ± 7.8	9.1 ± 13.1
<b>Heme iron (g/day)</b>	1.8 ± 0.9	0.1 ± 0.5	1.3 ± 1.4
<b>Survey nb</b>	7.5 ± 2	7.2 ± 1.8	7.5 ± 2.1
<b>Polyphenols<sup>b</sup></b>	1344.3 ± 567.8	1710.5 ± 755.9	1274.3 ± 448.8
<b>Ca (mg/day)</b>	911.9 ± 275.5	993.6 ± 342.3	945.7 ± 249.5
<b>Fibre (g/day)</b>	19 ± 6	29 ± 9.6	20.4 ± 6
<b>Sum vitamins<sup>c</sup></b>	212.7 ± 63.1	259.9 ± 112	206.9 ± 58.1
<b>Sum veg-fruits<sup>d</sup></b>	419 ± 196.5	618.4 ± 320.2	439.6 ± 181.9
<b>Physical activity</b>			
<b>Low</b>	25 (31.25)	37 (46.83)	26 (32.5)
<b>Moderate</b>	31 (38.75)	30 (37.97)	38 (47.5)
<b>High</b>	24 (30)	12 (15.18)	16 (20)
<b>Educational level</b>			
<b>Primary</b>	12 (15)	11 (13.75)	12 (15)
<b>Secondary</b>	17 (21.25)	11 (13.75)	12 (15)
<b>Superior</b>	51 (63.75)	58 (72.5)	56 (70)

<sup>a</sup>Sum of red and processed meat consumption in g/day

<sup>b</sup>Sum of polyphenol consumption in mg/day

<sup>c</sup>Sum of Vitamin C, E, Selenium and zinc consumption per day

<sup>d</sup>Sum of vegetable and fruit consumption in g/day

Table 1. Lifestyle and dietary characteristics of the study population. Values are in mean ± SD or n(%)

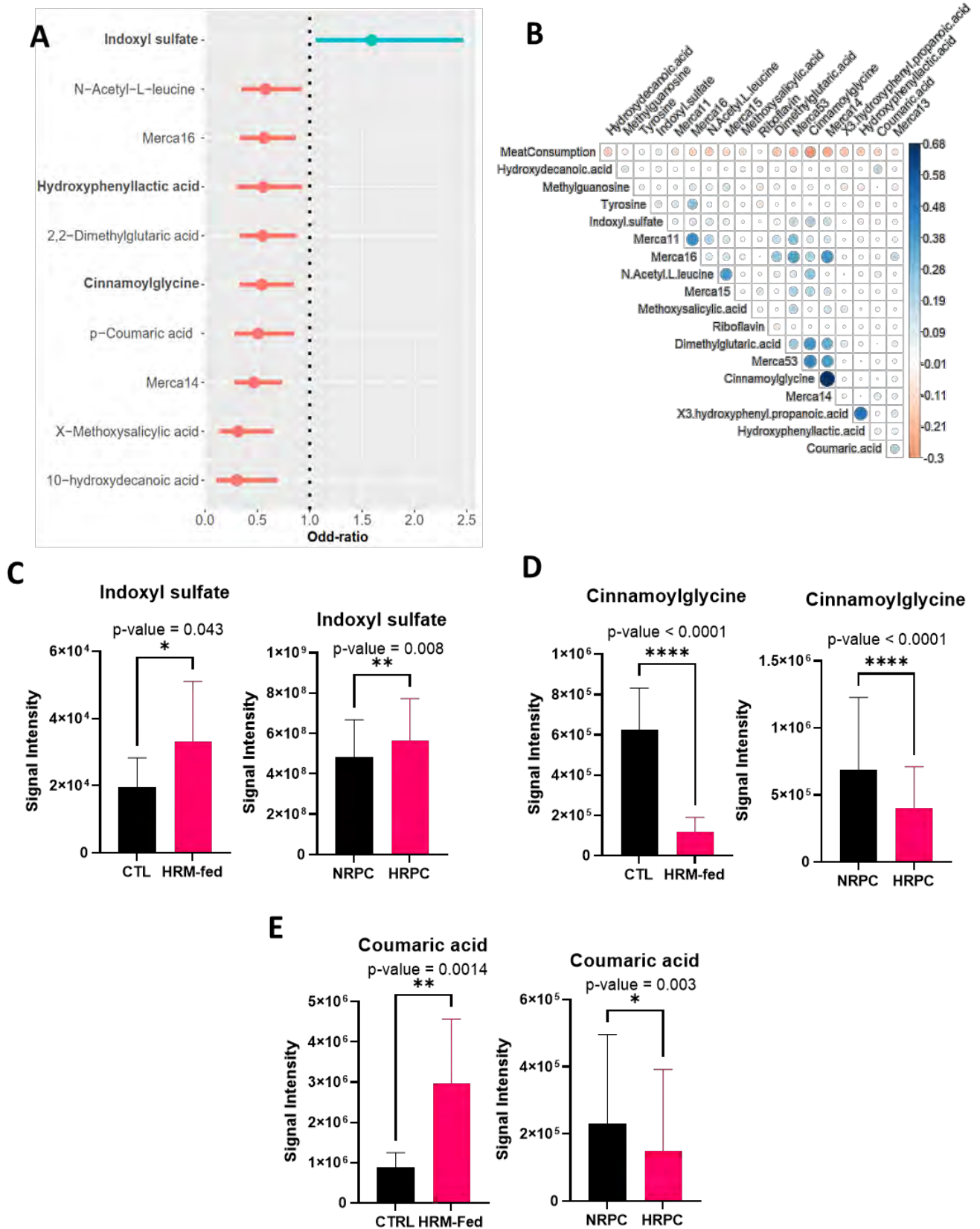


Figure 28 Odd-ratios of selected metabolites in the non-nutritionally adjusted model (A) Correlation matrix between red meat consumption and metabolite signal intensity for all metabolites selected in every calculated model (B) Indoxyl sulfate signal intensity measured in both animal experiment and human cohort samples (C) Cinnamoylglycine signal intensity measured in both animal experiment and human cohort samples (D) Coumaric acid signal intensity measured in both animal experiment and human cohort samples (E)

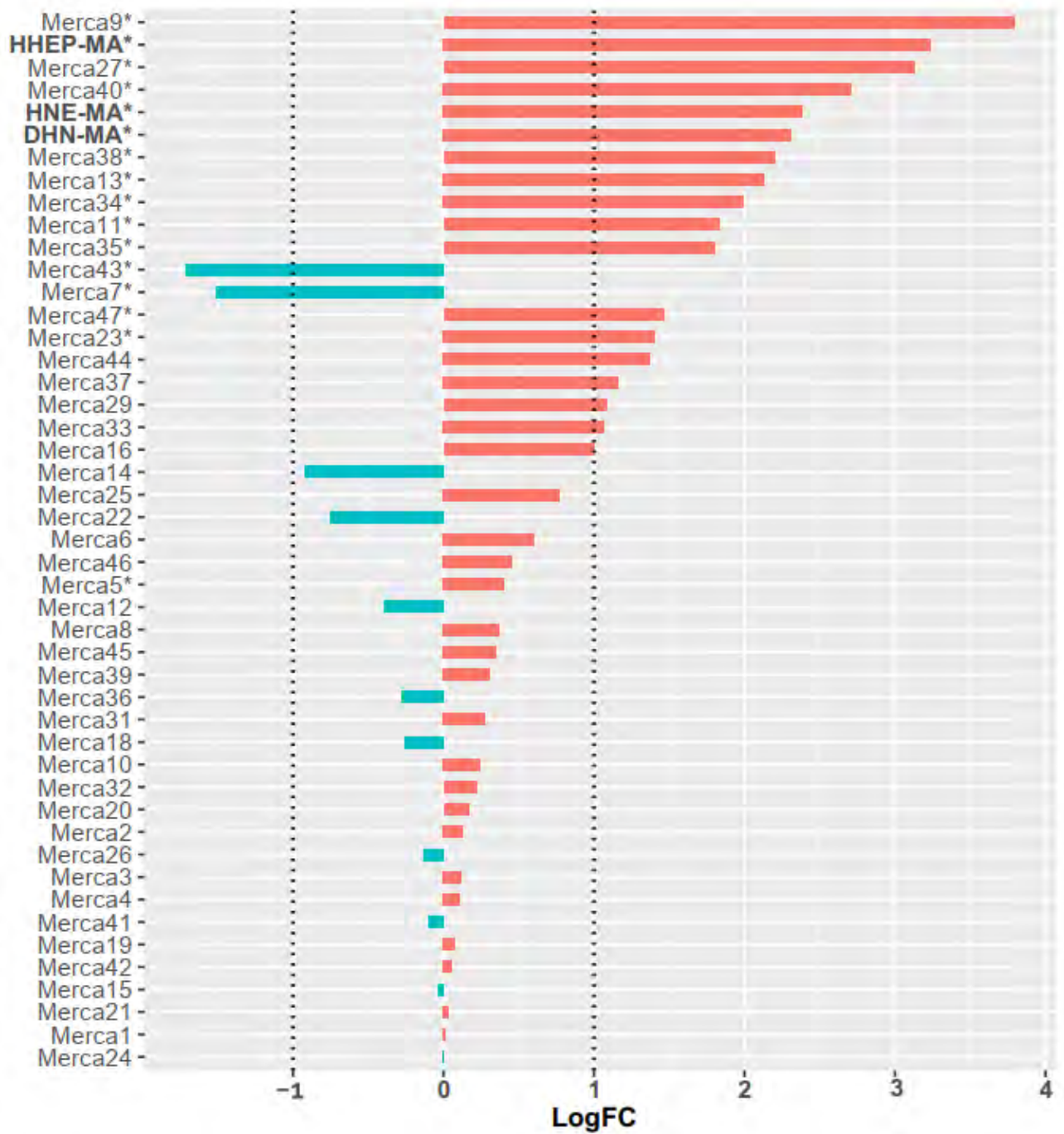


Figure 29 Log2 fold change of all 47 mercapturates detected in the animal experiment. « \* » correspond to mercapturates which VIP is > 1 in the OPLS-DA model

## References

- Abdul Rahim, M.B.H., Chilloux, J., Martinez-Gili, L., Neves, A.L., Myridakis, A., Gooderham, N., Dumas, M.-E., 2019. Diet-induced metabolic changes of the human gut microbiome: importance of short-chain fatty acids, methylamines and indoles. *Acta Diabetol.* 56, 493–500. <https://doi.org/10.1007/s00592-019-01312-x>
- Abete, I., Romaguera, D., Vieira, A.R., Lopez de Munain, A., Norat, T., 2014a. Association between total, processed, red and white meat consumption and all-cause, CVD and IHD mortality: a meta-analysis of cohort studies. *Br. J. Nutr.* 112, 762–775. <https://doi.org/10.1017/S000711451400124X>
- Abete, I., Romaguera, D., Vieira, A.R., Munain, A.L. de, Norat, T., 2014b. Association between total, processed, red and white meat consumption and all-cause, CVD and IHD mortality: a meta-analysis of cohort studies. *Br. J. Nutr.* 112, 762–775. <https://doi.org/10.1017/S000711451400124X>
- Andreeva, V.A., Deschamps, V., Salanave, B., Castetbon, K., Verdot, C., Kesse-Guyot, E., Hercberg, S., 2016. Comparison of Dietary Intakes Between a Large Online Cohort Study (Etude NutriNet-Santé) and a Nationally Representative Cross-Sectional Study (Etude Nationale Nutrition Santé) in France: Addressing the Issue of Generalizability in E-Epidemiology. *Am. J. Epidemiol.* 184, 660–669. <https://doi.org/10.1093/aje/kww016>
- Andreeva, V.A., Salanave, B., Castetbon, K., Deschamps, V., Vernay, M., Kesse-Guyot, E., Hercberg, S., 2015. Comparison of the sociodemographic characteristics of the large NutriNet-Santé e-cohort with French Census data: the issue of volunteer bias revisited. *J. Epidemiol. Community Health* 69, 893–898. <https://doi.org/10.1136/jech-2014-205263>
- Bastide, N.M., Chenni, F., Audebert, M., Santarelli, R.L., Taché, S., Naud, N., Baradat, M., Jouanin, I., Surya, R., Hobbs, D.A., Kuhnle, G.G., Raymond-Letron, I., Gueraud, F., Corpet, D.E., Pierre, F.H.F., 2015. A Central Role for Heme Iron in Colon Carcinogenesis Associated with Red Meat Intake. *Cancer Res.* 75, 870–879. <https://doi.org/10.1158/0008-5472.CAN-14-2554>
- Beloborodova, N., Bairamov, I., Olenin, A., Shubina, V., Teplova, V., Fedotcheva, N., 2012. Effect of phenolic acids of microbial origin on production of reactive oxygen species in mitochondria and neutrophils. *J. Biomed. Sci.* 19, 89. <https://doi.org/10.1186/1423-0127-19-89>
- Black, A.E., 2000. Critical evaluation of energy intake using the Goldberg cut-off for energy intake: basal metabolic rate. A practical guide to its calculation, use and limitations. *Int. J. Obes. Relat. Metab. Disord. J. Int. Assoc. Study Obes.* 24, 1119–1130. <https://doi.org/10.1038/sj.ijo.0801376>
- Bouvard, V., Loomis, D., Guyton, K.Z., Grosse, Y., Ghissassi, F.E., Benbrahim-Tallaa, L., Guha, N., Mattock, H., Straif, K., 2015. Carcinogenicity of consumption of red and processed meat. *Lancet Oncol.* 16, 1599–1600. [https://doi.org/10.1016/S1470-2045\(15\)00444-1](https://doi.org/10.1016/S1470-2045(15)00444-1)
- Carlsen, C.U., Møller, J.K.S., Skibsted, L.H., 2005. Heme-iron in lipid oxidation. *Coord. Chem. Rev., Inorganic Reaction Mechanisms, an appreciation of Henry Taube in his 90th year* 249, 485–498. <https://doi.org/10.1016/j.ccr.2004.08.028>
- Cheung, W., Keski-Rahkonen, P., Assi, N., Ferrari, P., Freisling, H., Rinaldi, S., Slimani, N., Zamora-Ros, R., Rundle, M., Frost, G., Gibbons, H., Carr, E., Brennan, L., Cross, A.J., Pala, V., Panico, S., Sacerdote, C., Palli, D., Tumino, R., Kühn, T., Kaaks, R., Boeing, H., Floegel, A., Mancini, F., Boutron-Ruault, M.-C., Baglietto, L., Trichopoulou, A., Naska, A., Orfanos, P., Scalbert, A., 2017. A metabolomic study of biomarkers of meat and fish intake. *Am. J. Clin. Nutr.* 105, 600–608. <https://doi.org/10.3945/ajcn.116.146639>
- Constante, M., Frago, G., Calvé, A., Samba-Mondonga, M., Santos, M.M., 2017. Dietary Heme Induces Gut Dysbiosis, Aggravates Colitis, and Potentiates the Development of Adenomas in Mice. *Front. Microbiol.* 8. <https://doi.org/10.3389/fmicb.2017.01809>
- Craig, C.L., Marshall, A.L., Sjöström, M., Bauman, A.E., Booth, M.L., Ainsworth, B.E., Pratt, M., Ekelund, U., Yngve, A., Sallis, J.F., Oja, P., 2003. International physical activity questionnaire: 12-country

- reliability and validity. *Med. Sci. Sports Exerc.* 35, 1381–1395. <https://doi.org/10.1249/01.MSS.0000078924.61453.FB>
- Di Iorio, B.R., Rocchetti, M.T., De Angelis, M., Cosola, C., Marzocco, S., Di Micco, L., di Bari, I., Accetturo, M., Vacca, M., Gobetti, M., Di Iorio, M., Bellasi, A., Gesualdo, L., 2019. Nutritional Therapy Modulates Intestinal Microbiota and Reduces Serum Levels of Total and Free Indoxyl Sulfate and P-Cresyl Sulfate in Chronic Kidney Disease (Medika Study). *J. Clin. Med.* 8. <https://doi.org/10.3390/jcm8091424>
- Ehrlich, A.M., Pacheco, A.R., Henrick, B.M., Taft, D., Xu, G., Huda, M.N., Mishchuk, D., Goodson, M.L., Slupsky, C., Barile, D., Lebrilla, C.B., Stephensen, C.B., Mills, D.A., Raybould, H.E., 2020. Indole-3-lactic acid associated with Bifidobacterium-dominated microbiota significantly decreases inflammation in intestinal epithelial cells. *BMC Microbiol.* 20, 357. <https://doi.org/10.1186/s12866-020-02023-y>
- Escudié, F., Auer, L., Bernard, M., Mariadassou, M., Cauquil, L., Vidal, K., Maman, S., Hernandez-Raquet, G., Combes, S., Pascal, G., 2018. FROGS: Find, Rapidly, OTUs with Galaxy Solution. *Bioinformatics* 34, 1287–1294. <https://doi.org/10.1093/bioinformatics/btx791>
- Farvid, M.S., Sidahmed, E., Spence, N.D., Mante Angua, K., Rosner, B.A., Barnett, J.B., 2021. Consumption of red meat and processed meat and cancer incidence: a systematic review and meta-analysis of prospective studies. *Eur. J. Epidemiol.* <https://doi.org/10.1007/s10654-021-00741-9>
- Gao, H., Liu, S., 2017. Role of uremic toxin indoxyl sulfate in the progression of cardiovascular disease. *Life Sci.* 185, 23–29. <https://doi.org/10.1016/j.lfs.2017.07.027>
- Giacomoni, F., Le Corguillé, G., Monsoor, M., Landi, M., Pericard, P., Pétéra, M., Duperier, C., Tremblay-Franco, M., Martin, J.-F., Jacob, D., Goullitquer, S., Thévenot, E.A., Caron, C., 2015. Workflow4Metabolomics: a collaborative research infrastructure for computational metabolomics. *Bioinformatics* 31, 1493–1495. <https://doi.org/10.1093/bioinformatics/btu813>
- González, A., Gálvez, N., Martín, J., Reyes, F., Pérez-Victoria, I., Dominguez-Vera, J.M., 2017. Identification of the key excreted molecule by *Lactobacillus fermentum* related to host iron absorption. *Food Chem.* 228, 374–380. <https://doi.org/10.1016/j.foodchem.2017.02.008>
- Guéraud, F., 2017. 4-Hydroxynonenal metabolites and adducts in pre-carcinogenic conditions and cancer. *Free Radic. Biol. Med.*, 4-Hydroxynonenal and Related Lipid Peroxidation Products 111, 196–208. <https://doi.org/10.1016/j.freeradbiomed.2016.12.025>
- Guéraud, F., Peiro, G., Bernard, H., Alary, J., Créminon, C., Debrauwer, L., Rathahao, E., Drumare, M.-F., Canlet, C., Wal, J.-M., Bories, G., 2006. Enzyme immunoassay for a urinary metabolite of 4-hydroxynonenal as a marker of lipid peroxidation. *Free Radic. Biol. Med.* 40, 54–62. <https://doi.org/10.1016/j.freeradbiomed.2005.08.011>
- Guéraud, F., Taché, S., Steghens, J.-P., Milkovic, L., Borovic-Sunjic, S., Zarkovic, N., Gaultier, E., Naud, N., Héliès-Toussaint, C., Pierre, F., Priymenko, N., 2015. Dietary polyunsaturated fatty acids and heme iron induce oxidative stress biomarkers and a cancer promoting environment in the colon of rats. *Free Radic. Biol. Med.* 83, 192–200. <https://doi.org/10.1016/j.freeradbiomed.2015.02.023>
- Hemeryck, L.Y., Rombouts, C., De Paepe, E., Vanhaecke, L., 2018. DNA adduct profiling of *in vitro* colonic meat digests to map red vs. white meat genotoxicity. *Food Chem. Toxicol. Int. J. Publ. Br. Ind. Biol. Res. Assoc.* 115, 73–87. <https://doi.org/10.1016/j.fct.2018.02.032>
- Hercberg, S., Castetbon, K., Czernichow, S., Malon, A., Mejean, C., Kesse, E., Touvier, M., Galan, P., 2010. The Nutrinet-Santé Study: a web-based prospective study on the relationship between nutrition and health and determinants of dietary patterns and nutritional status. *BMC Public Health* 10, 242. <https://doi.org/10.1186/1471-2458-10-242>
- Hinchman, C.A., Ballatori, N., 1994. Glutathione conjugation and conversion to mercapturic acids can occur as an intrahepatic process. *J. Toxicol. Environ. Health* 41, 387–409. <https://doi.org/10.1080/15287399409531852>
- Ijssennagger, N., Belzer, C., Hooiveld, G.J., Dekker, J., van Mil, S.W.C., Müller, M., Kleerebezem, M., van der Meer, R., 2015. Gut microbiota facilitates dietary heme-induced epithelial

- hyperproliferation by opening the mucus barrier in colon. *Proc. Natl. Acad. Sci.* 112, 10038–10043. <https://doi.org/10.1073/pnas.1507645112>
- Ijssennagger, N., Rijniere, A., de Wit, N.J.W., Boekschoten, M.V., Dekker, J., Schonewille, A., Muller, M., van der Meer, R., 2013. Dietary heme induces acute oxidative stress, but delayed cytotoxicity and compensatory hyperproliferation in mouse colon. *Carcinogenesis* 34, 1628–1635. <https://doi.org/10.1093/carcin/bgt084>
- Jamin, E.L., Costantino, R., Mervant, L., Martin, J.-F., Jouanin, I., Blas-Y-Estrada, F., Guéraud, F., Debrauwer, L., 2020. Global Profiling of Toxicologically Relevant Metabolites in Urine: Case Study of Reactive Aldehydes. *Anal. Chem.* 92, 1746–1754. <https://doi.org/10.1021/acs.analchem.9b03146>
- Ju, T., Kong, J.Y., Stothard, P., Willing, B.P., 2019. Defining the role of Parasutterella, a previously uncharacterized member of the core gut microbiota. *ISME J.* 13, 1520–1534. <https://doi.org/10.1038/s41396-019-0364-5>
- Keller, J., Baradat, M., Jouanin, I., Debrauwer, L., Guéraud, F., 2015. “Twin peaks”: Searching for 4-hydroxynonenal urinary metabolites after oral administration in rats. *Redox Biol.* 4, 136–148. <https://doi.org/10.1016/j.redox.2014.12.016>
- Keller, J., Chevolleau, S., Noguer-Meireles, M.-H., Pujos-Guillot, E., Delosière, M., Chantelauze, C., Joly, C., Blas-y-Estrada, F., Jouanin, I., Durand, D., Pierre, F., Debrauwer, L., Theodorou, V., Guéraud, F., 2020. Heme-Iron-Induced Production of 4-Hydroxynonenal in Intestinal Lumen May Have Extra-Intestinal Consequences Through Protein-Adduct Formation. *Antioxidants* 9, 1293. <https://doi.org/10.3390/antiox9121293>
- Khodorova, N.V., Rutledge, D.N., Oberli, M., Mathiron, D., Marcelo, P., Benamouzig, R., Tomé, D., Gaudichon, C., Pilard, S., 2019. Urinary Metabolomics Profiles Associated to Bovine Meat Ingestion in Humans. *Mol. Nutr. Food Res.* 63, 1700834. <https://doi.org/10.1002/mnfr.201700834>
- Kotronoulas, A., Pizarro, N., Serra, A., Robledo, P., Joglar, J., Rubió, L., Hernaéz, A., Tormos, C., Motilva, M.J., Fitó, M., Covas, M.-I., Solà, R., Farré, M., Saez, G., de la Torre, R., 2013. Dose-dependent metabolic disposition of hydroxytyrosol and formation of mercapturates in rats. *Pharmacol. Res.* 77, 47–56. <https://doi.org/10.1016/j.phrs.2013.09.001>
- Lapins, M., Eklund, M., Spjuth, O., Prusis, P., Wikberg, J.E., 2008. Proteochemometric modeling of HIV protease susceptibility. *BMC Bioinformatics* 9, 181. <https://doi.org/10.1186/1471-2105-9-181>
- Larsson, S.C., Orsini, N., 2014. Red meat and processed meat consumption and all-cause mortality: a meta-analysis. *Am. J. Epidemiol.* 179, 282–289. <https://doi.org/10.1093/aje/kwt261>
- Lassale, C., Castetbon, K., Laporte, F., Camilleri, G.M., Deschamps, V., Vernay, M., Faure, P., Hercberg, S., Galan, P., Kesse-Guyot, E., 2015. Validation of a Web-based, self-administered, non-consecutive-day dietary record tool against urinary biomarkers. *Br. J. Nutr.* 113, 953–962. <https://doi.org/10.1017/S0007114515000057>
- Lassale, C., Castetbon, K., Laporte, F., Deschamps, V., Vernay, M., Camilleri, G.M., Faure, P., Hercberg, S., Galan, P., Kesse-Guyot, E., 2016. Correlations between Fruit, Vegetables, Fish, Vitamins, and Fatty Acids Estimated by Web-Based Nonconsecutive Dietary Records and Respective Biomarkers of Nutritional Status. *J. Acad. Nutr. Diet.* 116, 427–438.e5. <https://doi.org/10.1016/j.jand.2015.09.017>
- Lassale, C., Péneau, S., Touvier, M., Julia, C., Galan, P., Hercberg, S., Kesse-Guyot, E., 2013. Validity of web-based self-reported weight and height: results of the Nutrinet-Santé study. *J. Med. Internet Res.* 15, e152. <https://doi.org/10.2196/jmir.2575>
- Le Leu, R.K., Young, G.P., Hu, Y., Winter, J., Conlon, M.A., 2013. Dietary red meat aggravates dextran sulfate sodium-induced colitis in mice whereas resistant starch attenuates inflammation. *Dig. Dis. Sci.* 58, 3475–3482. <https://doi.org/10.1007/s10620-013-2844-1>
- Lee, T.-H., Park, D., Kim, Y.J., Lee, I., Kim, S., Oh, C.-T., Kim, J.-Y., Yang, J., Jo, S.-K., 2020. Lactobacillus salivarius BP121 prevents cisplatin-induced acute kidney injury by inhibition of uremic toxins such as indoxyl sulfate and p-cresol sulfate *via* alleviating dysbiosis. *Int. J. Mol. Med.* 45, 1130–1140. <https://doi.org/10.3892/ijmm.2020.4495>

- Mafra, D., Borges, N.A., Cardozo, L.F.M. de F., Anjos, J.S., Black, A.P., Moraes, C., Bergman, P., Lindholm, B., Stenvinkel, P., 2018. Red meat intake in chronic kidney disease patients: Two sides of the coin. *Nutrition* 46, 26–32. <https://doi.org/10.1016/j.nut.2017.08.015>
- Mahé, F., Rognes, T., Quince, C., de Vargas, C., Dunthorn, M., 2014. Swarm: robust and fast clustering method for amplicon-based studies. *PeerJ* 2, e593. <https://doi.org/10.7717/peerj.593>
- Martin, O.C.B., Olier, M., Ellero-Simatos, S., Naud, N., Dupuy, J., Huc, L., Taché, S., Graillot, V., Levêque, M., Bézirard, V., Héliès-Toussaint, C., Estrada, F.B.Y., Tondereau, V., Lippi, Y., Naylies, C., Peyriga, L., Canlet, C., Davila, A.M., Blachier, F., Ferrier, L., Boutet-Robinet, E., Guéraud, F., Théodorou, V., Pierre, F.H.F., 2019a. Haem iron reshapes colonic luminal environment: impact on mucosal homeostasis and microbiome through aldehyde formation. *Microbiome* 7, 72. <https://doi.org/10.1186/s40168-019-0685-7>
- Martin, O.C.B., Olier, M., Ellero-Simatos, S., Naud, N., Dupuy, J., Huc, L., Taché, S., Graillot, V., Levêque, M., Bézirard, V., Héliès-Toussaint, C., Estrada, F.B.Y., Tondereau, V., Lippi, Y., Naylies, C., Peyriga, L., Canlet, C., Davila, A.M., Blachier, F., Ferrier, L., Boutet-Robinet, E., Guéraud, F., Théodorou, V., Pierre, F.H.F., 2019b. Haem iron reshapes colonic luminal environment: impact on mucosal homeostasis and microbiome through aldehyde formation. *Microbiome* 7, 72. <https://doi.org/10.1186/s40168-019-0685-7>
- Mervant, L., Tremblay-Franco, M., Jamin, E.L., Kesse-Guyot, E., Galan, P., Martin, J.-F., Guéraud, F., Debrauwer, L., 2021. Osmolality-based normalization enhances statistical discrimination of untargeted metabolomic urine analysis: results from a comparative study. *Metabolomics Off. J. Metabolomic Soc.* 17, 2. <https://doi.org/10.1007/s11306-020-01758-z>
- Obrenovich, M.E., Tima, M., Polinkovsky, A., Zhang, R., Emancipator, S.N., Donskey, C.J., 2017. Targeted Metabolomics Analysis Identifies Intestinal Microbiota-Derived Urinary Biomarkers of Colonization Resistance in Antibiotic-Treated Mice. *Antimicrob. Agents Chemother.* 61. <https://doi.org/10.1128/AAC.00477-17>
- Ohgaki, H., Hasegawa, H., Suenaga, M., Sato, S., Takayama, S., Sugimura, T., 1987. Carcinogenicity in mice of a mutagenic compound, 2-amino-3,8-dimethylimidazo[4,5-f]quinoxaline (MeIQx) from cooked foods. *Carcinogenesis* 8, 665–668. <https://doi.org/10.1093/carcin/8.5.665>
- Pan, A., Sun, Q., Bernstein, A.M., Schulze, M.B., Manson, J.E., Willett, W.C., Hu, F.B., 2011. Red meat consumption and risk of type 2 diabetes: 3 cohorts of US adults and an updated meta-analysis. *Am. J. Clin. Nutr.* 94, 1088–1096. <https://doi.org/10.3945/ajcn.111.018978>
- Patel, K.P., Luo, F.J.-G., Plummer, N.S., Hostetter, T.H., Meyer, T.W., 2012. The production of p-cresol sulfate and indoxyl sulfate in vegetarians versus omnivores. *Clin. J. Am. Soc. Nephrol. CJASN* 7, 982–988. <https://doi.org/10.2215/CJN.12491211>
- Philippe, C., Szabo de Edelenyi, F., Naudon, L., Druesne-Pecollo, N., Hercberg, S., Kesse-Guyot, E., Latino-Martel, P., Galan, P., Rabot, S., 2021. Relation between Mood and the Host-Microbiome Co-Metabolite 3-Indoxylsulfate: Results from the Observational Prospective NutriNet-Santé Study. *Microorganisms* 9, 716. <https://doi.org/10.3390/microorganisms9040716>
- Pierre, F., Taché, S., Petit, C.R., Van der Meer, R., Corpet, D.E., 2003. Meat and cancer: haemoglobin and haemin in a low-calcium diet promote colorectal carcinogenesis at the aberrant crypt stage in rats. *Carcinogenesis* 24, 1683–1690. <https://doi.org/10.1093/carcin/bgg130>
- Poesen, R., Mutsaers, H.A.M., Windey, K., van den Broek, P.H., Verweij, V., Augustijns, P., Kuypers, D., Jansen, J., Evenepoel, P., Verbeke, K., Meijers, B., Masereeuw, R., 2015. The Influence of Dietary Protein Intake on Mammalian Tryptophan and Phenolic Metabolites. *PloS One* 10, e0140820. <https://doi.org/10.1371/journal.pone.0140820>
- Samraj, A.N., Pearce, O.M.T., Läubli, H., Crittenden, A.N., Bergfeld, A.K., Banda, K., Gregg, C.J., Bingman, A.E., Secrest, P., Diaz, S.L., Varki, N.M., Varki, A., 2015. A red meat-derived glycan promotes inflammation and cancer progression. *Proc. Natl. Acad. Sci.* 112, 542–547. <https://doi.org/10.1073/pnas.1417508112>
- Sesink, A.L., Termont, D.S., Kleibeuker, J.H., Van der Meer, R., 2001. Red meat and colon cancer: dietary haem-induced colonic cytotoxicity and epithelial hyperproliferation are inhibited by calcium. *Carcinogenesis* 22, 1653–1659. <https://doi.org/10.1093/carcin/22.10.1653>



- Sesink, A.L.A., Termont, D.S.M.L., Kleibeuker, J.H., Meer, R.V. der, 1999. Red Meat and Colon Cancer: The Cytotoxic and Hyperproliferative Effects of Dietary Heme. *Cancer Res.* 59, 5704–5709.
- Sinha, R., Kullendorff, M., Gunter, M.J., Strickland, P., Rothman, N., 2005. Dietary benzo[a]pyrene intake and risk of colorectal adenoma. *Cancer Epidemiol. Biomark. Prev. Publ. Am. Assoc. Cancer Res. Cosponsored Am. Soc. Prev. Oncol.* 14, 2030–2034. <https://doi.org/10.1158/1055-9965.EPI-04-0854>
- Song, M., Chan, A.T., Sun, J., 2020. Influence of the Gut Microbiome, Diet, and Environment on Risk of Colorectal Cancer. *Gastroenterology* 158, 322–340. <https://doi.org/10.1053/j.gastro.2019.06.048>
- Sumner, L.W., Amberg, A., Barrett, D., Beale, M.H., Beger, R., Daykin, C.A., Fan, T.W.-M., Fiehn, O., Goodacre, R., Griffin, J.L., Hankemeier, T., Hardy, N., Harnly, J., Higashi, R., Kopka, J., Lane, A.N., Lindon, J.C., Marriott, P., Nicholls, A.W., Reily, M.D., Thaden, J.J., Viant, M.R., 2007. Proposed minimum reporting standards for chemical analysis. *Metabolomics* 3, 211–221. <https://doi.org/10.1007/s11306-007-0082-2>
- Touvier, M., Kesse-Guyot, E., Méjean, C., Pollet, C., Malon, A., Castetbon, K., Hercberg, S., 2011. Comparison between an interactive web-based self-administered 24 h dietary record and an interview by a dietitian for large-scale epidemiological studies. *Br. J. Nutr.* 105, 1055–1064. <https://doi.org/10.1017/S0007114510004617>
- Touvier, M., Méjean, C., Kesse-Guyot, E., Pollet, C., Malon, A., Castetbon, K., Hercberg, S., 2010. Comparison between web-based and paper versions of a self-administered anthropometric questionnaire. *Eur. J. Epidemiol.* 25, 287–296. <https://doi.org/10.1007/s10654-010-9433-9>
- Veettil, S.K., Wong, T.Y., Loo, Y.S., Playdon, M.C., Lai, N.M., Giovannucci, E.L., Chaiyakunapruk, N., 2021. Role of Diet in Colorectal Cancer Incidence. *JAMA Netw. Open* 4, e2037341. <https://doi.org/10.1001/jamanetworkopen.2020.37341>
- Wagner, S., Scholz, K., Donegan, M., Burton, L., Wingate, J., Völkel, W., 2006. Metabonomics and biomarker discovery: LC-MS metabolic profiling and constant neutral loss scanning combined with multivariate data analysis for mercapturic acid analysis. *Anal. Chem.* 78, 1296–1305. <https://doi.org/10.1021/ac051705s>
- Wang, X., Lin, X., Ouyang, Y.Y., Liu, J., Zhao, G., Pan, A., Hu, F.B., 2016. Red and processed meat consumption and mortality: dose-response meta-analysis of prospective cohort studies. *Public Health Nutr.* 19, 893–905. <https://doi.org/10.1017/S1368980015002062>
- WCRF, W., 2018. World Cancer Research Fund/American Institute for Cancer Research. Continuous Update Project Expert Report 2018. Meat, fish and dairy products and the risk of cancer.
- Wedekind, R., Kiss, A., Keski-Rahkonen, P., Viallon, V., Rothwell, J.A., Cross, A.J., Rostgaard-Hansen, A.L., Sandanger, T.M., Jakszyn, P., Schmidt, J.A., Pala, V., Vermeulen, R., Schulze, M.B., Kühn, T., Johnson, T., Trichopoulou, A., Peppas, E., La Vecchia, C., Masala, G., Tumino, R., Sacerdote, C., Wittenbecher, C., de Magistris, M.S., Dahm, C.C., Severi, G., Mancini, F.R., Weiderpass, E., Gunter, M.J., Huybrechts, I., Scalbert, A., 2020. A metabolomic study of red and processed meat intake and acylcarnitine concentrations in human urine and blood. *Am. J. Clin. Nutr.* 112, 381–388. <https://doi.org/10.1093/ajcn/nqaa140>
- Wikoff, W.R., Anfora, A.T., Liu, J., Schultz, P.G., Lesley, S.A., Peters, E.C., Siuzdak, G., 2009. Metabolomics analysis reveals large effects of gut microflora on mammalian blood metabolites. *Proc. Natl. Acad. Sci. U. S. A.* 106, 3698–3703. <https://doi.org/10.1073/pnas.0812874106>
- Wolk, A., 2017. Potential health hazards of eating red meat. *J. Intern. Med.* 281, 106–122. <https://doi.org/10.1111/joim.12543>
- Zheng, X., Xie, G., Zhao, A., Zhao, L., Yao, C., Chiu, N.H.L., Zhou, Z., Bao, Y., Jia, Weiping, Nicholson, J.K., Jia, Wei, 2011. The Footprints of Gut Microbial–Mammalian Co-Metabolism. *J. Proteome Res.* 10, 5512–5522. <https://doi.org/10.1021/pr2007945>
- Zhu, Y., Lin, X., Li, H., Li, Y., Shi, X., Zhao, F., Xu, X., Li, C., Zhou, G., 2016. Intake of Meat Proteins Substantially Increased the Relative Abundance of Genus *Lactobacillus* in Rat Feces. *PLOS ONE* 11, e0152678. <https://doi.org/10.1371/journal.pone.0152678>

- Zhu, Y., Lin, X., Zhao, F., Shi, X., Li, H., Li, Y., Zhu, W., Xu, X., Lu, C., Zhou, G., 2015. Meat, dairy and plant proteins alter bacterial composition of rat gut bacteria. *Sci. Rep.* 5, 15220. <https://doi.org/10.1038/srep15220>
- Zou, H., Hastie, T., 2005. Regularization and variable selection *via* the elastic net. *J. R. Stat. Soc. Ser. B Stat. Methodol.* 67, 301–320. <https://doi.org/10.1111/j.1467-9868.2005.00503.x>

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10	Model 11	Model 12
Methoxysalicylic acid	0.316							0.329	0.318	0.301		
Cinnamoylglycine	0.542	0.605			0.583		0.566			0.528		
2,2-Dimethylglutaric acid	0.551				0.574		0.480	0.588	0.509	0.477		0.500
p-Coumaric acid	0.508							0.541	0.531	0.500		
Hydroxyphenyllactic acid	0.556									0.545		
N-Acetyl-L-leucine	0.578	0.566	0.473	0.436	0.488	0.502	0.503	0.577	0.596	0.595	0.517	0.512
10-hydroxydecanoic acid	0.303	0.332				0.319		0.287	0.270	0.273	0.283	0.301
Indoxyl sulfate	1.593											
Merca14	0.466	0.561			0.551		0.531	0.520	0.494	0.443		
Tyrosine		0.597	0.584			0.520		0.585				
Merca15			0.425									
Merca16	0.565			0.586			0.499			0.478		
Merca53				0.525			0.575			0.562		
2-Methylguanosine					0.612							
3-(3-hydroxyphenyl)propanoic acid										0.375		

**Supplementary Table 1 Odd ratio values of each metabolite for all tested model. Odd ratio value was displayed only when  $p$ -value <0.05**

Model 1: adjusted on physical activity, BMI, education level, alcohol consumption, caloric intake

Model 2: Model 1 + adjusted on sum of vegetables and fruits

Model 3: Model 1 + adjusted on fibre intake

Model 4: Model 2 + adjusted on fibre intake

Model 5: Model 1 + adjusted on polyphenol intake

Model 6: Model 5 + adjusted on vitamins intake (sum of Vit C, Vit E, selenium, zinc)

Model 7: Model 5 + adjusted on calcium intake

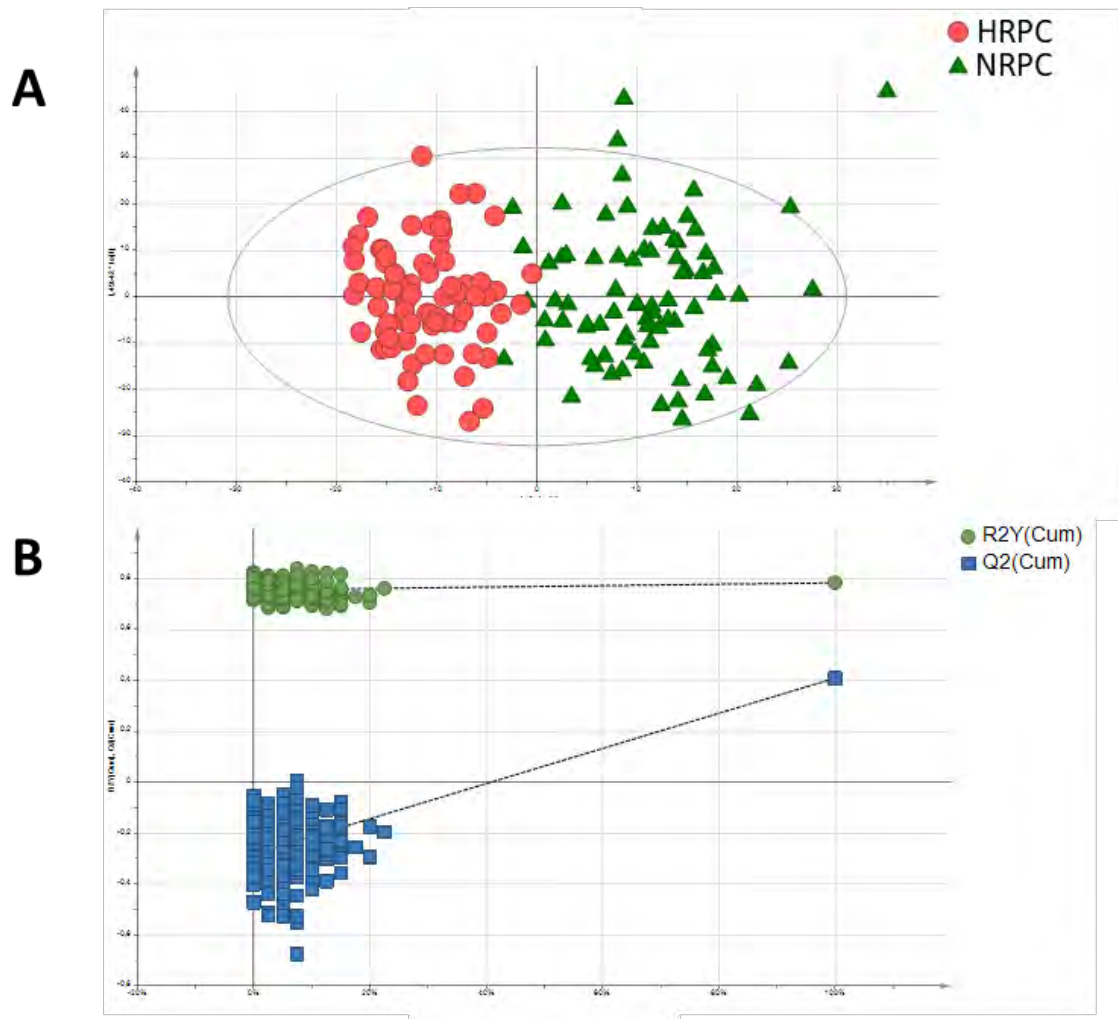
Model 8: Model 1 + adjusted on vitamins intake (sum of Vit C, Vit E, selenium, zinc)

Model 9: Model 8 + adjusted on calcium intake

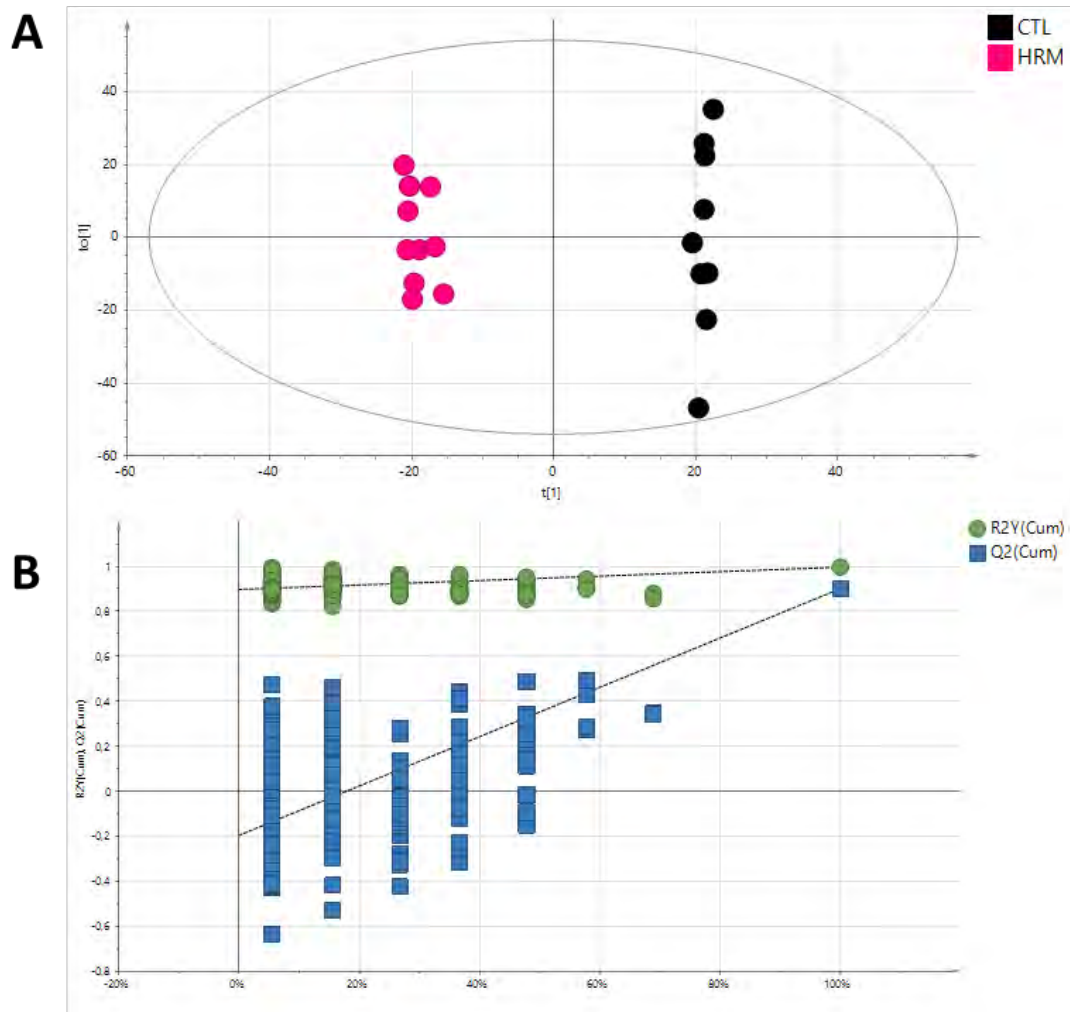
Model 10: Model 1 + adjusted on calcium intake

Model 11: Model 1 + adjusted non-heme iron intake

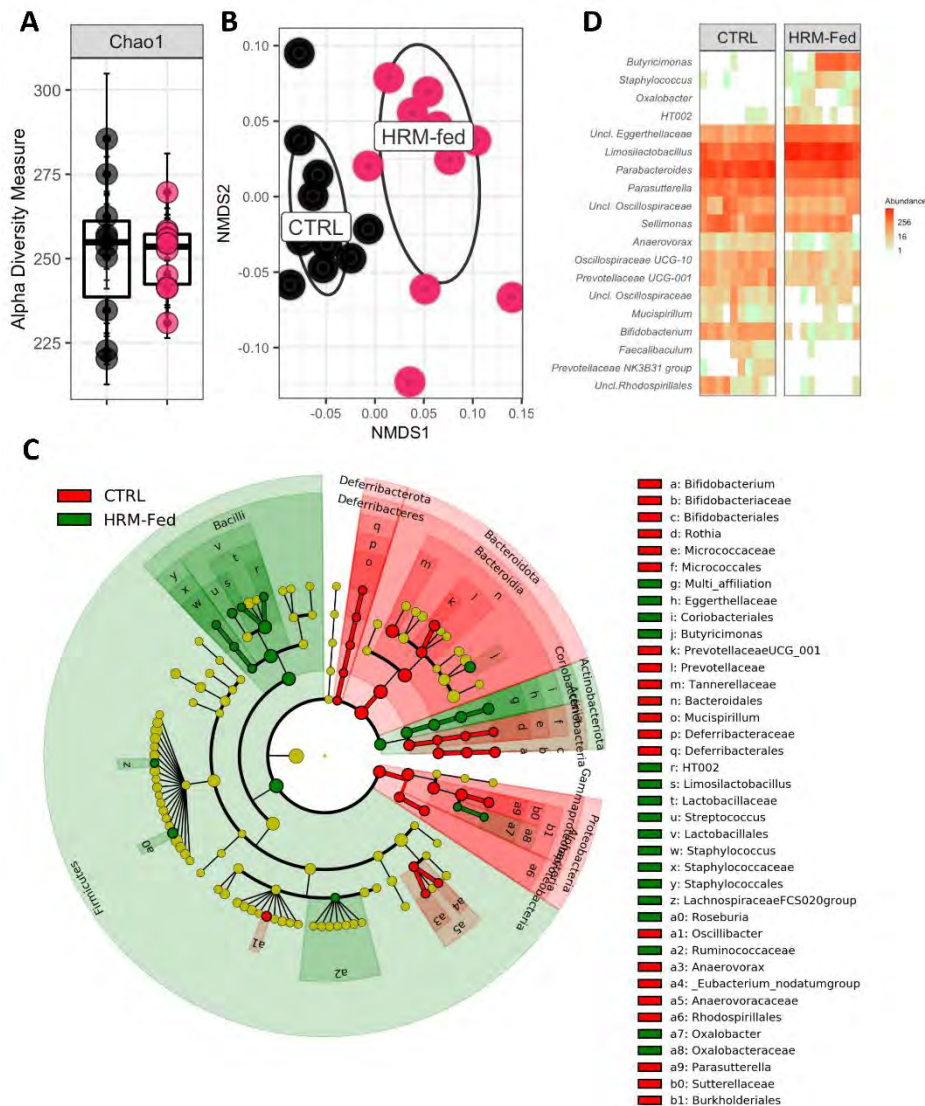
Model 12: Model 7 + adjusted on vitamins intake (sum of Vit C, Vit E, selenium, zinc)



**Supplementary Data 1 Two-dimensional OPLS-DA score plot based on all HRMS features from human experiment. Each symbol represent an individual data point projected onto the predictive (horizontal axis) and the orthogonal component (vertical axis) (A). Plot of permutation test Q<sup>2</sup> values corresponds to blue squares and R<sup>2</sup> to green circles (B)**



**Supplementary Data 2** Two-dimensional OPLS-DA score plot based on all HRMS *features* from *in vivo*. Each symbol represent an individual data point projected onto the predictive (horizontal axis) and the orthogonal component (vertical axis) (A). Plot of permutation test Q<sup>2</sup> values corresponds to blue squares and R<sup>2</sup> to green circles (B)



**Supplementary Data 3**  $\alpha$ -diversity determined by Chao 1 index (A)  $\beta$ -diversity the ordination based on UniFrac distances (B) Circular cladogram generated from LefSe analysis showing the most differentially abundant taxa enriched in microbiota from rats fed with control (red) or red meat (green) diets (C) Significant changes observed at the genus level between microbiota of rats fed with CTRL or HRM diet (D)

## **B. Principaux résultats, discussion et conclusions de l'article**

Cette étude réalisée sur des individus sains a permis de mettre en évidence des différences au niveau du métabolome urinaire au sein de la population des gros consommateurs de viande rouge et de charcuterie, en comparaison à la population ne consommant pas ces produits. Peu de différences ont cependant été constatées au niveau du profil des acides mercapturiques, ne permettant pas de conclure quant à une peroxydation lipidique plus élevée chez les individus consommant beaucoup de viande rouge et de charcuterie. En revanche, il a été déterminé que plusieurs métabolites dérivés du microbiote intestinal tels que la cinnamoyle glycine ou l'indoxyl sulfate étaient significativement associés à la consommation de viande rouge et de charcuterie. Il est notable que les métabolites microbiens négativement associés à la consommation de ces produits constituent des marqueurs de microbiote sain et diversifié tandis que l'indoxyl sulfate, dont l'abondance est plus élevée chez les gros consommateurs, est un métabolite associé à des effets délétères pour la santé tels qu'une augmentation du risque de maladies cardiovasculaires.

Ces résultats en faveur de l'hypothèse d'une dysbiose induite par la consommation de produits carnés ont été appuyés par des résultats similaires obtenus dans l'expérimentation nutritionnelle chez le rat. En effet, l'abondance relative des deux métabolites précédemment cités était significativement impactée par le type de régime alimentaire chez ces animaux. De plus, afin de renforcer cette hypothèse, un séquençage de l'ARN 16S a permis de mettre en évidence une diminution importante de l'abondance d'espèces bénéfiques telles que *Bifidobacterium* et *Parasutterella*. Il est important de noter que la dysbiose induite par la viande rouge, observée dans ce travail, semble différente de celle induite par les régimes à base d'hème purifié reportée notamment par Martin *et al.* ou Ijssennagger *et al.* (Ijssennagger *et al.*, 2012a; Martin *et al.*, 2019a) où une forte diminution d'espèces du genre *Lactobacillus* a été retrouvée.

Ainsi, si l'hème apparaît comme étant l'élément responsable d'une augmentation de la peroxydation lipidique induite par la viande rouge, la différence observée entre la signature microbienne induite par un régime riche en hème purifié et le régime riche en viande rouge utilisé dans cette étude suggère que ce composé ne constitue pas l'élément-clé de la dysbiose induite par la viande rouge. Cette observation est notamment appuyée par des études présentant des modifications similaires du microbiote lors de l'administration de régimes riches en viande rouge au cours d'expérimentations *in vivo* (Zhu *et al.*, 2016, 2015; Le Leu *et al.*, 2013).





## **Chapitre IV : Etude de l'impact de la teneur en fer héminique, du contexte lipidique et du statut en antioxydant**

### **A. Introduction**

Les travaux présentés dans cette partie visent à mieux caractériser les effets de la consommation de viande rouge, du contenu lipidique du régime et de l'ajout d'antioxydant dans la ration, sur des modèles *in vivo* de rats. L'étude de ces différents paramètres est motivée par l'hypothèse que la peroxydation lipidique hème-induite serait un élément clé des effets néfastes induits par la consommation de viande rouge et de charcuterie. Ainsi, l'ajout d'antioxydant dans les régimes pourrait permettre de réduire la peroxydation lipidique et atténuer les effets délétères de la consommation de ces produits. L'autre hypothèse également étudiée dans cette partie est liée à la nature lipidique des régimes. En effet, en fonction de la composition lipidique des régimes (taux de lipides insaturés, position des insaturations), les produits de peroxydation lipidique résultants peuvent varier et avoir des propriétés réactives plus ou moins importantes. De plus, de nombreuses études pointent l'impact de la composition lipidique, en particulier en acides gras  $\omega$ -3 ou  $\omega$ -6, des régimes alimentaires sur la santé. En effet, des régimes à haute teneur en acides gras  $\omega$ -6 ont été associés à des effets délétères tels qu'une augmentation de l'inflammation ou encore la promotion de certains cancer tels que le cancer colorectal (Huerta-Yépez *et al.*, 2016). A l'inverse, de nombreux travaux font état d'effets contraires des acides gras  $\omega$ -3 qui ont fréquemment été associés à une inhibition de la carcinogénèse colorectale ou de l'inflammation (Djuricic and Calder, 2021; Simopoulos, 2008).

Ainsi, dans ce chapitre, l'effet de ces différents paramètres a été étudié à plusieurs niveaux. Premièrement, l'impact sur la peroxydation lipidique a été étudié, par le profilage des acides mercapturiques dans les urines, mais aussi par l'analyse des contenus fécaux en réalisant une analyse « aldéhydrique » dont la méthode est décrite dans une publication disponible en fin de manuscrit (Annexe 16). En parallèle de ces analyses, le « *double-workflow* » proposé dans le Chapitre II : Développement et optimisation de la méthode d'analyse, de ce manuscrit a également permis d'évaluer l'impact de ces différents paramètres sur le métabolome urinaire. La réponse transcriptionnelle, en particulier des gènes impliqués dans le métabolisme des aldéhydes hème-induits ou dans l'induction de défenses antioxydantes, a également été étudiée. Pour finir, la réponse toxicologique associée aux différents régimes a été déterminée, en étudiant notamment la cytotoxicité des eaux fécales, les dommages à l'ADN au niveau colique ou encore l'effet promoteur sur la carcinogénèse colorectale.

## **B. Matériel et méthodes**

### **B.1. Réactifs et tampons utilisés**

Tous les solvants utilisés pour les analyses LC-MS sont de qualité LC-MS et proviennent de Fisher Scientific (Thermo Fisher Scientific, Illkirch, France). Le DHN-MA a été synthétisé au laboratoire d'après une méthode décrite précédemment (Alary *et al.*, 1998). La *N*-acétyl-S-phenylcystéine (NAC-Phe-Cys) provient de Sigma-Aldrich (Saint Quentin Fallavier, France).

### **B.2. Analyses LC-MS d'échantillons urinaires**

#### ***B.2.1. Préparation des échantillons urinaires pour analyse LC-MS***

L'osmolalité des échantillons est mesurée à l'aide d'un osmomètre à point de congélation (Loeser Messtechnik, Berlin, Germany). Les échantillons sont ensuite dilués avec de l'eau MilliQ afin d'uniformiser l'osmolalité, puis dilués au ¼ dans la phase mobile A (H<sub>2</sub>O-méthanol (95 :5, v/v) + 0,1 % d'acide acétique). Enfin, une centrifugation pendant 10 min à 10 000 g est réalisée, le surnageant est alors récupéré pour analyse.

#### ***B.2.2. Analyse LC-MS***

L'ensemble des analyses LC-MS a été réalisé selon les paramètres décrits dans le Chapitre II : Développement et optimisation de la méthode d'analyse. Seules quelques différences sont néanmoins à préciser : pour chaque analyse, des échantillons de QC-dilués ont été injectés en triplicat au début de chaque batch et des échantillons QC sont injectés régulièrement tout au long de la séquence (minimum de 8 par séquence). Pour chaque analyse des échantillons, des QC de groupe sont constitués (correspondant à un pool des échantillons d'un même groupe biologique) et sont analysés en triplicat en mode MS<sup>E</sup>.

#### ***B.2.3. Traitement des données LC-MS***

Les données issues des analyses LC-MS ont été traitées avec XCMS 3.0 disponible sous l'environnement collaboratif Workflow4Metabolomics (Giacomoni *et al.*, 2015). Le détail des paramètres utilisés est disponible en Annexe 1. A l'issue du traitement des données, toutes les *features* présentant un coefficient de corrélation avec les échantillons de QC-dilués < 0,6 ont été éliminées. Par la suite, les *features* dont le signal moyen dans les QC n'était pas 3 fois supérieur au signal dans les blancs ont été éliminés. Enfin, après l'application d'une correction LOESS, les variables présentant un coefficient de variation (CV) > 30 % ont été éliminées.

Les données issues des acquisitions MS<sup>E</sup> réalisées sur les QC de groupe ont été traitées à l'aide du

logiciel UNIFI™ (Waters, Manchester, UK) à l'aide de la « Méthode 4 » décrite dans le Chapitre II : Développement et optimisation de la méthode d'analyse. Seules les variables présentant la perte de neutre caractéristique des acides mercapturiques de 129,0426 Da détectés dans l'ensemble des répliqués pour un groupe donné ont été considérées. Par la suite, des acquisitions en MS/MS ciblées ont été réalisées, sur chacune des variables précédentes, afin de confirmer la perte de neutre détectée.

#### ***B.2.4. Analyses statistiques***

Afin d'identifier les métabolites et les conjugués mercapturiques d'intérêt, des analyses statistiques ont été réalisées pour déterminer les *features* impliqués dans la discrimination des différents groupes biologiques. Les analyses multivariées ACP et OPLS-DA ont été réalisées à l'aide du logiciel Simca-P 14.0 (Umetrics, Umea, Suède). Les données ont été log transformées (log10) et ont subi une standardisation Pareto avant analyse. La capacité prédictive ( $Q^2$ ) et des tests de permutations ont été utilisés pour déterminer la validité et la robustesse des modèles OPLS-DA. Seuls les modèles présentant une capacité prédictive supérieure à 0.4 et un  $Q^2$  intercept  $< 0.05$  (Lapins *et al.*, 2008) ont été considérés comme valides. Les *features* présentant un score VIP  $> 1$  ont été considérées comme statistiquement significatives. Des tests univariés non paramétriques, tels que Wilcoxon ou Kruskal-Wallis ont également été réalisés et les *features* présentant une  $p$ -value  $< 0.05$  étaient considérées comme significatives.

#### ***B.2.5. Identification des métabolites***

L'annotation des métabolites est réalisée par un *matching* entre les données issues du traitement XCMS et une base de données interne au laboratoire. Pour chaque annotation potentielle, un spectre de fragmentation MS/MS est réalisé puis comparé au spectre MS/MS du standard correspondant. Les variables présentant le même  $m/z$ , le même temps de rétention et un spectre de fragmentation similaire ont été considérées comme annotées au niveau 1 selon la classification proposée par Sumner *et al.* (Sumner *et al.*, 2007).

### **B.3. Analyses LC-MS d'échantillons d'eaux fécales**

#### ***B.3.1. Préparation des échantillons***

Les eaux fécales sont préparées par l'ajout de 1 mL d'eau distillé et de 50  $\mu$ L de BHT 0,45 M (2,6-Di-tert-butyl-4-méthyl-phénol) à 0,5 g de fèces (la collecte des fèces est détaillée dans la partie « Expérimentation animale »). Les échantillons sont ensuite homogénéisés dans des tubes Lysing Matrix E à l'aide d'un FastPrep puis centrifugés 20 min à 3200 g. Le surnageant est alors collecté et l'échantillon résultant est immédiatement soumis à une étape de dérivation pour stabiliser les aldéhydes.

Pour la dérivation, 60  $\mu\text{L}$  de tampon PIPES (0,1 M, pH 6,5) et 120  $\mu\text{L}$  de BBHA à 50 nmol/ $\mu\text{L}$  (1-((ammoniooxy)methyl)-2-bromobenzene chloride) sont ajoutés à 100  $\mu\text{L}$  d'eau fécale. Les échantillons sont ensuite conservés 1 h à 4 °C avant d'être placés à -20 °C pour conservation.

Enfin, une étape de purification par SPE est réalisée sur ces échantillons à l'aide de cartouches Agilent Bond Elut (100 mg, 1 mL). Les cartouches sont conditionnées par 1 mL de méthanol puis 1 mL d'eau milli-Q®. Les échantillons dérivatisés sont alors chargés sur la cartouche puis deux étapes de lavage successives sont réalisées par l'ajout de 1 mL de PIPES puis 1 mL d'un mélange d'acide trifluoroacétique à 0,05 % / méthanol (3:2, v / v). L'élution est ensuite réalisée à l'aide de 400  $\mu\text{L}$  de méthanol.

### ***B.3.2. Analyses LC-MS***

Les extraits sont analysés sur un système LC-MS constitué d'un système chromatographique HPLC Ultimate 3000 (Dionex, Thermo Scientific, Les Ulis, France) couplé à un spectromètre de masse LTQ-Orbitrap XL (Thermo Scientific, Les Ulis, France). La séparation chromatographique est effectuée à l'aide d'une colonne Kinetex C18 (150  $\times$  2,1 mm, 2,6  $\mu\text{m}$ ) (Phenomenex, Le Pecq, France) maintenue à 40 °C. Pour l'élution un gradient linéaire est réalisé à l'aide des phases A : H<sub>2</sub>O/Acétonitrile (95:5 (v/v)) + 0,1% d'acide formique et B : Acétonitrile / H<sub>2</sub>O (95:5 (v/v)) + 0,1% d'acide formique. Le gradient est programmé comme suit : 0 % de B de 0 à 3 min puis une étape linéaire de 0 % à 100 % de B de 3 à 15 min puis un plateau à 100 % de B pendant 10 min avant de revenir aux conditions initiales. Le débit est fixé à 0,2 mL.min<sup>-1</sup>.

Les spectres MS ont été acquis sur une gamme de  $m/z$  de 80 à 800 en ionisation positive à une résolution de 30 000 en utilisant les paramètres suivants : voltage du spray à 4,5 kV, température du capillaire 300 °C, les débits des *sheath gas* et *auxiliary gas* sont fixés respectivement à 35 et 5 unités arbitraires. Les analyses MS/MS, conduites pour confirmer les identifications, ont été réalisées en fixant l'énergie de collision normalisée (NCE) à 25 %.

### ***B.3.3. Traitement des données***

Les données issues des analyses LC-MS ont été traitées avec XCMS 3.0 disponible sous l'environnement collaboratif Workflow4Metabolomics (Giacomoni *et al.*, 2015). Le détail des paramètres utilisés est disponible en Annexe 1. Les *features* correspondant aux molécules dérivatisées ont été extraites à partir du massif isotopique de l'atome de brome contenu dans le dérivé BBHA à l'aide de la stratégie suivante : chaque ion  $[M+H]^+$  possédant un isotopologue d'intensité équivalente (ratio compris entre 0,667 et 1,33) et présentant une différence de masse  $\Delta M = 1.998 (\pm 0.001)$  correspondant à la différence de masse entre deux isotopes du brome a été considéré.

#### ***B.3.4. Analyses statistiques***

Les paramètres et outils utilisés pour les analyses statistiques sont décrits ci-dessus dans la partie « Analyse LC-MS d'échantillon urinaires ».

#### ***B.3.5. Identification des métabolites***

La stratégie et les critères utilisés pour l'identification des métabolites sont dans ci-dessus dans la partie « Analyse LC-MS d'échantillon urinaires ».

### **B.4. Expérimentations animales et régimes**

#### ***B.4.1. Expérimentation animale nutritionnelle***

Un total de 90 (n = 10 rats par groupes) rats mâles F344 âgés de 4 semaines fournis par Charles River (Allemagne) ont été hébergés en cage conventionnelle par groupe de 3 ou 4 par cage dans une salle à température contrôlée avec un cycle photopériodique de 12 h. Cette expérimentation animale a été autorisée par le Ministère de l'enseignement supérieur, de l'innovation et de la recherche en accord avec le comité d'éthique local (Projet APAFiS #22200-2019093015339804v2). Après réception des animaux, une période d'acclimatation de 10 jours a été observée avant le début de l'expérimentation durant laquelle l'ensemble des animaux ont été nourris avec le régime contrôle utilisé lors de l'expérimentation. Les animaux ont bénéficié d'un accès illimité à l'eau et la nourriture durant toute la durée de l'expérimentation. Afin de limiter l'oxydation des aliments durant la journée, les régimes étaient distribués quotidiennement en fin de journée. Après 16 jours de régimes expérimentaux, les animaux ont été placés dans des cages métaboliques individuelles pendant 24 h. Un suivi de la consommation d'eau et de nourriture ainsi que la collecte de fèces et d'urine ont été réalisés après 24 h d'hébergement. Des prélèvements de fèces à l'anus, destinés au séquençage du microbiote, ont également été réalisés aux premiers et derniers jours de la distribution des régimes expérimentaux. A J22 après le démarrage des régimes expérimentaux les rats ont été sacrifiés.

La composition des différents régimes expérimentaux est détaillée dans le Tableau 4 et la composition des farines utilisées pour la constitution des régimes expérimentaux est renseignée en Annexe 5. La notation utilisée pour décrire les régimes tout au long de ce manuscrit est la suivante : **BO = bœuf ; PU = poulet ; SA = huile de carthame ; FO = huile de poisson ; R = antioxydants** (mélange ascorbate tocophérol) et **CON = contrôle** (régime sans viande type AIN-76), par exemple la dénomination PUSA correspond ainsi au régime à base de poulet et d'huile de carthame.

Tableau 4 Composition des régimes expérimentaux. \* Toutes les quantités sont exprimées en g pour 100 g eq sec

	CON	BOSA	BOSAR	BOFO	BOFOR	PUSA	PUSAR	PUFO	PUFOR
Farine régime contrôle *	73.34	-	-	-	-	-	-	-	-
Farine complément *	0	50	50	50	50	50	50	50	50
Huile de poisson (FO) *	-	-	-	5	5	-	-	5	5
Huile de carthame (SA) *	5	5	5	-	-	5	5	-	-
Eau*	-	-	-	-	-	-	-	-	-
Saindoux *	7.1	-	-	-	-	5.5	5.5	5.5	5.5
Bœuf (BO) *	-	45	45	45	45	-	-	-	-
Poulet (PU) *	-	-	-	-	-	40	40	40	40
Ascorbate*	-	-	0.042	-	0.042	-	0.042	-	0.042
Tocophérol *	-	-	0.1	-	0.1	-	0.1	-	0.1
Caséine *	6.56	-	-	-	-	-	-	-	-
Méthionine *	0.084	-	-	-	-	-	-	-	-
Tryptophane *	0.038	-	-	-	-	-	-	-	-
Gélatine de porc *	8	-	-	-	-	-	-	-	-

#### B.4.2. Expérimentation animale de carcinogénèse

Un total de 72 rats F344 âgés de 4-5 semaines fournis par Charles River (Allemagne) et répartis en 5 groupes expérimentaux (BOFO (n = 12), BOSA (n = 18), BOSAR (n = 12), PUFO (n = 12) et PUSA (n= 18)) ont été hébergés en cage conventionnelle en binôme dans une salle à température contrôlée avec un cycle photopériodique de 12 h. Cette expérimentation animale a été autorisée par le Ministère de l'enseignement supérieur, de l'innovation et de la recherche en accord avec le comité d'éthique local (Project APAFIS 27180\_2020091017209910\_v2). Une semaine après réception, les animaux sont soumis à une injection intrapéritonéale d'AOM à une dose de 20 mg/kg pour initier la cancérogenèse colique. Après une période d'acclimatation de 12 jours, les régimes expérimentaux ont été distribués quotidiennement durant 100 jours pour étudier leur effet sur la promotion de la cancérogenèse. Après 80 jours de régimes expérimentaux, les animaux ont été placés dans des cages métaboliques individuelles pendant 24 h pour suivre leur consommation d'eau et de nourriture ainsi que pour collecter les fèces et l'urine. Au bout de 100 jours d'expérimentation les animaux ont été sacrifiés. La composition des régimes expérimentaux utilisés est décrite dans le Tableau 4.

#### B.5. Dosages du DHN-MA et 8-OH-2-deoxyguanosine par immunodosage

Les urines collectées lors du passage en cage à métabolisme puis normalisées selon l'osmolalité ont été utilisées pour déterminer la concentration de DHN-MA urinaire. Pour cela une méthode immuno-

enzymatique compétitive développée par Guéraud *et al.* (Guéraud *et al.*, 2006) a été utilisée. Le principe de la méthode repose sur la compétition pour la liaison à un anticorps monoclonal entre le composé d'intérêt (ici le DHN-MA) et un traceur (DHN-MA couplé à l'acétylcholinestérase). Un total de 50 µL d'urine normalisée selon l'osmolalité (puis diluée au 1/100<sup>ème</sup> ou 1/1000<sup>ème</sup> en fonction du groupe expérimental), 50 µL d'anticorps anti DHN-MA et 50 µL de traceur ont été déposés sur une plaque contenant un anticorps immobilisé dirigé contre l'anticorps anti DHN-MA. Après 18 h d'incubation les plaques sont lavées, puis 200 µL de réactif d'Ellman (substrat de l'acétylcholinestérase) sont ajoutés afin de procéder à la révélation. Après 2 h d'incubation, les plaques sont lues à l'aide d'un lecteur de plaque à la longueur d'onde 414 nm.

Le même procédé est utilisé pour la mesure de la concentration de la 8-OH-2-deoxyguanosine. Les urines normalisées sont diluées au 1/20<sup>ème</sup> puis le dosage est réalisé selon les spécifications du kit utilisé (Cayman Chemical, DNA/RNA Oxidative Damage (Clone 7E6.9) ELISA Kit).

### **B.6. Dosage des TBARS dans les eaux fécales**

Le protocole utilisé est basé sur une méthode décrite par Ohkawa *et al.* (Ohkawa *et al.*, 1979). Le principe de ce dosage est basé sur la formation d'un conjugué entre l'acide thiobarbiturique (TBA) et des substances réactives telles que les aldéhydes issus de la peroxydation lipidique. Les diènes résultant de cette réaction sont quantifiables par spectrophotométrie car ils émettent un signal à 532 nm. Ainsi, 10 µL d'eau fécale ont été déposés en duplicat dans des plaques 96 puits avec soit 190 µL d'une solution « mix » (préparée comme suit : 20 mL d'eau distillée, 4,2 mL d'acide acétique et 1,75 mL de SDS 20%) soit 190 µL d'une solution « mix + TBA » (préparée en dissolvant 175 mg de TBA dans 25 mL de solution « mix »). Une gamme de calibration de malondialdéhyde (MDA) de 0 à 400 µM a également été déposée en suivant le même procédé. Après 1 h d'incubation à 95 °C les plaques sont lues au lecteur de plaque à 532 nm.

### **B.7. Cytotoxicité des eaux fécales**

#### ***B.7.1. Préparation des eaux fécales***

Les eaux fécales préparées à partir des fèces récoltées en cage à métabolisme ont été diluées au 1/20<sup>ème</sup> dans du DMEM (*Dulbecco/Vogt modified Eagle's minimal essential medium*) sans sérum de veau foetal (SVF) puis filtrées à 0,22 µm en condition stérile.

#### ***B.7.2. Culture cellulaire et cytotoxicité***

L'activité cytotoxique des eaux fécales a été mesurée sur une lignée de cellules murines d'épithélium colique sain provenant d'Immortomouse® croisées avec une lignée de souris saine C57BL6. Les



Immortomouse sont des souris modifiées génétiquement portant une forme thermosensible et inductible en présence d'interféron gamma (INF- $\gamma$ ) d'un gène suppresseur de p53 : le gène « *SV40 large T antigen* » (Jat *et al.*, 1991). Ainsi, en condition dites « permissives » *i.e.* à 33 °C et en présence d'INF- $\gamma$ , p53 est inactivé ce qui conduit à une prolifération des cellules de manière indéfinies. A l'inverse à 37 °C et sans INF- $\gamma$  les cellules présentent un phénotype de primo-culture.

Les cellules sont cultivées sur plaque 96 puits dans du DMEM (auquel est ajouté 10 % de SVF, 1 % de glutamine et 1 % de pénicilline/streptomycine). Les cellules sontensemencées à 10 000 cellules/puits et cultivées en condition permissives à 33 °C en présence d'interféron gamma et de facteur de croissance (EGF) jusqu'à atteindre une confluence d'environ 70 %. Après 24 h en conditions non-permissives (37 °C sans interféron ni facteur de croissance), 100  $\mu$ L d'eaux fécales sont déposées sur les cellules qui sont ensuite incubées 24 h à 37 °C. La mesure de la cytotoxicité est réalisée sur un lecteur de plaque après ajout de Wst-1, réactif métabolisé par les mitochondries conduisant à la formation d'un chromogène émettant à 440 nm.

### **B.8. Expression génique dans la muqueuse colique**

Un total de 1 mL de Tri reagent (Molecular Research Center) a été ajouté à muqueuse colique prélevée lors de la dissection des rats afin d'en extraire l'ARN. Les échantillons ont ensuite été homogénéisés dans des tubes Lysing Matrix D à l'aide d'un FastPrep, puis 200  $\mu$ L de dichlorométhane ont été ajoutés. Une fois la phase aqueuse collectée, 500  $\mu$ L d'isopropanol sont ajoutés afin de précipiter l'ARN. Les culots d'ARN sont lavés avec 1 mL d'éthanol à 75 % avant d'être dilués dans l'eau. Les extraits d'ARN totaux (1  $\mu$ g) sont ensuite reverse-transcrits à l'aide du iScript™ Reverse Transcription Supermix (Bio-Rad) afin d'être analysés par qPCR en temps réel sur un instrument ViiA 7 Real-Time PCR System (Applied Biosystems). La liste des amorces utilisées est disponible en Annexe 6. Les données sont analysées à l'aide du logiciel LinRegPCR v.11 et normalisées par rapport au niveau d'expression du gène Polymerase II Subunit A (POLR2A).

### **B.9. Comptage histologique des lésions prénéoplasiques**

Lors de la dissection des rats de l'expérimentation de carcinogénèse, les colons ont été prélevés, ouverts longitudinalement puis étalés à plat avant d'être fixés et conservés dans de la formaline à 10 %. Avant lecture, les colons ont été rincés à l'eau distillée puis plongés dans du bleu de méthylène 0,05 % afin d'être colorés. La lecture et le comptage des ACF a été réalisé à l'aide d'un microscope optique (grossissement x40). Le nombre d'ACF ainsi que la multiplicité de ces derniers (nombre de cryptes aberrantes par ACF) ont été enregistrés. Le nombre de tumeurs a également été reporté.

A l'issue du comptage des ACF, les colons ont été coloré à l'aide d'une coloration HID-AB (*high iron*

*diamine/alcian blue*) (Spicer, 1965). Le nombre de MDF a été dénombré à l'aide d'un microscope optique (grossissement x40).

### **B.10. Dosage de l'hème dans les eaux fécales**

La méthode utilisée pour le dosage de l'hème dans les eaux fécales est basée sur celle décrite par Van den Berg *et al.* (van den Berg *et al.*, 1988). Cette méthode de quantification de l'hème repose sur la transformation, par du sulfate de fer ferreux en milieu acide, de la protoporphyrine en porphyrine fluorescente quantifiable par spectrophotométrie. Après précipitation des protéines, par ajout de 125  $\mu\text{L}$  d'isopropanol acidifié à 120  $\mu\text{L}$  d'eaux fécales, puis centrifugation, 25  $\mu\text{L}$  de la phase supérieure ont été récupérés. Puis, 500  $\mu\text{L}$  d'acide acétique et 25  $\mu\text{L}$  d'un mélange sulfate de fer / acide chlorhydrique ont été ajoutés aux échantillons ensuite incubés 30 min à 60 °C. Un total de 1 mL d'isopropanol à 50 % a ensuite été ajouté afin de stopper la réaction. Les échantillons ont ensuite été déposés en plaque 96 puits afin de procéder à la quantification de la fluorescence émise à 605 nm à l'aide d'un lecteur de plaque automatique. Parallèlement à la préparation des échantillons une gamme étalon d'hémoglobine (de 0 à 140  $\mu\text{M}$ ) a également été préparée en suivant le même protocole.

### **B.11. Détermination du taux de phosphorylation de l'histone H2AX**

Pour chaque rat, un morceau de côlon proximal prélevé lors de la dissection a été fixé dans la formaline à 10 % avant d'être inclus dans la paraffine. Les blocs de paraffine ont été coupés en section de 4  $\mu\text{m}$ .

Les sections de paraffine ont été réhydratées et l'activité peroxydase endogène a été inactivée par une incubation de 10 minutes dans une solution BLOXALL® (Vector Laboratories). Après lavage des lames, le démasquage des sites antigéniques a été effectué par chauffage au micro-ondes (14min) dans un tampon citrate 0,01 M (pH 6,0). Après incubation dans du sérum de blocage à 2,5 % (ImmPRESS® HRP Goat anti-Rabbit IgG Polymer Detection Kit, Vector Laboratories), les sections ont été incubées une nuit à 4 °C avec de l'anti- $\gamma\text{H2Ax}$  de lapin (1/200) dilué dans une solution Animal-Free Blocker® (Vector Laboratories). Après lavage, les lames ont été incubées pendant 30 minutes avec le réactif polymère ImmPRESS® pour IHC qui intègre d'anticorps secondaire et d'enzymes pour une détection IHC de haute sensibilité. Le kit de substrat DAB (3,3'-Diaminobenzidine) pour la peroxydase (ImmPACT®DAB, Vector laboratories) a été utilisé comme chromogène. Les sections ont ensuite été contre-colorées et montées. L'ensemble des lames a été analysé en aveugle par le même expérimentateur.

## **B.12. Imagerie par spectrométrie de masse**

### ***B.12.1. Préparation d'échantillons***

Un total de 10 colons (n = 5/groupes) de rats issus de l'expérimentation de carcinogénèse ont été prélevés puis congelés sous forme de *swiss-roll* sur de l'isopentane refroidi. Les tissus ont été coupés à 12 µm à l'aide d'un cryostat et montés sur lame en verre conductrice ITO 75\*25 mm (« *indium tin oxide* »). Selon le mode d'ionisation, deux matrices MALDI ont été déposées avec un sprayer automatique HTX TM-Sprayer (HTX Imaging, Chapel Hill, NC, USA). Pour le mode positif la matrice DHB (acide 2,5-dihydroxy benzoïque) à 15 mg/mL dans ACN/H<sub>2</sub>O (90:10) + 0,1% TFA a été déposée selon les conditions suivantes : température 60 °C, 14 passes, débit 125 µL/min, vitesse 1200 mm/min, track spacing 3 mm, pression d'azote 10 psi, hauteur de la buse 40 mm et gas flow 2 L/min. Pour le mode négatif la matrice 9-AA (9-aminoacridine) à 10 mg/mL dans EtOH/H<sub>2</sub>O (70:30, v/v) a été déposée selon les conditions suivantes : température 70 °C, 8 passes, débit 100 µL/min, vitesse 1200 mm/min, track spacing 3 mm, pression d'azote 10 psi, hauteur de la buse 40 mm, gas flow 3 L/min et temps de séchage entre chaque passe 10 s.

Avant analyse chaque lame est séchée au dessiccateur au minimum 1 h.

### ***B.12.2. Acquisition et traitement des données***

Les données ont été acquises sur un spectromètre de masse TIMS ToF Flex (Bruker, Bremen, Germany) équipé d'une double source ESI-MALDI avec un laser Smartbeam II 355nm. Chaque spectre MALDI pour chaque position est le résultat d'un scan acquis sur 200 tirs lasers consécutifs. Les images ont été acquises à une résolution spatiale de 100 µm.

Les images ont été retraitées à l'aide du logiciel SCiLS Lab Pro (Bruker). Les données ont été normalisées sous SCiLS par utilisation de la méthode « root min square ». Afin de différencier les différentes zones du tissu, la segmentation entre la partie épithéliale et la partie musculaire du colon a été réalisée *via* SCiLS et seule la partie épithéliale des colons a été étudiée.

Enfin, afin de déterminer les métabolites discriminants entre les deux conditions analysées, un test basé sur une courbe ROC (*receiving operator curve*) a été réalisé. Les métabolites présentant une AUC > 0.65 ou < 0.35 ont été considéré comme discriminants.

### ***B.12.3. Identification des métabolites***

L'identification préliminaire des métabolites a été réalisée en comparant les rapports *m/z* obtenus avec la base de données HMDB en utilisant une erreur de masse de ± 10 ppm. Pour conforter ces annotations putatives des spectres MS/MS ont été acquis en MALDI puis comparés à des spectres expérimentaux disponibles sous HMDB. Les spectres MS/MS ont été acquis en utilisant une fragmentation CID à des énergies comprises entre 10 et 35 eV et une fenêtre d'isolation de ± 2 Da.

### C. Résultats

Dans cette partie, l'effet de différents régimes alimentaires, riches en viande rouge ou blanche, enrichis en antioxydants ou non, ou encore contenant des acides gras  $\omega$ -3 ou  $\omega$ -6, sera étudié sur la base de plusieurs critères tels que la peroxydation lipidique, la cytotoxicité des eaux fécales ou encore le nombre de lésions préneoplasiques. Notre hypothèse principale reposant sur des effets néfastes induits par la viande rouge directement liés à la teneur de cet aliment en fer héminique, la teneur en hème dans les eaux fécales des animaux a été déterminée. Les résultats, présentés en Figure 30 A et B montrent une très nette différence entre les régimes riches en viande rouge en comparaison aux régimes à base de viande blanche. Pour l'expérimentation nutritionnelle (Figure 30B), la teneur en hème du régime sans viande type AIN-76 est similaire à celle des groupes poulet (moyenne : 17,18  $\mu$ M ; données non-montrées). Ainsi, une plus forte teneur en hème dans les fèces des animaux soumis à des régimes riches en viande rouge induit une plus forte exposition à l'hème au niveau colique et intestinal.

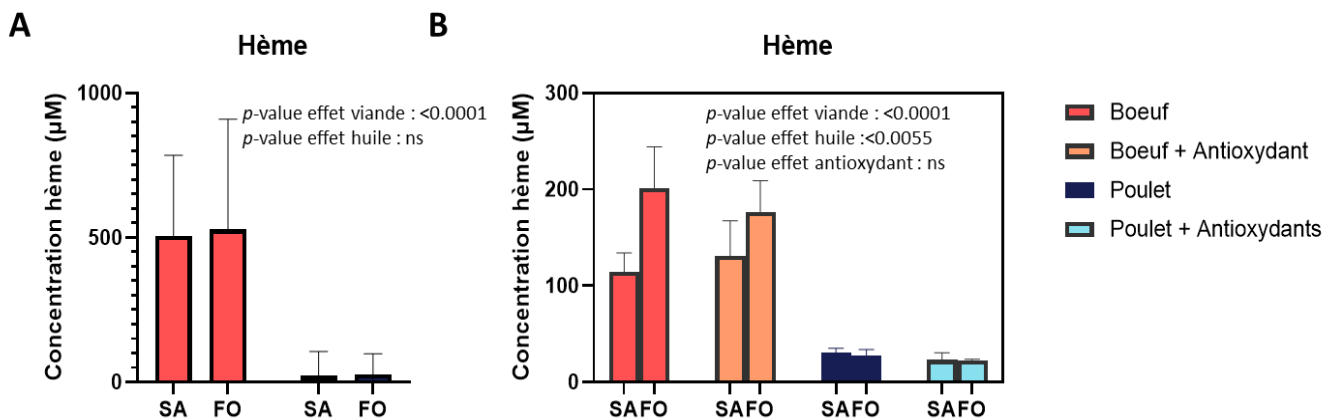


Figure 30 Concentration en hème dans les eaux fécales (A) dans l'expérimentation de carcinogénèse (B) dans l'expérimentation nutritionnelle. Les valeurs correspondent à des moyennes  $\pm$  écart-type. La significativité des différents facteurs a été déterminée par (A) ANOVA à deux voies et (B) ANOVA à 3 voies

**C.1. Les régimes riches en viande rouge induisent une augmentation de la peroxydation lipidique**

**C.1.1. Résultats de l'expérimentation nutritionnelle.**

**C.1.1.1. Profilage des acides mercapturiques comme témoins de peroxydation lipidique**

Dans cette partie, la méthode d'analyse non-ciblée des mercapturates présentée au Chapitre II : Développement et optimisation de la méthode d'analyse, a été utilisée afin de déterminer l'impact des différents régimes alimentaires sur la peroxydation lipidique au travers de la caractérisation de l'abondance relative et de la diversité des mercapturates. Suite à l'analyse des données acquises en LC-HRMS, 3470 variables ont été conservées. L'analyse DIA en LC-HRMS<sup>E</sup> sur les échantillons QC de groupe a permis de détecter 219 potentiels acides mercapturiques dont 26 communs aux 5 QC différents (Figure 31). De façon intéressante, la Figure 31 montre que les recouvrements les plus importants concernent les groupes « BOFO/PUFO » dont les régimes contiennent de l'huile de poisson et « BOSA/PUSA/CON » dont les régimes contiennent de l'huile de carthame. Ceci confirme que la nature des acides mercapturiques détectés est bien influencée par la nature des acides gras constituant le régime. Les 219 expériences MS/MS réalisées ont permis de confirmer 86 mercapturates (soit environ 40 % de vrais positifs). Parmi les 86 mercapturates validés, 47 ont été retrouvés dans le jeu de données issu du traitement XCMS, et seuls ces 47 composés seront considérés pour la suite des

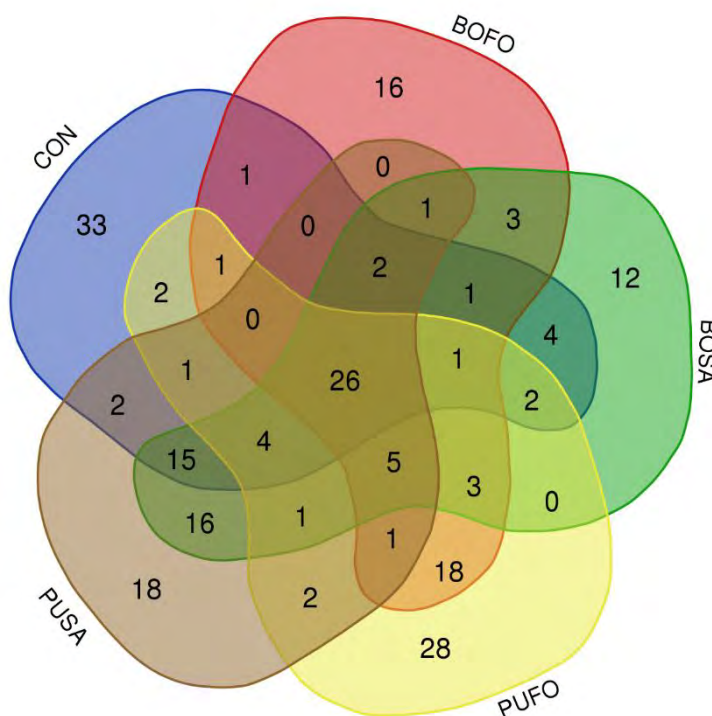


Figure 31 Nombre de mercapturates potentiels détectés dans chaque QC de groupe

résultats. Enfin, au sein de ces 47 composés, 22 annotations putatives ont pu être proposées et trois mercapturates ont été identifiés au niveau 1 en comparant le spectre MS/MS expérimental avec le spectre MS/MS du standard correspondant : le DHN-MA, le 4-hydroxy-nonanal mercapturic acid (HNE-MA) et le 4-hydroxy-heptanol mercapturic acid (HHEP-MA).

Afin de comparer l'impact du régime alimentaire sur l'abondance des acides mercapturiques, les moyennes des intensités de chaque composé ont été déterminées pour chaque groupe expérimental et comparées. Les comparaisons des groupes « BOSA » vs. « CON » et « BOSA » vs. « PUSA » présentées respectivement en Figure 32A et B témoignent d'une augmentation globale de l'abondance des acides mercapturiques dans les urines des rats soumis à un régime riche en viande rouge combiné à l'huile de carthame. Pour conclure quant à la significativité des différences observées, un test univarié (Wilcoxon + fdr) a été réalisé sur l'ensemble des 3470 variables (résultats reportés en Annexe 7 pour l'ensemble des comparaisons). De manière inattendue, une diminution significative de l'abondance des acides mercapturiques a été constatée en comparant les régimes BOFO et PUFO, à l'exception du DHN-MA (voir Annexe 8Annexe ).

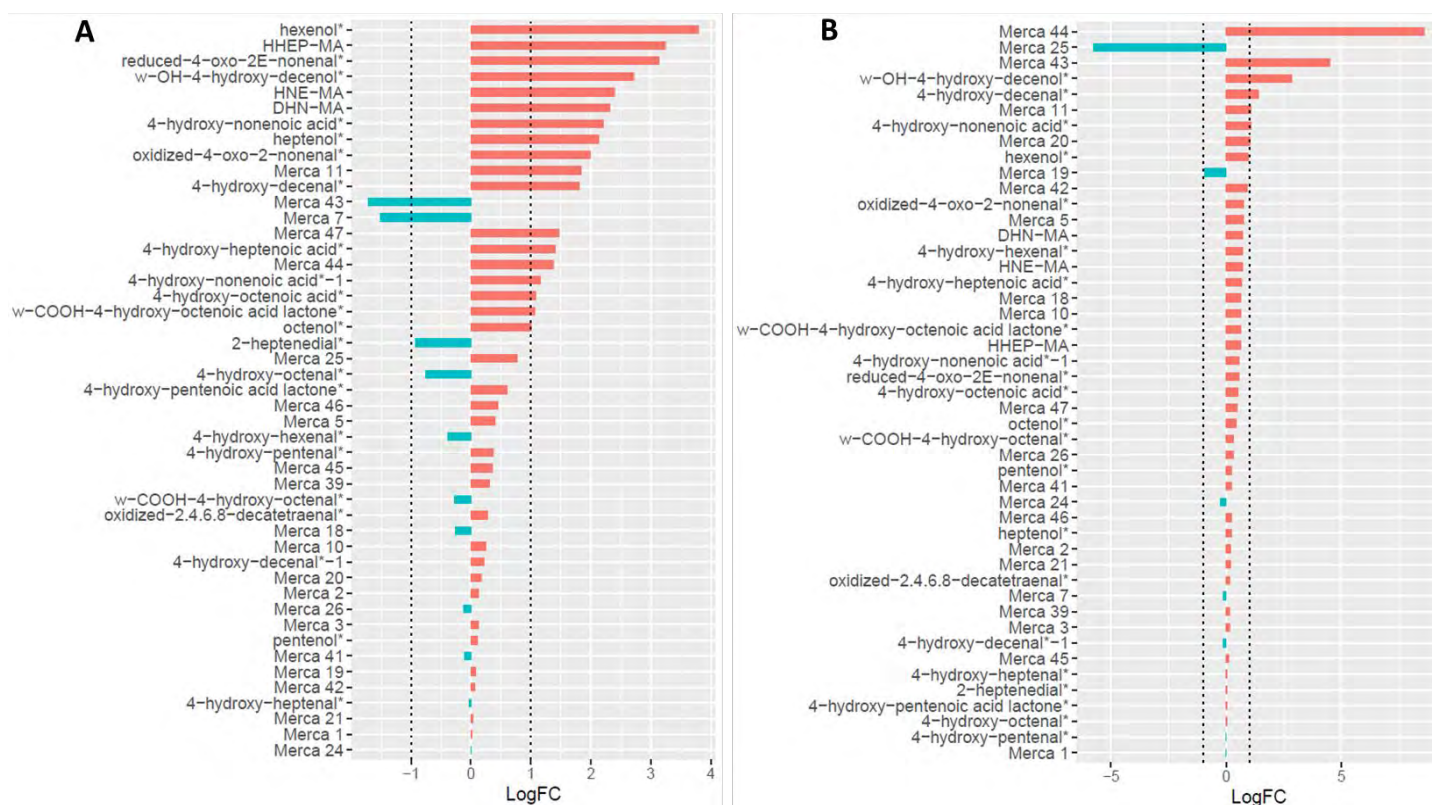


Figure 32 Fold change des 47 acides mercapturiques (A) entre les groupes BOSA et CON (B) entre les groupes BOSA et PUSA. Les valeurs sont exprimées en Log<sub>2</sub>(Fold change) une valeur positive correspondant à une augmentation dans le groupe BOSA par rapport au groupe comparé. Une « \* » signifie une annotation putative

L'ajout d'antioxydant dans les régimes alimentaires semble induire une diminution de l'abondance des acides mercapturiques dans le groupe BOSAR en comparaison au groupe BOSA (Figure 33), en particulier pour le DHN-MA et le HHEP-MA, suggérant une probable diminution de l'abondance de leur

précurseur, le HNE et le 4-OH-heptenal, respectivement. Il est à noter que les composés nommés Merca 45 à 47 semblent provenir des antioxydants eux-mêmes et non pas de la détoxification de produits de peroxydation lipidique. En effet, leur signal est abondant dans l'ensemble des groupes supplémentés en antioxydant (données non-montrées). En revanche, l'ajout d'antioxydant dans le régime BOFO n'induit pas de diminution de l'abondance de ces composés (voir Annexe 9).

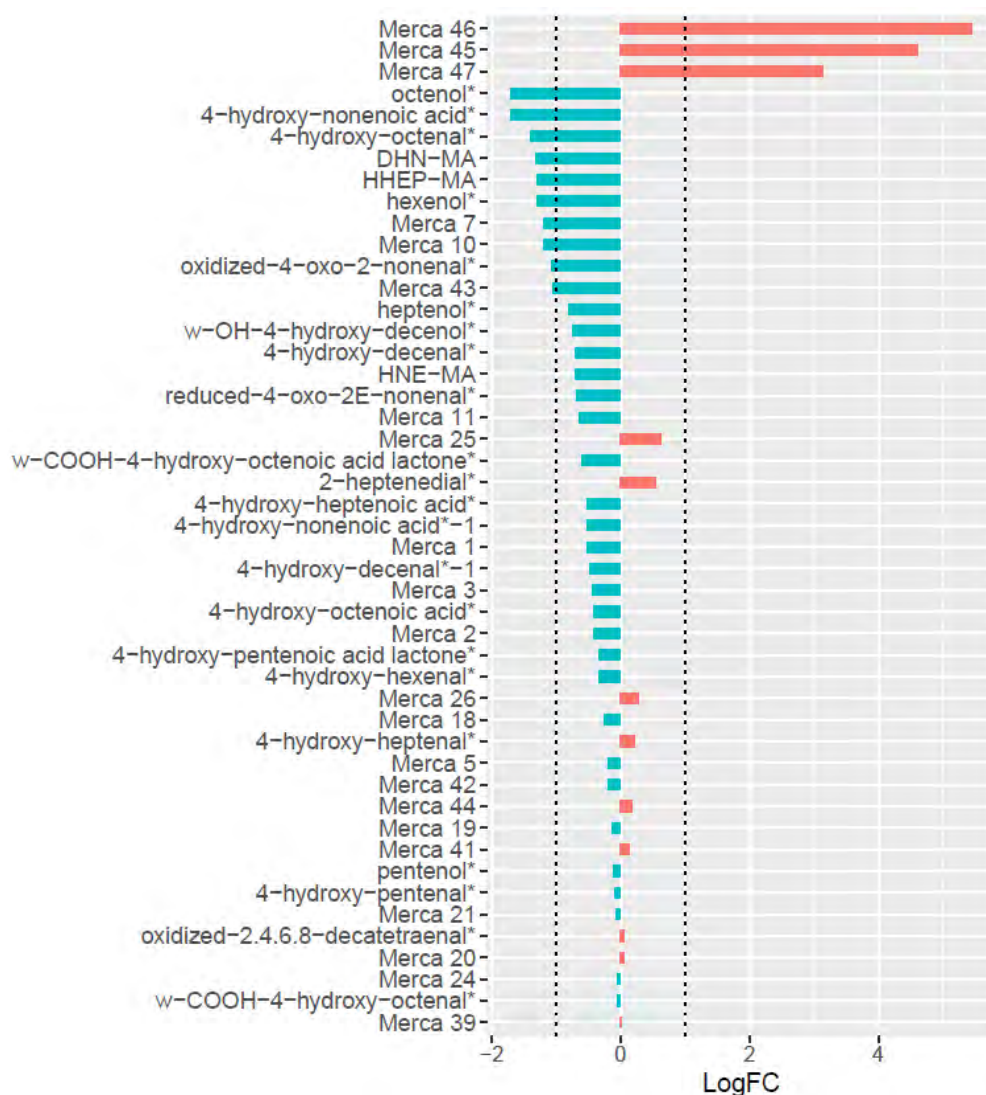


Figure 33 Fold change des 47 acides mercapturiques entre les groupes BOSAR et BOSA. Les valeurs sont exprimées en Log<sub>2</sub>(Fold change), une valeur négative correspondant à une diminution dans le groupe BOSAR par rapport au groupe BOSA. Une « \* » signifie une annotation putative

### Dosages complémentaires

Afin d'appuyer les données issues de l'analyse des acides mercapturiques et de renforcer l'hypothèse d'une peroxydation lipidique induite par les régimes riches en viande rouge, des dosages du DHN-MA urinaire ainsi que des TBARS fécaux ont été réalisés. Les résultats présentés en Figure 34A montrent une augmentation significative des TBARS fécaux dans les échantillons issus des régimes riches en viande rouge. L'ajout d'antioxydant dans le régime a également permis de réduire l'abondance de ces

espèces. En revanche, le facteur huile ne semble pas impacter l'abondance des TBARS fécaux. La concentration en TBARS fécaux a également été augmentée par le régime BOSA en comparaison au régime CON ( $p$ -value : 0.0241, données non-montrées). Les résultats obtenus par immuno-dosage du DHN-MA urinaire ont montré une augmentation significative de la quantité de DHN-MA urinaire dans les régimes riches en viande rouge en comparaison aux régimes poulet et sont en accord avec les données issues de l'analyse LC-MS (Figure 34B). Une réduction de la quantité de DHN-MA urinaire par l'ajout d'antioxydant a été observée en particulier pour les régimes combinant bœuf et huile de carthame. La nette différence entre les régimes SA et FO s'explique par le fait que le DHN-MA dérive du HNE, et donc de l'oxydation des acides gras  $\omega$ -6, constituants majoritaires de l'huile SA, tandis que l'huile FO est riche en acides gras  $\omega$ -3.

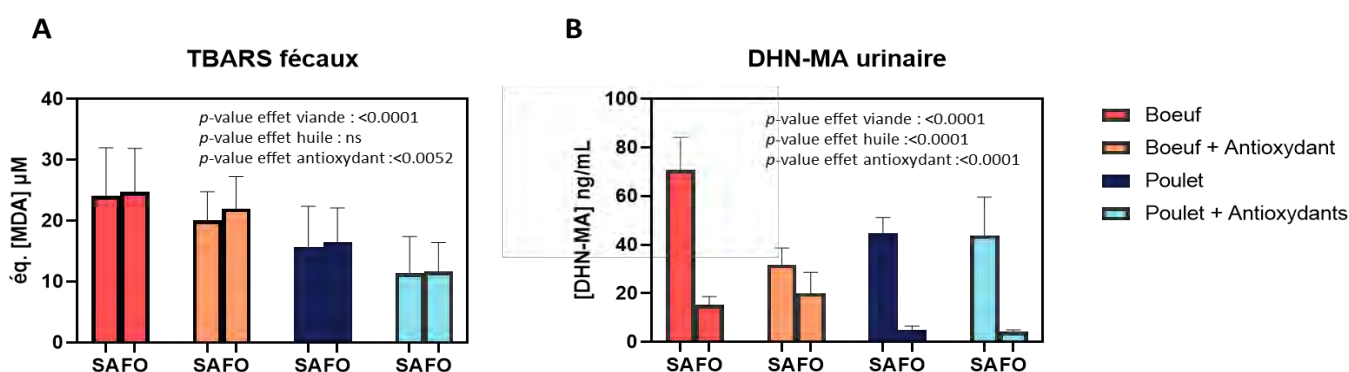


Figure 34 Evaluation de la peroxydation lipidique induite par les régimes alimentaires par dosage de (A) TBARS dans les eaux fécales (B) DHN-MA urinaire. La significativité est déterminée par ANOVA à 3 facteurs

#### C.1.1.2. Profilage aldéhydique des contenus fécaux

Dans le but d'estimer la diversité et l'abondance de l'ensemble des aldéhydes présents dans les échantillons fécaux une analyse qualifiée d'aldéhydique, dont le développement a fait l'objet d'une publication disponible en fin de manuscrit (Annexe 16), a été réalisée. Cette méthodologie permet, grâce à une dérivation sélective des aldéhydes à l'aide d'une sonde bromée, d'extraire les signaux correspondant aux aldéhydes dans un jeu de données de métabolomique non-ciblée. A l'issue de l'analyse LC-MS et du traitement des données à l'aide d'XCMS, 3332 variables ont été conservées dont 1531 correspondant à des ions contenant du brome. Afin d'éviter la redondance d'information, seuls les ions correspondant à l'isotopomère 79 Br ont été conservés. Au total, 716 variables correspondant à des ions contenant au moins un atome de brome ont été conservées pour la suite des analyses. L'annotation de ces variables a permis l'identification de 24 composés dont 9 annotés au niveau 1 (voir Annexe 10). Parmi ces variables, des aldéhydes issus de la peroxydation lipidique tels que le MDA ont pu être annotés mais aussi d'autres types d'aldéhydes (alkanals, aldoses...) ou encore certains acides carboxyliques. L'ACP présentée en Figure 35B réalisée à partir des 716 variables montre une bonne discrimination des différents groupes en comparaison à l'ACP réalisée



sur l'ensemble des 3332 variables (Figure 35A) où les effets « huile » et « antioxydant » semblent prédominer. Il est notable que bien que les analyses statistiques (OPLS-DA et test de Wilcoxon) aient révélé des différences significatives entre les différents groupes, les aldéhydes issus de la peroxydation lipidique semblent peu affectés par les différents régimes, à l'exception du MDA et du 8-carboxy-4.5-dihydroxy-octenal (voir Annexe 10). De plus, parmi les aldéhydes annotés, seuls ces deux composés

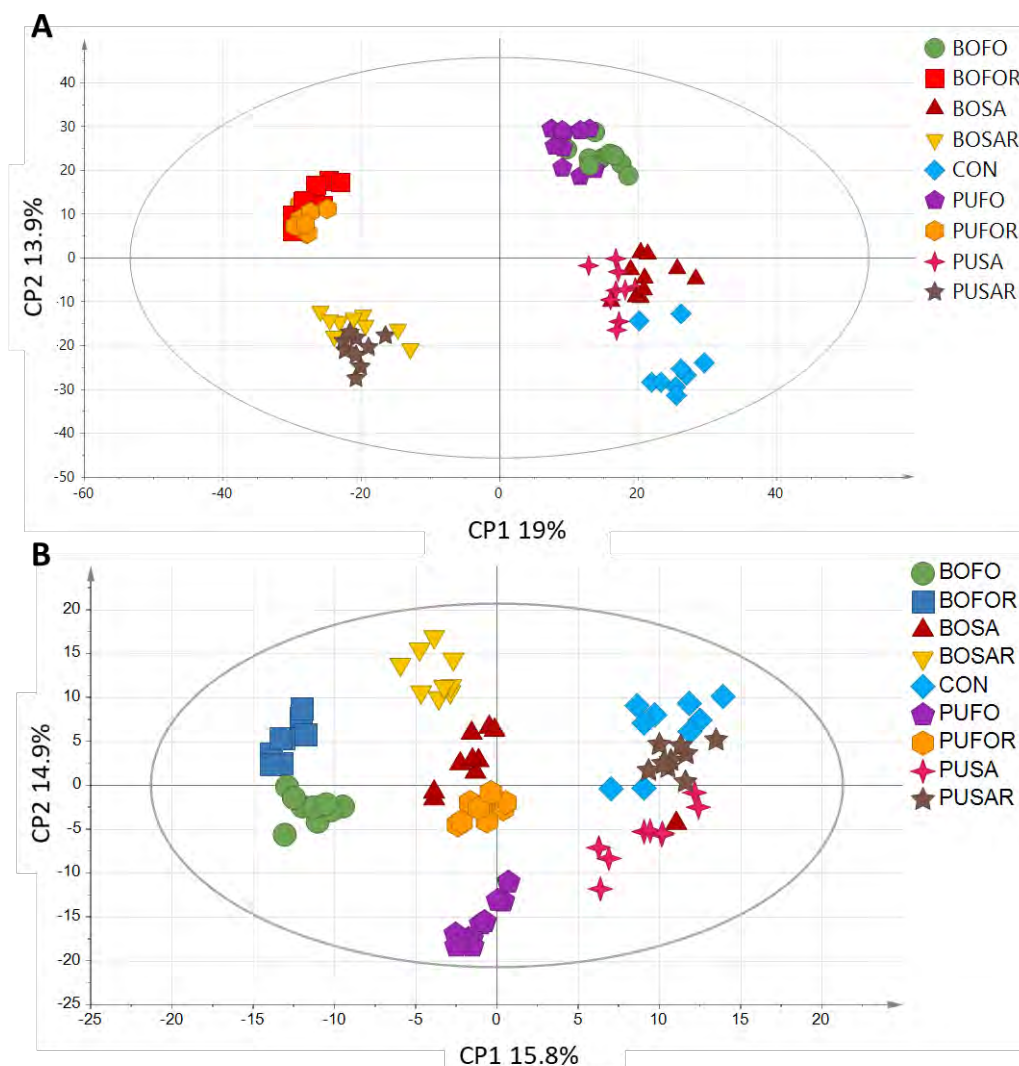


Figure 35 Analyse en composante principale à partir de (A) l'ensemble des variables détectées *via* XMCS (R<sup>2</sup>=32,9%) (B) l'ensemble des variables correspondant aux aldéhydes (R<sup>2</sup>=30,7%). Chaque point représente un individu

présentent une corrélation importante avec les mercapturates HHEP-MA, HNE-MA et DHN-MA (par exemple  $r = 0.69$  entre le HHEP-MA et le 8-carboxy-4.5-dihydroxy-octenal).

### C.1.2. Résultats de l'expérimentation de cancérogène

#### C.1.2.1. Profilage des acides mercapturiques

Dans cette expérimentation, 4717 variables ont été conservées à l'issue de l'analyse des données acquises en LC-HRMS. Après traitement des données LC-HRMS<sup>E</sup>, 583 mercapturates potentiels ont été retenus dont 142 ont pu être retrouvés parmi les 4717 variables. Les expériences de MS/MS ciblée ont permis de confirmer la présence de 71 mercapturates qui seront considérés pour la suite. Au sein de ces composés, 39 ont pu être putativement annotés et deux mercapturates ont été annotés au niveau 1 : le DHN-MA et le HNE-MA (le HHEP-MA n'ayant pas été retrouvé dans cette expérimentation). De la même manière que dans l'expérimentation nutritionnelle, une augmentation significative (test Wilcoxon + fdr) de nombreux mercapturates est observée en comparant les régimes BOSA et PUSA ainsi qu'une diminution de l'abondance de ces produits lors de la supplémentation en antioxydants (Figure 36A et B).

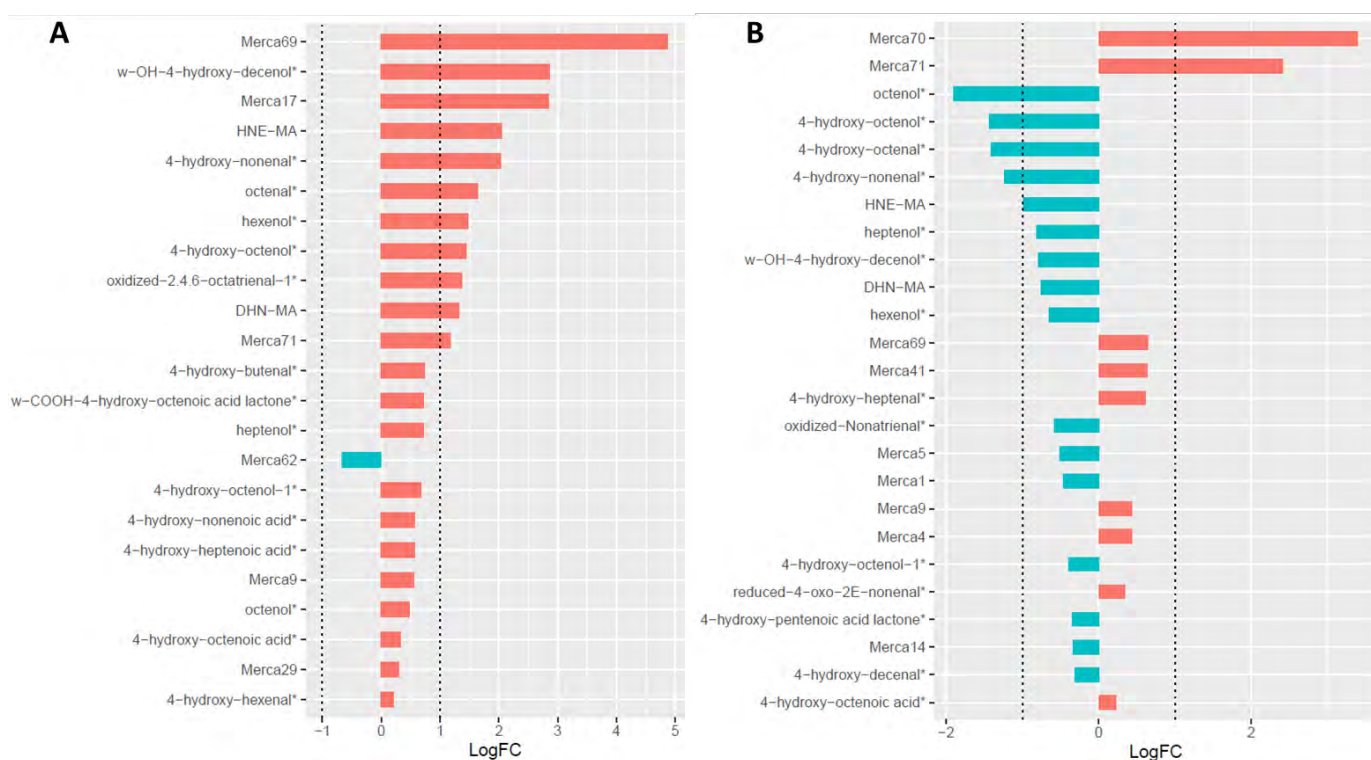


Figure 36 Fold change des mercapturates retrouvés significatifs (Wilcoxon + fdr) lors des comparaison (A) BOSA/PUSA (B) BOSAR/BOSA. Les valeurs sont exprimées en Log<sub>2</sub>(Fold change). Une « \* » signifie une annotation putative

Les données pour l'ensemble des mercapturates et des comparaisons sont disponibles en Annexe 11. Il est notable que l'augmentation de l'abondance relative des mercapturates entre les groupes BOSA et PUSA semble être encore accentuée par la durée de l'expérimentation, en effet les *fold-change* du DHN-MA et HNE-MA passent respectivement de 1.64 et 1.61 dans l'expérimentation nutritionnelle à 2.5 et 4.1 dans l'expérimentation de carcinogénèse. De plus, contrairement à ce qui a été observé dans l'étude nutritionnelle, l'augmentation de l'abondance relative entre les groupes BOFO et PUFO de

certaines mercapturates tels que le DHN-MA ou le HNE-MA (*fold-change* respectifs de 3.6 et 11) suggère une augmentation de la peroxydation lipidique dans le régime bœuf.

#### C.1.2.2. Dosages complémentaires

Conformément aux résultats obtenus lors de l'expérimentation nutritionnelle, une augmentation des TBARS et du DHN-MA urinaire a été observée dans cette expérimentation. Les résultats présentés en Figure 37A et B montrent une augmentation significative des TBARS et DHN-MA avec les régimes riches en viande rouge en comparaison aux régimes à base de viande blanche. Cependant, contrairement à ce qui a été observé lors de l'expérimentation nutritionnelle, l'ajout d'antioxydant n'a pas conduit à une diminution de la quantité de TBARS dans le groupe BOSAR en comparaison au groupe BOSA (t-test,  $p$ -value : 0.98 ; données non-montrées). En revanche, et conformément aux données obtenues par LC-MS, l'ajout d'antioxydant a conduit à une diminution significative du DHN-MA urinaire dans le groupe BOSAR en comparaison au groupe BOSA (t-test,  $p$ -value : <0.0001 ; données non-montrées).

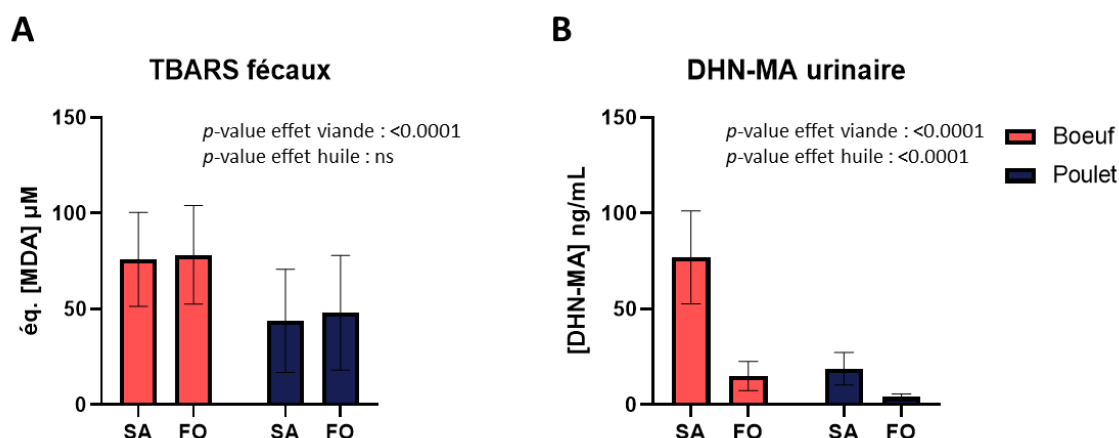


Figure 37 Dosage de (A) TBARS dans les eaux fécales (B) DHN-MA urinaire. La significativité est déterminée à l'aide d'une ANOVA à deux facteurs.

## C.2. Les régimes riches en viande rouge induisent une activation transcriptionnelle de gènes de détoxification

Afin d'évaluer l'impact des différents régimes et de la peroxydation lipidique hème-induite sur la transcription, au sein de l'épithélium colique, de gènes impliqués dans la détoxification des aldéhydes issus de la peroxydation lipidique dans le catabolisme de l'hème ou dans l'induction de défense antioxydantes, des expériences de qPCR ont été réalisées. Les résultats présentés en Figure 38

montrent une réponse différentielle en fonction de la durée de l'expérimentation, bien qu'un effet lié à l'injection d'AOM ne soit pas à exclure.

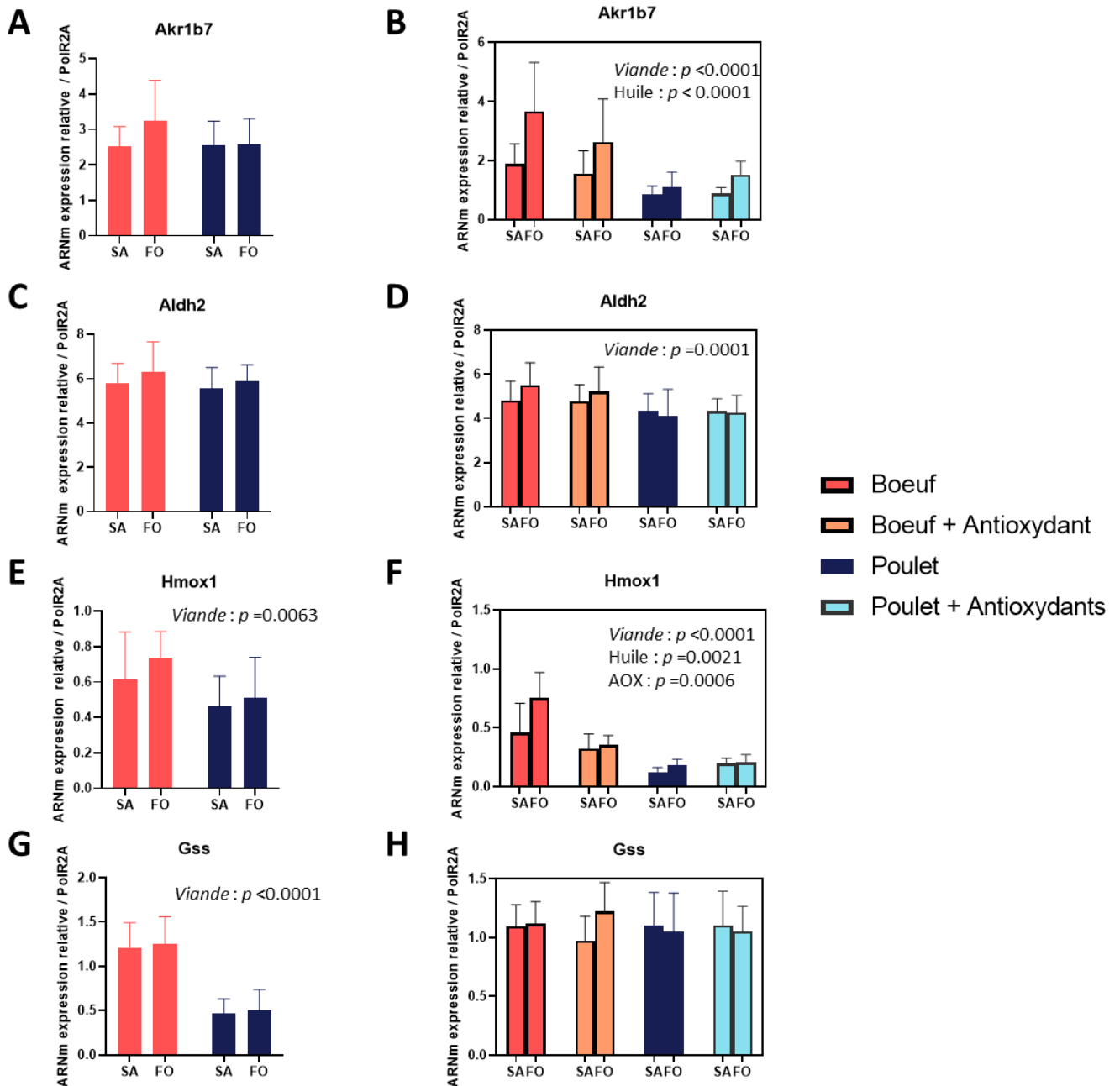


Figure 38 Expression relative des gènes impliqués dans la détoxification des aldéhydes (A et B : Aldo keto reductase family 1 member ; B et D : Aldehyde Dehydrogenase 1 Family Member A2) du catabolisme de l'hème (E et F : Heme oxygenase 1) et dans la synthèse de glutathion (G et H : Glutathione synthetase). Les valeurs correspondent à des moyennes  $\pm$  écart-type. La significativité des différents facteurs a été déterminée par ANOVA à deux voies pour l'expérimentation de carcinogénèse et ANOVA à 3 voies pour l'expérimentation nutritionnelle. Seules les différences significatives sont indiquées

En effet, pour l'expérimentation nutritionnelle, une activation de la transcription des gènes de détoxification des aldéhydes a été observée (Figure 38B et D) lors d'un régime riche en viande rouge tandis qu'aucune différence significative n'a été observée lors de l'expérimentation de carcinogénèse (Figure 38A et C). Il est notable que le statut lipidique du régime semble également avoir un impact sur

la transcription de ces gènes, notamment sur Akr1b7 pour lequel une augmentation de la transcription dans les régimes à base d'huile de poisson a été observée (Figure 38B). De manière attendue, les régimes riches en viande rouge ont induit une activation de la transcription du gène « Heme oxygénase 1 » responsable du catabolisme de l'hème dans les deux expérimentations (Figure 38E et F). Il est intéressant de noter que l'ajout d'antioxydant dans le régime a conduit à une diminution de la transcription de ce gène en particulier pour les régimes bœuf dans l'expérimentation nutritionnelle (Figure 38F) mais qu'aucun changement n'a été observé après 100 jours de régime. Enfin, le glutathion représentant l'élément clé de la voie des mercapturates et une augmentation de la quantité des acides mercapturiques ayant été reportée dans les urines des rats soumis aux régimes riches en viande rouge, il a paru pertinent de s'intéresser aux gènes impliqués dans la synthèse de ce composé. Pour cela, la transcription du gène « Glutathione synthetase » (Gss) a été étudiée. De la même manière que pour les autres gènes, un résultat différent en fonction de la durée de l'expérimentation a été observé. Après 100 jours de régime, une activation de la transcription de Gss a été observée tandis qu'aucune différence n'a été notée dans l'expérimentation nutritionnelle (Figure 38G et H).

L'ensemble de ces résultats met en lumière à la fois l'impact du régime alimentaire sur la transcription des gènes impliqués dans le métabolisme rédox mais également l'adaptation de la réponse en fonction de la durée du régime.

### C.3. Impact sur le métabolome urinaire

La méthode d'analyse développée et présentée au Chapitre II : Développement et optimisation de la méthode d'analyse, de ce manuscrit permettant à la fois l'exploration non-ciblée de l'ensemble des acides mercapturiques mais aussi du métabolome urinaire, il a paru pertinent de s'intéresser aux répercussions des différents régimes sur les métabolites urinaires. A l'issue des analyses, respectivement 40 et 59 annotations au niveau 1 ont pu être confirmées dans l'expérimentation nutritionnelle et dans l'expérimentation de carcinogénèse. Les métabolites identifiés ainsi que leur significativité (métabolites présentant un  $VIP > 1$  et une  $p$ -value  $< 0.05$ ) dans les principales comparaisons sont disponibles en Annexe 12 et Annexe 13. Pour chaque expérimentation, les ACP réalisées à partir de l'ensemble des variables montrent une bonne discrimination des différents groupes avec un effet marqué du facteur « huile » sur la première composante, tandis que l'ajout d'antioxydant dans le régime semble être responsable de la séparation observée sur la deuxième composante en particulier pour l'expérimentation nutritionnelle (Figure 39A et B). Il est notable que la discrimination entre les régimes bœuf et poulet semble plus marquée après 100 jours de régime (Figure 39B). Pour compléter ces observations, des analyses supervisées OPLS-DA ont été réalisées et

des modèles robustes et valides ( $Q^2 > 0.4$  avec un intercept  $< 0.05$ ) ont été obtenus pour toutes les comparaisons que ce soit deux à deux ou par facteur.

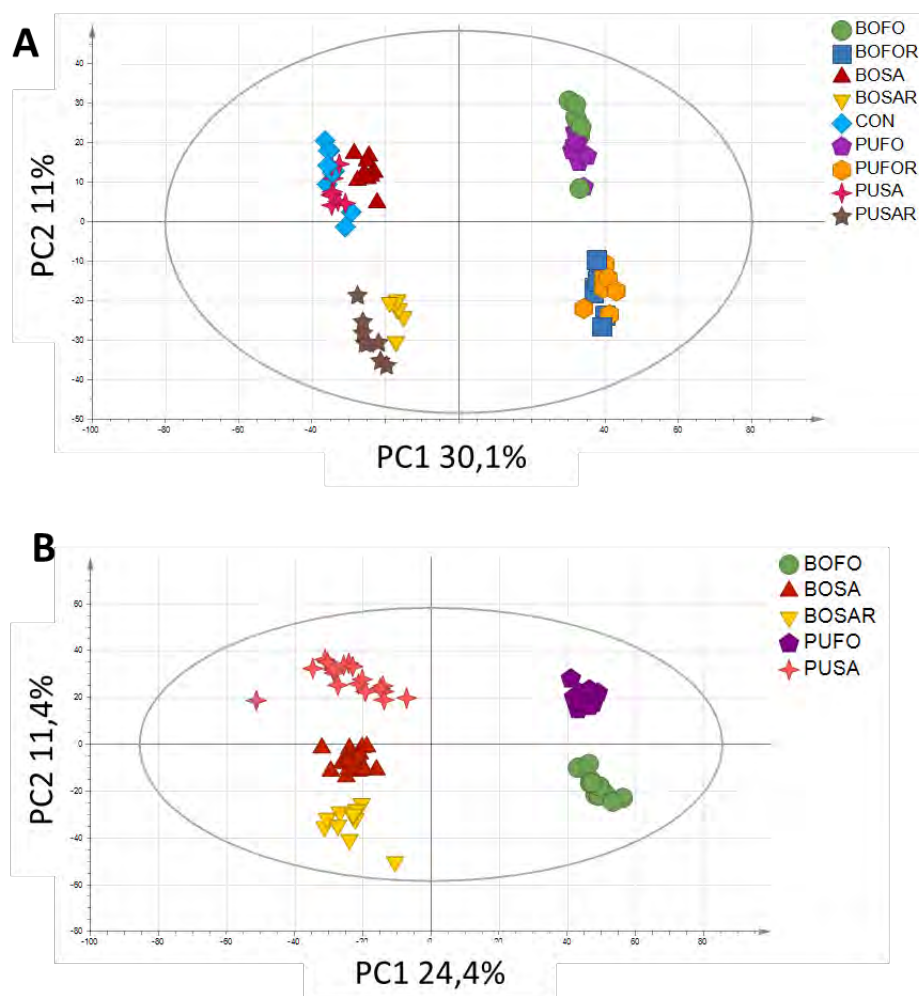


Figure 39 ACP réalisée sur l'ensemble des variables détectées dans les urines des rats issues de (A) l'expérimentation nutritionnelle (B) l'expérimentation de carcinogénèse. Chaque point représente un individu.

Afin d'évaluer le potentiel lien entre les modifications du métabolome urinaire observées et la peroxydation lipidique, les corrélations entre les mercapturates annotés au niveau 1 et les métabolites discriminants pour l'effet viande (comparaison de tous les groupes bœuf vs poulet) ont été étudiées. Il est intéressant de constater qu'après une longue période de régime de fortes corrélations (positives ou négatives) sont observées entre les niveaux des métabolites et les mercapturates, tandis que ces corrélations sont plus faibles dans l'expérimentation nutritionnelle (Figure 40A et B). En particulier, plusieurs métabolites présentent des corrélations négatives importantes avec les mercapturates : le *N*-acetyl glutamic acid, la *N*-acetyl-leucine, le pyroglutamic acid, l' $\alpha$ -ketoglutarate et le pantothenic acid. Cependant, malgré une corrélation élevée avec les acides mercapturiques, il est important de souligner que la diminution ou l'augmentation des niveaux de certains métabolites, en particulier la 3-

methyl-histidine, n'est probablement pas liée à la peroxydation lipidique. En effet, ce métabolite est largement décrit dans la littérature comme biomarqueur de consommation de volaille (Giesbertz *et al.*, 2020).

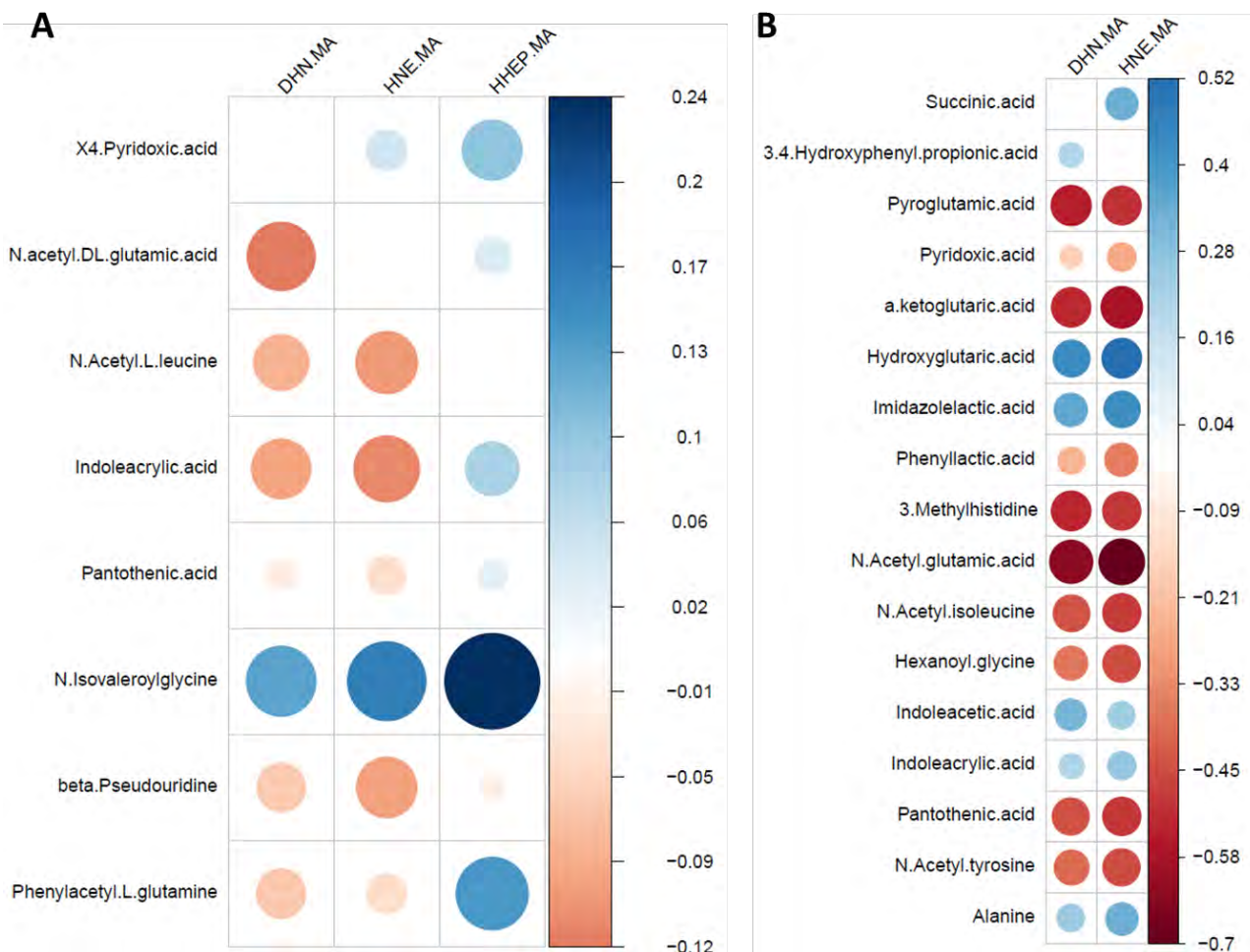


Figure 40 Corrélation entre les mercapturates annotés au niveau 1 et les métabolites discriminants entre les groupes bœuf et poulet (A) expérimentation nutritionnelle (B) expérimentation de carcinogénèse

#### C.4. Impact d'un régime riche en viande rouge/ fer héminique sur l'épithélium colique

Afin d'évaluer le potentiel impact des régimes riches en viande rouge sur le métabolisme de l'épithélium colique, une étude préliminaire sur 9 rats issus de l'expérience de carcinogénèse (5 rats du groupe BOSA et 4 du groupe PUSA) a été réalisée en imagerie par spectrométrie de masse. Le but et l'intérêt de l'utilisation de cette technique repose sur le fait que seule la partie épithéliale colique (correspondant aux cryptes du colon) a été analysée. En effet, étant donné que cette partie du colon est en contact direct avec les constituants présents dans la lumière colique, le métabolisme du tissu épithélial pourrait être impacté par le régime alimentaire. A l'issue de l'analyse, 24 métabolites discriminants putativement annotés (présentant une AUC > 0.65 ou < 0.35) ont été mis en évidence

(voir Annexe 14). Parmi ces métabolites, l'abondance relative du glutathion dans la partie épithéliale du colon semble être impactée par le type de régime. En effet, une différence statistiquement significative (AUC = 0.3) a été observée lors de la comparaison du signal de ce composé dans les colons du groupe « BOSA » en comparaison au groupe « PUSA » (Figure 41). La valeur d'AUC déterminée pour cette comparaison indique une plus faible intensité du glutathion au niveau de l'épithélium colique du groupe « BOSA ». Même si les données spectrales obtenues sont précises, il est cependant à noter que l'annotation du glutathion et des autres métabolites discriminants reste putative et doit être confirmée par comparaison avec un standard.

Ces résultats mettent en avant de possibles changements métaboliques induits par la consommation de viande rouge au niveau de l'épithélium colique. Dans cette expérience, l'utilisation de l'imagerie par spectrométrie de masse a permis d'étudier la réponse métabolique du seul tissu épithélial, en contact direct avec la lumière colique, et par conséquent avec les composés formés au sein de cet environnement. La diminution de l'abondance du glutathion au niveau colique pourrait être liée à une surconsommation de ce métabolite afin de détoxifier les produits de peroxydation lipidiques hème-induit formés au sein de l'environnement colique.

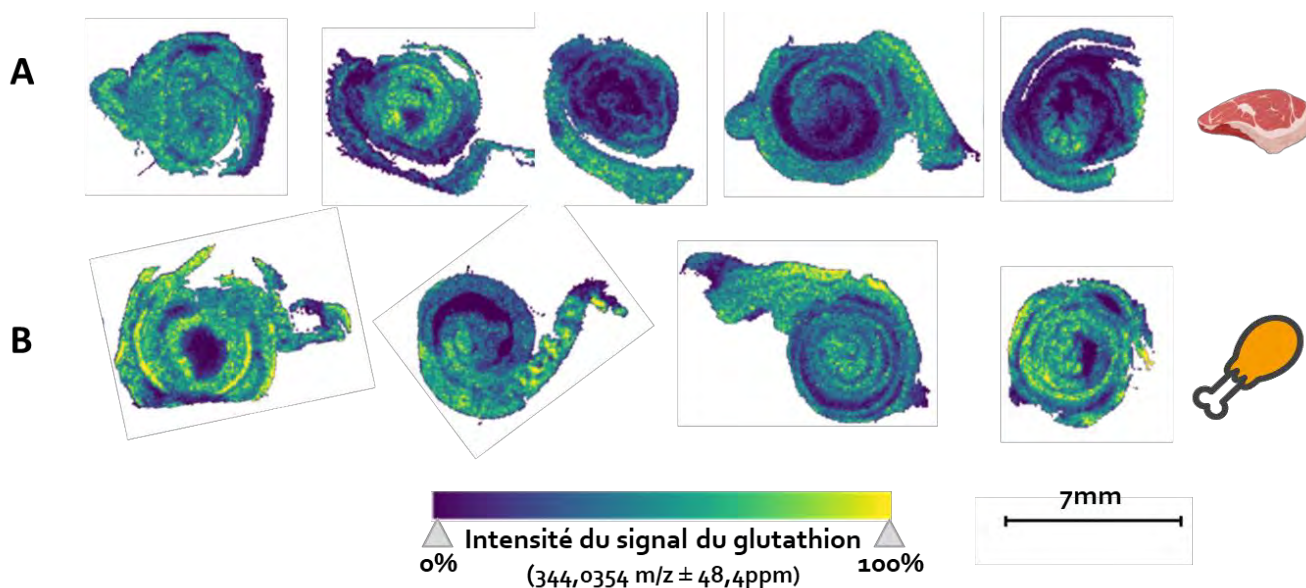


Figure 41 Intensité du signal du glutathion [M+K-2H]- (A) dans les colons du groupe BOSA (B) dans les colons du groupe PUSA)

### C.5. Réponse toxicologique associée

L'ensemble des résultats précédents ont permis de mettre en évidence une peroxydation lipidique augmentée dans les régimes riches en viande rouge en comparaison aux régimes à base de viande blanche ou sans viande. Afin de déterminer si l'accroissement du stress électrophile et de la



peroxydation lipidique ont conduit à des effets toxiques sur l'organisme, plusieurs facteurs tels que la cytotoxicité des eaux fécales, des marqueurs de stress oxydant ou encore le dénombrement de lésions pré-cancéreuses dans les colons ont été étudiés et sont présentés dans cette partie.

### C.5.1. Cytotoxicité des eaux fécales

Dans un premier temps, l'activité cytotoxique des eaux fécales provenant des rats soumis aux différents régimes expérimentaux a été déterminée. Pour cela des cellules épithéliales coliques murines immortalisées ont été exposées aux eaux fécales issues de chaque individu durant 24 h. Le pourcentage de cellules viables au bout de ce temps a ensuite été déterminé. Les résultats présentés en Figure 42A et B montrent qu'aucune différence de cytotoxicité n'a été observée entre les régimes riches en viande rouge et les régimes riches en viande blanche. Dans l'expérimentation nutritionnelle, un effet huile en faveur d'une meilleure viabilité cellulaire avec les régimes contenant de l'huile de poisson a été reportée, bien que cet effet n'ait pas été retrouvé dans l'expérimentation de carcinogénèse. Cependant, il est intéressant de noter que les régimes à base de viande (bœuf ou poulet) semblent induire une diminution de la viabilité cellulaire en comparaison au régime contrôle sans viande type AIN-76 (Figure 42C). Par ailleurs, aucune amélioration de la viabilité cellulaire n'a été observée dans l'expérimentation nutritionnelle lors de l'ajout d'antioxydant dans le régime. De la même manière, dans l'expérimentation de carcinogénèse, une tendance à la diminution de la viabilité a été observée lors de l'ajout d'antioxydant lorsque les groupes BOSA et BOSAR ont été comparés (t-test,  $p$ -value 0.064 ; données non-montrées).

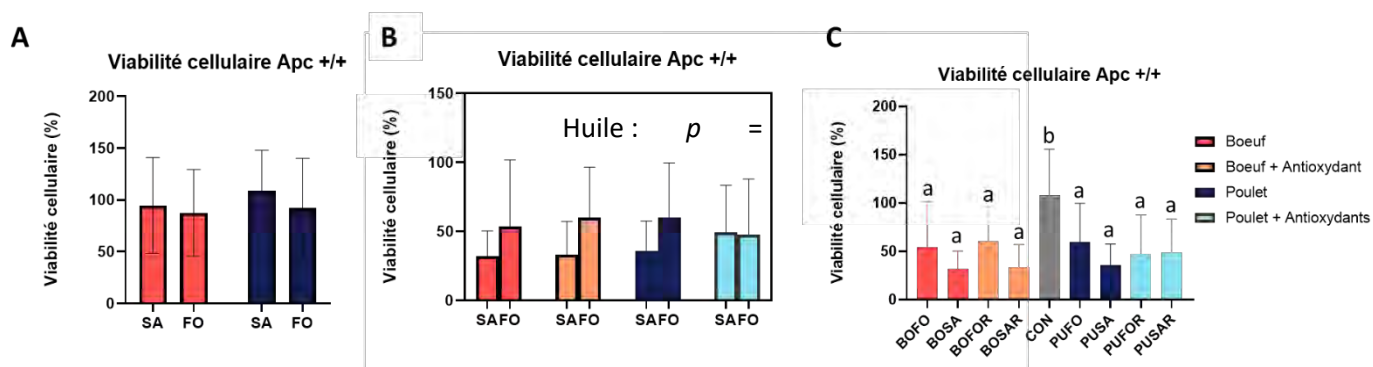


Figure 42 Viabilité cellulaire des cellules Apc +/+ exposées aux eaux fécales (A) issues de l'expérimentation de carcinogénèse (B et C) issues de l'expérimentation nutritionnelle. Des lettres identiques signifient aucune différence significative entre les groupes

### C.5.2. Dosage de la base oxydée 8-OH-2-deoxyguanosine

Afin d'estimer l'impact des régimes riches en viande rouge sur les dommages à l'ADN pouvant être induits par le stress oxydant un dosage urinaire de la base oxydée de l'ADN (8-OH-2-deoxyguanosine) a été réalisé. Ce marqueur, témoin de l'attaque de l'ADN par des radicaux libres, est aujourd'hui largement utilisé pour évaluer le stress oxydatif et évaluer le risque de cancer suite à l'exposition à divers composés (Valavanidis *et al.*, 2009). Les résultats présentés en Figure 43 montrent qu'aucune différence significative en fonction du type de régime n'a été observée concernant ce marqueur dans aucune des deux expérimentations. Aucune différence significative n'a également été notée en comparaison au groupe « CON » dans l'expérimentation nutritionnelle.

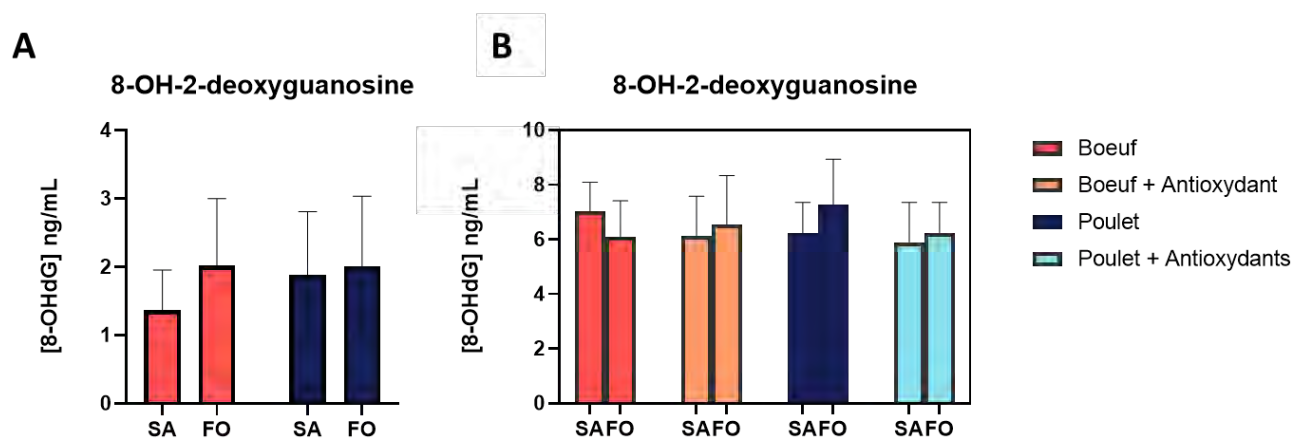


Figure 43 Dosage urinaire de 8-OH-2-deoxyguanosine (A) dans l'expérimentation de carcinogénèse (B) dans l'expérimentation nutritionnelle

### C.5.3. Détermination du taux de phosphorylation de l'histone H2AX

Afin d'obtenir des données complémentaires au dosage systémique de la 8-OH-dG, et de mieux évaluer les effets des différents régimes, au niveau local, sur d'éventuels dommages à l'ADN, le taux de phosphorylation de l'histone H2AX dans le tissu épithélial colique des animaux de l'expérimentation nutritionnelle a été déterminé par immunofluorescence. La forme phosphorylée de l'histone H2AX sur le résidu sérine 139 (forme appelée  $\gamma$ -H2AX) est un marqueur largement utilisé pour déterminer les dommages à l'ADN (Sharma *et al.*, 2012). En effet, lors d'une cassure double-brin la phosphorylation de H2AX constitue le premier évènement du recrutement des protéines de réparation à l'ADN. Ainsi quantifier la phosphorylation de H2AX permet de traduire la réponse à une lésion génotoxique. Les résultats présentés en Figure 44 montrent des effets contraires à ceux attendus avec un pourcentage de cellules positives à  $\gamma$ -H2AX plus important dans les régimes riches en viande blanche en comparaison aux régimes riches en viande rouge. Cette observation pourrait être liée à l'effet inducteur de l'hème sur Nrf2, dont l'activation conduirait à une induction des systèmes de réparation de l'ADN (Jayakumar *et al.*, 2015). Il est également notable que les régimes contenant de l'huile de poisson ont conduit à un taux de phosphorylation de H2AX plus important, en particulier pour le

groupe BOFO. Comme pour les autres marqueurs de toxicité étudiés, l'ajout d'antioxydant dans le régime n'a pas eu d'effet significatif. Il est également intéressant de constater que le régime sans viande ne présentait pas de différences significatives en comparaison à l'ensemble des autres régimes (moyenne groupe « CON » : 3,34 %).

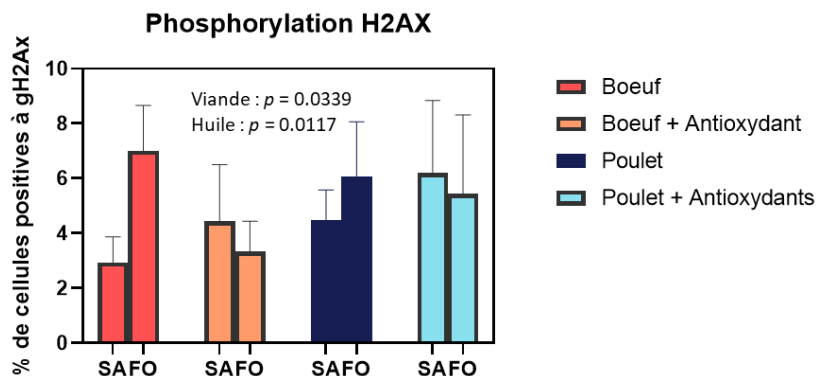


Figure 44 Taux de phosphorylation de l'histone H2AX en fonction du type de régime dans l'expérimentation nutritionnelle.

#### C.5.4. Dénombrement des lésions prénéoplasiques

L'impact de la consommation de viande rouge sur la promotion de la carcinogénèse colorectale a été déterminé par dénombrement des lésions prénéoplasiques dans les colons des rats chimio-induits de l'expérimentation de carcinogénèse. Deux types de lésions ont été quantifiés, les ACF et les MDF. Pour chaque type de lésion le nombre de lésions ainsi que leur multiplicité (nombre de cryptes altérés dans un ACF ou un MDF donné) ont été déterminés comme illustré sur la Figure 45A et B.

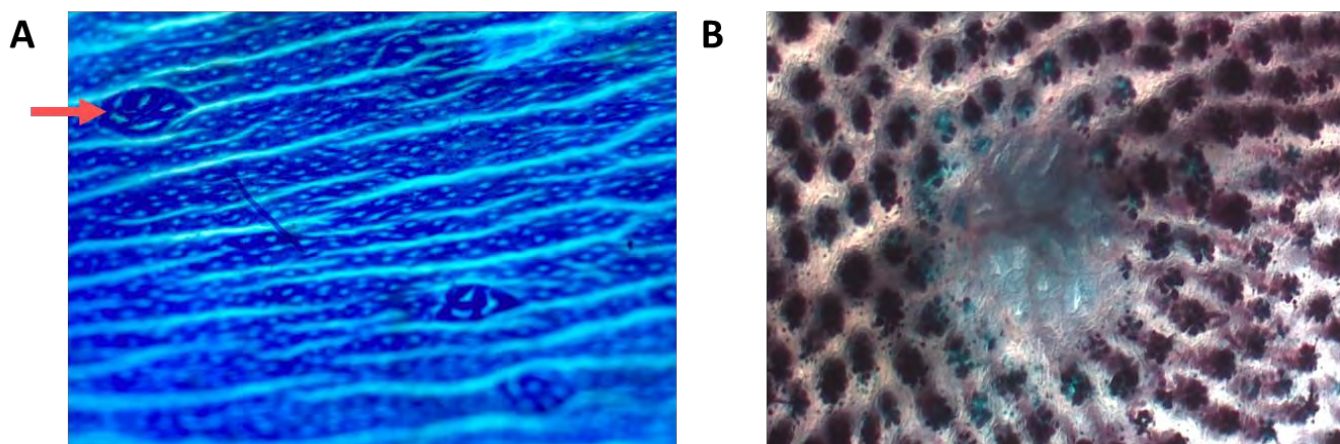


Figure 45 Lésions prénéoplasiques observées au microscope optique (A) ACF présentant une multiplicité de 5 cryptes (B) MDF présentant une multiplicité de 13 cryptes.

Les résultats issus du dénombrement des lésions ont montré une augmentation significative du nombre d'ACF dans les colons des rats soumis aux régimes riches en viande rouge en comparaison aux régimes riches en viande blanche (Figure 46A). Il est intéressant de noter que l'augmentation du nombre d'ACF est indépendante du facteur « huile » et apparaît être régionalisée avec un effet significatif seulement au niveau du colon proximal alors qu'aucune différence n'a été constatée au

niveau du colon distal (Figure 46A et B). Concernant la multiplicité, aucune différence significative n'a été notée entre les groupes bœuf et poulet (données non-montrées). En revanche, une augmentation significative du nombre d'ACF des colons des rats du groupe « BOSAR » a été constatée (Figure 46C) par rapport au groupe BOSA. De la même manière, une différence significative en matière de multiplicité a également été observée avec un nombre d'ACF > 4 cryptes plus élevé dans les colons des rats BOSAR en comparaison aux rats BOSA (t-test,  $p$ -value = 0.031 ; données non-montrées).

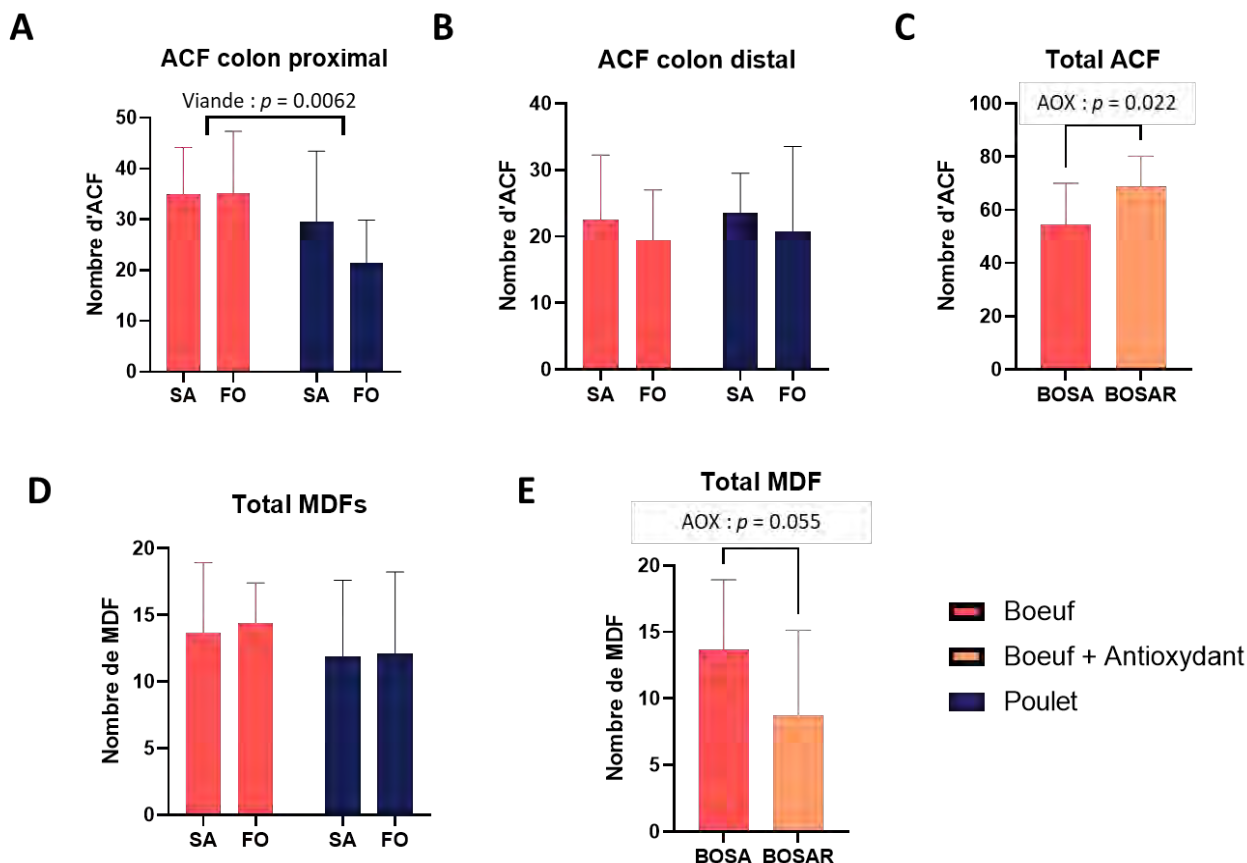


Figure 46 Dénombrement des lésions préneoplasiques dans les colons des rats chimio induits (A) Nombre d'ACF au niveau du colon proximal (B) Nombre d'ACF au niveau du colon distal (C) Comparaison du nombre total d'ACF entre les groupes BOSAR et BOSA (D) Nombre total de MDF (E) Comparaison du nombre total de MDF entre les groupes BOSAR et BOSA. La significativité des comparaisons est déterminée par ANOVA à 2 facteurs pour les comparaisons A, B et D et par t-test pour les comparaisons C et E

Concernant les MDF, aucune différence significative n'a été observée entre les régimes riches en viande rouge et les régimes riches en viande blanche (Figure 46E). Cependant, pour ce marqueur, contrairement à ce qui a été observé pour les ACF, l'ajout d'antioxydant dans le régime semble présenter un effet protecteur quant à l'apparition de ce type de lésions dans le colon ( $p$ -value = 0.055). Cette absence de différence significative au niveau des MDF entre les groupes bœuf et poulet est renforcé par un résultat obtenu dans l'expérimentation nutritionnelle où le nombre de cellules sécrétrices de mucus (identifiées à l'aide d'une coloration au bleu alcian) ont été dénombrées dans les colons des rats et aucune différence significative n'a été observée (données non-montrées). Ces résultats indiquent une réponse différentielle au niveau de la nature des lésions préneoplasiques.

En effet, la consommation de viande rouge augmente le nombre d'ACF, laissant supposer une augmentation de la prolifération des cellules de épithéliales coliques. En revanche, aucune différence significative n'a été observée au niveau des MDF, laissant supposer un impact moindre de la consommation de viande rouge sur les cellules sécrétrices de mucus.

#### **D. Discussion et conclusion**

Les travaux réalisés dans ce chapitre ont permis de renforcer l'hypothèse d'une peroxydation lipidique augmentée lors de la consommation de viande rouge. En effet, les résultats obtenus ont montré une augmentation des acides mercapturiques urinaires et des TBARS fécaux chez les rats nourris avec les régimes bœuf. D'autres effets, reliés ou non à cette augmentation de la peroxydation lipidique, ont également été observés comme une activation des gènes de détoxification ou encore des changements au niveau du métabolome urinaire. Enfin, dans l'objectif de mieux caractériser le lien entre consommation de viande rouge et promotion du cancer colorectal, plusieurs réponses toxicologiques ont été étudiées dont le nombre de lésions préneoplasiques dans le colon des rats chimio-induit.

Dans les deux expérimentations *in vivo*, une augmentation de l'abondance relative de nombreux mercapturates a été observée dans les urines des rats soumis aux régimes riches en viande rouge. Bien que tous les acides mercapturiques ne dérivent pas de la détoxification de produits de peroxydation lipidique, de nombreuses annotations putatives ainsi que des annotations confirmées (DHN-MA, HNE-MA et HHEP-MA) semblent indiquer qu'une part importante des acides mercapturiques retrouvés augmentés chez les rats soumis aux régimes riches en viande rouge dériveraient de produits issus de la peroxydation lipidique. Ces résultats sont en accord avec de nombreux résultats précédents de l'équipe PPCA où une augmentation de la quantité de DHN-MA urinaire a été constatée chez le rat ou l'Homme après ingestion de régimes riches en viande rouge ou riches en hème (sous forme d'hème purifiée ou d'hémoglobine) (Keller et al., 2020b; Martin et al., 2018; Pierre et al., 2006). Cependant, en complément de ces études, ces résultats permettent, par le nombre et la diversité des mercapturates détectés, de mettre en lumière la probable diversité des produits de peroxydation lipidique hème-induits jusqu'alors peu soulignée dans la littérature. Bien que l'augmentation des acides mercapturiques urinaires ne constitue qu'un reflet indirect de l'exposition à des composés réactifs tels que des produits de peroxydation lipidique, les résultats sont renforcés par l'augmentation significative des TBARS fécaux observée dans les deux expérimentations traduisant ainsi une augmentation de la quantité d'espèces réactives dans le contenu colique. Des résultats similaires ont été obtenus par l'équipe PPCA, mais aussi par d'autres équipes qui ont notamment montré une augmentation de la quantité de MDA (produit d'oxydation des acides gras polyinsaturés) dans le colon

de rats nourris avec de la viande rouge en comparaison à un régime poulet (Van Hecke *et al.*, 2016). Il est cependant à noter que, bien que le dosage des TBARS et l'augmentation des acides mercapturiques dans les urines semblent indiquer une plus forte peroxydation lipidique, les résultats obtenus lors de l'analyse des eaux fécales en spectrométrie de masse semblent moins contrastés et aucune différence significative n'a été observée quant à l'abondance du MDA entre les groupes bœuf et poulet. De plus, contrairement à nos attentes, nous n'avons pas observé d'augmentation du MDA ou des TBARS avec les régimes FO (riches en  $\omega$ -3) alors que ce phénomène avait été observé précédemment en comparant des groupes « hème + huile de poisson » et « hème + huile de carthame (Guéraud *et al.*, 2015b).

Les résultats obtenus dans ces études montrent également que l'augmentation de la peroxydation lipidique hème-induite a pour conséquence une activation, à court terme, de la transcription de gènes impliqués dans le métabolisme et la détoxification des aldéhydes, alors que cet effet n'est plus retrouvé après 100 jours de régime. Ces résultats appuient de précédentes observations de l'équipe PPCA où une activation des gènes impliqués dans le métabolisme des aldéhydes (tels que l'Aldh3A1) a été observée après exposition de cellules Apc +/+ à des eaux fécales de rats nourris avec un régime riche en bœuf (Surya *et al.*, 2016). Dans une étude de 2013, le groupe de Van der Meer a reporté des résultats similaires en mettant également en avant une temporalité de la réponse transcriptionnelle (Ijssennagger *et al.*, 2013). En effet, les auteurs ont observé, chez des souris nourries avec un régime riche en hème purifié, une activation des gènes Akr et Aldh dès le début de l'expérimentation tandis que les gènes impliqués dans le métabolisme du glutathion n'étaient activés que plus tardivement (j7 à j14 après le début des régimes expérimentaux). Bien que nos observations aient été réalisées à j21 et j100 après les débuts des régimes expérimentaux, il est probable que les réponses observées soient moins aiguës et donc moins rapides que dans l'étude précédemment citée où les comparaisons étaient réalisées entre un régime à base d'hème purifié et un régime non supplémenté en hème.

Ces résultats suggèrent donc une adaptation du métabolisme au cours du temps en induisant une augmentation de la production de glutathion afin de lutter contre l'augmentation de la peroxydation lipidique hème-induite. Cette hypothèse est notamment appuyée par les résultats obtenus suite à l'analyse du métabolome urinaire et l'analyse des colons en imagerie par spectrométrie de masse. En effet, le pyroglutamate (ou 5-oxoproline) qui est l'un des métabolites urinaires présentant une forte corrélation négative avec les acides mercapturiques (voir Figure 40B) est un métabolite du cycle de la synthèse et dégradation du glutathion. Ce métabolite formé à partir du précurseur du glutathion, la  $\gamma$ -glutamylcystéine, ou lors du recyclage du glutathion est retrouvé fortement diminué dans les urines des rats soumis aux régimes riches en viande rouge. Cet effet pourrait être lié à une consommation élevée de la réserve de glutathion chez ces animaux afin de détoxifier les produits de peroxydation

lipidique *via* la voie des acides mercapturiques, limitant ainsi le recyclage de la  $\gamma$ -glutamylcystéine ou du glutathion en 5-oxoproline (Figure 47). L'hypothèse d'une surconsommation du glutathion au niveau colique malgré une activation des gènes de synthèse est appuyée par les résultats de l'étude préliminaire en imagerie par spectrométrie de masse où une déplétion de ce composé a été observée dans les colons des rats BOSA en comparaison aux rats PUSA. Bien que plusieurs résultats d'analyses complémentaires tendent à renforcer cette hypothèse mécanistique, des expérimentations additionnelles sont nécessaires afin de valider cette hypothèse. En effet, afin de confirmer l'hypothèse d'un effet « temps » dans la mise en place de ce phénomène, réaliser des prélèvements d'urines réguliers permettrait d'étudier les niveaux de glutathion et d'oxoproline au cours du temps. De plus, la quantification absolue du glutathion permettrait également de déterminer si les animaux soumis aux régimes riches en viande rouge présentent des niveaux anormalement bas de ce métabolite.

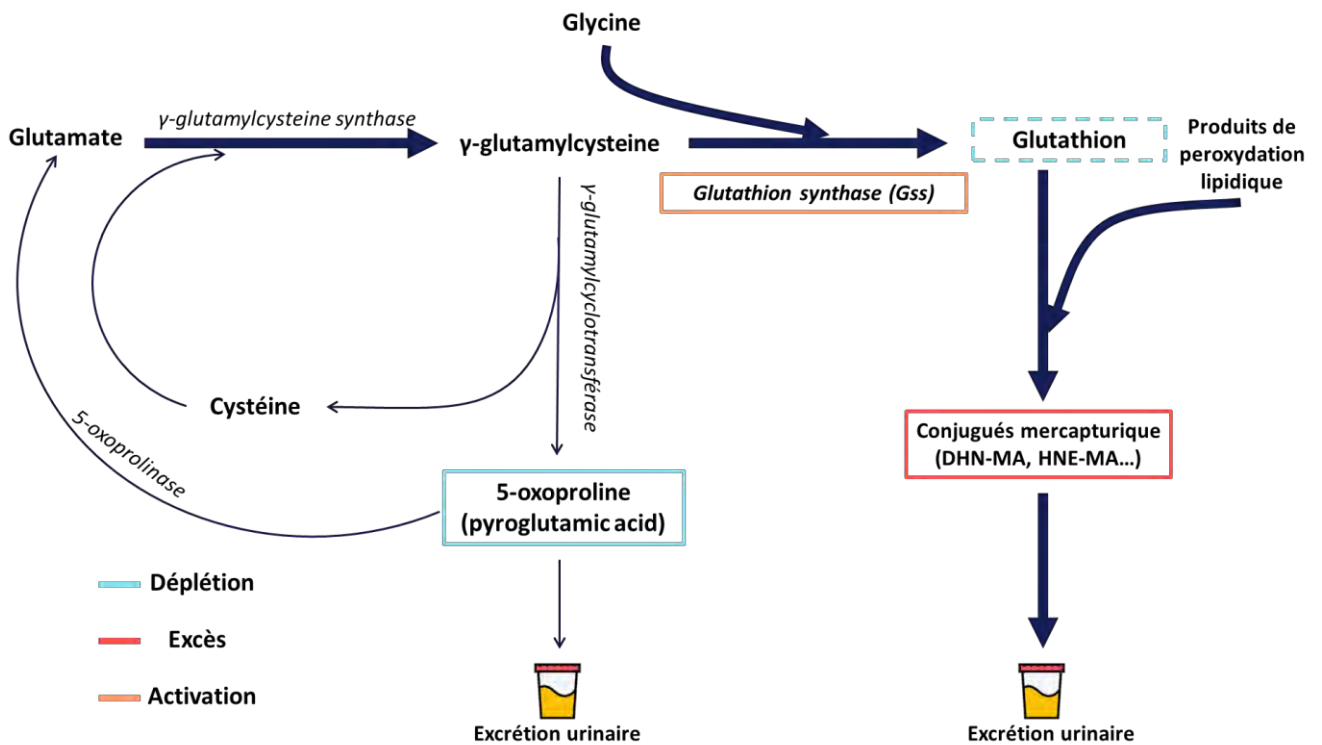


Figure 47 Mécanisme proposé quant à l'impact de la consommation de viande rouge sur le métabolisme du glutathion

Bien qu'une augmentation de la peroxydation lipidique chez les rats soumis aux régimes riches en viande rouge ait été démontrée et que l'hypothèse d'un impact sur le métabolisme du glutathion ait été formulée, les résultats obtenus concernant les potentiels effets toxiques de cette consommation semblent moins contrastés. En effet, les résultats concernant la cytotoxicité des eaux fécales n'indiquent aucune différence significative entre les régimes bœuf et poulet. Cependant, les résultats obtenus dans l'expérimentation nutritionnelle sont en accord avec des résultats précédents de l'équipe PPCA où une augmentation de la cytotoxicité a été reportée pour un régime bœuf en comparaison à un régime contrôle type AIN-76 (Pierre et al., 2008b). Il est notable que parmi les

nombreuses études ayant démontré une augmentation de la cytotoxicité des eaux fécales issues d'animaux soumis à des régimes riches en viande rouge ou riche en hème (sous forme purifiée ou hémoglobine), la plupart des régimes contrôles sont des régimes sans viande (de Vogel et al., 2005b; Sesink et al., 2001c) et que peu d'études ont effectué des comparaisons avec d'autres types de régimes (type viande blanche). Néanmoins, parmi les études ayant comparé des régimes viande rouge à des régimes viande blanche, il est possible de citer une étude de 2017 visant à comparer la cytotoxicité des eaux fécales issues de souris nourries avec différents régimes et qui n'a pas reporté de différence majeure en terme de cytotoxicité entre les régimes bœuf et poulet (Steppeler *et al.*, 2017). Une étude de l'équipe PPCA a cependant reporté une cytotoxicité supérieure des eaux fécales de rats ayant été nourris avec du bœuf en comparaison à un régime poulet (Pierre et al., 2004b).

Les résultats obtenus pour le dosage de la 8-OHdG et la détermination du taux de phosphorylation de l'histone H2AX n'ont pas non plus permis de conclure de façon claire quant à un effet génotoxique plus important induit par les régimes riches en viande rouge. En particulier pour la détermination du taux de cassures double-brin de l'ADN *via* le taux de  $\gamma$ -H2AX, les résultats ne permettent pas d'extraire un effet évident quant à l'impact du type de viande ou d'huile utilisé sur le taux de cassure double-brin bien qu'il semblerait que ce dernier soit augmenté avec la viande blanche. De la même manière que pour les études de cytotoxicité, peu de travaux ont comparé l'impact de régimes riches en viande blanche ou en viande rouge sur le nombre de cassures double-brin de l'ADN. Néanmoins, et contrairement à nos observations, Toden *et al.* ont reporté un taux plus important de cassures double-brin chez des rats soumis à des régimes riches en viande rouge en comparaison à des régimes riches en viande blanche (Toden *et al.*, 2007). En accord avec cette précédente étude, des résultats récents de l'équipe PPCA ont également montré une augmentation du nombre de cellules positives à  $\gamma$ -H2AX lors de régimes riches en hème (Beslay, 2021).

Enfin, malgré des effets limités sur la cytotoxicité ou les dommages à l'ADN, les régimes riches en viande rouge ont cependant conduit à une augmentation du nombre de lésions prénéoplasiques dans le colon proximal des rats nourris avec les régimes riches en viande rouge. Plusieurs travaux ont reporté des résultats similaires en comparant des rats soumis à des régimes riches en hème à des régimes contrôles sans viande (de Moura et al., 2019b; Pierre et al., 2008b, 2004b). Cependant, des différences moins marquées ont été notées en comparant des régimes bœuf et des régimes poulet où aucune différence significative n'a été observée quant au nombre d'ACF ou de MDF (Pierre et al., 2004b). Les travaux proposés dans ce manuscrit viennent appuyer les précédents résultats obtenus par l'équipe PPCA quant à l'effet promoteur de la viande rouge sur la carcinogénèse colique. Cependant, il est notable que bien que la multiplicité des études de l'équipe PPCA renforce cette hypothèse, plusieurs travaux ne sont pas parvenus aux mêmes conclusions et n'ont pas noté d'effet ni



délétère ni protecteur de la consommation de viande rouge ou d'hème (Steppeler et al., 2017; J. M. Winter et al., 2014). La stratégie antioxydante proposée dans ces travaux a également apporté des résultats contrastés en induisant une augmentation du nombre d'ACF et une diminution du nombre de MDF et ce résultat reste à ce jour inexpliqué. Une précédente étude fait également état d'une diminution du nombre de MDF en ajoutant des antioxydants dans un régime riche en jambon alors qu'aucun effet n'est observé sur les ACF (Bastide et al., 2017b).

En définitive, les résultats présentés dans cette partie ont montré une augmentation de la peroxydation lipidique dans les régimes en viande rouge, sans effet notable du statut lipidique du régime. Cependant, les résultats obtenus sur les divers paramètres toxicologiques étudiés posent question quant au rôle des produits de peroxydation lipidique sur l'induction de la cytotoxicité ou sur l'augmentation du nombre d'ACF. En effet, si l'ajout d'antioxydant dans le régime a semblé conduire à une diminution de l'abondance des acides mercapturiques, peu de changements au niveau de la réponse toxicologique ont été constatés suite à cet ajout. Ainsi, le lien causal entre peroxydation lipidique hème-induite et promotion de la carcinogénèse ne semble pas évident au regard de ces résultats bien qu'un effet promoteur de la viande rouge ait été observé en comparaison à la viande blanche. Une expérimentation animale jusqu'au stade tumeur serait un atout afin de conclure quant à l'effet promoteur de la viande rouge. En effet, comme présenté en introduction de ce manuscrit les lésions prénéoplasiques ne conduisent pas toujours à la formation d'une tumeur et c'est d'ailleurs ce qu'ont rapporté de Moura *et al.* qui ont observé une augmentation des ACF mais pas du nombre de tumeurs chez des rats nourris avec un régime riche en hème (de Moura et al., 2019b). Il est malgré tout notable que lors du comptage des ACF dans l'expérimentation de carcinogénèse, 4 tumeurs ont été dénombrées, toutes chez des rats ayant été soumis aux régimes riches en viande rouge.

## **Conclusions et perspectives**

### *Conclusions*

Les travaux présentés dans cette thèse proposent une méthode d'analyse par LC-HRMS permettant l'exploration conjointe du métabolome urinaire et de composés d'intérêt toxicologique : les acides mercapturiques. La volonté d'explorer les acides mercapturiques de manière non-ciblée a émergé d'un constat selon lequel, jusqu'à présent, et malgré un intérêt toxicologique évident, peu de travaux visant à la caractérisation de l'ensemble de ces composés ont été menés. Ainsi, la méthode proposée dans cette thèse se voulait la plus exhaustive possible pour la caractérisation de ces composés. Pour cela, l'utilisation du mode MS<sup>E</sup> permettant d'acquérir, en théorie, des informations structurales sur l'ensemble des ions *via* une fragmentation non-sélective a été préférée à d'autres modes d'acquisition en spectrométrie de masse. Cependant, les travaux présentés dans cette thèse ont mis en lumière un certain nombre de défis inhérents à l'utilisation de ce type d'approche, avec notamment de nombreuses erreurs concernant la déconvolution des spectres MS<sup>E</sup>, conduisant ainsi à des erreurs quant à l'association des ions parents et de leurs fragments. Ainsi, la première partie de cette thèse a été consacrée à l'optimisation de la méthode afin de faire face notamment aux limitations du mode d'acquisition MS<sup>E</sup>, mais aussi aux écueils de l'analyse d'urine en spectrométrie de masse. Si les optimisations proposées dans ce manuscrit ont conduit à une meilleure détection des acides mercapturiques, ces dernières n'ont en revanche pas permis d'améliorer de manière significative la qualité de l'attribution parent / fragment lors du traitement des données MS<sup>E</sup>. Face à cela, des stratégies palliatives associant injections de réplicats et acquisitions MS/MS ciblées ont été proposées, et semblent, à ce jour, nécessaires pour assurer la qualité des données issues de ce type d'acquisition.

Dans ce travail de thèse, le développement de cette méthodologie a été mis au service de la toxicologie alimentaire afin d'étudier le lien entre consommation de viande rouge et de charcuterie et la promotion du CCR. Compte-tenu de l'implication suspectée des produits de peroxydation lipidique dans le processus de promotion de la carcinogénèse colorectale et du caractère électrophile de ces composés, la caractérisation de l'exposition globale à ces composés *via* l'analyse de leurs produits de détoxification, leurs conjugués à l'acide mercapturique, apparaissait comme une démarche pertinente et complémentaire aux dosages de certains acides mercapturiques déjà existants. L'augmentation globale des acides mercapturiques urinaires, observée dans des modèles *in vivo* de rats nourris avec des régimes riches en viande rouge, soutient l'hypothèse d'une peroxydation lipidique augmentée par la viande rouge ou ses constituants. L'utilisation de dosages complémentaires et bien établis tels que les TBARS ou le DHN-MA ont permis de conforter l'utilisation du profilage des acides mercapturiques comme témoins de la peroxydation lipidique.

Dans l'optique de limiter la peroxydation lipidique induite par la viande rouge, l'ajout d'antioxydants dans les régimes alimentaires a également été testé dans ces travaux. L'analyse des mercapturates a révélé une diminution du pouvoir oxydant de la viande rouge lors de l'ajout d'antioxydants dans les régimes.

Les travaux présentés dans cette thèse proposent également une étude de l'impact de la consommation de viande rouge, du contenu lipidique des régimes et du statut en antioxydant, allant au-delà de la peroxydation lipidique. La combinaison d'analyses qPCR, d'imagerie par spectrométrie de masse et de l'exploration du métabolome urinaire, par l'emploi du « *double-workflow* » mis en place au cours de cette thèse, nous a permis de formuler l'hypothèse selon laquelle l'augmentation de la peroxydation lipidique induirait une consommation accrue de glutathion au niveau colique induisant ainsi une réponse transcriptionnelle au niveau des gènes de synthèse de ce composé.

Dans l'optique de lier les précédentes observations à un « effet santé », la réponse toxicologique associée aux différents régimes alimentaires a été étudiée et a révélé des résultats hétérogènes. Si aucune différence significative n'a été constatée au niveau de la cytotoxicité des eaux fécales issues des régimes riches en viande rouge comparativement aux régimes riches en viande blanche, l'augmentation significative du nombre de lésions précancéreuses semble soutenir l'hypothèse d'une promotion de la carcinogénèse par la viande rouge. En revanche, l'augmentation du nombre de ces lésions dans le régime supplémenté en antioxydant pose question quant au rôle des produits issus de la peroxydation lipidique sur la promotion du CCR.

Pour finir, l'analyse des échantillons issus de la cohorte NutriNet-Santé a permis d'étudier de manière plus directe les effets d'une consommation élevée de viande rouge et de charcuterie chez l'Homme. Bien que les résultats issus de cette analyse n'aient pas permis d'étayer l'hypothèse d'une peroxydation lipidique induite par la consommation de produits carnés, les modifications du métabolome urinaire observées nous ont permis de formuler l'hypothèse d'une probable modification du microbiote par les régimes alimentaires. Cette hypothèse a été renforcée par des observations similaires dans l'expérimentation nutritionnelle chez le rat au niveau des métabolites microbiens, mais aussi par un séquençage du microbiote révélant une modulation du microbiote par la consommation de viande rouge chez ces animaux.

En définitive, cette thèse propose une méthode d'exploration du métabolome urinaire avec une focalisation sur les acides mercapturiques. Les développements et optimisations réalisés ont conduit à l'obtention d'une méthode robuste permettant de pallier les obstacles inhérents au mode d'acquisition employé. L'application de cette approche à la problématique du lien entre consommation de viande rouge et CCR a permis de confirmer l'hypothèse de départ suspectant une large diversité

des produits de peroxydation lipidique, ne pouvant être résumée au dosage du seul DHN-MA. La combinaison des techniques de chimie analytique avec d'autres types d'expérimentation complémentaires a permis d'obtenir une vue globale des effets de la consommation de viande rouge, générant ainsi de nouvelles hypothèses non exclusivement centrées sur les produits de peroxydation lipidique. Enfin, cette thèse appuie l'hypothèse d'une promotion de la carcinogénèse induite par la viande rouge, cependant, compte-tenu des résultats proposés, le rôle de produits de peroxydation lipidique ou du microbiote dans ce phénomène reste à approfondir.

### *Perspectives*

#### Améliorer l'identification des inconnus pour mieux étudier leurs effets.

L'une des limites principales de l'approche proposée réside dans le faible nombre de mercapturates formellement identifiés. En effet, dans l'ensemble des expérimentations réalisées seuls 3 composés ont été annotés au niveau 1. Cependant, les approches utilisées ont mis en lumière de nombreux mercapturates putativement identifiés ou inconnus, qui semblent dériver de produits de peroxydation lipidique. Ces identifications putatives doivent être confirmées par la comparaison avec des standards de mercapturates afin de confirmer leur identité. Néanmoins, comme de nombreux métabolites issus de métabolisation, peu de standards commerciaux sont disponibles et le recours à la synthèse organique semble nécessaire afin de constituer une base de standards de conjugués mercapturiques. S'il apparaît irréaliste de synthétiser des standards de mercapturates pour l'ensemble des produits de peroxydation lipidique, il serait envisageable de rationaliser les efforts en se concentrant, dans un premier temps, sur les mercapturates putatifs retrouvés significativement augmentés dans plusieurs expérimentations. L'utilisation des réseaux moléculaires pourrait également apporter des informations sur l'origine des différents mercapturates détectés. En effet, les mercapturates ne provenant pas uniquement des produits de peroxydation lipidique, il est probable que les mercapturates issus d'une même classe chimique produisent des fragments communs et qu'il soit possible de distinguer différentes « familles » de mercapturates.

Ainsi, l'identification formelle de nombreux mercapturates, et par conséquent de leurs précurseurs, pourrait permettre de mieux caractériser les effets biologiques de l'exposition aux produits de peroxydation lipidique. En effet, si la cytotoxicité du 4-HNE est largement connue, les résultats obtenus dans ces travaux laissent supposer qu'une exposition à un mélange de produits de peroxydation lipidique est une vision plus conforme à la réalité. Une connaissance accrue des produits de peroxydation lipidique formés lors de la consommation de viande rouge permettrait de planifier des expériences de cytotoxicité faisant intervenir des mélanges de composés, permettant ainsi d'étudier de potentiels effets synergiques. En effet, si des effets cocktail additifs ou synergiques ont déjà été

décrits pour des expositions aux pesticides, par exemple, il est probable que ce genre de phénomène puisse s'appliquer aux produits issus de la peroxydation lipidique (Cedergreen, 2014).

#### Standardiser les pratiques pour des conclusions plus robustes

Dans ces travaux une augmentation de la peroxydation lipidique a été reportée dans les modèles *in vivo* soumis aux régimes riches en viande rouge, mais pas lors de l'analyse des échantillons issus de la cohorte NutriNet-Santé. Bien que l'hypothèse que les quantités employées dans les expérimentations *in vivo* puissent induire des effets non-observables dans des conditions plus « physiologiques » ne soit pas à exclure, il est probable que cet effet n'ait pas été observé en raison d'un manque de standardisation quant au prélèvement des urines dans la cohorte humaine. En effet, compte tenu du caractère transitoire de la formation des acides mercapturiques et de leur élimination, un prélèvement idéal se situerait entre 6 et 24 h après ingestion d'un repas (Keller et al., 2015). De plus, ce prélèvement devrait être réalisé au même moment de la journée pour éviter les biais liés au cycle circadien. Dans l'étude NutriNet-Santé, les habitudes alimentaires sont déterminées à partir de questionnaires, reflétant ainsi des moyennes de consommation, mais aucune information n'est disponible sur les repas consommés durant les 24 h précédant le prélèvement. Ce biais pourrait ainsi induire des différences entre les quantités consommées habituellement et déclarées dans les questionnaires, et la quantité réelle ingérée avant les prélèvements ayant généré les échantillons. Ainsi, collecter cette information lors du prélèvement des échantillons permettrait d'éliminer les individus dont les repas précédant le prélèvement diffèrent significativement des habitudes déclarées.

La mise en place d'une étude interventionnelle pourrait également permettre de conclure quant à la peroxydation lipidique induite par des régimes carnés. Pour cela, l'étude clinique MeaTlc dans laquelle est impliquée l'équipe PPCA et où un prélèvement d'urine est prévu après 3 mois de régimes standardisés pourrait être utilisée (Sofi et al., 2019).

#### Combiner études interventionnelles et observationnelles pour mieux caractériser la modulation du microbiote induite par la viande rouge

L'analyse métabolomique réalisée sur les échantillons de la cohorte NutriNet-Santé et la comparaison avec l'expérimentation nutritionnelle *in vivo* a mis en lumière une probable modulation du microbiote chez les gros consommateurs de viande rouge. L'analyse du microbiote dans une étude interventionnelle, telle que l'étude MeaTlc mentionnée ci-avant, permettrait de caractériser les modulations à court-terme induites par la consommation de viande rouge, tandis que l'analyse du microbiote des individus de la cohorte NutriNet-Santé appartenant au plus haut quintile de consommation de viande rouge et de charcuterie permettrait de mettre en avant des changements plus profonds. En effet, si plusieurs études ont montré que des changements de régime alimentaire induisaient des modifications très rapides du microbiote (de l'ordre de quelques jours), d'autres

travaux ont également mis en avant des changements à plus long terme (pouvant se répercuter sur les générations suivantes) (David et al., 2014; Leeming et al., 2019; Sonnenburg et al., 2016). Ainsi, la combinaison d'études interventionnelles et observationnelles permettrait d'établir un « profil microbien » des gros consommateurs de viande rouge et de charcuterie. L'établissement d'un tel profil permettrait alors d'étudier plus finement les effets d'un microbiote caractéristique des gros consommateurs de viande rouge et de charcuterie sur la santé de cette population en implantant, par exemple, ce microbiote dans des animaux axéniques.

#### Conduire une expérimentation jusqu'au stade tumoral

Les résultats de l'expérimentation de carcinogénèse réalisée au cours de cette thèse montrent une augmentation significative du nombre d'ACF chez les rats nourris avec de la viande rouge, tandis qu'aucune différence significative n'a été observée quant au nombre de MDF. En considérant le fait que le nombre d'ACF n'est pas toujours corrélé avec le nombre d'adénome, la réalisation d'une expérimentation de carcinogénèse jusqu'au stade tumoral apporterait des conclusions plus robustes quant à la promotion du CCR induite par la viande rouge.

#### Appliquer la méthodologie analytique à d'autres thématiques et conjugués

La méthodologie analytique proposée dans ces travaux permet une exploration non ciblée des mercapturates, reflétant ainsi notre exposition à des composés électrophiles. Cette caractérisation globale des mercapturates urinaires réalisée au cours de cette thèse pourrait être appliquée à d'autres problématiques que celle étudiée dans ces travaux, par exemple pour répondre à un besoin de suivi de l'exposition à des composés toxiques *via* une exposition professionnelle ou liée à un environnement particulier. Par exemple, Frigerio *et al.* ont quantifié 17 mercapturates urinaires dans une étude comparant 50 individus travaillant dans une usine de production de coke (combustible utilisé dans la sidérurgie), à 50 individus contrôle, et ont révélé une augmentation significative de certains mercapturates chez les travailleurs (Frigerio et al., 2020). L'application à ce type d'étude de la méthodologie développée au cours de cette thèse permettrait d'obtenir des informations complémentaires et une vue plus exhaustive de l'ensemble des mercapturates.

De plus, cette stratégie pourrait être étendue à d'autres types de conjugués que le seul conjugué *N*-actetyl-cystéine. En effet, la recherche des conjugués glutathion ou cystéine permettrait d'augmenter la couverture de la voie des mercapturates bien que cela nécessiterait un nombre plus important de MS/MS de confirmation. S'agissant en particulier de la matrice urinaire, et bien qu'ils ne soient pas forcément représentatifs d'une exposition à des composés électrophiles et /ou toxiques, les conjugués sulfates ou glucuronides issus des réactions de métabolisme de phase II pourraient également être *screenés* en s'inspirant de cette méthodologie.

Dans la même optique, le *monitoring* de l'exposition à des composés électrophiles, et autres composés réactifs, pourrait également être réalisé *via* l'étude des adduits protéiques, en particulier les adduits à l'hémoglobine ou l'albumine, en appliquant une stratégie similaire. Bien que, d'un point de vue analytique, la méthode proposée ne permette pas directement d'analyser les adduits protéiques (ceci implique l'utilisation d'une matrice différente et un protocole de préparation d'échantillon spécifique), l'application d'une stratégie similaire à cette problématique permettrait de révéler des expositions sur une période de temps plus longue, compte-tenu de la durée de vie de ces protéines dans l'organisme (120 jours pour l'hémoglobine par exemple) (Sabbioni and Turesky, 2017).

**Annexes***Annexe 1*

		Données QToF	Données Orbitrap
<b>XCMS find ChromPeaks</b>	Extraction method for peaks detection Max tolerated ppm $m/z$ deviation in consecutive scans in ppm Min,Max peak width in seconds Signal to Noise ratio cutoff Prefilter step for for the first analysis step (ROI detection)  Name of the function to calculate the $m/z$ center of the chromatographic peak  Minimum difference in $m/z$ for peaks with overlapping retention times Noise filter	CentWave 10 5, 40 10 3 , 1000  intensity weighted mean of the peak's $m/z$ values 0.01 10000	CentWave 10 5, 40 10 4 , 1000  intensity weighted mean of the peak's $m/z$ values 0.01 5000
<b>XCMS groupChromPeaks (group)</b>	Method to use for grouping Bandwidth Minimum fraction of samples Minimum number of samples Width of overlapping $m/z$ slices Maximum number of groups to identify in a single $m/z$ slice	PeakDensity 10 0.7 1 0.01 50	PeakDensity 10 0.7 1 0.01 50
<b>XCMS adjustRtime (retcor)</b>	Method to use for retention time correction Minimum required fraction of samples in which peaks for the peak group were identified Maximal number of additional peaks for all samples to be assigned to a peak group for retention time correction Smooth method	PeakGroups 0.7 1 loess	PeakGroups 0.7 1 loess
<b>XCMS groupChromPeaks (group)</b>	Method to use for grouping Bandwidth Minimum fraction of samples Minimum number of samples Width of overlapping $m/z$ slices Maximum number of groups to identify in a single $m/z$ slice	PeakDensity 7.5 0.7 1 0.01 50	PeakDensity 7.5 0.7 1 0.01 50
<b>XCMS fillChromPeaks (fillPeaks)</b>	Number of decimal places for mass values reported in ions' identifiers. Number of decimal places for retention time values reported in ions' identifiers. Reported intensity values	4 0 into	4 0 into

Annexe 1 Paramètres utilisés pour le traitement des données *via* XCMS sous l'environnement W4M



**Annexe 2**

		Method 1
	Rt range (min)	2.5-17
<b>High energy</b>	Intensity threshold	200
	<i>m/z</i> start	10
	<i>m/z</i> end	530
<b>Low energy</b>	Intensity threshold	2500
	<i>m/z</i> start	10
	<i>m/z</i> end	530
	Background filter	High
<b>3D isotope clustering and high to low energy association</b>	Fraction of the chromatographic peak width to apply during cluster creation <sup>a</sup>	0.25
	Fraction of the chromatographic peak width to apply during high to low energy association <sup>b</sup>	0.25
<b>Common Neural Loss</b>	Tolerance (mDa) <sup>c</sup>	5

Annexe 2 Paramètres du logiciel UNIFI pour le traitement des données MS<sup>E</sup>

<sup>a</sup> Correspond à la largeur du pic à considérer lors de la recherche des isotopes

<sup>b</sup> Correspond à la largeur du pic à considérer lors de l'attribution parent/fragment entre le scan de haute énergie et le scan basse énergie

<sup>c</sup> Correspond à la tolérance en masse appliquée lors de la recherche de la perte de neutre caractéristiques des acides mercapturiques (129,0426Da)

**Annexe 3**

			<b>Tune 1</b>	<b>Tune 2</b>	<b>Tune 3</b>
StepWave	Step Wave 1	<b>Wave Velocity (m/s)</b>	<b>20</b>	<b>20</b>	<b>300</b>
		<b>Wave height</b>	<b>15</b>	<b>15</b>	<b>10</b>
	StepWaveDC	<b>StepWave2 Offset</b>	<b>25</b>	<b>25</b>	<b>16</b>
		Diff Aperture 1	3	3	3
		Diff Aperture 2	0	0	0
	RF Settings	<b>StepWave</b>	<b>100</b>	<b>100</b>	<b>150</b>
		<b>Ion Guide</b>	<b>100</b>	<b>100</b>	<b>200</b>
TriWave DC	Trap DC	<b>Entrance</b>	<b>1</b>	<b>0</b>	<b>1</b>
		<b>Bias</b>	<b>2</b>	<b>1</b>	<b>1</b>
		<b>Trap DC</b>	<b>-2</b>	<b>0</b>	<b>-1</b>
		Exit	0	0	0
	IMS DC	Entrance	-20	-20	-20
		Helium Cell DC	1	1	1
		Helium Exit	-20	-20	-20
		<b>Bias</b>	<b>2</b>	<b>1</b>	<b>1</b>
		Exit	20	20	20
	Transfer DC	Entrance	5	5	5
		Exit	15	15	15

Annexe 3 Paramètres de transfert des ions au niveau du StepWave et TriWave

**Annexe 4**

		Méthode 1	Méthode 2	Méthode 3	Méthode 4	Méthode 5	Méthode 6
	Rt range (min)	2,5-20	2,5-20	2,5-20	2,5-20	2,5-20	2,5-20
<b>High energy</b>	Intensity treshold	200	200	200	20	500	20
	<i>m/z</i> start	10	10	10	10	10	10
	<i>m/z</i> end	800	800	800	800	800	800
<b>Low energy</b>	Intensity treshold	2500	2500	2500	2500	5000	5000
	<i>m/z</i> start	230	230	230	230	230	230
	<i>m/z</i> end	800	800	800	800	800	800
	Background filter	High	High	High	High	High	High
<b>3D isotope clustering and high to low energy association</b>	Fraction of the chromatographic peak width to apply during cluster creation <sup>a</sup>	0,25	0,25	0,33	0,33	0,33	0,33
	Fraction of the chromatographic peak width to apply during high to low energy association <sup>b</sup>	0,25	0,25	0,33	0,33	0,33	0,33
<b>Common Neural Loss</b>	Tolerance (mDa) <sup>c</sup>	2	4	5	5	5	5

**Annexe 4 Ensemble des jeux de paramètres testés pour le traitement des données MSE avec le logiciel UNIFI**

<sup>a</sup> Correspond à la largeur du pic à considérer lors de la recherche des isotopes

<sup>b</sup> Correspond à la largeur du pic à considérer lors de l'attribution parent/fragment entre le scan de haute énergie et le scan basse énergie

<sup>c</sup> Correspond à la tolérance en masse appliquée lors de la recherche de la perte de neutre caractéristiques des acides mercapturiques (129,0426Da)

**Annexe 5**

	Farine régime contrôle	Farine complément
<b>Caseine</b>	13.7	0
<b>C.Starch</b>	17.12	30
<b>Sucrose</b>	57.08	49.44
<b>Cellulose</b>	5.71	10
<b>Methionine</b>	0.59	0.6
<b>Min. ss Ca</b>	4	7
<b>Vitamine</b>	1.14	2
<b>Choline Bitartrate</b>	0.23	0.4
<b>Ca phos.</b>	0.34	0.56
<b>Citrate de fer</b>	0	0
<b>Tryptophane</b>	0.096	0

Annexe 5 Composition des farines utilisées pour les expérimentations animales. Toutes les quantités sont exprimées en g pour 100 g

**Annexe 6**

Gene	Alias	sequence primer Forward	sequence primer Reverse	rapport F/R (en nM)
<b>Aldh2</b>	<b>ALDH2</b>	GCCGCAGACCGTGGTTACT	CGATGGTCATGCCATCTTTG	100/100
<b>Akr1b7</b>	<b>AKR1B7</b>	CTTATTCAGGAGCCTTTCGATCTG	GGGCTGGTGAAAGCTTTGG	100/100
<b>Hmox1</b>	<b>HO1</b>	CAACCCACCAAGTTCAAACA	AGGCGGTCTTAGCCTCTTCTG	100/100
<b>Gss</b>	<b>GSHS</b>	GCATCTCGGAGCTGGGTATT	TCACAAGTGTGTTCCCTGTCT	100/100

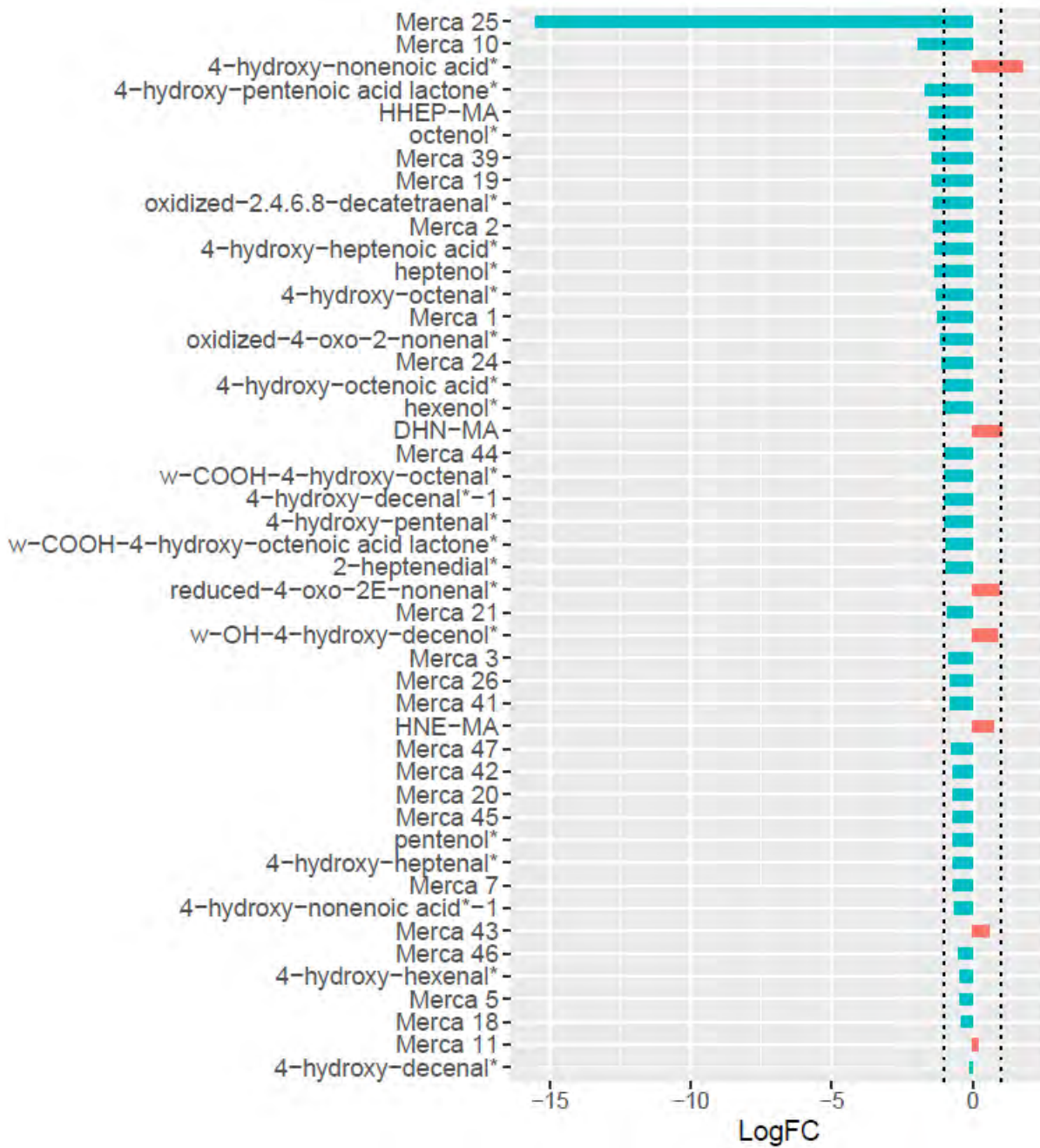
Annexe 6 Primers utilisés pour les expériences de qPCR sur le tissu épithélial colique dans les expérimentations *in vivo*

## Annexe 7

Annotation	m/z	Temps de rétention (min)	Fold BOSA/CON	Fold BOSA/PUSA	Fold BOFO/PUFO	Fold BOSAR/BOSA	Fold BOFOR/BOFO	Fold PUSAR/PUSA	Fold PUFOR/PUFO
Merca 1	230.1031	4.41	1.01	0.99	<b>0.43</b>	0.70	0.84	0.63	<b>0.56</b>
Merca 2	230.1030	4.71	1.09	1.14	<b>0.38</b>	0.75	0.73	0.56	<b>0.51</b>
Merca 3	244.1187	8.33	1.08	1.09	<b>0.56</b>	0.74	0.87	0.74	<b>0.60</b>
pentenol*	248.0957	4.86	1.08	1.18	<b>0.62</b>	0.93	1.38	1.10	1.00
Merca 5	250.0716	5.41	1.33	1.67	0.73	0.88	0.91	0.82	0.83
4-hydroxy-pentenoic acid lactone*	260.0584	5.35	1.52	1.02	<b>0.31</b>	0.79	0.90	0.75	0.72
Merca 7	261.0877	5.54	<b>0.35</b>	0.90	<b>0.62</b>	<b>0.44</b>	0.81	<b>0.54</b>	<b>0.60</b>
4-hydroxy-pentenal*	262.0749	5.88	1.30	0.99	<b>0.51</b>	0.94	0.72	0.59	<b>0.57</b>
hexenol*	262.1115	9.16	<b>13.86</b>	<b>1.96</b>	<b>0.49</b>	<b>0.41</b>	1.04	1.59	0.73
Merca 10	272.1502	12.69	1.18	1.55	<b>0.26</b>	0.44	2.17	1.01	0.94
Merca 11	276.0890	7.69	<b>3.56</b>	<b>2.12</b>	1.14	0.64	0.87	1.01	0.87
4-hydroxy-hexenal*	276.0905	9.25	0.76	1.64	<b>0.73</b>	0.80	1.18	1.08	0.89
heptenol*	276.1274	12.06	<b>4.39</b>	1.17	<b>0.40</b>	0.57	1.24	1.48	0.72
2-heptenedial*	288.0908	8.66	0.53	1.03	0.52	1.45	1.62	<b>1.92</b>	1.01
4-hydroxy-heptenal*	290.1062	9.96	0.98	1.03	<b>0.62</b>	1.15	1.73	1.32	1.36
octenol*	290.1423	14.73	<b>2.00</b>	1.33	<b>0.35</b>	<b>0.31</b>	1.06	1.04	<b>0.31</b>
HHEP-MA	292.1217	4.88	<b>9.42</b>	1.53	<b>0.35</b>	<b>0.41</b>	3.01	1.51	0.76
Merca 18	298.0728	9.24	0.83	1.55	0.75	0.84	1.02	0.86	0.84
Merca 19	300.0986	9.40	1.05	0.52	<b>0.36</b>	0.91	1.14	1.05	0.96
Merca 20	300.0909	13.30	1.12	2.02	<b>0.61</b>	1.04	1.04	1.49	0.76
Merca 21	303.0973	6.67	1.02	1.14	<b>0.54</b>	0.95	0.93	0.97	0.88
4-hydroxy-octenal*	304.1221	14.78	0.59	1.02	<b>0.40</b>	<b>0.38</b>	0.84	0.55	<b>0.47</b>
4-hydroxy-heptenoic acid*	306.1014	4.98	<b>2.64</b>	1.60	<b>0.39</b>	0.69	0.94	1.00	0.60
Merca 24	307.0928	4.28	1.00	0.85	<b>0.47</b>	0.96	0.97	0.89	0.65
Merca 25	312.0653	2.55	1.71	<b>0.02</b>	<b>0.00</b>	1.53	637.09	0.88	0.72
Merca 26	314.1409	14.16	0.92	1.22	0.57	1.21	1.23	1.12	1.08
reduced-4-oxo-2E-nonenal*	318.1376	12.67	<b>8.73</b>	1.44	1.88	0.62	1.04	1.12	0.73
HNE-MA	318.1376	13.05	<b>5.23</b>	1.62	1.68	0.62	1.49	1.37	0.67
4-hydroxy-octenoic acid*	320.1171	6.84	<b>2.12</b>	1.44	<b>0.48</b>	0.75	1.46	1.40	0.87
DHN-MA	320.1536	12.48	<b>4.97</b>	1.65	<b>2.03</b>	<b>0.40</b>	1.42	1.24	0.75
oxidized-2.4.6.8-decatetraenal*	326.1065	16.27	1.21	1.11	<b>0.38</b>	1.04	0.78	1.06	0.63
4-hydroxy-decenal*-1	332.1531	17.08	1.16	0.92	0.50	0.72	0.98	0.78	0.48
w-COOH-4-hydroxy-octenoic acid lactone*	332.0805	4.71	<b>2.09</b>	1.54	<b>0.51</b>	0.66	1.28	1.29	0.74
oxidized-4-oxo-2-nonenal*	332.1165	7.94	<b>3.99</b>	1.70	<b>0.46</b>	0.48	1.03	1.08	0.80
4-hydroxy-decenal*	332.1532	14.10	<b>3.49</b>	<b>2.61</b>	0.94	0.61	2.39	2.39	0.63
w-COOH-4-hydroxy-octenal*	334.0973	6.70	0.82	1.25	<b>0.50</b>	0.97	0.89	0.76	0.66
4-hydroxy-nonenic acid*-1	334.1327	8.97	<b>2.24</b>	1.45	<b>0.64</b>	0.69	1.28	1.17	0.79
4-hydroxy-nonenic acid*	334.1325	12.16	<b>4.60</b>	2.11	3.43	<b>0.31</b>	1.66	1.17	0.66
Merca 39	348.0882	16.26	1.24	1.10	<b>0.36</b>	1.01	0.99	1.02	0.58
w-OH-4-hydroxy-decenol*	350.1638	10.14	<b>6.56</b>	<b>7.10</b>	1.81	0.60	<b>2.52</b>	<b>4.47</b>	0.50
Merca 41	377.9647	7.30	0.93	1.18	0.58	1.09	1.60	1.11	1.34
Merca 42	381.1117	7.11	1.04	1.86	<b>0.61</b>	0.88	1.09	3.12	1.11
Merca 43	384.1301	4.49	<b>0.31</b>	<b>22.35</b>	1.51	0.49	1.53	13.94	2.62
Merca 44	395.1640	9.68	2.59	383.34	<b>0.50</b>	1.13	1.06	65.93	0.76
Merca 45	422.1640	17.39	1.27	1.06	0.61	<b>24.61</b>	<b>35.65</b>	<b>38.31</b>	<b>26.08</b>
Merca 46	444.1459	17.39	1.37	1.18	0.70	<b>43.89</b>	<b>44.67</b>	<b>75.56</b>	<b>36.62</b>
Merca 47	460.1391	14.20	2.76	1.36	<b>0.60</b>	8.81	<b>11.65</b>	<b>30.19</b>	<b>17.41</b>

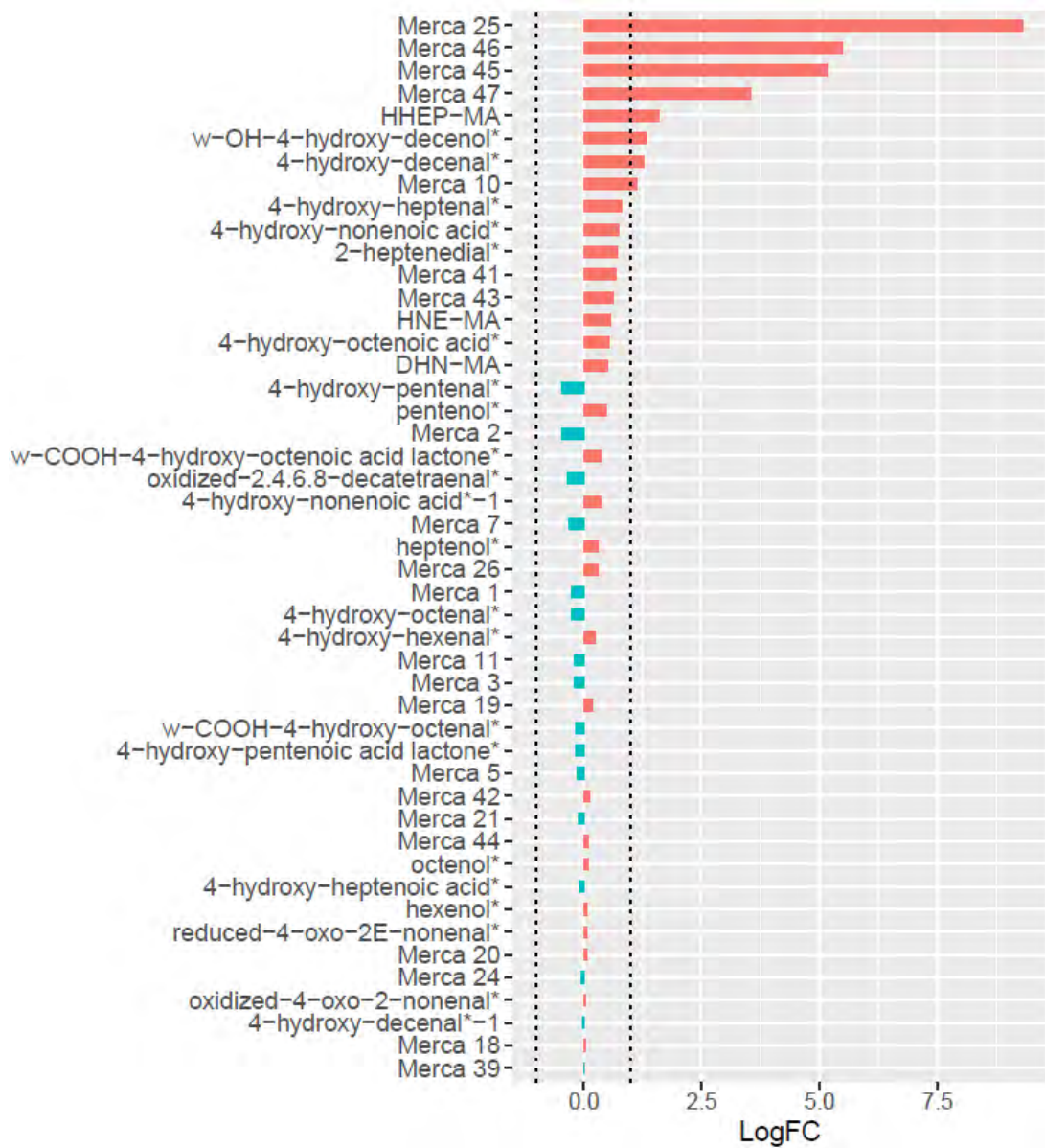
Annexe 7 Mercapturates détectés dans l'expérimentation nutritionnelle. Les valeurs en rouge et gras correspondent aux mercapturates significatifs (Wilcoxon + fdr, p-value 0.05) pour chaque comparaison

Annexe 8



Annexe 8 Fold change des 47 acides mercapturiques entre les groupes BOFO et PUFO. Les résultats sont exprimés en Log(2), une valeur positive correspondant à une augmentation dans le groupe BOFO par rapport au groupe PUFO. Une « \* » signifie une annotation putative

**Annexe 9**



Annexe 9 Fold change des 47 acides mercapturiques entre les groupes BOFOR et BOFO. Les résultats sont exprimés en Log(2), un valeur positive correspondant à une augmentation dans le groupe BOFOR par rapport au groupe BOFO. Une « \* » signifie une annotation putative



## Annexe 10

Annotation	m/z [M+H] <sup>+</sup>	Temps de rétention (min)	Fold BOSA/CON	FOLD BOSA/PUSA	Fold BOFO/PUFO	Fold BOSAR/BOSA	Niveau d'identification
Acetone	242.0178	16.42	0.96	0.98	0.98	0.99	1
Butanal	256.0334	17.36	1.01	0.92	1.02	1.10	2
Butanal	256.0336	18.08	0.89	1.01	0.96	1.07	2
lactaldehyde	258.0123	12.71	1.44	0.85	<b>0.53</b>	1.20	3
Pentanal	270.0490	18.02	<b>0.38</b>	0.41	0.92	1.90	1
Hexanal	284.0648	18.90	0.99	0.98	1.01	1.01	2
Heptanal	298.0802	19.48	0.36	0.79	0.80	7.86	2
acide 5 oxopentanoïque	300.0232	14.79	1.42	0.53	0.42	0.88	1
oxo-hexanoic acid	314.0384	16.23	0.73	0.44	0.50	1.57	3
acide 7 oxoheptanoïque	328.0545	15.97	0.53	0.84	<b>0.60</b>	1.27	1
xylose	334.0284	11.55	1.45	1.11	1.09	0.99	1
rhamnose	348.0439	11.90	1.57	0.83	1.42	0.97	1
galactose	364.0390	11.31	<b>2.90</b>	1.69	1.52	0.90	1
4,5-dihydroxy-decenal	370.1015	18.44	0.61	0.55	0.49	0.79	3
8-carboxy-4,5-dihydroxy-octenal	372.0442	13.65	<b>1.73</b>	1.03	<b>0.72</b>	0.78	3
10-carboxy-4-oxo-decenoic acid	398.0593	14.19	1.84	0.70	0.40	0.52	3
4-hydroxy-pentadecenal	424.1847	19.87	0.76	0.99	0.79	0.78	3
12-carboxy-4,5-dihydroxy-dodecenal	428.1064	15.33	2.12	0.25	0.54	0.43	3
pyruvaldehyde	438.9653	20.01	1.19	0.88	0.81	<b>0.63</b>	1
MDA	438.9656	18.55	<b>2.41</b>	1.70	0.81	<b>0.28</b>	1
oxo-octadecanoic acid	482.2263	18.52	<b>0.32</b>	<b>0.36</b>	<b>0.22</b>	1.31	3
3-oxostearic acid	482.2267	21.53	0.94	<b>0.53</b>	<b>0.35</b>	1.29	2
eicosapentaenoic acid	486.2000	19.84	<b>26.91</b>	2.63	0.79	0.75	2
oxo-nonadecanoic acid	496.2422	24.34	<b>1.59</b>	<b>0.51</b>	<b>0.36</b>	0.81	2

Annexe 10 Aldéhydes et autres composés annotés et détectés suite à l'analyse des fèces de l'expérimentation nutritionnelle *in vivo*. Les valeurs en rouge et en gras correspondent aux composés significatifs (ViP > 1 + p-value < 0.05) pour chaque comparaison

## Annexe 11

Annotation	<i>m/z</i>	Temps de rétention (min)	Fold BOSA/PUSA	Fold BOFO/PUFO	Fold BOSAR/BOSA
Merca1	230.1027	4.77	0.97	1.02	<b>0.73</b>
Merca2	230.1028	5.07	1.10	1.07	0.76
Merca3	242.1039	8.02	0.95	<b>0.62</b>	1.02
Merca4	244.1190	7.76	1.38	1.15	<b>1.35</b>
Merca5	244.1186	8.74	0.95	0.98	<b>0.70</b>
Merca6	246.0973	4.24	1.09	0.95	0.95
4-hydroxy-butenal*	248.0591	4.24	<b>1.67</b>	<b>1.64</b>	0.99
pentenol*	248.0957	5.21	1.03	1.10	1.19
Merca9	249.0877	4.49	<b>1.46</b>	<b>1.35</b>	<b>1.36</b>
Merca10	256.1006	13.62	0.94	1.10	1.23
4-hydroxy-pentenoic acid lactone*	260.0583	5.67	0.88	0.93	<b>0.79</b>
Merca12	260.1127	5.70	0.99	0.72	1.05
Merca13	259.9128	7.62	0.89	<b>0.65</b>	1.07
Merca14	261.0878	5.77	0.96	<b>1.23</b>	<b>0.79</b>
Merca15	261.9119	7.62	0.89	<b>0.45</b>	0.83
hexenol*	262.1110	9.54	<b>2.76</b>	1.05	<b>0.64</b>
Merca17	261.9098	11.38	<b>7.15</b>	4.44	0.45
Merca18	263.9102	7.62	0.86	0.50	0.87
Merca19	267.0809	5.78	1.20	1.26	1.30
Merca20	272.1500	13.05	0.94	0.77	0.91
4-hydroxy-hexenal*	276.0905	9.61	<b>1.15</b>	<b>1.36</b>	1.08
heptenol*	276.1269	12.42	<b>1.64</b>	0.94	<b>0.57</b>
4-hydroxy-heptenoic acid lactone*	288.0904	8.86	0.91	<b>0.86</b>	1.11
octenal*	288.1267	13.06	<b>3.11</b>	0.95	0.65
4-hydroxy-heptenal*	290.1057	10.27	0.85	0.86	<b>1.53</b>
octenol*	290.1423	15.07	<b>1.40</b>	1.07	<b>0.27</b>
4-hydroxy-hexenoic acid*	292.0851	3.84	1.05	1.38	0.94
Merca28	293.1121	6.72	0.67	0.74	1.72
Merca29	298.0724	9.60	<b>1.23</b>	<b>1.55</b>	1.00
oxidized-2.4.6-octatrienal-1*	300.0908	12.17	<b>2.58</b>	1.03	1.07
oxidized-2.4.6-octatrienal*	300.0906	13.62	0.99	0.98	1.29
nonenal-1*	302.1427	12.63	1.09	1.08	1.02
nonenal-2*	302.1426	13.96	1.11	1.09	1.06
nonenal-3*	302.1406	14.50	1.13	0.94	0.71
oxidized-2-heptenedial*	304.0851	4.28	1.29	1.05	0.84
4-hydroxy-octenal-1*	304.1222	7.67	0.85	0.77	0.77
4-hydroxy-octenal*	304.1221	15.10	0.91	0.97	<b>0.38</b>
4-hydroxy-heptenoic acid*	306.1012	5.28	<b>1.48</b>	1.16	0.77

4-hydroxy-octenol-1*	306.1368	7.45	1.59	1.24	0.76
4-hydroxy-octenol*	306.1368	9.93	2.72	4.24	0.37
Merca41	312.0888	10.28	0.78	0.41	1.56
oxidized-Nonatrienal*	314.1065	9.81	1.01	0.86	0.67
reduced-4-oxo-2E-nonenal*	318.1371	10.93	1.10	1.11	1.27
HNE-MA	318.1376	13.05	4.13	11.05	0.50
4-hydroxy-nonenal*	318.1374	13.43	4.08	5.13	0.42
4-hydroxy-octenoic acid*	320.1168	7.15	1.25	1.04	1.17
DHN-MA	320.1532	12.84	2.50	3.62	0.59
Merca48	322.0722	13.61	0.95	0.94	1.27
oxidized-decatetraenal-1*	326.1062	9.63	0.92	0.82	0.92
oxidized-decatetraenal*	326.1040	16.55	0.83	0.71	1.11
dolichodial*	328.1220	17.88	0.72	1.00	1.10
4-hydroxy-decenal*	332.1529	17.38	0.94	0.72	0.80
w-COOH-4-hydroxy-octenoic acid lactone*	332.0798	5.02	1.65	1.35	0.75
w-OH-4-hydroxy-nonenic acid lactone*	332.1160	8.25	1.25	0.86	0.73
Merca55	333.1179	5.21	1.04	0.72	0.94
4-hydroxy-nonenic acid*	334.1328	9.27	1.49	1.17	0.89
10-oxo-5.8-decadienoic acid*	344.1160	10.24	0.84	0.37	1.08
w-OH-4-hydroxy-decenoic acid lactone-1*	346.1322	10.30	0.83	1.13	0.96
w-OH-4-hydroxy-decenoic acid lactone*	346.1317	10.64	0.92	1.13	0.77
4-hydroxy-decenoic acid*	348.1475	11.88	0.92	0.85	1.23
Merca61	348.0886	16.55	0.85	0.59	1.00
Merca62	350.1041	17.88	0.63	0.17	1.16
w-OH-4-hydroxy-decenol*	350.1625	10.56	7.23	15.90	0.58
w-COOH-4-hydroxy-decenal*	362.1288	11.40	0.78	0.85	1.10
Merca65	364.0979	4.45	1.06	1.09	1.06
Merca66	384.1081	11.41	0.75	0.68	0.93
Merca67	387.0079	7.61	0.71	0.22	0.94
Merca68	393.1517	12.24	0.99	0.98	1.11
Merca69	395.1630	12.44	29.25	1.16	1.57
Merca70	422.1634	17.69	1.17	1.12	10.53
Merca71	438.1560	14.57	2.26	2.79	5.34

Annexe 11 Mercapturates détectés dans l'expérimentation de carcinogénèse. Les valeurs en rouge et gras correspondent aux mercapturates significatifs (Wilcoxon + fdr, p-value 0.05) pour chaque comparaison.

## Annexe 12

Annotation	m/z	Temps de rétention (min)	Attribution	Fold BOFO PUFO	Fold BOSA PUSA	Fold BOSAR BOSA	Fold BOSA CON	Fold BOFOR BOFO	Fold Boeuf/Poulet
2-Methylglutaric acid	101.0605	4.59	[(M-CO2)-H]-	<b>0.28</b>	1.29	0.72	1.45	0.63	0.72
Resorcinol	109.0292	2.58	[M-H]-	0.75	1.30	0.97	0.53	1.11	0.81
3-Methyladipic acid	115.0761	5.55	[(M-CO2)-H]-	0.81	1.62	1.16	1.69	1.18	0.95
2,2-Dimethylglutaric acid	115.0759	7.28	[(M-CO2)-H]-	0.40	1.57	1.03	1.82	2.50	0.98
Indole-3-carboxylic acid	116.0501	8.16	[(M-H2O)-H]-	<b>0.49</b>	1.16	0.76	0.93	0.89	0.74
p-Coumaric acid	119.0498	6.87	[(M-CO2)-H]-	0.56	1.04	1.09	<b>3.40</b>	0.99	0.90
Indoleacetic acid	130.0658	10.00	[(M-CO2)-H]-	0.60	0.94	0.98	1.65	1.56	0.78
Pimelic acid	141.0552	5.68	[(M-H2O)-H]-	0.66	1.10	0.98	1.14	1.95	1.10
3-(4-Hydroxyphenyl)propionic acid	147.0449	6.87	[(M-H2O)-H]-	0.62	1.07	1.06	<b>3.78</b>	0.99	0.93
N-Acetyl-4-aminosalicylic acid	150.0557	5.81	[(M-CO2)-H]-	1.07	1.22	1.17	<b>0.40</b>	1.26	1.33
3-Methyl-2-oxovaleric acid	151.0399	2.99	[(M-H+Na)-H]-	1.03	1.62	1.52	0.48	2.14	1.46
Pimelic acid	159.0660	5.68	[M-H]-	0.68	1.29	1.23	1.10	1.98	1.29
22-Dimethylglutaric acid	159.0658	7.29	[M-H]-	<b>0.40</b>	1.57	0.99	2.02	1.34	0.93
Phenylpyruvic acid	163.0398	4.37	[M-H]-	0.51	0.92	1.06	1.10	1.05	0.78
4-Pyridoxic acid	164.0352	1.99	[(M-H2O)-H]-	<b>0.22</b>	0.64	0.98	<b>3.90</b>	1.11	<b>0.43</b>
3-(4-Hydroxyphenyl)propionic acid	165.0555	6.87	[M-H]-	0.59	1.06	1.11	<b>3.64</b>	1.06	0.93
N acetyl DL glutamic acid	170.0455	1.09	[(M-H2O)-H]-	<b>0.09</b>	<b>0.25</b>	0.80	0.90	2.23	<b>0.20</b>
N-Acetyl-L-leucine	172.0978	7.01	[M-H]-	<b>0.09</b>	<b>0.23</b>	1.40	1.24	2.61	<b>0.22</b>
Suberic acid	173.0816	8.67	[M-H]-	<b>0.29</b>	1.75	0.40	2.60	0.69	0.67
4-Methylhippuric acid	174.0559	8.76	[(M-H2O)-H]-	0.55	0.82	1.35	1.61	1.53	0.73
3-isopropylmalic acid	175.0608	4.38	[M-H]-	0.54	3.09	0.61	2.14	0.67	1.37
Indoleacrylic acid	186.0555	13.94	[M-H]-	<b>0.13</b>	0.43	1.50	1.67	<b>5.25</b>	<b>0.33</b>
N acetyl L glutamine	187.0723	1.01	[M-H]-	<b>0.47</b>	0.81	0.83	0.69	1.93	0.74
Azelaic acid	187.0976	11.38	[M-H]-	0.50	1.19	0.95	1.60	0.96	0.92
Pantothenic acid	200.0925	2.93	[(M-H2O)-H]-	<b>0.18</b>	<b>0.23</b>	1.04	0.81	0.98	<b>0.25</b>
Indolelactic acid	204.0662	7.36	[M-H]-	0.67	1.22	1.06	1.23	0.92	1.03
Cinnamoylglycine	204.0658	8.55	[M-H]-	1.39	2.69	1.22	<b>0.19</b>	1.39	1.87
N-Acetyl-L-phenylalanine	206.0816	7.45	[M-H]-	<b>0.47</b>	1.33	1.19	1.75	2.03	0.98
Jasmonic acid	209.1177	13.55	[M-H]-	<b>0.32</b>	1.15	0.84	1.03	2.04	1.01
undecanedioic acid	215.1284	15.57	[M-H]-	0.97	<b>60.45</b>	0.60	<b>16.70</b>	1.26	1.60
Sebacic acid	223.0955	13.93	[(M-H+Na)-H]-	0.69	1.14	1.15	0.88	1.38	1.02
N-Isovaleroylglycine	240.0865	5.00	[(M-H+Na)+CH3CO2]-	<b>0.19</b>	0.76	1.10	1.37	0.66	<b>0.42</b>
beta-Pseudouridine	243.0618	1.01	[M-H]-	0.76	0.66	0.78	<b>0.33</b>	2.01	<b>0.73</b>
N-acetyl-DL-tryptophan	245.0938	7.80	[M-H]-	<b>0.40</b>	0.71	1.33	1.55	1.45	0.63
Phenylacetyl-L-glutamine	263.1034	4.73	[M-H]-	<b>0.46</b>	0.75	0.98	1.71	1.26	<b>0.57</b>
10-hydroxydecanoic acid	269.1388	16.10	[(M-H+Na)+CH3CO2]-	0.71	<b>3.06</b>	0.86	1.97	1.07	0.88
Ferulic acid	275.0543	8.22	[(M-H+Na)+CH3CO2]-	0.85	<b>11.27</b>	0.31	4.12	0.60	0.99
Guanosine	304.0688	2.69	[(M-H+Na)-H]-	0.56	0.05	0.86	0.00	1.01	0.58
Riboflavin	375.1309	6.70	[M-H]-	0.61	1.06	1.14	1.06	1.05	0.89
2-Aminobenzoic acid	92.0502	5.73	[(M-CO2)-H]-	0.72	0.74	1.09	0.90	0.78	0.82

Annexe 12 Métabolites détectés et annotés au niveau 1 dans l'expérimentation de nutritionnelle. Les valeurs en rouge et gras correspondent aux métabolites significatifs (ViP > 1 + p-value < 0.05) pour chaque comparaison.

## Annexe 13

Annotation	m/z	Temps de rétention (min)	Attribution	Fold BOFO/PUFO	Fold BOSAR/BOSA	Fold BOSA/PUSA	Fold Boeuf/Poulet
3methyl-adipic	115.0760	5.98	[(M-CO2)-H]-	1.26	1.30	0.99	1.06
Succinic acid	117.0187	1.44	[M-H]-	<b>5.17</b>	<b>3.27</b>	<b>4.19</b>	<b>6.83</b>
p-Coumaric acid	119.0497	7.63	[(M-CO2)-H]-	<b>0.43</b>	0.96	0.59	0.46
3-(4-Hydroxyphenyl)propionic acid	121.0651	7.63	[(M-CO2)-H]-	<b>0.39</b>	1.33	0.44	<b>0.40</b>
4-Methylcatechol	123.0448	6.52	[M-H]-	0.46	1.15	1.01	0.61
L-Pyroglutamic acid	128.0348	1.18	[M-H]-	<b>0.63</b>	1.01	<b>0.18</b>	<b>0.25</b>
3-Methyl-2-oxovaleric acid	129.0554	3.79	[M-H]-	0.88	1.11	0.63	0.72
2-Ketohexanoic acid	129.0551	4.24	[M-H]-	0.89	0.74	0.87	0.81
Suberic acid	129.0918	8.97	[(M-CO2)-H]-	0.94	0.92	1.02	0.98
D-2-Hydroxyglutaric acid	129.0187	1.27	[(M-H2O)-H]-	1.62	1.34	0.99	1.27
Ehtylmalonic acid	131.0344	2.75	[M-H]-	1.12	1.00	1.15	1.12
Indoxyl sulfate	132.0452	3.71	[(M-SO3)-H]-	0.97	1.54	0.90	1.07
Homovanillic acid	137.0603	5.86	[(M-CO2)-H]-	1.05	1.06	1.04	1.11
4-Pyridoxic acid	138.0554	1.86	[(M-CO2)-H]-	<b>0.58</b>	1.14	<b>0.51</b>	<b>0.55</b>
4-Acetamidobutanoic acid	144.0658	1.83	[M-H]-	0.78	0.96	0.70	0.72
N-Isobutyrylglycine	144.0660	2.48	[M-H]-	1.21	1.19	1.18	1.25
Adipic acid	145.0503	3.80	[M-H]-	1.36	0.88	1.50	1.31
Monomethyl glutaric acid	145.0502	4.87	[M-H]-	1.05	0.75	0.96	0.91
a-Ketoglutaric acid	145.0137	1.08	[M-H]-	<b>1.64</b>	0.89	<b>1.88</b>	<b>1.73</b>
D-2-Hydroxyglutaric acid	147.0292	0.98	[M-H]-	1.34	0.94	1.66	<b>1.49</b>
DL-Mandelic acid	151.0398	4.88	[M-H]-	0.88	1.02	0.84	0.86
Imidazolelactic acid	155.0458	0.94	[M-H]-	<b>0.45</b>	<b>0.55</b>	<b>0.42</b>	<b>0.39</b>
Indole-3-carboxylic acid	160.0401	8.43	[M-H]-	1.06	0.94	0.99	1.00
3-Hydroxy-3-methylglutaric acid	161.0448	1.56	[M-H]-	1.26	1.01	1.10	1.16
Phenylalanine	164.0710	2.49	[M-H]-	1.03	1.34	0.96	1.06
3-(4-Hydroxyphenyl)propionic acid	165.0551	7.23	[M-H]-	0.90	1.05	1.20	1.11
Phenyllactic acid	165.0548	7.63	[M-H]-	<b>0.40</b>	1.23	<b>0.44</b>	<b>0.40</b>
3-Methylhistidine	168.0772	0.92	[M-H]-	<b>0.17</b>	1.06	<b>0.17</b>	<b>0.17</b>
Azelaic acid	169.0864	11.65	[(M-H2O)-H]-	1.20	0.99	0.97	1.05
Azelaic acid	169.0864	12.42	[(M-H2O)-H]-	1.31	0.75	1.41	1.25
N-Acetyl-L-glutamic acid	170.0451	1.17	[(M-H2O)-H]-	<b>0.24</b>	1.56	<b>0.10</b>	<b>0.15</b>
N-Acetyl-D-allo-isoleucine	172.0976	6.90	[M-H]-	0.75	1.61	0.49	0.67
N-Acetyl-D-allo-isoleucine	172.0973	7.37	[M-H]-	<b>0.64</b>	1.44	<b>0.41</b>	<b>0.54</b>
Hexanoyl glycine	172.0976	7.98	[M-H]-	1.44	1.23	1.21	<b>1.44</b>
Indoleacetic acid	174.0559	9.06	[M-H]-	<b>1.66</b>	1.32	1.52	<b>1.68</b>
L-Ascorbic acid 2-sulfate	175.0241	1.07	[(M-SO3)-H]-	1.14	1.05	1.22	1.21
Indoleacrylic acid	186.0553	14.18	[M-H]-	<b>0.58</b>	1.50	<b>0.48</b>	<b>0.57</b>
N acetyl L glutamine	187.0717	1.08	[M-H]-	1.05	1.54	0.71	0.92
10-hydroxydecanoic acid	187.1332	14.37	[M-H]-	0.80	1.05	0.84	0.85
N-Acetyl-L-methionine	190.0535	3.89	[M-H]-	0.65	1.64	0.45	0.59

Annexes

Phenylacetylglucine	192.0661	5.59	[M-H]-	0.98	1.12	1.01	1.03
3-Methylhippuric acid	192.0659	7.67	[M-H]-	1.08	0.70	1.18	0.97
Sebacic acid	201.1127	14.20	[M-H]-	1.14	1.31	0.84	1.02
N-Acetyl-L-phenylalanine	206.0815	7.81	[M-H]-	1.21	1.27	1.19	1.28
Jasmonic acid	209.1174	14.21	[M-H]-	0.88	1.29	0.89	1.13
Pantothenic acid	218.1028	3.14	[M-H]-	<b>0.59</b>	0.95	0.65	<b>0.62</b>
N-Acetyl-L-tyrosine	222.0762	3.80	[M-H]-	<b>2.00</b>	1.18	1.68	<b>1.88</b>
N-acetyl-DL-tryptophan	245.0946	8.11	[M-H]-	0.80	0.95	0.76	0.76
3-(4-Hydroxyphenyl)propionic acid	247.0575	7.24	[(M-H+Na)+CH <sub>3</sub> CO <sub>2</sub> ]-	0.76	0.97	1.00	0.92
L-Ascorbic acid 2-sulfate	254.9814	1.07	[M-H]-	1.13	1.00	1.20	1.17
Phenylacetyl-L-glutamine	263.1030	5.09	[M-H]-	0.70	1.24	0.69	0.74
4-Hydroxy-3-methoxyphenylglycol sulfate	263.0226	1.47	[M-H]-	1.05	1.05	1.03	1.06
Inosine	267.0728	1.68	[M-H]-	1.16	0.91	0.99	1.02
N-Acetyl-L-carnosine	267.1090	1.04	[M-H]-	<b>0.60</b>	1.23	0.63	0.66
N4-Acetylcytidine	284.0880	2.65	[M-H]-	0.80	0.92	0.84	0.81
Riboflavin	375.1305	6.97	[M-H]-	0.75	1.28	0.72	0.79
3-Methyl-2-oxovaleric acid	85.0657	3.90	[(M-CO <sub>2</sub> )-H]-	1.33	1.26	1.09	1.26
2-Ketohexanoic acid	85.0656	4.56	[(M-CO <sub>2</sub> )-H]-	0.89	1.23	0.83	0.92
Alanine	88.0401	0.92	[M-H]-	<b>2.87</b>	1.34	<b>2.51</b>	<b>2.83</b>

Annexe 13 Métabolites détectés et annotés au niveau 1 dans l'expérimentation de carcinogénèse. Les valeurs en rouge et gras correspondent aux métabolites significatifs (ViP > 1 + p-value < 0.05) pour chaque comparaison.

**Annexe 14**

<i>m/z</i> exp	AUC cryptes BOSA vs PUSA	Annotation	Formule brute	Adduit	Erreur de mesure (ppm)
134.048	0.28	Adenine	C5H5N5	[M-H] <sup>-</sup>	2.8
135.031	0.68	Hypoxanthine	C5H4N4O	[M-H] <sup>-</sup>	1.2
151.025	0.72	Xanthine	C5H4N4O2	[M-H] <sup>-</sup>	8.9
267.072	0.76	Inosine	C10H12N4O5	[M-H] <sup>-</sup>	5.6
280.022	0.23	Glucosamine 6-phosphate	C6H14NO8P	[M+Na-2H] <sup>-</sup>	5.8
328.042	0.27	Adenosine monophosphate	C10H14N5O7P	[M-H2O-H] <sup>-</sup>	8.7
344.035	0.3	Glutathion	C10H17N3O6S	[M+K-2H] <sup>-</sup>	6.9
362.054	0.29	Guanosine monophosphate	C10H14N5O8P	[M-H] <sup>-</sup>	8.2
402.992	0.29	Uridine diphosphate	C9H14N2O12P2	[M-H] <sup>-</sup>	6.5
418.957	0.34	Inositol 1,4,5 triphosphate	C6H15O15P3	[M-H] <sup>-</sup>	3.6
426.023	0.25	Adenosine diphosphate	C10H15N5O10P2	[M-H] <sup>-</sup>	1.1
442.014	0.34	Guanosine diphosphate	C10H15N5O11P2	[M-H] <sup>-</sup>	6.2
465.307	0.66	Cholesterol sulfate	C27H46O4S	[M-H] <sup>-</sup>	4.5
482.964	0.32	Uridine triphosphate	C9H15N2O15P3	[M-H] <sup>-</sup>	5.7
505.99	0.31	Adenosine triphosphate	C10H16N5O13P3	[M-H] <sup>-</sup>	3
604.071	0.31	GDP-mannose	C16H25N5O16P2	[M-H] <sup>-</sup>	1.7
611.15	0.35	Oxidized glutathione	C20H32N6O12S2	[M-H] <sup>-</sup>	7.9
322.046	0.35	Cytidine monophosphate	C9H14N3O8P	[M-H] <sup>-</sup>	2.9
664.122	0.73	NADH	C21H29N7O14P2	[M-H] <sup>-</sup>	6.9
465.307	0.34	Cholesterol sulfate	C27H46O4S	[M-H] <sup>-</sup>	4.5
170.093	0.32	3-methyl-histidine	C7H11N3O2	[M+H] <sup>+</sup>	5.3
146.118	0.67	Acetylcholine	C7H16NO2	[M+H] <sup>+</sup>	0.4
524.373	0.75	LPC(18:0)	C26H54NO7P	[M+H] <sup>+</sup>	3.5
796.529	0.29	PC(34:2)	C42H80NO8P	[M+K] <sup>+</sup>	4.1

**Annexe 14 Métabolites discriminants entre les groupes BOSA et PUSA suite à l'analyse de l'épithélium colique par imagerie par spectrométrie de masse. Toutes les annotations correspondent à des annotations putatives au niveau 2. Une valeur d'AUC < 0.35 signifie une diminution de l'abondance dans le groupe BOSA.**

## Annexe 15

### Principe et fonctionnement du mode PASEF

Le mode PASEF, pour *Parallel Accumulation Serial Fragmentation*, est une méthode de fragmentation haut débit (> 100 Hz) disponible depuis 2017 sur les spectromètres de masse de types TimsTOF proposés par Bruker. Elle repose sur la synchronisation entre une cellule de mobilité de type TIMS (Trapped Ion Mobility Spectrometry) et une cellule de collision.

La mobilité ionique consiste en la séparation et la caractérisation d'ions en phase gazeuse en fonction de leur comportement lors de leur transport à travers un gaz inerte en présence d'un champ électrique. Classiquement le fonctionnement des cellules de mobilité ionique est basé sur l'application d'un courant électrique qui guide les ions au travers d'un flux de gaz à contre courant. Les ions présentant la plus grande résistance au flux de gaz (CCS les plus élevées) sortent ainsi plus tard de la cellule. A l'inverse, dans une cellule de mobilité ionique de type TIMS, les ions sont entraînés par un flux de gaz (Michelmann *et al.*, 2015). Une rampe de champ électrique va être appliquée le long du funnel pour permettre de contrecarrer la force exercée par le flux. L'abaissement de la force du champ électrique appliqué conduit ainsi à la libération séquentielle des ions en fonction de leur mobilité, les ions les plus mobiles sont alors entraînés en premier. Une représentation schématique du fonctionnement d'une cellule de mobilité classique et d'une cellule TIMS est présentée en Figure 48.

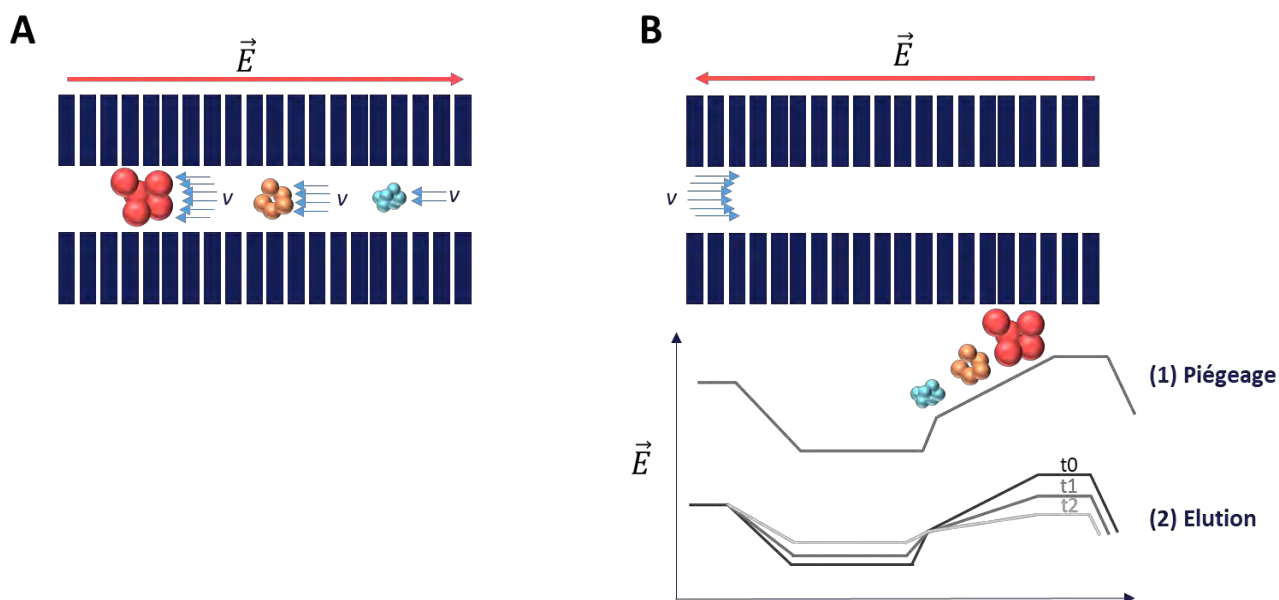


Figure 48 Représentation schématique du fonctionnement de (A) une cellule de mobilité classique et (B) une cellule TIMS.  $\vec{E}$  correspond au champ électrique appliqué et  $v$  à la vitesse du gaz.

Ainsi, dans le cas du TimsTOF flex, le mode PASEF repose sur la synchronisation entre la cellule TIMS (constituée de deux funnels), le quadrupole et la cellule de collision située avant le détecteur ToF comme présenté en Figure 49.





Figure 49 Schéma de l'architecture du spectromètre de masse TIMS ToF

Dans le premier *funnel*, les ions sont accumulés et piégés par l'application d'un champ électrique. Ces ions sont ensuite transférés dans le deuxième *funnel* où le potentiel électrique est graduellement abaissé afin d'éluer les ions de la cellule de mobilité. Pendant ce temps d'élution, les ions suivants sont accumulés dans le *funnel 1* (*duty cycle*). La première étape du cycle PASEF consiste à réaliser un scan Full-MS pour détecter les ions présents. Ensuite, durant chaque scan de la cellule Tims, chaque précurseur est sélectionné dans le quadrupole puis fragmenté dans la cellule de collision. Le nombre de cycles MS/MS et d'expériences MS/MS par cycle avant le scan MS suivant sont déterminés par l'utilisateur. Le schéma général d'un cycle PASEF est présenté en Figure 50. La vitesse d'acquisition élevée de l'appareil (>100 Hz) permet la fragmentation et l'analyse de nombreux ions en un temps réduit. Ainsi durant chaque scan de la cellule TIMS, dont la durée est définie par l'expérimentateur (à l'aide du paramètre *ramp time* et en général de l'ordre de 100 ms) un nombre *n* de précurseurs peut être fragmenté. Dans une étude récente visant à évaluer l'utilisation de ce mode pour des analyses lipidomiques une moyenne de 15 MS/MS par scan TIMS de 100 ms a été reportée, résultant en l'acquisition de plus de 180 000 spectres MS/MS pour une analyse de 30 min (Vasilopoulou *et al.*, 2020).

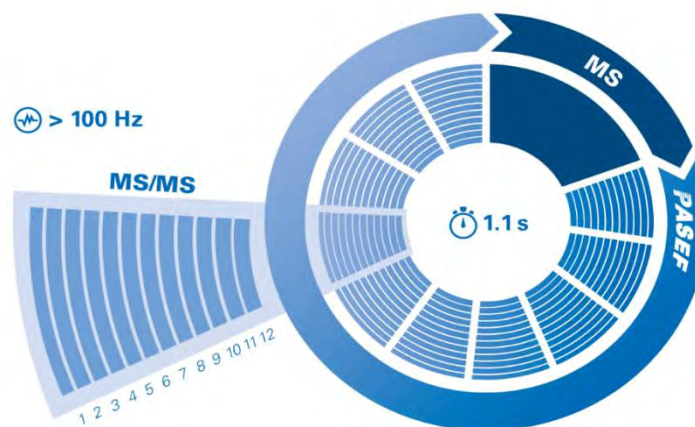


Figure 50 Schéma général d'une acquisition en mode PASEF (Bruker©)



## Article

# Towards Aldehydomics: Untargeted Trapping and Analysis of Reactive Diet-Related Carbonyl Compounds Formed in the Intestinal Lumen

Sylvie Chevolleau <sup>1,2,\*</sup>, Maria-Helena Nogueira-Meireles <sup>1,2</sup>, Loïc Mervant <sup>1,2</sup>, Jean-François Martin <sup>1,2</sup>, Isabelle Jouanin <sup>1,2</sup>, Fabrice Pierre <sup>1</sup>, Nathalie Naud <sup>1</sup>, Françoise Guéraud <sup>1,†</sup> and Laurent Debrauwer <sup>1,2,†</sup>

<sup>1</sup> Toxalim, Research Centre in Food Toxicology, Université de Toulouse, INRAE, ENVT, INP-Purpan, UPS, CEDEX 3, 31027 Toulouse, France; mariahelena.nogueira@gmail.com (M.-H.N.-M.); loic.mervant@inrae.fr (L.M.); jean-francois.martin@inrae.fr (J.-F.M.); isabelle.jouanin@inrae.fr (I.J.); fabrice.pierre@inrae.fr (F.P.); nathalie.naud.31@inrae.fr (N.N.); francoise.gueraud@inrae.fr (F.G.); laurent.debrauwer@inrae.fr (L.D.)

<sup>2</sup> Metatoul-AXIOM Platform, National Infrastructure for Metabolomics and Fluxomics, MetaboHUB, CEDEX 3, 31027 Toulouse, France

\* Correspondence: sylvie.chevolleau@inrae.fr

† These authors contributed equally to this work.



**Citation:** Chevolleau, S.; Nogueira-Meireles, M.-H.; Mervant, L.; Martin, J.-F.; Jouanin, I.; Pierre, F.; Naud, N.; Guéraud, F.; Debrauwer, L. Towards Aldehydomics: Untargeted Trapping and Analysis of Reactive Diet-Related Carbonyl Compounds Formed in the Intestinal Lumen. *Antioxidants* **2021**, *10*, 1261. <https://doi.org/10.3390/antiox10081261>

Academic Editors: Nikolaus Bresgen and Werner Siems

Received: 11 June 2021

Accepted: 3 August 2021

Published: 6 August 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Abstract:** Lipid peroxidation and subsequent formation of toxic aldehydes, such as 4-hydroxynonenal, is known to be involved in numerous pathophysiological processes, possibly including the development of colorectal cancer. This work aimed at the development of an untargeted approach using high-performance liquid chromatography coupled with high-resolution mass spectrometry (HPLC–HRMS) for tracking aldehydes in both suspect screening and untargeted methods in fecal water, representing the aqueous environment of colon epithelial cells. This original approach is based on the introduction of a characteristic isotopic labeling by selective derivatization of the carbonyl function using a brominated reagent. Following a metabolomics workflow, the developed methodology was applied to the characterization of aldehyde compounds formed by lipid peroxidation in rats fed two different diets differentially prone to lipoperoxidation. Derivatized aldehydes were first selectively detected on the basis of their isotopic pattern, then annotated and finally identified by tandem mass spectrometry. This original approach allowed us to evidence the occurrence of expected aldehydes according to their fatty acid precursors in the diet, and to characterize other aldehydes differentiating the different diets.

**Keywords:** metabolomic; untargeted trapping; aldehydes; 4-hydroxy-nonenal; 4-hydroxy-hexenal; malondialdehyde; HPLC–HRMS; isotopic labelling; fecal water; lipid peroxidation

## 1. Introduction

Recent epidemiological studies have shown that diets rich in red or processed meat are associated with an elevated risk of colorectal cancer (CRC), whereas those rich in fish and white meat are not [1–3]. As of its 2007 report [4], and more recently in its 2018 Continuous Update Project [5], the World Cancer Research Fund (WCRF) panel recommended limiting red meat and avoiding processed meat consumption to limit the risk of colorectal cancer. One of the major hypotheses of this association involves dietary heme iron, the concentration of which is much higher in red meat than in white meat, and which could promote CRC by inducing or enhancing luminal dietary lipid oxidation [6,7].

The oxidation of dietary polyunsaturated fatty acids (PUFAs) is a non-enzymatic free-radical-mediated peroxidation induced by, for instance, ions of metals such as iron. It occurs in three steps: (1) the initiation step, in which the abstraction of a hydrogen atom from the fatty acid chain gives rise to a lipid radical; (2) the propagation step, in which the lipid radical reacts with dioxygen to produce a lipoperoxyl radical that, in turn,

reacts with another lipid chain; and (3) the termination step, in which radical species combine to produce non-propagating species. All of these lipid intermediates can undergo  $\beta$ -scission and give rise to the reactive formation of secondary lipid oxidation products—such as aldehydes, and in particular 4-hydroxy-alk-2,3-enals—which display electrophilic features, making them reactive towards nucleophilic biomolecules—especially proteins, but also lipids and DNA [8,9]. Indeed, several secondary lipoperoxidation products, including these aldehydes, can be formed not only in the diet (particularly in heme-iron- and polyunsaturated-fat-rich foods), but also during digestion in the intestinal lumen, with pro-oxidative conditions favoring their production [10].

Aldehydes, and in particular hydroxyl-alkenals such as (2E)-4-hydroxy-non-2-enal (HNE) and (2E)-4-hydroxy-hex-2-enal (HHE)—formed from  $\omega$ 6- (e.g., linoleic and arachidonic acid) and  $\omega$ 3- (e.g.,  $\alpha$ -linolenic acid) PUFAs, respectively—are known to be cytotoxic and genotoxic. When present in the bioavailable fraction of feces (fecal water), such aldehydes may play a promoting role in the development of CRC, by selecting precancerous colon epithelial cells [6,11].

The detection and quantification of such species in the intestinal lumen is thus of particular interest. Due to their volatility (especially for short-chain aldehydes) and their reactive properties (e.g., binding, decomposition), the analysis of aldehydes remains challenging. Several analytical approaches have been developed for the sensitive detection and characterization of aldehydes as lipid peroxidation products. As a general feature, due to their physicochemical properties, a derivatization step is required for their gas chromatography (GC) analysis. This is also the case for their liquid chromatography (LC) analysis, since they display poor UV or fluorometric responses as well as poor MS ionization yields, and in most cases they first need to be stabilized. For aldehyde derivatization, several reactants have been used, mainly involving hydrazine and hydroxylamine chemistry. The most popular methods use pentafluorobenzyl hydroxylamine (PFBHA) for GC analysis [12] and dinitrophenylhydrazine (DNPH) for LC analysis [13]. However, the selectivity of the DNPH reagent towards carbonyl compounds has been questioned [14].

Mass spectrometry coupled with GC or LC is now the most popular technique for the selective and sensitive screening and quantification of molecules in various matrices, particularly with the use of triple-quadrupole mass spectrometers operating in the multiple reaction monitoring (MRM) mode. This technique currently provides the most sensitive detection method when applied to known targeted compounds, and has been extensively used for the analysis of several oxidation products—such as aldehydes and ketones—in various matrices [15–19]. However, depending on the chain length and the unsaturation degree and position of the fatty acids, lipid peroxidation is known to give rise to numerous and various carbonyl compounds. Although specific to their precursor fatty acids, these compounds are still not fully identified, and current targeted detection methods may not provide complete coverage of aldehyde formation as a result of lipid peroxidation in the intestinal lumen. In order to get the most complete coverage, the use of full-scan untargeted or semi-targeted metabolomics approaches performed in high-resolution mass spectrometry (HRMS) represents one of the most promising alternative solutions. Although the HRMS acquisition mode remains less sensitive than the MRM mode, time-of-flight (TOF) or Orbitrap mass analyzers are now more and more frequently used for screening and biomonitoring applications [20–22].

In the context of aldehyde profiling, the derivatization strategy should aim at (1) incorporating chemical groups that allow stabilization of the aldehyde function and enhancement of the ionization yields in order to improve sensitivity, and (2) incorporating stable isotopes to work on derivatives displaying characteristic isotopic patterns, thus allowing the untargeted but specific screening of derivatives by means of signal filtering. Benzenaminium halides proposed by Eggink et al. fulfill the first criterion, and can even fulfill both when considering the corresponding isotopically labelled reagents introduced either via a halogen atom or a mixture of “normal” and “heavy” reagents [23–26]. Hydrazinotriazines were also proposed as derivatization reagents for fatty aldehydes [27],

and could also be used for profiling by monitoring specific neutral losses using deuterated hydrazinotriazines. More recently, another method has been proposed for isotope-coded derivatization of aldehydes using  $^{14}\text{N}/^{15}\text{N}$ -ammonium acetate and 9,10-phenanthrene quinone [28].

In this work, we present the development of an “aldehydomics” workflow based on the untargeted aldehyde characterization method in fecal water, using aldehyde stabilization by quantitative in situ derivatization under mild conditions before LC–HRMS analysis. In this approach, we selectively labeled aldehydes with a bromine atom to take advantage of the specific detection of its characteristic  $^{79}\text{Br}/^{81}\text{Br}$  isotopic pattern, as already used for pharmaceutical compounds or chemical contaminants [29,30]. For this, we propose here the use of the brominated derivatizing reagent 1-((ammoniooxy)methyl)-2-bromobenzene chloride (BBHA), which has already proven to be well suited for the targeted analysis of HNE and HHE [31]. To validate our proof of concept, this approach was then applied to the tracking of free aldehydes in fecal waters prepared from rats fed two diets containing saturated (namely, coconut oil (COHE)) or unsaturated (namely, linseed oil (LIHE)) fatty acids with added pro-oxidant heme iron. Aldehydes whose formation could be expected according to the diet could be characterized, and several other brominated signals corresponding to potentially unknown aldehydes were also evidenced.

## 2. Materials and Methods

### 2.1. Chemicals

Methanol (HPLC grade/Optima LC/MS grade) and acetonitrile (Optima LC/MS grade) were purchased from Fisher (Illkirch, France), while formic acid was purchased from Sigma-Aldrich (Saint-Quentin-Fallavier, France). Ultrapure water was obtained using a Milli-Q system (Millipore, Saint-Quentin-en-Yvelines, France).

4-(2-((4-Bromophenethyl)dimethylammonio)ethoxy) benzenaminium dibromide (APEBA) was prepared and used as described by Eggink et al. [25]. 1-((Ammoniooxy)methyl)-2-bromobenzene chloride (BBHA) was purchased from Interchim (Montluçon, France). Piperazine- $\text{N},\text{N}'$ -bis(2-ethanesulfonic acid) (PIPES) and trifluoroacetic acid (TFA) were purchased from Acros Organics (Geel, Belgium).

Standard HHE, HNE, 4-hydroxynonanal (OHN), and  $\omega$ -oxo-carboxylic acids (5-oxo-pentanoic acid (5-OPA), 6-oxo-hexanoic acid (6-OHA), 7-oxo-heptanoic acid (7-OHA), 8-oxo-octanoic acid (8-OOA), 9-oxo-nonanoic acid (9-ONA), 10-oxo-decanoic acid (10-ODA)) were synthesized in house according to previously published methods [32,33]. All other standard aldehydes and other standard compounds were purchased from Sigma-Aldrich (Saint-Quentin-Fallavier, France), except for 9-oxo-10,12-octadecadienoic acid (9-OxoODE), 9-Oxo-10,12,15-octadecatrienoic acid (9-OxoOTrE), and  $\alpha$ -linolenic and eicosapentaenoic acid which were purchased from Bertin Bioreagent (Montigny-le-Bretonneux, France).

### 2.2. Animal Experiment and Fecal Water Preparation

Animal experiments were carried out with seven-week-old male Fisher 344 rats (6 rats/group) purchased from Charles River Laboratories (Les Oncins, France). Animal care was in accordance with our local ethics committee (number TOXCOM/006/FG). Rats were housed in individual plastic metabolic cages equipped for fecal collection. Their diets were based on a modified AIN-76 diet prepared and formulated in a powdered form by UPAE (Jouy-en-Josas, France). Each diet was formulated with 5% oil—either hydrogenated coconut oil (COHE, MP Biomedicals, Solon, OH, USA) or raw linseed oil (LIHE, MP Biomedicals, Solon, OH, USA)—and 0.94 g of hemin (Sigma-Aldrich, Saint-Quentin-Fallavier, France) per kg of diet. To minimize diet lipid peroxidation during storage, oils and hemin were added to the powdered diet just before the start of the experiments. After preparation, each diet was aliquoted into individual portions for each experimental day, and stored at  $-20\text{ }^{\circ}\text{C}$  until use. Rats were fed their respective experimental diet ad libitum, at the end of each day, just before their active period. Diet consumption was recorded every two days. Feces were collected over 24 h on day 13

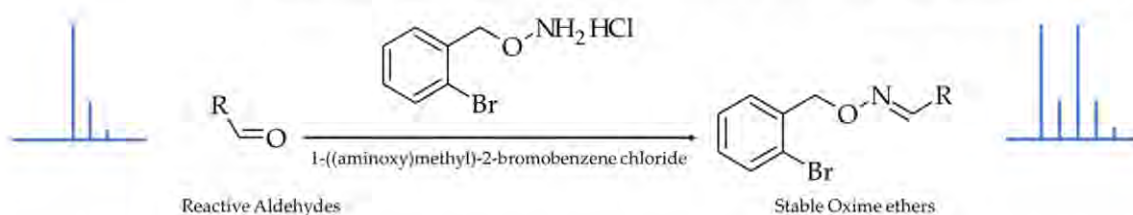
of the experimental diets. Fecal water was prepared by adding 0.5 g of feces to 1 mL of distilled water and 50  $\mu$ L of 0.45 M BHT in ethanol (2,6-Di-tert-butyl-4-methyl-phenol). The resulting samples were then processed using FastPrep technology and then centrifuged at  $3200\times g$  for 20 min. After centrifugation, supernatants were sampled and immediately submitted to derivatization for aldehyde stabilization.

### 2.3. Sample Treatment

Derivatization with APEBA was carried out as described by Eggink et al. [25]. In brief, fecal water was derivatized with APEBA at pH 5.7 ( $\text{NH}_4\text{OAc}$  150 mM), in the presence of  $\text{NaBH}_3\text{CN}$ , for 3 h at 8  $^\circ\text{C}$ . Aldehyde-APEBA derivatives were then extracted with dichloromethane, evaporated to dryness, and reconstituted in a  $\text{CH}_3\text{OH-H}_2\text{O}$  (1:1) mixture for LC-HRMS analysis.

For BBHA derivatization, reaction conditions were investigated in a previous work [33] on several aldehydes and ketones. Special attention was given to testing the reactivity of BBHA towards carbonyl compounds at low temperatures and a slightly acidic pH to preserve the integrity of the aldehyde function. According to these adjusted conditions, fecal water samples (100  $\mu$ L) were added to 60  $\mu$ L of PIPES buffer (0.1 M, pH 6.5) and derivatized with 120  $\mu$ L of BBHA (50 nmol/ $\mu$ L) in PIPES buffer. The samples were stirred at 6–8  $^\circ\text{C}$  for one hour.

The derivatization reaction and the chemical structures of the BBHA derivatives of HNE and HHE—the two most representative hydroxy-alkenals—are shown in Figure 1.



**Figure 1.** Derivatization reaction and chemical structures of the resulting oxime ethers, showing the introduction of the characteristic brominated isotopic pattern.

SPE purification of the derivatized extracts was conducted on a Visiprep SPE Vacuum Manifold (Supelco, Saint-Quentin-Fallavier, France), using Agilent Bond Elut C18 (100 mg, 1 mL) cartridges. The sorbent was conditioned with 1 mL of  $\text{CH}_3\text{OH}$ , then 1 mL of water. The derivatized samples were vortexed and then deposited on the SPE cartridge. Washing was first performed with 1 mL of PIPES (rinsing the container and depositing on the cartridge), and then with 0.05% TFA/ $\text{CH}_3\text{OH}$  (3/2). The cartridge was then dried under vacuum (1 min), eluted with 400  $\mu$ L of  $\text{CH}_3\text{OH}$ , and collected in glass tubes. Finally, the volume was adjusted to exactly 400  $\mu$ L with  $\text{CH}_3\text{OH}$ , and the extracts were stored at  $-20^\circ\text{C}$  until analysis.

### 2.4. Liquid Chromatography–Mass Spectrometry

Sample extracts were analyzed via high-performance liquid chromatography coupled with high-resolution mass spectrometry (HPLC-HRMS). The HPLC system consisted of an Ultimate 3000 RS pump fitted with the Ultimate 3000 autosampler (Dionex, Thermo Scientific, Les Ulis, France). The chromatographic pump was operated at a flow rate of 0.200 mL/min with the following elution gradient program—0 min: 0% of B; from 3 min to 15 min: 100% of B; and from 15 to 25 min: 100% of B. Mobile phases were composed of (A)  $\text{H}_2\text{O/CH}_3\text{CN/HCOOH}$  95/5/0.1 (v/v/v) and (B)  $\text{CH}_3\text{CN/H}_2\text{O/HCOOH}$  95/5/0.1 (v/v/v). Five microliters of sample was injected into a Kinetex C18 (150  $\times$  2.1 mm, 2.6  $\mu\text{m}$ ) column (Phenomenex, Le Pecq, France) maintained at 40  $^\circ\text{C}$ . Detection was achieved on an LTQ-Orbitrap XL hybrid mass spectrometer (Thermo Scientific, Les Ulis, France) equipped with an electrospray ionization source used in the positive mode. Ionization

parameters were set at +4.5 kV for the spray voltage, 35 arbitrary units (au) for the sheath gas flow rate ( $N_2$ ), 5 au for the auxiliary gas flow rate ( $N_2$ ), and 300 °C for the capillary temperature. High-resolution mass spectra were acquired at a resolution power of 30,000 from  $m/z$  80 to 800 in centroid mode. Identifications were carried out via tandem mass spectrometry experiments (MS/MS) using the ion trap mass analyzer of the LTQ-Orbitrap mass spectrometer (NCE = 25%). Solutions of synthesized standard HHE-BBHA and HNE-BBHA at different concentrations were used to characterize the method in terms of linearity of response, repeatability, and sensitivity.

### 2.5. Data Processing and Statistical Treatment

Features corresponding to brominated compounds were extracted from raw data using XCMS software. The centWave algorithm was used with a scan-to-scan deviation value set to 10 ppm. The chromatographic peak width was set between 10 and 40 s. For the alignment we used an  $m/z$  range of 0.01 (mzwid) and a retention window of 20 s (bandwidth). Based on the HRMS signal of the exact mass of each  $[M+H]^+$  ion, features were extracted according to a mass measurement error of  $\pm 5$  ppm, and to the occurrence of two signals of equivalent intensities (ratio of intensities between 0.667 and 1.333), with  $\Delta M = 1.998$  ( $\pm 0.001$ ) corresponding to the mass difference between the two bromine isotopes [34]. Following their detection, structural characterization of potential aldehydes was carried out by processing  $MS^n$  spectra using Xcalibur Qual Browser (Thermo Scientific, Les Ulis, France).

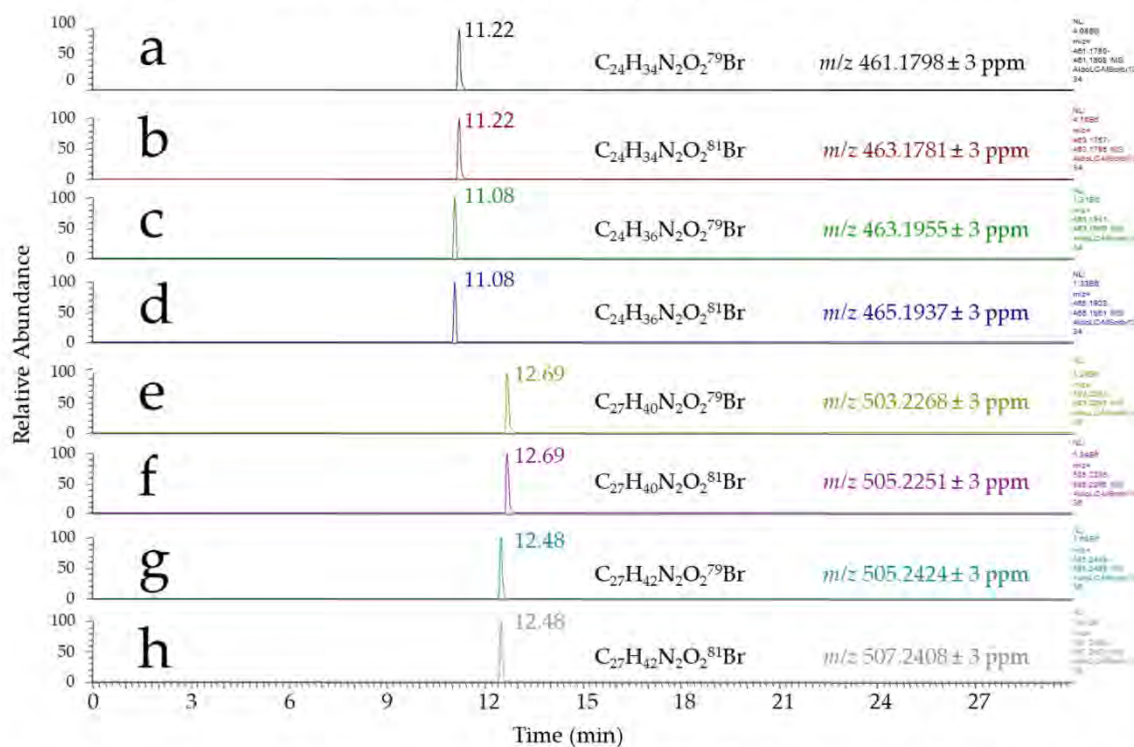
In order to target aldehydes of interest, we used statistical treatment to discover a list of features differentiating the studied diets. Both multivariate unsupervised principal component analysis (PCA) and supervised partial least squares discriminant analysis (PLS-DA) were used. PLS-DA aims to find the most discriminating features using the variable importance in projection (VIP) criterion. A VIP greater than 1 gives a first list of potential discriminant features. A univariate, non-parametric Kruskal–Wallis test was eventually carried out, and features with a Kruskal–Wallis  $p$ -value lower than 0.05 define the list of potential aldehyde ions.

## 3. Results and Discussion

### 3.1. Set-Up of the Derivatization Method

Based on the work of Eggink et al., we first investigated 4-(2-((4-bromophenethyl)dime-thylammonio)ethoxy) benzenaminium dibromide (APEBA) for derivatizing aldehydes in fecal water. This reactant was designed for a selective derivatization of aldehydes under reductive conditions using  $NaBH_3CN$ , and has been successfully applied to the detection of saturated aldehydes (i.e., pentanal, hexanal, heptanal, octanal, nonanal, and decanal) [25]. Using APEBA, the expected responses for saturated aldehydes were observed at nominal  $m/z$  of, for example, 461/463, 475/477, and 489/491 for heptanal, octanal, and nonanal, respectively, in agreement with previously reported results. Indeed, since those aldehydes represent lipid peroxidation products of low toxicological concern, we focused our attention on more reactive species, among which HHE and HNE were considered as reference compounds for the setup of our method. Thus, APEBA derivatives of standard HHE and HNE were synthesized and analyzed by LC–HRMS. Figure 2 shows the chromatograms obtained from the injection of standard HHE and HNE derivatized using APEBA as described by Eggink et al. [25]. Along with the expected HHE-APEBA and HNE-APEBA derivatives observed at  $m/z$  461.1798/463.1781 and 503.2268/505.2251, respectively (Figure 2a,b,e,f), the occurrence of additional signals corresponding to brominated compounds was also noticed, most likely arising from aldehyde-APEBA derivatives, at  $m/z$  463.1955/465.1937 and 505.2424/507.2408, respectively (Figure 2c,d,g,h). These  $m/z$  ratios corresponded to the chemical formulae  $C_{24}H_{36}N_2O_2Br$  and  $C_{27}H_{42}N_2O_2Br$ , i.e., two additional hydrogen atoms with respect to HHE and HNE derivatives, respectively. These signals represented ~10–20% of the signals measured for HHE-APEBA and HNE-APEBA, and very likely resulted from the partial reduction of the unsaturated HNE or HHE aldehydes—although the chemical

explanation for this reduction remains unclear, since  $\text{NaBH}_3\text{CN}$  is expected to selectively reduce the imine function formed by the reaction of the aldehyde with APEBA.



**Figure 2.** LC-HRMS-extracted chromatograms: (a,b) ion chromatograms of the  $^{79}\text{Br}$  and  $^{81}\text{Br}$  isotopomers of HHE-APEBA; (e,f) ion chromatograms of the  $^{79}\text{Br}$  and  $^{81}\text{Br}$  isotopomers of HNE-APEBA; and (c,d,g,h) ion chromatograms of their corresponding reduction byproducts.

These kinds of byproducts were also observed when derivatizing standard 4-oxo-hex-2-enal (OHE) and 4-oxo-non-2-enal (ONE), for which signals corresponding to HHE-APEBA and HNE-APEBA were present (data not shown), suggesting that the 4-oxo group was reduced to 4-hydroxyl. Note that unsaturated aldehydes such as 2-heptenal and 2-nonenal were also tested under the same conditions, and produced APEBA derivatives of both heptenal and heptanal in the first case, and both nonenal and nonanal in the second case (data not shown). However, to the best of our knowledge, the APEBA derivatization methodology has only been applied to linear saturated aldehydes or aldehydes that are not prone to reductive processes. This is also the case in the work recently published by Yu et al. [26], using APC—a reagent similar to APEBA previously developed by Eggink et al. [25]. Among other aldehydes, HNE and HHE were also analyzed as APC derivatives by these authors [24], but the search for potential reduction byproducts of these two hydroxy-alkenals is not mentioned.

Since this feature could yield unexpected byproducts and lead to false positive results for several compounds of interest due to an artefactual reduction process, several other reaction conditions were assessed to avoid the formation of such reduction byproducts using APEBA, but all of our trials failed. Therefore, the use of a reagent allowing the selective derivatization of aldehydes under non-reductive, mild, cold, and neutral conditions had to be considered. We selected a simple, commercially available reagent—namely, 1-((ammoniooxy)methyl)-2-bromobenzene chloride (BBHA), which reacts with free carbonyl compounds [33] in one step, with no co-reagent, following a very simple and quick protocol, under mild conditions (no degradation due to chemical conditions) and at low

temperature (important because of the volatility of some aldehydes). Similarly to the case of APEBA, high-resolution mass spectrometric analysis allowed a sensitive and selective detection of derivatized aldehydes on the basis of the  $^{79}\text{Br}$  and  $^{81}\text{Br}$  isotopomers of the BBHA derivatives (isotopic pattern, exact mass, mass defect).

Although method validation was not our main scope, standard HHE- and HNE-BBHA derivatives were used to evaluate the performances of the method in terms of linearity, repeatability, and sensitivity of detection. For that, standard mixtures of HHE-BBHA and HNE-BBHA (10, 20, 50, 100, 250, and 500 pg/ $\mu\text{L}$ ) were analyzed. The retention time RSD was found to be <1% for both derivatives ( $n = 3$ ). The RSD values calculated from the peak areas of both derivatives were found to be <15% over the whole concentration range ( $n = 3$  for each concentration point). The LOD was reached for 10–15 pg (30–40 fmol) injected into the column based on HHE and HNE, which is around 5–10 times higher than the LOD obtained by targeted analysis of these two compounds using the same derivatization reagent. Compared to other published aldehyde screening methods, these values are in the same range as the LOD reported by Eggink et al. [24], and slightly higher than that reported by Yu et al. [26] with the APC reagent, likely due to the fact that (1) BBHA is not a charged reagent as compared to APC and APEBA, and (2) the response was split over the two bromine isotopomers of the BBHA derivatives as compared to APC. The linearity of the response was checked over the 10–500 pg/ $\mu\text{L}$  range, and gave  $R^2$  values of 0.994 and 0.997 for HHE-BBHA and HNE-BBHA, respectively, enabling the use of this method at least in a semi-quantitative way.

The selectivity of the derivatization method has been assessed with several carbonyl compounds, including aldehydes and ketones, as well as carbohydrates. Carbohydrates that could be present in the intestinal lumen can also be derivatized with BBHA when occurring in their reductive form. As expected, those species bearing several hydroxyl groups yielded characteristic MS/MS spectra with intense  $[\text{MH}-\text{H}_2\text{O}]^+$  ions, weak signals at  $m/z$  168.9647, and a common fragment ion at  $m/z$  255.9972 (data not shown). Conversely, no BBHA derivative was formed from any standard fatty acid tested in the derivatization medium.

### 3.2. MS/MS Fragmentation of Aldehyde BBHA Derivatives

The synthesis and spectral characterization ( $^1\text{H}$  NMR, exact molecular mass) of 11 BBHA derivatives of standard carbonyl compounds (9 aldehydes and 2 ketones) has been published elsewhere [33]. To enable efficient suspect screening data mining from high-resolution full-scan acquisitions, other compounds of interest were further selected and synthesized, and MS/MS spectra were acquired for a total of 16 derivatized carbonyl compounds. Standard derivatized aldehydes are listed in Table 1 together with their molecular formulae, calculated exact mass, and the major fragment ions observed in MS/MS experiments.

The MS/MS spectra of all derivatives display the  $m/z$  168.9646 fragment ion ( $\text{C}_7\text{H}_6\text{Br}$ ) as base peak, resulting from the cleavage of the C–O bond of the oxime ether, leading to the bromo-benzyl carbocation. This diagnostic fragment ion could be used as tracer for untargeted unknown searching using, for example, all-ion (DIA) MS/MS strategies. Interestingly, hydroxy-alkenals such as HHE and HNE display an intense fragment ion at  $m/z$  224.0069, which may be considered characteristic of hydroxy-alkenals. This particular  $m/z$  ratio fits the chemical formula  $\text{C}_{10}\text{H}_{11}\text{NBr}$ . The hydroperoxide homolog of HNE (HPNE) displays the same type of fragment ion, but at  $m/z$  240.0019 ( $\text{C}_{10}\text{H}_{11}\text{NOBr}$ )—i.e., shifted by 15.995 AMU—corresponding to one oxygen atom, suggesting that these compounds behave in the same way towards collisional activation. The formation of this fragment ion must involve a rearrangement in which the oxygen atom of the ether oxime function is eliminated. However, this fragmentation mechanism was not investigated further. Furthermore, the MS/MS spectrum of HPNE also displays a fragment ion at  $m/z$  322.0807, which fits the elimination of  $\text{H}_2\text{O}_2$  from the  $\text{MH}^+$  parent ion.



**Table 1.** List of prepared standard BBHA derivatives and their respective molecular mass, exact mass of the [M+H]<sup>+</sup> ion, main MS/MS fragment ions, and chromatographic retention times.

Compound	Abbreviation	Chemical Formula BBHA Derivative	Molecular Mass	Calculated Exact Mass ( <sup>79</sup> Br/ <sup>81</sup> Br) or ( <sup>79</sup> Br/ <sup>79</sup> Br/ <sup>79</sup> Br/ <sup>81</sup> Br/ <sup>81</sup> Br) [M+H] <sup>+</sup>	Main Observed Fragment Ions (m/z) from MS/MS Carried Out on the <sup>79</sup> Br Isotopomer	Retention Time (min.)
4-OH-hexenal	IIIIE	C <sub>13</sub> H <sub>16</sub> O <sub>2</sub> NBr	297/299	298.04372/300.04167	168.9645; 224.0069	15.6; 15.7
4-Oxo-hexenal	OIII E	C <sub>13</sub> H <sub>14</sub> O <sub>2</sub> NBr	295/297	296.02807/298.02602	168.9645; 240.0020; 268.0329; 221.9912	16.8
4-OH-nonenal	HNE	C <sub>16</sub> H <sub>22</sub> O <sub>2</sub> NBr	339/341	340.09067/342.08862	168.9645; 224.0069	17.6; 17.8
4-Oxo-nonenal	ONE	C <sub>16</sub> H <sub>20</sub> O <sub>2</sub> NBr	337/339	338.07502/340.07297	168.9646; 240.0021; 221.9914; 310.0803	18.5
4,5-Epoxy-2-decenal	EDE	C <sub>17</sub> H <sub>22</sub> O <sub>2</sub> NBr	351/353	352.09067/354.08862	168.9645; 254.0174; 201.9860	17.8
2,4-Decadienal	DDI E	C <sub>17</sub> H <sub>22</sub> ONBr	335/337	336.09575/338.09371	168.9649; 150.1279; 248.0072; 256.1698	20.3
4-Hydroperoxynonenal	HPNE	C <sub>16</sub> H <sub>22</sub> O <sub>3</sub> NBr	355/357	356.08558/358.08354	240.0022; 168.9648; 170.1176; 322.0807	17.7; 17.9
4-OH-nonanal	OHN	C <sub>16</sub> H <sub>22</sub> O <sub>2</sub> NBr	341/343	342.10632/344.10427	324.0960; 168.9642; 141.1273; 254.0175	17.6
Malondialdehyde	MDA	C <sub>17</sub> H <sub>16</sub> O <sub>2</sub> N <sub>2</sub> Br <sub>2</sub> <sup>1</sup>	438/440/442	438.96513/440.96308/442.96104	168.9645; 254.0048; 237.0021	18.7
Pyruvaldehyde	PRA	C <sub>17</sub> H <sub>16</sub> O <sub>2</sub> N <sub>2</sub> Br <sub>2</sub> <sup>1</sup>	438/440/442	438.96513/440.96308/442.96104	168.9650; 235.9948; 336.9224	19.8
5-Oxo-pentanoic acid	5-OPA	C <sub>12</sub> H <sub>14</sub> O <sub>3</sub> NBr	299/301	300.02298/302.02094	168.9649; 282.0128; 96.0044	14.5
6-Oxo-hexanoic acid	6-OIIA	C <sub>13</sub> H <sub>16</sub> O <sub>3</sub> NBr	313/315	314.03863/316.03659	168.9649; 296.0286; 183.9758; 110.0601	15.0
7-Oxo-heptanoic acid	7-OIIA	C <sub>14</sub> H <sub>18</sub> O <sub>3</sub> NBr	327/329	328.05428/330.05224	168.9648; 310.0441; 183.9761; 124.0755	15.5
8-Oxo-octanoic acid	8-OOA	C <sub>15</sub> H <sub>20</sub> O <sub>3</sub> NBr	341/343	342.06993/344.06789	168.9649; 324.0599; 183.9758; 138.0914	16.1
9-Oxo-nonanoic acid	9-ONA	C <sub>16</sub> H <sub>22</sub> O <sub>3</sub> NBr	355/357	356.08558/358.08354	168.9646; 338.0751; 183.9756; 152.1070	16.7
10-Oxo-decanoic acid	10-ODA	C <sub>17</sub> H <sub>24</sub> O <sub>3</sub> NBr	369/371	370.10123/372.09919	168.9648; 352.0910; 183.9758; 166.1228	17.2

<sup>1</sup> bis-BBHA derivative.

Two 4-oxo alkenals were also analyzed—namely, OHE and ONE. As expected, the  $m/z$  169 ion was also the base peak for both compounds. As reported in Table 1, both compounds displayed a loss of CO as well as an  $m/z$  221.9913 fragment ion in agreement with the chemical formula  $C_{10}H_9NBr$ , suggesting that 4-oxo alkenals undergo the same rearrangement as their 4-hydroxy analogs. However, in the case of 4-oxo alkenals such as OHE and ONE, a fragment ion was also observed at  $m/z$  240.0020, corresponding to  $C_{10}H_{11}ONBr$ , which could be attributed to an  $\alpha$ -cleavage relative to the CO group.

The saturated hydroxylated aldehyde 4-hydroxynonanal (OHN), as well as  $\omega$ -oxo carboxylic acids (5-OPA, 6-OHA, 7-OHA, 8-OOA, 9-ONA, and 10-ODA), all display an  $[MH-H_2O]^+$  ion as one of the most intense fragments. These compounds with a non-functionalized carbon chain then undergo a breakage of the N–O bond from the ether-oxime function of their BBHA derivative, giving rise to the  $m/z$  96.0044, 110.0601, 124.0755, 138.0914, 152.1070, and 166.1228 ions for 5-OPA, 6-OHA, 7-OHA, 8-OOA, 9-ONA, and 10-ODA, respectively. 2,4-Decadienal (DDE)—another non-functionalized compound—also undergoes this process, yielding the  $C_{10}H_{16}N$  fragment ion at  $m/z$  150.1279. However, the same process was observed for 4-hydroperoxynonenal (HPNE), leading to the  $C_9H_{16}O_2N$  ion ( $m/z$  170.1176), meaning that this cleavage may also occur at significant levels for different collision energies. Surprisingly, DDE was the only BBHA derivative to eliminate HBr from the  $MH^+$  precursor ion. Upon MS/MS, 4,5-epoxy-2-decenal (EDE) yielded two main fragment ions at  $m/z$  254.0174 ( $C_{11}H_{13}NOBr$ ) and  $m/z$  201.9860 ( $C_7H_9NOBr$ ), which could be attributed to cleavages at the epoxy function and the nitrogen atom, respectively.

As expected, the two dialdehydes—namely, malondialdehyde (MDA) and pyruvaldehyde (PRA)—were detected as their bis-BBHA derivatives, giving a characteristic isotopic pattern with ions at  $m/z$  438.9651, 440.9631, and 442.9610, corresponding to their  $^{79}Br^{79}Br$ ,  $^{79}Br^{81}Br$ , and  $^{81}Br^{81}Br$  isotopologs, respectively. Despite their close chemical structure, these two derivatives yielded different fragment ions. Finally, it should be noted that all  $\omega$ -oxo carboxylic acids except for 5-OPA gave a fragment ion at  $m/z$  183.9756 fitting the molecular formula  $C_7H_7NBr$ , which appears to be specific to this series; its formation may also involve a complex rearrangement process, which was not investigated further.

Altogether, these MS/MS features show that although fragmentation processes are different depending on the structure of the BBHA derivatives, compounds belonging to the same class (e.g., HHE and HNE; OHE and ONE;  $\omega$ -oxo carboxylic acids) show similar behavior towards MS/MS collisional activation, and that the corresponding fragment ions can be used for structural confirmation as well as for providing clues as to the class of compounds to which they belong. Furthermore, as previously mentioned, the  $m/z$  168.9645 can be used for the tracking and confirmation of the BBHA derivatives using all-ion MS/MS strategies.

### 3.3. Profiling of Diet-Related Peroxidation Carbonyl Compounds Formed in the Intestinal Lumen

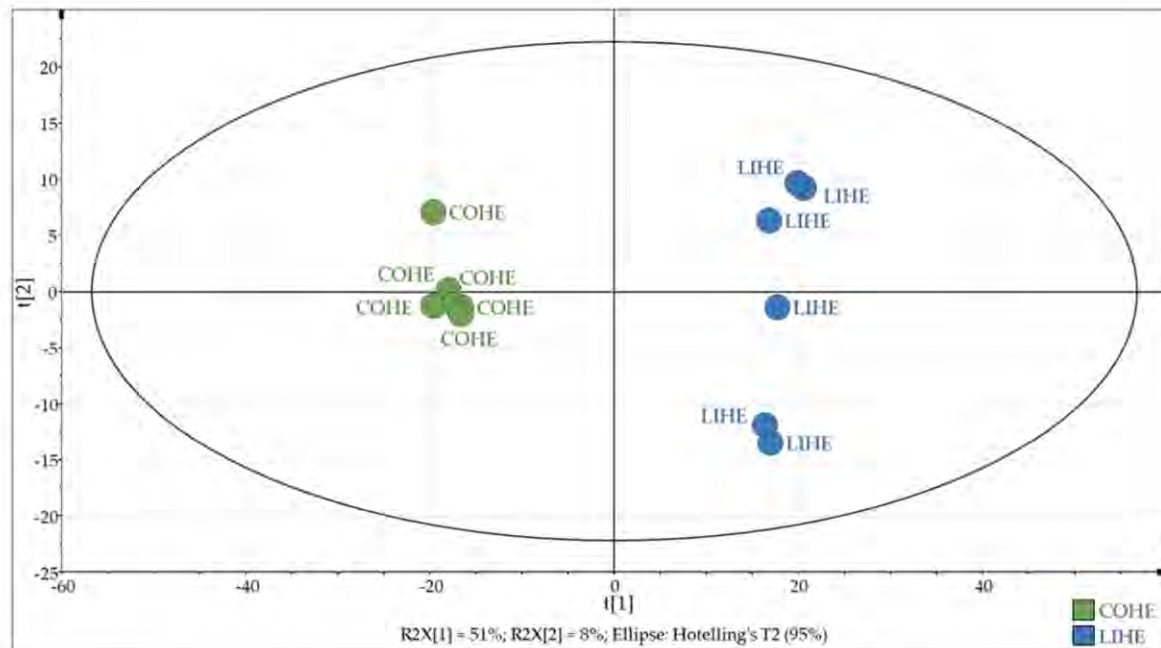
In order to assess the potential of the developed method for the untargeted profiling of lipid peroxidation aldehydes, fecal waters prepared from rats ( $n = 6$ ) fed two distinct diets—namely, hydrogenated coconut oil (COHE) and linseed oil (LIHE)—containing different precursor fatty acid profiles, both with added heme iron, were derivatized with BBHA and analyzed via LC–HRMS. A post-analysis raw data-mining protocol was then set up to extract signals corresponding to brominated compounds (e.g., exact mass, isotopic pattern) coming from the derivatized aldehydes. Since many aldehydes produced by lipid oxidation are specific to their fatty acid precursors, several expected aldehyde derivatives (including, in particular, toxicologically relevant species such as hydroxy- and oxo-alkenals) were searched on the basis of the exact mass of their BBHA derivatives (i.e., both the  $^{79}Br$  and  $^{81}Br$  isotopic peaks) in a suspect screening approach. When a suspect signal was detected, confirmation was achieved using the standard BBHA derivatives synthesized in our previous works and used for the setup of the derivatization method. Signal intensities and fold changes between the two diets observed for the 16 compounds prepared as BBHA derivative standards in this suspect screening approach are reported in Supplementary

Table S1. As expected, the occurrence of HHE was detected in fecal waters from rats fed the  $\omega$ -3-fatty-acid-rich linseed oil, whereas it was not detected in fecal waters from rats fed hydrogenated coconut oil, mainly composed of saturated fatty acids. To a lesser extent, HNE and OHN were also detected from the linseed oil diet, which also contains  $\omega$ -6 fatty acids, and not (OHN) or very weakly (HNE, fold change 10) detected in the coconut oil diet. In addition, linseed oil rich in  $\alpha$ -linolenic acid gave important amounts of malondialdehyde (MDA, which was detected as its bis-BBHA derivative), whereas it was detected only in very weak amounts in the samples from rats fed coconut oil, leading to a fold change of 500. The other dialdehyde listed in Table 1—namely, PRA—was detected in similar amounts in both groups. This compound is known to result from carbohydrate fermentation, and is not considered to be a lipid peroxidation product [35]. Conversely, OHE, ONE, and EDE were not detected in any of the samples. Among the  $\omega$ -oxo carboxylic acids, and as far as they can be considered to behave similarly towards electrospray ionization, 5-OPA and 9-ONA were the most abundantly detected (see Supplementary Table S1); although present in less abundance in the hydrogenated coconut oil diet, they were both detected in all samples. 9-ONA is known to be one of the main peroxidation products of  $\omega$ -6 fatty acids [36], and its presence in all of the samples could be tentatively explained by basal lipid peroxidation of intestinal cells [37]. Finally, 8-OOA was the  $\omega$ -oxo carboxylic acid displaying the highest fold change between the two diets. This observation illustrates the added value provided by untargeted approaches to get early warning messages to prioritize specific studies, which could be focused in this precise case on the origin of 8-OOA.

#### 3.4. Untargeted “Aldehydomics” Approach: Statistical Data Analysis and Identification of Aldehydes

From the whole dataset acquired via full-scan HRMS from individual rat fecal waters obtained from the two diets differing in precursor fatty acid profiles, signals corresponding to brominated species were first extracted according to a mass measurement error of  $\pm 5$  ppm, using a homemade R script for filtering MS signals on the basis of (1) the mass difference between the  $^{79}\text{Br}$  and  $^{81}\text{Br}$  isotopes—i.e., 1.998 Da—and (2) the equivalent intensities of the two signals (see Materials and Methods). This process led to the extraction of 774 features containing only variables related to brominated compounds. This dataset was then subjected to multivariate statistical analysis using a classical global metabolomics method. As shown in Figure 3, the two diets were very well discriminated by principal component analysis (PCA), without the need for a supervised analysis, which confirms that BBHA-derivatized species are responsible for the observed separation.

After PLS-DA statistical analyses, 413 discriminating variables out of the 774 extracted features (corresponding to filtered brominated compounds) were validated with a VIP score  $> 1$ . A total of 84 features could be putatively annotated based on their exact mass using several databases (i.e., HMDB, LIPID MAPS, homemade database), among which 36 could be linked to lipids and their peroxidation products (the annotation table is reported in Supplementary Table S2). For several annotated compounds (namely, oxo-hexanoic acid, oxo-heptanoic acid, and oxo-eicosadienoic acid), standard compounds could be purchased and their corresponding BBHA derivatives synthesized. Unfortunately, the mass spectrometric features of the standard compounds did not fit those of the compounds detected in fecal waters. This enabled us to rule out only one of the possible isomeric structures for the annotated compounds, which remained at identification level 3 according to the Metabolomics Standards Initiative [38]. Finally, seven compounds were formally identified at identification level 1 thanks to the comparison with authentic standard compounds. These seven compounds are reported in Table 2.



**Figure 3.** PCA performed on the 774 extracted BBHA-derivatized features. COHE: coconut oil supplemented with heme iron. LIHE: linseed oil supplemented with heme iron.

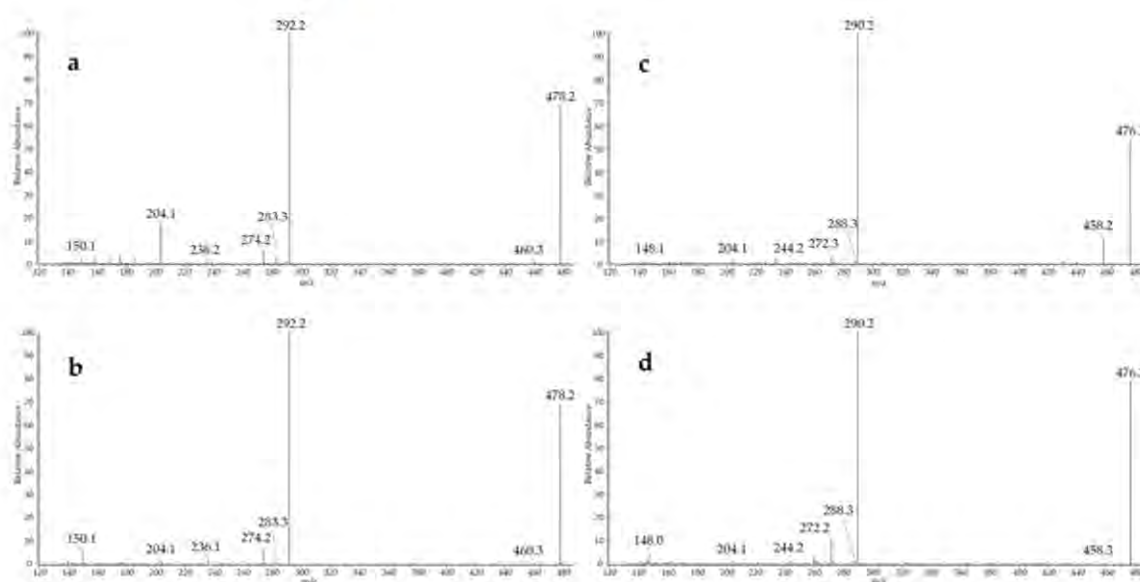
**Table 2.** Lipid oxidation products and fatty acids identified at level 1 by the untargeted aldehydomics strategy.

Compound	$m/z$ [M+H] <sup>+</sup>	Rt (min)	Identification Level	FoldLIHE/COHE
Malondialdehyde (MDA)	438.96519	18.60	1	202.27
4-Hydroxy-hex-2-enal (HHE)	298.04369	15.68	1	11.46
9-Oxo-octadecadienoic acid (9-oxoODE)	478.19490	20.51	1	18.49
9-Oxo-octadecatrienoic acid (9-oxoOTrE)	476.17916	19.93	1	78.39
4-Hydroxy-non-2-enal (HNE)	340.09047	17.65	1	3.27
$\alpha$ -Linolenic acid	462.20024	19.90	1	371.22
Eicosapentaenoic acid	486.19979	19.84	1	59.18

As expected, MDA and HHE were identified among the variables exhibiting the highest VIP values (see Table 2). MDA was detected at  $m/z$  438.9652 (Rt 18.6 min.) as its bis-BBHA derivative, whereas HHE was detected as its single-BBHA derivative at  $m/z$  298.0437 (Rt 15.7 min.). Both compounds were confirmed at level 1 by comparison with the corresponding authentic standards. Although not representative of a single-precursor fatty acid, MDA represented the most significant annotated variable, and ranked third in the whole PLS-DA analysis. It displayed a fold change of 202, showing that the peroxidation rate of linseed-oil-fed rats' feces was much higher than that of hydrogenated-coconut-oil-fed rats. With a fold change of 11, HHE was also among the most significant variables, confirming that it represents a major marker of  $\omega$ -3 fatty acid peroxidation. HNE ( $m/z$  340.0905, Rt 17.7 min.) was also identified at level 1 among the significant variables. The VIP and the fold change of this compound were not as elevated as for HHE, which was expected, since linseed oil is known to contain a majority of  $\omega$ -3 fatty acids and a lesser proportion of  $\omega$ -6 fatty acids. Surprisingly, some BBHA derivatives of fatty acids were also identified or annotated, although our previous trials in the derivatization buffer (without fecal water matrix) never produced any BBHA derivatives from fatty acid standards. Among them, eicosapentaenoic acid and  $\alpha$ -linolenic acid were identified at level 1. It is noteworthy that in order to be able to make confirmation trials on these fatty acids, the derivatization process had to be carried out by spiking fecal waters, since no reaction

occurs in the buffer medium used for the derivatization of standard compounds. One hypothesis for this observation may be found in the very high proportion of unchanged fatty acids coming from the oil-based diet (5% hydrogenated coconut or linseed oil in the diet) in the feces compared to lipid peroxidation products, which could explain the formation of their BBHA derivatives in a detectable abundance despite a very low reactivity. Another hypothesis may be found in the particular composition of fecal water that might have accelerated the reaction between BBHA and fatty acids, which are not prone to derivatization under “classical” chemical conditions. Note that a PCA statistical analysis carried out after removing these variables yielded the same result, showing that those unoxidized diet-sourced fatty acids were not influential for the discrimination of the two groups (data not shown).

In addition to the expected MDA, HHE and, to a lesser extent, HNE, the developed method allowed for the identification of two additional lipid peroxidation compounds—namely, 9-oxo-octadecadienoic acid (9-oxo-ODE) and 9-oxo-octadecatrienoic acid (9-oxo-OTrE), which were not included in our first suspects. Their identity was confirmed at level 1 by comparison of their retention time, exact mass, and MS/MS spectra with those of the BBHA derivatives prepared from the authentic standard of the precursor oxo-fatty acids (Figure 4).



**Figure 4.** MS/MS spectra of the BBHA derivatives of 9-oxo-octadecadienoic acid obtained from (a) rat fecal water and (b) standard compound, and of the BBHA derivatives of 9-oxo-octadecatrienoic acid from (c) rat fecal water and (d) standard compound.

These oxo-fatty acids belong to the class of oxylipins—a family of compounds that come from the enzymatic or non-enzymatic oxidation of fatty acids, and have been reported to have biological activity [39]. In the case of two other compounds that could be annotated (see Supplementary Table S2) as oxo-eicosadienoic acid ( $m/z$  506.2262) and oxo-octadecanoic acid ( $m/z$  482.2206), the standard 15-oxo-eicosadienoic acid ( $m/z$  506.2262) and 3-oxo-octadecanoic acid could be purchased for structural confirmation. Unfortunately, neither of these two compounds fit the retention time of our discriminating variables. This was also the case for the oxo-hexanoic acid and oxo-heptanoic acid, which could not be confirmed as the  $\omega$ -oxo-acids by comparison with the BBHA standard derivatives, and for which an isomeric structure should also be considered.

#### 4. Conclusions

The formation of lipid peroxidation products such as aldehydes is one of the most convincing hypotheses for linking increased CRC risk with red meat consumption. Aside from well-known reactive hydroxy-alkenals (e.g., HHE and HNE), several other lipid peroxidation products are likely to be formed, and a significant proportion of them are still unknown. Although metabolomics has become a powerful tool to identify biomarkers from various biological situations, it still suffers from a lack of capacity for the reliable annotation of several classes of metabolites. Moreover, in the case of lipid peroxidation products such as aldehydes, a specific approach has to be adopted in order to overcome the unstable character of these compounds, which prevents their analysis in their native form.

In this context, we developed an original analytical strategy based on the introduction of a bromine atom by derivatization. This allowed, on the one hand, the stabilization of aldehydes that are prone to degradation, and on the other hand, their selective detection after untargeted high-resolution mass spectrometry data acquisition, thanks to the selective filtration of the characteristic bromine isotopic pattern of derivatives, and therefore the selective filtration of aldehyde species. This allows access to semi-quantitative information, thereby comparing different situations by means of fold changes between signal intensities for a given compound. The developed method has been tested on fecal water samples from rats fed two contrasting diets according to their fatty acid composition. The results showed that the method allowed for (1) the detection of several expected aldehydes, including hydroxyl-alkenals such as HHE and HNE, by using a suspect screening approach, and (2) the detection of several other brominated signals, corresponding to potentially unknown aldehydes, among which several compounds could be identified in an untargeted manner. After this proof of concept developed on the basis of fecal water from rats fed with highly contrasting diets, we have shown that this method can be applied for the non-targeted detection of aldehydes for a much wider field of application. This opens the way for the easier identification of new carbonylated species resulting from lipid peroxidation.

**Supplementary Materials:** The following are available online at <https://www.mdpi.com/article/10.3390/antiox10081261/s1>, Table S1: List of screened carbonyl compounds, with their respective intensities in both diets, and fold changes between diets; Table S2: Annotation table of lipid peroxidation products from LC–HRMS analysis of rat fecal waters after BBHA derivatization.

**Author Contributions:** Conceptualization, S.C., I.J., F.G. and L.D.; methodology, S.C., M.-H.N.-M., L.M., I.J., N.N., J.-F.M. and F.G.; investigation, S.C., M.-H.N.-M., F.G. and L.D.; writing—original draft preparation, S.C., M.-H.N.-M., J.-F.M., F.G. and L.D.; writing—review and editing, S.C., L.M., F.G. and L.D.; supervision, S.C., F.P., F.G. and L.D.; project administration, F.G. and L.D.; funding acquisition, F.P. and F.G. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by French National Agency for Research (ANR, grant ANR-10-ALIA-014) and by the INRAE-AlimH research division.

**Institutional Review Board Statement:** The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Local ethical committee TOXCOM/006/FG.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data presented in this study are available in article and supplementary material.

**Acknowledgments:** The authors would like to thank Thaïs Hautbergue for preliminary works during her master's internship, and Florence Blas-y-Estrada (INRA Toxalim, Toulouse, France) for assistance with the animal experiments. The Région Occitanie and the INRAE-AlimH research division are thanked for granting the PhD thesis of L.M. This work was achieved within the MetaToul-AXIOM analytical platform, partner of the MetaboHUB infrastructure for metabolomics (ANR 11-INBS-0010). It was supported by INRAE-AlimH research division and by the Securiviande ANR grant (ANR 10-ALIA-014). L.D. and S.C. would like to dedicate this article to the memory of Francis Marty, Emeritus Professor at the University of Burgundy, France.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Bastide, N.M.; Pierre, F.H.F.; Corpet, D.E. Heme iron from meat and risk of colorectal cancer: A meta-analysis and a review of the mechanisms involved. *Cancer Prev. Res.* **2011**, *4*, 177–184. [[CrossRef](#)]
2. Vieira, A.R.; Abar, L.; Chan, D.S.M.; Vingeliene, S.; Polemiti, E.; Stevens, C.; Greenwood, D.; Norat, T. Foods and beverages and colorectal cancer risk: A systematic review and meta-analysis of cohort studies, an update of the evidence of the WCRF-AICR Continuous Update Project. *Ann. Oncol.* **2017**, *28*, 1788–1802. [[CrossRef](#)]
3. Lippi, G.; Mattiuzzi, C.; Cervellin, G. Meat consumption and cancer risk: A critical review of published meta-analyses. *Crit. Rev. Oncol. Hematol.* **2016**, *97*, 1–14. [[CrossRef](#)]
4. WCRF World Cancer Research Fund/American Institute for Cancer Research. *Food, Nutrition, Physical Activity, and the Prevention of Cancer: A Global Perspective*; AICR: Washington, DC, USA, 2007.
5. WCRF World Cancer Research Fund/American Institute for Cancer Research. *Diet, Nutrition, Physical Activity and Cancer: A Global Perspective Continuous Update Project Expert Report*; AICR: Washington, DC, USA, 2018.
6. Bastide, N.M.; Chenni, F.; Audebert, M.; Santarelli, R.L.; Tache, S.; Naud, N.; Baradat, M.; Jouanin, I.; Surya, R.; Hobbs, D.A.; et al. A central role for heme iron in colon carcinogenesis associated with red meat intake. *Cancer Res.* **2015**, *75*, 870–879. [[CrossRef](#)]
7. Sesink, A.L.A.; Termont, D.S.M.L.; Kleibeuker, J.H.; Van der Meer, R. Red meat and colon cancer: The cytotoxic and hyperproliferative effects of dietary heme meat and colon cancer. *Cancer Res.* **1999**, *59*, 5704–5709.
8. Esterbauer, H.; Schaur, R.J.; Zollner, H. Chemistry and biochemistry of 4-hydroxynonenal, malonaldehyde and related aldehydes. *Free Radic. Biol. Med.* **1991**, *11*, 81–128. [[CrossRef](#)]
9. Guéraud, F.; Atalay, M.; Bresgen, N.; Cipak, A.; Eckl, P.M.; Huc, L.; Jouanin, I.; Siems, W.; Uchida, K. Chemistry and biochemistry of lipid peroxidation products. *Free Radic. Res.* **2010**, *44*, 1098–1124. [[CrossRef](#)] [[PubMed](#)]
10. Keller, J.; Chevolleau, S.; Noguer-Meireles, M.-H.; Pujos-Guillot, E.; Delosière, M.; Chantelauze, C.; Joly, C.; Blas-y-Estrada, F.; Jouanin, I.; Durand, D.; et al. Heme-iron-induced production of 4-hydroxynonenal in intestinal lumen may have extra-intestinal consequences through protein-adduct formation. *Antioxidants* **2020**, *9*, 1293. [[CrossRef](#)]
11. Baradat, M.; Jouanin, I.; Dalleau, S.; Taché, S.; Gieules, M.; Debrauwer, L.; Canlet, C.; Huc, L.; Dupuy, J.; Pierre, F.H.F.; et al. 4-Hydroxy-2(E)-nonenal metabolism differs in apc+/+ cells and in apcmin/+ cells: It may explain colon cancer promotion by heme iron. *Chem. Res. Toxicol.* **2011**, *24*, 1984–1993. [[CrossRef](#)] [[PubMed](#)]
12. Kawai, Y.; Takeda, S.; Terao, J. Lipidomic analysis for lipid peroxidation-derived aldehydes using gas chromatography-mass spectrometry. *Chem. Res. Toxicol.* **2007**, *20*, 99–107. [[CrossRef](#)] [[PubMed](#)]
13. Uchiyama, S.; Inaba, Y.; Kunugita, N. Derivatization of carbonyl compounds with 2,4-dinitrophenylhydrazine and their subsequent determination by high-performance liquid chromatography. *J. Chromatogr. B* **2011**, *879*, 1282–1289. [[CrossRef](#)]
14. Dalle-Donne, I.; Carini, M.; Orioli, M.; Vistoli, G.; Regazzoni, L.; Colombo, G.; Rossi, R.; Milzani, A.; Aldini, G. Protein carbonylation: 2,4-dinitrophenylhydrazine reacts with both aldehydes/ketones and sulfenic acids. *Free Radic. Biol. Med.* **2009**, *46*, 1411–1419. [[CrossRef](#)]
15. Douny, C.; Bayram, P.; Brose, F.; Degand, G.; Scippo, M.-L. Development of an LC-MS/MS Analytical Method for the simultaneous measurement of aldehydes from polyunsaturated fatty acids degradation in animal feed. *Drug Test. Anal.* **2016**, *8*, 458–464. [[CrossRef](#)]
16. Hu, Y.-N.; Chen, D.; Zhang, T.-Y.; Ding, J.; Feng, Y.-Q. Use of ammonium sulfite as a post-column derivatization reagent for rapid detection and quantification of aldehydes by LC-MS. *Talanta* **2020**, *206*, 120172. [[CrossRef](#)]
17. Cao, Y.; Guan, Q.; Sun, T.; Qi, W.; Guo, Y. Charged tag founded in n-(1-chloroalkyl)pyridinium quaternization for quantification of fatty aldehydes. *Anal. Chim. Acta* **2016**, *937*, 80–86. [[CrossRef](#)]
18. Guo, N.; Peng, C.-Y.; Zhu, Q.-F.; Yuan, B.-F.; Feng, Y.-Q. Profiling of carbonyl compounds in serum by stable isotope labeling-double precursor ion scan-mass spectrometry analysis. *Anal. Chim. Acta* **2017**, *967*, 42–51. [[CrossRef](#)] [[PubMed](#)]
19. Cao, J.; Zou, X.-G.; Deng, L.; Fan, Y.-W.; Li, H.; Li, J.; Deng, Z.-Y. Analysis of nonpolar lipophilic aldehydes/ketones in oxidized edible oils using HPLC-QqQ-MS for the evaluation of their parent fatty acids. *Food Res. Int.* **2014**, *64*, 901–907. [[CrossRef](#)]
20. Schymanski, E.L.; Singer, H.P.; Slobodnik, J.; Ipolyi, I.M.; Oswald, P.; Krauss, M.; Schulze, T.; Haglund, P.; Letzel, T.; Grosse, S.; et al. Non-target screening with high-resolution mass spectrometry: Critical review using a collaborative trial on water analysis. *Anal. Bioanal. Chem.* **2015**, *407*, 6237–6255. [[CrossRef](#)] [[PubMed](#)]
21. Jamin, E.L.; Costantino, R.; Mervant, L.; Martin, J.-F.; Jouanin, I.; Blas-Y-Estrada, F.; Guéraud, F.; Debrauwer, L. Global profiling of toxicologically relevant metabolites in urine: Case study of reactive aldehydes. *Anal. Chem.* **2020**, *92*, 1746–1754. [[CrossRef](#)]
22. Pourchet, M.; Debrauwer, L.; Oberacher, M.; Krauss, M.; Vlaanderen, A.; Covaci, A.; Damont, A.; Klanova, J.; Lamoree, M.; Sarigiannis, D.; et al. Suspect and non-targeted screening of chemicals of emerging concern for human biomonitoring and environmental health studies: Current capabilities, promises, and challenges. *Environ. Int.* **2020**, *139*, 105545. [[CrossRef](#)] [[PubMed](#)]
23. Eggink, M.; Wijtman, M.; Ekkebus, R.; Lingeman, H.; de Esch, I.J.P.; Kool, J.; Niessen, W.M.A.; Irth, H. Development of a selective esi-ms derivatization reagent: Synthesis and optimization for the analysis of aldehydes in biological mixtures. *Anal. Chem.* **2008**, *80*, 9042–9051. [[CrossRef](#)] [[PubMed](#)]

24. Eggink, M.; Charret, S.; Wijtmans, M.; Lingeman, H.; Kool, J.; Niessen, W.M.A.; Irth, H. Development of an on-line weak-cation exchange liquid chromatography–tandem mass spectrometric method for screening aldehyde products in biological matrices. *J. Chromatogr. B* **2009**, *877*, 3937–3945. [[CrossRef](#)] [[PubMed](#)]
25. Eggink, M.; Wijtmans, M.; Kretschmer, A.; Kool, J.; Lingeman, H.; de Esch, I.J.P.; Niessen, W.M.A.; Irth, H. Targeted LC–MS derivatization for aldehydes and carboxylic acids with a new derivatization agent 4-APEBA. *Anal. Bioanal. Chem.* **2010**, *397*, 665–675. [[CrossRef](#)] [[PubMed](#)]
26. Yu, L.; Liu, P.; Wang, Y.-L.; Yu, Q.-W.; Yuan, B.-F.; Feng, Y.-Q. Profiling of aldehyde-containing compounds by stable isotope labelling-assisted mass spectrometry analysis. *Analyst* **2015**, *140*, 5276–5286. [[CrossRef](#)]
27. Tie, C.; Hu, T.; Jia, Z.-X.; Zhang, J.-L. Derivatization strategy for the comprehensive characterization of endogenous fatty aldehydes using hplc-multiple reaction monitoring. *Anal. Chem.* **2016**, *88*, 7762–7768. [[CrossRef](#)]
28. El-Maghrabey, M.; Kishikawa, N.; Kuroda, N. Novel isotope-coded derivatization method for aldehydes using <sup>14</sup>N/<sup>15</sup>N-ammonium acetate and 9,10-phenanthrenequinone. *Anal. Chem.* **2018**, *90*, 13867–13875. [[CrossRef](#)]
29. Cuyckens, F.; Balcaen, L.I.L.; De Wolf, K.; De Samber, B.; Van Looveren, C.; Hurkmans, R.; Vanhaecke, F. Use of the bromine isotope ratio in HPLC-ICP-MS and HPLC-ESI-MS analysis of a new drug in development. *Anal. Bioanal. Chem.* **2008**, *390*, 1717–1729. [[CrossRef](#)] [[PubMed](#)]
30. Bierla, K.; Riu, A.; Debrauwer, L.; Zalko, D.; Bouyssiere, B.; Szpunar, J. Screening for polybrominated diphenyl ethers in biological samples by reversed-phase fast HPLC-ICP MS. *J. Anal. At. Spectrom.* **2010**, *25*, 889. [[CrossRef](#)]
31. Chevolleau, S.; Nogueira-Meireles, M.-H.; Jouanin, I.; Naud, N.; Pierre, F.; Gueraud, F.; Debrauwer, L. Development and validation of an ultra high performance liquid chromatography–electrospray tandem mass spectrometry method using selective derivatisation, for the quantification of two reactive aldehydes produced by lipid peroxidation, HNE (4-Hydroxy-2( E )-Nonenal) and HHE (4-Hydroxy-2( E )-Hexenal) in faecal water. *J. Chromatogr. B* **2018**, *1083*, 171–179. [[CrossRef](#)]
32. Chandra, A.; Srivastava, S.K. A Synthesis of 4-hydroxy-2-trans-nonenal and 4-(3h) 4-hydroxy-2-trans-nonenal. *Lipids* **1997**, *32*, 779–782. [[CrossRef](#)]
33. Jouanin, I.; Chevolleau, S.; Canlet, C.; Lorber, C.; Pierre, F.; Guéraud, F.; Debrauwer, L. facile oxime ether synthesis: Free carbonyl compound derivatization by a brominated *o*-benzylhydroxylamine. *Synth. Commun.* **2015**, *45*, 1585–1591. [[CrossRef](#)]
34. Tautenhahn, R.; Böttcher, C.; Neumann, S. Highly sensitive feature detection for high resolution LC/MS. *BMC Bioinform.* **2008**, *9*, 504. [[CrossRef](#)] [[PubMed](#)]
35. Campbell, A.K.; Matthews, S.B.; Vassel, N.; Cox, C.D.; Naseem, R.; Chaichi, J.; Holland, I.B.; Green, J.; Wann, K.T. Bacterial metabolic ‘toxins’: A new mechanism for lactose and food intolerance, and irritable bowel syndrome. *Toxicology* **2010**, *278*, 268–276. [[CrossRef](#)] [[PubMed](#)]
36. Spickett, C.M. The lipid peroxidation product 4-hydroxy-2-nonenal: Advances in chemistry and analysis. *Redox Biol.* **2013**, *1*, 145–152. [[CrossRef](#)]
37. Guéraud, F.; Taché, S.; Steghens, J.-P.; Milkovic, L.; Borovic-Sunjic, S.; Zarkovic, N.; Gaultier, E.; Naud, N.; Héliers-Toussaint, C.; Pierre, F.; et al. Dietary polyunsaturated fatty acids and heme iron induce oxidative stress biomarkers and a cancer promoting environment in the colon of rats. *Free Radic. Biol. Med.* **2015**, *83*, 192–200. [[CrossRef](#)]
38. Sumner, L.W.; Amberg, A.; Barrett, D.; Beale, M.H.; Beger, R.; Daykin, C.A.; Fan, T.W.-M.; Fiehn, O.; Goodacre, R.; Griffin, J.L.; et al. Proposed minimum reporting standards for chemical analysis: Chemical analysis working group (CAWG) Metabolomics Standards Initiative (MSI). *Metabolomics* **2007**, *3*, 211–221. [[CrossRef](#)]
39. Gabbs, M.; Leng, S.; Devassy, J.G.; Monirujjaman, M.; Aukema, H.M. Advances in our understanding of oxylipins derived from dietary PUFAs. *Adv. Nutr.* **2015**, *6*, 513–540. [[CrossRef](#)]



## **Références**

- Abdul Rahim, M.B.H., Chilloux, J., Martinez-Gili, L., Neves, A.L., Myridakis, A., Gooderham, N., Dumas, M.-E., 2019. Diet-induced metabolic changes of the human gut microbiome: importance of short-chain fatty acids, methylamines and indoles. *Acta Diabetol.* 56, 493–500. <https://doi.org/10.1007/s00592-019-01312-x>
- Abete, I., Romaguera, D., Vieira, A.R., Lopez de Munain, A., Norat, T., 2014a. Association between total, processed, red and white meat consumption and all-cause, CVD and IHD mortality: a meta-analysis of cohort studies. *Br. J. Nutr.* 112, 762–775. <https://doi.org/10.1017/S000711451400124X>
- Abete, I., Romaguera, D., Vieira, A.R., Munain, A.L. de, Norat, T., 2014b. Association between total, processed, red and white meat consumption and all-cause, CVD and IHD mortality: a meta-analysis of cohort studies. *Br. J. Nutr.* 112, 762–775. <https://doi.org/10.1017/S000711451400124X>
- Agnantis, N., Goussia, A.C., 2012. La carcinogénèse colorectale. *Bull. Académie Natl. Médecine* 196, 705–716. [https://doi.org/10.1016/S0001-4079\(19\)31807-2](https://doi.org/10.1016/S0001-4079(19)31807-2)
- Alary, J., Bravais, F., Cravedi, J.P., Debrauwer, L., Rao, D., Bories, G., 1995a. Mercapturic acid conjugates as urinary end metabolites of the lipid peroxidation product 4-hydroxy-2-nonenal in the rat. *Chem. Res. Toxicol.* 8, 34–39. <https://doi.org/10.1021/tx00043a004>
- Alary, J., Bravais, F., Cravedi, J.P., Debrauwer, L., Rao, D., Bories, G., 1995b. Mercapturic acid conjugates as urinary end metabolites of the lipid peroxidation product 4-hydroxy-2-nonenal in the rat. *Chem. Res. Toxicol.* 8, 34–39. <https://doi.org/10.1021/tx00043a004>
- Alary, J., Debrauwer, L., Fernandez, Y., Paris, A., Cravedi, J.P., Dolo, L., Rao, D., Bories, G., 1998. Identification of novel urinary metabolites of the lipid peroxidation product 4-hydroxy-2-nonenal in rats. *Chem. Res. Toxicol.* 11, 1368–1376. <https://doi.org/10.1021/tx980068g>
- Alonso, A., Marsal, S., Julià, A., 2015. Analytical Methods in Untargeted Metabolomics: State of the Art in 2015. *Front. Bioeng. Biotechnol.* 3, 23. <https://doi.org/10.3389/fbioe.2015.00023>
- Alwis, K.U., Blount, B.C., Britt, A.S., Patel, D., Ashley, D.L., 2012. Simultaneous analysis of 28 urinary VOC metabolites using ultra high performance liquid chromatography coupled with electrospray ionization tandem mass spectrometry (UPLC-ESI/MSMS). *Anal. Chim. Acta*, 750th Anniversary Volume 750, 152–160. <https://doi.org/10.1016/j.aca.2012.04.009>
- Amos-Landgraf, J.M., Kwong, L.N., Kendzierski, C.M., Reichelderfer, M., Torrealba, J., Weichert, J., Haag, J.D., Chen, K.-S., Waller, J.L., Gould, M.N., Dove, W.F., 2007. A target-selected Apc-mutant rat kindred enhances the modeling of familial human colon cancer. *Proc. Natl. Acad. Sci. U. S. A.* 104, 4036–4041. <https://doi.org/10.1073/pnas.0611690104>
- Andreeva, V.A., Deschamps, V., Salanave, B., Castetbon, K., Verdot, C., Kesse-Guyot, E., Hercberg, S., 2016. Comparison of Dietary Intakes Between a Large Online Cohort Study (Etude NutriNet-Santé) and a Nationally Representative Cross-Sectional Study (Etude Nationale Nutrition Santé) in France: Addressing the Issue of Generalizability in E-Epidemiology. *Am. J. Epidemiol.* 184, 660–669. <https://doi.org/10.1093/aje/kww016>
- Andreeva, V.A., Salanave, B., Castetbon, K., Deschamps, V., Vernay, M., Kesse-Guyot, E., Hercberg, S., 2015. Comparison of the sociodemographic characteristics of the large NutriNet-Santé e-cohort with French Census data: the issue of volunteer bias revisited. *J. Epidemiol. Community Health* 69, 893–898. <https://doi.org/10.1136/jech-2014-205263>
- Andres, D.A., Young, L.E.A., Veeranki, S., Hawkinson, T.R., Levitan, B.M., He, D., Wang, C., Satin, J., Sun, R.C., 2020. Improved workflow for mass spectrometry-based metabolomics analysis of the heart. *J. Biol. Chem.* 295, 2676–2686. <https://doi.org/10.1074/jbc.RA119.011081>
- Arnold, M., Sierra, M.S., Laversanne, M., Soerjomataram, I., Jemal, A., Bray, F., 2017. Global patterns and trends in colorectal cancer incidence and mortality. *Gut* 66, 683–691. <https://doi.org/10.1136/gutjnl-2015-310912>

- Artati, A., Prehn, C., Adamski, J., 2019. LC-MS/MS-Based Metabolomics for Cell Cultures. *Methods Mol. Biol.* Clifton NJ 1994, 119–130. [https://doi.org/10.1007/978-1-4939-9477-9\\_10](https://doi.org/10.1007/978-1-4939-9477-9_10)
- Ashby, J., Tennant, R.W., 1991. Definitive relationships among chemical structure, carcinogenicity and mutagenicity for 301 chemicals tested by the U.S. NTP. *Mutat. Res.* 257, 229–306. [https://doi.org/10.1016/0165-1110\(91\)90003-e](https://doi.org/10.1016/0165-1110(91)90003-e)
- Balder, H.F., Vogel, J., Jansen, M.C.J.F., Weijenberg, M.P., van den Brandt, P.A., Westenbrink, S., van der Meer, R., Goldbohm, R.A., 2006. Heme and chlorophyll intake and risk of colorectal cancer in the Netherlands cohort study. *Cancer Epidemiol. Biomark. Prev. Publ. Am. Assoc. Cancer Res. Cosponsored Am. Soc. Prev. Oncol.* 15, 717–725. <https://doi.org/10.1158/1055-9965.EPI-05-0772>
- Baradat, M., Jouanin, I., Dalleau, S., Taché, S., Gieules, M., Debrauwer, L., Canlet, C., Huc, L., Dupuy, J., Pierre, F.H.F., Guéraud, F., 2011. 4-Hydroxy-2(E)-nonenal metabolism differs in Apc(+/+) cells and in Apc(Min/+) cells: it may explain colon cancer promotion by heme iron. *Chem. Res. Toxicol.* 24, 1984–1993. <https://doi.org/10.1021/tx2003036>
- Barbier Saint Hilaire, P., Rousseau, K., Seyer, A., Dechaumet, S., Damont, A., Junot, C., Fenaille, F., 2020. Comparative Evaluation of Data Dependent and Data Independent Acquisition Workflows Implemented on an Orbitrap Fusion for Untargeted Metabolomics. *Metabolites* 10, E158. <https://doi.org/10.3390/metabo10040158>
- Barupal, D.K., Fan, S., Fiehn, O., 2018. Integrating bioinformatics approaches for a comprehensive interpretation of metabolomics datasets. *Curr. Opin. Biotechnol.* 54, 1–9. <https://doi.org/10.1016/j.copbio.2018.01.010>
- Barzegar, F., Kamankesh, M., Mohammadi, A., 2019. Heterocyclic aromatic amines in cooked food: A review on formation, health risk-toxicology and their analytical techniques. *Food Chem.* 280, 240–254. <https://doi.org/10.1016/j.foodchem.2018.12.058>
- Bashir, S., Fezeu, L.K., Leviatan Ben-Arye, S., Yehuda, S., Reuven, E.M., Szabo de Edelenyi, F., Fellah-Hebia, I., Le Tourneau, T., Imbert-Marcille, B.M., Drouet, E.B., Touvier, M., Roussel, J.-C., Yu, H., Chen, X., Hercberg, S., Cozzi, E., Soullillou, J.-P., Galan, P., Padler-Karavani, V., 2020. Association between Neu5Gc carbohydrate and serum antibodies against it provides the molecular link to cancer: French NutriNet-Santé study. *BMC Med.* 18, 262. <https://doi.org/10.1186/s12916-020-01721-8>
- Bastide, N., Morois, S., Cadeau, C., Kangas, S., Serafini, M., Gusto, G., Dossus, L., Pierre, F.H., Clavel-Chapelon, F., Boutron-Ruault, M.-C., 2016. Heme Iron Intake, Dietary Antioxidant Capacity, and Risk of Colorectal Adenomas in a Large Cohort Study of French Women. *Cancer Epidemiol. Biomark. Prev. Publ. Am. Assoc. Cancer Res. Cosponsored Am. Soc. Prev. Oncol.* 25, 640–647. <https://doi.org/10.1158/1055-9965.EPI-15-0724>
- Bastide, N.M., Chenni, F., Audebert, M., Santarelli, R.L., Taché, S., Naud, N., Baradat, M., Jouanin, I., Surya, R., Hobbs, D.A., Kuhnle, G.G., Raymond-Letron, I., Gueraud, F., Corpet, D.E., Pierre, F.H.F., 2015. A Central Role for Heme Iron in Colon Carcinogenesis Associated with Red Meat Intake. *Cancer Res.* 75, 870–879. <https://doi.org/10.1158/0008-5472.CAN-14-2554>
- Bastide, N.M., Naud, N., Nassy, G., Vendevre, J.-L., Taché, S., Guéraud, F., Hobbs, D.A., Kuhnle, G.G., Corpet, D.E., Pierre, F.H.F., 2017a. Red Wine and Pomegranate Extracts Suppress Cured Meat Promotion of Colonic Mucin-Depleted Foci in Carcinogen-Induced Rats. *Nutr. Cancer* 69, 289–298. <https://doi.org/10.1080/01635581.2017.1263745>
- Bastide, N.M., Naud, N., Nassy, G., Vendevre, J.-L., Taché, S., Guéraud, F., Hobbs, D.A., Kuhnle, G.G., Corpet, D.E., Pierre, F.H.F., 2017b. Red Wine and Pomegranate Extracts Suppress Cured Meat Promotion of Colonic Mucin-Depleted Foci in Carcinogen-Induced Rats. *Nutr. Cancer* 69, 289–298. <https://doi.org/10.1080/01635581.2017.1263745>
- Bastide, N.M., Pierre, F.H.F., Corpet, D.E., 2011. Heme Iron from Meat and Risk of Colorectal Cancer: A Meta-analysis and a Review of the Mechanisms Involved. *Cancer Prev. Res. (Phila. Pa.)* 4, 177–184. <https://doi.org/10.1158/1940-6207.CAPR-10-0113>
- Basu, S., Duren, W., Evans, C.R., Burant, C.F., Michailidis, G., Karnovsky, A., 2017. Sparse network modeling and metscape-based visualization methods for the analysis of large-scale

- metabolomics data. *Bioinforma. Oxf. Engl.* 33, 1545–1553. <https://doi.org/10.1093/bioinformatics/btx012>
- Bechaux, J., de La Pomélie, D., Théron, L., Santé-Lhoutellier, V., Gatellier, P., 2018. Iron-catalysed chemistry in the gastrointestinal tract: Mechanisms, kinetics and consequences. A review. *Food Chem.* 268, 27–39. <https://doi.org/10.1016/j.foodchem.2018.06.018>
- Begou, O., Gika, H.G., Theodoridis, G.A., Wilson, I.D., 2018. Quality Control and Validation Issues in LC-MS Metabolomics, in: Theodoridis, G.A., Gika, H.G., Wilson, I.D. (Eds.), *Metabolic Profiling: Methods and Protocols, Methods in Molecular Biology*. Springer, New York, NY, pp. 15–26. [https://doi.org/10.1007/978-1-4939-7643-0\\_2](https://doi.org/10.1007/978-1-4939-7643-0_2)
- Beloborodova, N., Bairamov, I., Olenin, A., Shubina, V., Teplova, V., Fedotcheva, N., 2012. Effect of phenolic acids of microbial origin on production of reactive oxygen species in mitochondria and neutrophils. *J. Biomed. Sci.* 19, 89. <https://doi.org/10.1186/1423-0127-19-89>
- Beslay, M., 2021. Effet de l'anxiété sur l'association entre viande rouge et risque de cancer colorectal : une approche multidisciplinaire (These de doctorat). Toulouse, INPT.
- Bessette, E.E., Yasa, I., Dunbar, D., Wilkens, L.R., Le Marchand, L., Turesky, R.J., 2009. Biomonitoring of carcinogenic heterocyclic aromatic amines in hair: a validation study. *Chem. Res. Toxicol.* 22, 1454–1463. <https://doi.org/10.1021/tx900155f>
- Bi, H., Guo, Z., Jia, X., Liu, H., Ma, L., Xue, L., 2020. The key points in the pre-analytical procedures of blood and urine samples in metabolomics studies. *Metabolomics Off. J. Metabolomic Soc.* 16, 68. <https://doi.org/10.1007/s11306-020-01666-2>
- Bird, R.P., 1987. Observation and quantification of aberrant crypts in the murine colon treated with a colon carcinogen: preliminary findings. *Cancer Lett.* 37, 147–151. [https://doi.org/10.1016/0304-3835\(87\)90157-1](https://doi.org/10.1016/0304-3835(87)90157-1)
- Black, A.E., 2000. Critical evaluation of energy intake using the Goldberg cut-off for energy intake: basal metabolic rate. A practical guide to its calculation, use and limitations. *Int. J. Obes. Relat. Metab. Disord. J. Int. Assoc. Study Obes.* 24, 1119–1130. <https://doi.org/10.1038/sj.ijo.0801376>
- Bloch, R., Schütze, S.-E., Müller, E., Röder, S., Lehmann, I., Brack, W., Krauss, M., 2019. Non-targeted mercapturic acid screening in urine using LC-MS/MS with matrix effect compensation by postcolumn infusion of internal standard (PCI-IS). *Anal. Bioanal. Chem.* 411, 7771–7781. <https://doi.org/10.1007/s00216-019-02166-6>
- Boccard, J., Rudaz, S., 2016. Exploring Omics data from designed experiments using analysis of variance multiblock Orthogonal Partial Least Squares. *Anal. Chim. Acta* 920, 18–28. <https://doi.org/10.1016/j.aca.2016.03.042>
- Bonvallot, N., Tremblay-Franco, M., Chevrier, C., Canlet, C., Warembourg, C., Cravedi, J.-P., Cordier, S., 2013. Metabolomics tools for describing complex pesticide exposure in pregnant women in Brittany (France). *PloS One* 8, e64433. <https://doi.org/10.1371/journal.pone.0064433>
- Boudah, S., Olivier, M.-F., Aros-Calt, S., Oliveira, L., Fenaille, F., Tabet, J.-C., Junot, C., 2014. Annotation of the human serum metabolome by coupling three liquid chromatography methods to high-resolution mass spectrometry. *J. Chromatogr. B Analyt. Technol. Biomed. Life. Sci.* 966, 34–47. <https://doi.org/10.1016/j.jchromb.2014.04.025>
- Bouvard, V., Loomis, D., Guyton, K.Z., Grosse, Y., Ghissassi, F.E., Benbrahim-Tallaa, L., Guha, N., Mattock, H., Straif, K., 2015. Carcinogenicity of consumption of red and processed meat. *Lancet Oncol.* 16, 1599–1600. [https://doi.org/10.1016/S1470-2045\(15\)00444-1](https://doi.org/10.1016/S1470-2045(15)00444-1)
- Bradbury, K.E., Murphy, N., Key, T.J., 2020. Diet and colorectal cancer in UK Biobank: a prospective study. *Int. J. Epidemiol.* 49, 246–258. <https://doi.org/10.1093/ije/dyz064>
- Brambilla, G., Sciabà, L., Faggini, P., Maura, A., Marinari, U.M., Ferro, M., Esterbauer, H., 1986. Cytotoxicity, DNA fragmentation and sister-chromatid exchange in Chinese hamster ovary cells exposed to the lipid peroxidation product 4-hydroxynonenal and homologous aldehydes. *Mutat. Res. Toxicol.* 171, 169–176. [https://doi.org/10.1016/0165-1218\(86\)90051-0](https://doi.org/10.1016/0165-1218(86)90051-0)

- Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R.L., Torre, L.A., Jemal, A., 2018. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA. Cancer J. Clin.* 68, 394–424. <https://doi.org/10.3322/caac.21492>
- Broadhurst, D., Goodacre, R., Reinke, S.N., Kuligowski, J., Wilson, I.D., Lewis, M.R., Dunn, W.B., 2018. Guidelines and considerations for the use of system suitability and quality control samples in mass spectrometry assays applied in untargeted clinical metabolomic studies. *Metabolomics* 14, 72. <https://doi.org/10.1007/s11306-018-1367-3>
- Bujak, R., Dagher-Wojtkowiak, E., Kaliszan, R., Markuszewski, M.J., 2016. PLS-Based and Regularization-Based Methods for the Selection of Relevant Variables in Non-targeted Metabolomics Data. *Front. Mol. Biosci.* 3, 35. <https://doi.org/10.3389/fmolb.2016.00035>
- Burton, L., Ivosev, G., Tate, S., Impey, G., Wingate, J., Bonner, R., 2008. Instrumental and experimental effects in LC–MS-based metabolomics. *J. Chromatogr. B, Hyphenated Techniques for Global Metabolite Profiling* 871, 227–235. <https://doi.org/10.1016/j.jchromb.2008.04.044>
- Cabaton, N.J., Poupin, N., Canlet, C., Tremblay-Franco, M., Audebert, M., Cravedi, J.-P., Riu, A., Jourdan, F., Zalko, D., 2018. An Untargeted Metabolomics Approach to Investigate the Metabolic Modulations of HepG2 Cells Exposed to Low Doses of Bisphenol A and 17 $\beta$ -Estradiol. *Front. Endocrinol.* 9, 571. <https://doi.org/10.3389/fendo.2018.00571>
- Caderni, G., Femia, A.P., Giannini, A., Favuzza, A., Luceri, C., Salvadori, M., Dolara, P., 2003. Identification of mucin-depleted foci in the unsectioned colon of azoxymethane-treated rats: correlation with carcinogenesis. *Cancer Res.* 63, 2388–2392.
- Cancer Genome Atlas Network, 2012. Comprehensive molecular characterization of human colon and rectal cancer. *Nature* 487, 330–337. <https://doi.org/10.1038/nature11252>
- Carlsen, C.U., Møller, J.K.S., Skibsted, L.H., 2005. Heme-iron in lipid oxidation. *Coord. Chem. Rev., Inorganic Reaction Mechanisms, an appreciation of Henry Taube in his 90th year* 249, 485–498. <https://doi.org/10.1016/j.ccr.2004.08.028>
- Carpenter, C.E., Mahoney, A.W., 1992. Contributions of heme and nonheme iron to human nutrition. *Crit. Rev. Food Sci. Nutr.* 31, 333–367. <https://doi.org/10.1080/10408399209527576>
- Carr, P.R., Weigl, K., Jansen, L., Walter, V., Erben, V., Chang-Claude, J., Brenner, H., Hoffmeister, M., 2018. Healthy Lifestyle Factors Associated With Lower Risk of Colorectal Cancer Irrespective of Genetic Risk. *Gastroenterology* 155, 1805-1815.e5. <https://doi.org/10.1053/j.gastro.2018.08.044>
- Carvunis, A.-R., Gomez, E., Thierry-Mieg, N., Trilling, L., Vidal, M., 2009. Biologie systémique - Des concepts d'hier aux découvertes de demain. *médecine/sciences* 25, 578–584. <https://doi.org/10.1051/medsci/2009256-7578>
- Catalá, A., 2009. Lipid peroxidation of membrane phospholipids generates hydroxy-alkenals and oxidized phospholipids active in physiological and/or pathological conditions. *Chem. Phys. Lipids* 157, 1–11. <https://doi.org/10.1016/j.chemphyslip.2008.09.004>
- Cedergreen, N., 2014. Quantifying synergy: a systematic review of mixture toxicity studies within environmental toxicology. *PLoS One* 9, e96580. <https://doi.org/10.1371/journal.pone.0096580>
- Cediel, G., Olivares, M., Gaitán, D., Flores, S., Brito, A., Pizarro, F., 2012. Effect of trypsin and mucin on heme iron bioavailability in humans. *Biol. Trace Elem. Res.* 150, 37–41. <https://doi.org/10.1007/s12011-012-9483-9>
- Chaleckis, R., Meister, I., Zhang, P., Wheelock, C.E., 2019. Challenges, progress and promises of metabolite annotation for LC–MS-based metabolomics. *Curr. Opin. Biotechnol., Analytical Biotechnology* 55, 44–50. <https://doi.org/10.1016/j.copbio.2018.07.010>
- Chanana, S., Thomas, C.S., Zhang, F., Rajski, S.R., Bugni, T.S., 2020. hcapca: Automated Hierarchical Clustering and Principal Component Analysis of Large Metabolomic Datasets in R. *Metabolites* 10, E297. <https://doi.org/10.3390/metabo10070297>
- Chawade, A., Alexandersson, E., Levander, F., 2014. Normalyzer: A Tool for Rapid Evaluation of Normalization Methods for Omics Data Sets. *J. Proteome Res.* 13, 3114–3120. <https://doi.org/10.1021/pr401264n>

- Chekmeneva, E., Dos Santos Correia, G., Gómez-Romero, M., Stamler, J., Chan, Q., Elliott, P., Nicholson, J.K., Holmes, E., 2018. Ultra-Performance Liquid Chromatography-High-Resolution Mass Spectrometry and Direct Infusion-High-Resolution Mass Spectrometry for Combined Exploratory and Targeted Metabolic Profiling of Human Urine. *J. Proteome Res.* 17, 3492–3502. <https://doi.org/10.1021/acs.jproteome.8b00413>
- Chetwynd, A.J., Abdul-Sada, A., Hill, E.M., 2015. Solid-phase extraction and nanoflow liquid chromatography-nanoelectrospray ionization mass spectrometry for improved global urine metabolomics. *Anal. Chem.* 87, 1158–1165. <https://doi.org/10.1021/ac503769q>
- Cheung, W., Keski-Rahkonen, P., Assi, N., Ferrari, P., Freisling, H., Rinaldi, S., Slimani, N., Zamora-Ros, R., Rundle, M., Frost, G., Gibbons, H., Carr, E., Brennan, L., Cross, A.J., Pala, V., Panico, S., Sacerdote, C., Palli, D., Tumino, R., Kühn, T., Kaaks, R., Boeing, H., Floegel, A., Mancini, F., Boutron-Ruault, M.-C., Baglietto, L., Trichopoulou, A., Naska, A., Orfanos, P., Scalbert, A., 2017. A metabolomic study of biomarkers of meat and fish intake. *Am. J. Clin. Nutr.* 105, 600–608. <https://doi.org/10.3945/ajcn.116.146639>
- Chevolleau, S., Bouville, A., Debrauwer, L., 2020. Development and validation of a modified QuEChERS protocol coupled to UHPLC-APCI-MS/MS for the simple and rapid quantification of 16 heterocyclic aromatic amines in cooked beef. *Food Chem.* 316, 126327. <https://doi.org/10.1016/j.foodchem.2020.126327>
- Chevolleau, S., Nogueir-Meireles, M.-H., Jouanin, I., Naud, N., Pierre, F., Gueraud, F., Debrauwer, L., 2018. Development and validation of an ultra high performance liquid chromatography-electrospray tandem mass spectrometry method using selective derivatisation, for the quantification of two reactive aldehydes produced by lipid peroxidation, HNE (4-hydroxy-2(E)-nonenal) and HHE (4-hydroxy-2(E)-hexenal) in faecal water. *J. Chromatogr. B Analyt. Technol. Biomed. Life. Sci.* 1083, 171–179. <https://doi.org/10.1016/j.jchromb.2018.03.002>
- Chindarkar, N.S., Wakefield, M.R., Stone, J.A., Fitzgerald, R.L., 2014. Liquid chromatography high-resolution TOF analysis: investigation of MSE for broad-spectrum drug screening. *Clin. Chem.* 60, 1115–1125. <https://doi.org/10.1373/clinchem.2014.222976>
- Cillard, J., Cillard, P., 2006. Mécanismes de la peroxydation lipidique et des anti-oxydations. *Ol. Corps Gras Lipides* 13, 24–29. <https://doi.org/10.1051/ocl.2006.6666>
- Clapper, M.L., Chang, W.-C.L., Cooper, H.S., 2020. Dysplastic Aberrant Crypt Foci: Biomarkers of Early Colorectal Neoplasia and Response to Preventive Intervention. *Cancer Prev. Res. Phila. Pa* 13, 229–240. <https://doi.org/10.1158/1940-6207.CAPR-19-0316>
- Constante, M., Fragoso, G., Calvé, A., Samba-Mondonga, M., Santos, M.M., 2017a. Dietary Heme Induces Gut Dysbiosis, Aggravates Colitis, and Potentiates the Development of Adenomas in Mice. *Front. Microbiol.* 8, 1809. <https://doi.org/10.3389/fmicb.2017.01809>
- Constante, M., Fragoso, G., Calvé, A., Samba-Mondonga, M., Santos, M.M., 2017b. Dietary Heme Induces Gut Dysbiosis, Aggravates Colitis, and Potentiates the Development of Adenomas in Mice. *Front. Microbiol.* 8. <https://doi.org/10.3389/fmicb.2017.01809>
- Corpet, D.E., Pierre, F., 2005. How good are rodent models of carcinogenesis in predicting efficacy in humans? A systematic review and meta-analysis of colon chemoprevention in rats, mice and men. *Eur. J. Cancer, Cancer Chemoprevention - An Update on a Novel and Exciting Field of Oncology* 41, 1911–1922. <https://doi.org/10.1016/j.ejca.2005.06.006>
- Cottret, L., Frainay, C., Chazalviel, M., Cabanettes, F., Gloaguen, Y., Camenen, E., Merlet, B., Heux, S., Portais, J.-C., Poupin, N., Vinson, F., Jourdan, F., 2018. MetExplore: collaborative edition and exploration of metabolic networks. *Nucleic Acids Res.* 46, W495–W502. <https://doi.org/10.1093/nar/gky301>
- Craig, C.L., Marshall, A.L., Sjöström, M., Bauman, A.E., Booth, M.L., Ainsworth, B.E., Pratt, M., Ekelund, U., Yngve, A., Sallis, J.F., Oja, P., 2003. International physical activity questionnaire: 12-country reliability and validity. *Med. Sci. Sports Exerc.* 35, 1381–1395. <https://doi.org/10.1249/01.MSS.0000078924.61453.FB>
- Creek, D.J., Dunn, W.B., Fiehn, O., Griffin, J.L., Hall, R.D., Lei, Z., Mistrik, R., Neumann, S., Schymanski, E.L., Sumner, L.W., Trengove, R., Wolfender, J.-L., 2014. Metabolite identification: are you

- sure? And how do your peers gauge your confidence? *Metabolomics* 10, 350–353. <https://doi.org/10.1007/s11306-014-0656-8>
- Criscuolo, A., Zeller, M., Fedorova, M., 2020. Evaluation of Lipid In-Source Fragmentation on Different Orbitrap-based Mass Spectrometers. *J. Am. Soc. Mass Spectrom.* 31, 463–466. <https://doi.org/10.1021/jasms.9b00061>
- Cross, A.J., Ferrucci, L.M., Risch, A., Graubard, B.I., Ward, M.H., Park, Y., Hollenbeck, A.R., Schatzkin, A., Sinha, R., 2010. A large prospective study of meat consumption and colorectal cancer risk: an investigation of potential mechanisms underlying this association. *Cancer Res.* 70, 2406–2414. <https://doi.org/10.1158/0008-5472.CAN-09-3929>
- Cross, A.J., Pollock, J.R.A., Bingham, S.A., 2003. Haem, not protein or inorganic iron, is responsible for endogenous intestinal N-nitrosation arising from red meat. *Cancer Res.* 63, 2358–2360.
- David, L.A., Maurice, C.F., Carmody, R.N., Gootenberg, D.B., Button, J.E., Wolfe, B.E., Ling, A.V., Devlin, A.S., Varma, Y., Fischbach, M.A., Biddinger, S.B., Dutton, R.J., Turnbaugh, P.J., 2014. Diet rapidly and reproducibly alters the human gut microbiome. *Nature* 505, 559–563. <https://doi.org/10.1038/nature12820>
- Davis, M.E., Lisowyj, M.P., Zhou, L., Wisecarver, J.L., Gulizia, J.M., Shostrom, V.K., Naud, N., Corpet, D.E., Mirvish, S.S., 2012. Induction of colonic aberrant crypts in mice by feeding apparent N-nitroso compounds derived from hot dogs. *Nutr. Cancer* 64, 342–349. <https://doi.org/10.1080/01635581.2012.650777>
- de La Pomélie, D., Santé-Lhoutellier, V., Gatellier, P., 2017. Mechanisms and kinetics of tryptophan N-nitrosation in a gastro-intestinal model. *Food Chem.* 218, 487–495. <https://doi.org/10.1016/j.foodchem.2016.08.131>
- de Moura, N.A., Caetano, B.F.R., Bidinotto, L.T., Rodrigues, M.A.M., Barbisan, L.F., 2019a. Dietary heme promotes colonic preneoplastic lesions and DNA damage but not tumor development in a medium-term model of colon carcinogenesis in rats. *Mutat. Res. Toxicol. Environ. Mutagen.* <https://doi.org/10.1016/j.mrgentox.2019.07.006>
- de Moura, N.A., Caetano, B.F.R., Bidinotto, L.T., Rodrigues, M.A.M., Barbisan, L.F., 2019b. Dietary heme promotes colonic preneoplastic lesions and DNA damage but not tumor development in a medium-term model of colon carcinogenesis in rats. *Mutat. Res. Genet. Toxicol. Environ. Mutagen.* 846, 403076. <https://doi.org/10.1016/j.mrgentox.2019.07.006>
- de Vogel, J., Jonker-Termont, D.S.M.L., Katan, M.B., van der Meer, R., 2005a. Natural chlorophyll but not chlorophyllin prevents heme-induced cytotoxic and hyperproliferative effects in rat colon. *J. Nutr.* 135, 1995–2000. <https://doi.org/10.1093/jn/135.8.1995>
- de Vogel, J., Jonker-Termont, D.S.M.L., Katan, M.B., van der Meer, R., 2005b. Natural chlorophyll but not chlorophyllin prevents heme-induced cytotoxic and hyperproliferative effects in rat colon. *J. Nutr.* 135, 1995–2000. <https://doi.org/10.1093/jn/135.8.1995>
- Dearfield, K.L., Abernathy, C.O., Ottley, M.S., Brantner, J.H., Hayes, P.F., 1988. Acrylamide: its metabolism, developmental and reproductive effects, genotoxicity, and carcinogenicity. *Mutat. Res.* 195, 45–77. [https://doi.org/10.1016/0165-1110\(88\)90015-2](https://doi.org/10.1016/0165-1110(88)90015-2)
- Defossez, E., Bourquin, J., von Reuss, S., Rasmann, S., Glauser, G., 2021. Eight key rules for successful data-dependent acquisition in mass spectrometry-based metabolomics. *Mass Spectrom. Rev.* <https://doi.org/10.1002/mas.21715>
- DellaValle, C.T., Xiao, Q., Yang, G., Shu, X.O., Aschebrook-Kilfoy, B., Zheng, W., Li, H.L., Ji, B.-T., Rothman, N., Chow, W.-H., Gao, Y.-T., Ward, M.H., 2014. Dietary nitrate and nitrite intake and risk of colorectal cancer in the Shanghai Women’s Health Study. *Int. J. Cancer* 134, 2917–2926. <https://doi.org/10.1002/ijc.28612>
- Demeyer, D., Mertens, B., De Smet, S., Ulens, M., 2016. Mechanisms Linking Colorectal Cancer to the Consumption of (Processed) Red Meat: A Review. *Crit. Rev. Food Sci. Nutr.* 56, 2747–2766. <https://doi.org/10.1080/10408398.2013.873886>
- Di Iorio, B.R., Rocchetti, M.T., De Angelis, M., Cosola, C., Marzocco, S., Di Micco, L., di Bari, I., Accetturo, M., Vacca, M., Gobetti, M., Di Iorio, M., Bellasi, A., Gesualdo, L., 2019. Nutritional Therapy Modulates Intestinal Microbiota and Reduces Serum Levels of Total and Free Indoxyl Sulfate

- and P-Cresyl Sulfate in Chronic Kidney Disease (Medika Study). *J. Clin. Med.* 8. <https://doi.org/10.3390/jcm8091424>
- Diallo, A., Deschasaux, M., Latino-Martel, P., Hercberg, S., Galan, P., Fassier, P., Allès, B., Guéraud, F., Pierre, F.H., Touvier, M., 2018. Red and processed meat intake and cancer risk: Results from the prospective NutriNet-Santé cohort study. *Int. J. Cancer* 142, 230–237. <https://doi.org/10.1002/ijc.31046>
- Dieterle, F., Ross, A., Schlotterbeck, G., Senn, H., 2006. Probabilistic Quotient Normalization as Robust Method to Account for Dilution of Complex Biological Mixtures. Application in <sup>1</sup> H NMR Metabonomics. *Anal. Chem.* 78, 4281–4290. <https://doi.org/10.1021/ac051632c>
- Djuricic, I., Calder, P.C., 2021. Beneficial Outcomes of Omega-6 and Omega-3 Polyunsaturated Fatty Acids on Human Health: An Update for 2021. *Nutrients* 13, 2421. <https://doi.org/10.3390/nu13072421>
- Doll, R., Peto, R., 1981. The causes of cancer: quantitative estimates of avoidable risks of cancer in the United States today. *J. Natl. Cancer Inst.* 66, 1191–1308.
- Dunn, W.B., Broadhurst, D., Begley, P., Zelena, E., Francis-McIntyre, S., Anderson, N., Brown, M., Knowles, J.D., Halsall, A., Haselden, J.N., Nicholls, A.W., Wilson, I.D., Kell, D.B., Goodacre, R., Human Serum Metabolome (HUSERMET) Consortium, 2011. Procedures for large-scale metabolic profiling of serum and plasma using gas chromatography and liquid chromatography coupled to mass spectrometry. *Nat. Protoc.* 6, 1060–1083. <https://doi.org/10.1038/nprot.2011.335>
- Dunn, W.B., Broadhurst, D., Ellis, D.I., Brown, M., Halsall, A., O'Hagan, S., Spasic, I., Tseng, A., Kell, D.B., 2008. A GC-TOF-MS study of the stability of serum and urine metabolomes during the UK Biobank sample collection and preparation protocols. *Int. J. Epidemiol.* 37 Suppl 1, i23-30. <https://doi.org/10.1093/ije/dym281>
- Dunn, W.B., Ellis, David.I., 2005. Metabolomics: Current analytical platforms and methodologies. *TrAC Trends Anal. Chem.* 24, 285–294. <https://doi.org/10.1016/j.trac.2004.11.021>
- Dunn, W.B., Wilson, I.D., Nicholls, A.W., Broadhurst, D., 2012. The importance of experimental design and QC samples in large-scale and MS-driven untargeted metabolomic studies of humans [WWW Document]. <http://dx.doi.org/10.4155/bio.12.204>. <https://doi.org/10.4155/bio.12.204>
- Eaton, D., Klaassen, C.D., 2008. Principles of toxicology in Casarett and Doull's Toxicology: The Basis Science of Poisons. *Princ. Toxicol. Casarett Doulls Toxicol. Basic Sci. Poisons* 11–43.
- Eckl, P.M., Ortner, A., Esterbauer, H., 1993. Genotoxic properties of 4-hydroxyalkenals and analogous aldehydes. *Mutat. Res. Mol. Mech. Mutagen.* 290, 183–192. [https://doi.org/10.1016/0027-5107\(93\)90158-C](https://doi.org/10.1016/0027-5107(93)90158-C)
- Edwards, L.M., 2017. Metabolic systems biology: a brief primer. *J. Physiol.* 595, 2849–2855. <https://doi.org/10.1113/JP272275>
- Ehrlich, A.M., Pacheco, A.R., Henrick, B.M., Taft, D., Xu, G., Huda, M.N., Mishchuk, D., Goodson, M.L., Slupsky, C., Barile, D., Lebrilla, C.B., Stephensen, C.B., Mills, D.A., Raybould, H.E., 2020. Indole-3-lactic acid associated with Bifidobacterium-dominated microbiota significantly decreases inflammation in intestinal epithelial cells. *BMC Microbiol.* 20, 357. <https://doi.org/10.1186/s12866-020-02023-y>
- Eichholzer, M., Rohrmann, S., Barbir, A., Hermann, S., Teucher, B., Kaaks, R., Linseisen, J., 2012. Polymorphisms in heterocyclic aromatic amines metabolism-related genes are associated with colorectal adenoma risk. *Int. J. Mol. Epidemiol. Genet.* 3, 96–106.
- Einaudi, L., Courbiere, B., Tassistro, V., Prevot, C., Sari-Minodier, I., Orsiere, T., Perrin, J., 2014. In vivo exposure to benzo(a)pyrene induces significant DNA damage in mouse oocytes and cumulus cells. *Hum. Reprod. Oxf. Engl.* 29, 548–554. <https://doi.org/10.1093/humrep/det439>
- Emwas, A.-H., Roy, R., McKay, R.T., Tenori, L., Saccenti, E., Gowda, G.A.N., Raftery, D., Alahmari, F., Jaremko, L., Jaremko, M., Wishart, D.S., 2019. NMR Spectroscopy for Metabolomics Research. *Metabolites* 9, E123. <https://doi.org/10.3390/metabo9070123>

- Emwas, A.-H.M., 2015. The Strengths and Weaknesses of NMR Spectroscopy and Mass Spectrometry with Particular Focus on Metabolomics Research, in: Bjerrum, J.T. (Ed.), *Metabonomics: Methods and Protocols*, Methods in Molecular Biology. Springer, New York, NY, pp. 161–193. [https://doi.org/10.1007/978-1-4939-2377-9\\_13](https://doi.org/10.1007/978-1-4939-2377-9_13)
- Erdrich, J., Zhang, X., Giovannucci, E., Willett, W., 2015. Proportion of Colon Cancer Attributable to Lifestyle in a Cohort of US Women. *Cancer Causes Control CCC* 26, 1271–1279. <https://doi.org/10.1007/s10552-015-0619-z>
- Escudié, F., Auer, L., Bernard, M., Mariadassou, M., Cauquil, L., Vidal, K., Maman, S., Hernandez-Raquet, G., Combes, S., Pascal, G., 2018. FROGS: Find, Rapidly, OTUs with Galaxy Solution. *Bioinformatics* 34, 1287–1294. <https://doi.org/10.1093/bioinformatics/btx791>
- Espejo-Herrera, N., Gràcia-Lavedan, E., Boldo, E., Aragonés, N., Pérez-Gómez, B., Pollán, M., Molina, A.J., Fernández, T., Martín, V., La Vecchia, C., Bosetti, C., Tavani, A., Polesel, J., Serraino, D., Gómez Acebo, I., Altzibar, J.M., Ardanaz, E., Burgui, R., Pisa, F., Fernández-Tardón, G., Tardón, A., Peiró, R., Navarro, C., Castaño-Vinyals, G., Moreno, V., Righi, E., Aggazzotti, G., Basagaña, X., Nieuwenhuijsen, M., Kogevinas, M., Villanueva, C.M., 2016. Colorectal cancer risk and nitrate exposure through drinking water and diet. *Int. J. Cancer* 139, 334–346. <https://doi.org/10.1002/ijc.30083>
- Esterbauer, H., Zollern, H., 1989. Methods for determination of aldehydic lipid peroxidation products. *Free Radic. Biol. Med.* 7, 197–203. [https://doi.org/10.1016/0891-5849\(89\)90015-4](https://doi.org/10.1016/0891-5849(89)90015-4)
- Farvid, M.S., Sidahmed, E., Spence, N.D., Mante Angua, K., Rosner, B.A., Barnett, J.B., 2021. Consumption of red meat and processed meat and cancer incidence: a systematic review and meta-analysis of prospective studies. *Eur. J. Epidemiol.* <https://doi.org/10.1007/s10654-021-00741-9>
- Fearon, E.R., 2011. Molecular genetics of colorectal cancer. *Annu. Rev. Pathol.* 6, 479–507. <https://doi.org/10.1146/annurev-pathol-011110-130235>
- Fearon, E.R., Vogelstein, B., 1990. A genetic model for colorectal tumorigenesis. *Cell* 61, 759–767. [https://doi.org/10.1016/0092-8674\(90\)90186-I](https://doi.org/10.1016/0092-8674(90)90186-I)
- Femia, A.P., Caderni, G., 2008. Rodent Models of Colon Carcinogenesis for the Study of Chemopreventive Activity of Natural Products. *Planta Med.* 74, 1602–1607. <https://doi.org/10.1055/s-2008-1074577>
- Femia, A.P., Dolara, P., Giannini, A., Salvadori, M., Biggeri, A., Caderni, G., 2007. Frequent mutation of Apc gene in rat colon tumors and mucin-depleted foci, preneoplastic lesions in experimental colon carcinogenesis. *Cancer Res.* 67, 445–449. <https://doi.org/10.1158/0008-5472.CAN-06-3861>
- Femia, A.P., Paulsen, J.E., Dolara, P., Alexander, J., Caderni, G., 2008. Correspondence between flat aberrant crypt foci and mucin-depleted foci in rodent colon carcinogenesis. *Anticancer Res.* 28, 3771–3775.
- Ferlay, J., Colombet, M., Soerjomataram, I., Mathers, C., Parkin, D.M., Piñeros, M., Znaor, A., Bray, F., 2019. Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods. *Int. J. Cancer* 144, 1941–1953. <https://doi.org/10.1002/ijc.31937>
- Fernández-Peralbo, M.A., Luque de Castro, M.D., 2012. Preparation of urine samples prior to targeted or untargeted metabolomics mass-spectrometry analysis. *TrAC Trends Anal. Chem.* 41, 75–85. <https://doi.org/10.1016/j.trac.2012.08.011>
- Ferrucci, L.M., Sinha, R., Huang, W.-Y., Berndt, S.I., Katki, H.A., Schoen, R.E., Hayes, R.B., Cross, A.J., 2012. Meat consumption and the risk of incident distal colon and rectal adenoma. *Br. J. Cancer* 106, 608–616. <https://doi.org/10.1038/bjc.2011.549>
- Fidler, M.M., Soerjomataram, I., Bray, F., 2016. A global view on cancer incidence and national levels of the human development index. *Int. J. Cancer* 139, 2436–2446. <https://doi.org/10.1002/ijc.30382>
- Fiehn, O., 2002. Metabolomics — the link between genotypes and phenotypes, in: Town, C. (Ed.), *Functional Genomics*. Springer Netherlands, Dordrecht, pp. 155–171. [https://doi.org/10.1007/978-94-010-0448-0\\_11](https://doi.org/10.1007/978-94-010-0448-0_11)



- Flisikowska, T., Merkl, C., Landmann, M., Eser, S., Rezaei, N., Cui, X., Kurome, M., Zakhartchenko, V., Kessler, B., Wieland, H., Rottmann, O., Schmid, R.M., Schneider, G., Kind, A., Wolf, E., Saur, D., Schnieke, A., 2012. A porcine model of familial adenomatous polyposis. *Gastroenterology* 143, 1173-1175.e7. <https://doi.org/10.1053/j.gastro.2012.07.110>
- Frainay, C., Jourdan, F., 2017. Computational methods to identify metabolic sub-networks based on metabolomic profiles. *Brief. Bioinform.* 18, 43–56. <https://doi.org/10.1093/bib/bbv115>
- Frigerio, G., Mercadante, R., Campo, L., Polledri, E., Boniardi, L., Olgiati, L., Missineo, P., Fustinoni, S., 2020a. Urinary biomonitoring of subjects with different smoking habits. Part I: Profiling mercapturic acids. *Toxicol. Lett.* 327, 48–57. <https://doi.org/10.1016/j.toxlet.2020.03.010>
- Frigerio, G., Mercadante, R., Campo, L., Polledri, E., Boniardi, L., Olgiati, L., Missineo, P., Nash, W.J., Dunn, W.B., Fustinoni, S., 2020b. Urinary biomonitoring of subjects with different smoking habits. Part II: an untargeted metabolomic approach and the comparison with the targeted measurement of mercapturic acids. *Toxicol. Lett.* 329, 56–66. <https://doi.org/10.1016/j.toxlet.2020.03.020>
- Fu, Z., Shrubsole, M.J., Smalley, W.E., Wu, H., Chen, Z., Shyr, Y., Ness, R.M., Zheng, W., 2011. Association of meat intake and meat-derived mutagen exposure with the risk of colorectal polyps by histologic type. *Cancer Prev. Res. Phila. Pa* 4, 1686–1697. <https://doi.org/10.1158/1940-6207.CAPR-11-0191>
- Fuqua, B.K., Vulpe, C.D., Anderson, G.J., 2012. Intestinal iron absorption. *J. Trace Elem. Med. Biol., IX ISTERH Conference. Trace elements in health and disease: Essentiality, toxicity* 26, 115–119. <https://doi.org/10.1016/j.jtemb.2012.03.015>
- Gagnebin, Y., Jaques, D.A., Rudaz, S., de Seigneux, S., Boccard, J., Ponte, B., 2020. Exploring blood alterations in chronic kidney disease and haemodialysis using metabolomics. *Sci. Rep.* 10, 19502. <https://doi.org/10.1038/s41598-020-76524-1>
- Gagnebin, Y., Tonoli, D., Lescuyer, P., Ponte, B., de Seigneux, S., Martin, P.-Y., Schappler, J., Boccard, J., Rudaz, S., 2017. Metabolomic analysis of urine samples by UHPLC-QTOF-MS: Impact of normalization strategies. *Anal. Chim. Acta* 955, 27–35. <https://doi.org/10.1016/j.aca.2016.12.029>
- Gagnière, J., Raisch, J., Veziat, J., Barnich, N., Bonnet, R., Buc, E., Bringer, M.-A., Pezet, D., Bonnet, M., 2016. Gut microbiota imbalance and colorectal cancer. *World J. Gastroenterol.* 22, 501–518. <https://doi.org/10.3748/wjg.v22.i2.501>
- Gaitán, D., Flores, S., Saavedra, P., Miranda, C., Olivares, M., Arredondo, M., López de Romaña, D., Lönnerdal, B., Pizarro, F., 2011. Calcium Does Not Inhibit the Absorption of 5 Milligrams of Nonheme or Heme Iron at Doses Less Than 800 Milligrams in Nonpregnant Women. *J. Nutr.* 141, 1652–1656. <https://doi.org/10.3945/jn.111.138651>
- Galiatsatos, P., Foulkes, W.D., 2006. Familial Adenomatous Polyposis. *Am. J. Gastroenterol.* 101, 385–398.
- Gao, H., Liu, S., 2017. Role of uremic toxin indoxyl sulfate in the progression of cardiovascular disease. *Life Sci.* 185, 23–29. <https://doi.org/10.1016/j.lfs.2017.07.027>
- Gechev, T.S., Benina, M., Obata, T., Tohge, T., Sujeeth, N., Minkov, I., Hille, J., Temanni, M.-R., Marriott, A.S., Bergström, E., Thomas-Oates, J., Antonio, C., Mueller-Roeber, B., Schippers, J.H.M., Fernie, A.R., Toneva, V., 2013. Molecular mechanisms of desiccation tolerance in the resurrection glacial relic *Haberlea rhodopensis*. *Cell. Mol. Life Sci. CMLS* 70, 689–709. <https://doi.org/10.1007/s00018-012-1155-6>
- Geiger, T., Cox, J., Mann, M., 2010. Proteomics on an Orbitrap benchtop mass spectrometer using all-ion fragmentation. *Mol. Cell. Proteomics MCP* 9, 2252–2261. <https://doi.org/10.1074/mcp.M110.001537>
- Giacomini, F., Le Corguillé, G., Monsoor, M., Landi, M., Pericard, P., Pétéra, M., Duperier, C., Tremblay-Franco, M., Martin, J.-F., Jacob, D., Goulitquer, S., Thévenot, E.A., Caron, C., 2015. Workflow4Metabolomics: a collaborative research infrastructure for computational metabolomics. *Bioinformatics* 31, 1493–1495. <https://doi.org/10.1093/bioinformatics/btu813>

- Gibis, M., 2016. Heterocyclic Aromatic Amines in Cooked Meat Products: Causes, Formation, Occurrence, and Risk Assessment. *Compr. Rev. Food Sci. Food Saf.* 15, 269–302. <https://doi.org/10.1111/1541-4337.12186>
- Giesbertz, P., Brandl, B., Lee, Y.-M., Hauner, H., Daniel, H., Skurk, T., 2020. Specificity, Dose Dependency, and Kinetics of Markers of Chicken and Beef Intake Using Targeted Quantitative LC-MS/MS: A Human Intervention Trial. *Mol. Nutr. Food Res.* 64, e1900921. <https://doi.org/10.1002/mnfr.201900921>
- Gika, H.G., Theodoridis, G.A., Wilson, I.D., 2008. Liquid chromatography and ultra-performance liquid chromatography–mass spectrometry fingerprinting of human urine: Sample stability under different handling and storage conditions for metabonomics studies. *J. Chromatogr. A*, 31st International Symposium on High Performance Liquid Phase Separations and Related Techniques 1189, 314–322. <https://doi.org/10.1016/j.chroma.2007.10.066>
- Gika, H.G., Zisi, C., Theodoridis, G., Wilson, I.D., 2016. Protocol for quality control in metabolic profiling of biological fluids by U(H)PLC-MS. *J. Chromatogr. B* 1008, 15–25. <https://doi.org/10.1016/j.jchromb.2015.10.045>
- Gillet, L.C., Navarro, P., Tate, S., Röst, H., Selevsek, N., Reiter, L., Bonner, R., Aebersold, R., 2012. Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis. *Mol. Cell. Proteomics MCP* 11, O111.016717. <https://doi.org/10.1074/mcp.O111.016717>
- Giskeødegård, G.F., Davies, S.K., Revell, V.L., Keun, H., Skene, D.J., 2015. Diurnal rhythms in the human urine metabolome during sleep and total sleep deprivation. *Sci. Rep.* 5, 14843. <https://doi.org/10.1038/srep14843>
- Glade, M.J., 1999. Food, nutrition, and the prevention of cancer: a global perspective. American Institute for Cancer Research/World Cancer Research Fund, American Institute for Cancer Research, 1997. *Nutr. Burbank Los Angel. Cty. Calif* 15, 523–526. [https://doi.org/10.1016/s0899-9007\(99\)00021-0](https://doi.org/10.1016/s0899-9007(99)00021-0)
- Godzien, J., Ciborowski, M., Angulo, S., Ruperez, F.J., Paz Martínez, M., Señorans, F.J., Cifuentes, A., Ibañez, E., Barbas, C., 2011. Metabolomic Approach with LC-QTOF to Study the Effect of a Nutraceutical Treatment on Urine of Diabetic Rats. *J. Proteome Res.* 10, 837–844. <https://doi.org/10.1021/pr100993x>
- Gonçalves-Dias, C., Morello, J., Semedo, V., Correia, M.J., Coelho, N.R., Monteiro, E.C., Antunes, A.M.M., Pereira, S.A., 2019. The Mercapturomic Profile of Health and Non-Communicable Diseases. *High-Throughput* 8, 10. <https://doi.org/10.3390/ht8020010>
- González, A., Gálvez, N., Martín, J., Reyes, F., Pérez-Victoria, I., Dominguez-Vera, J.M., 2017. Identification of the key excreted molecule by *Lactobacillus fermentum* related to host iron absorption. *Food Chem.* 228, 374–380. <https://doi.org/10.1016/j.foodchem.2017.02.008>
- Goupy, P., Vulcain, E., Caris-Veyrat, C., Dangles, O., 2007. Dietary antioxidants as inhibitors of the heme-induced peroxidation of linoleic acid: Mechanism of action and synergism. *Free Radic. Biol. Med.* 43, 933–946. <https://doi.org/10.1016/j.freeradbiomed.2007.06.013>
- Grant, D.M., 1991. Detoxification pathways in the liver. *J. Inherit. Metab. Dis.* 14, 421–430. <https://doi.org/10.1007/BF01797915>
- Gromski, P.S., Muhamadali, H., Ellis, D.I., Xu, Y., Correa, E., Turner, M.L., Goodacre, R., 2015. A tutorial review: Metabolomics and partial least squares-discriminant analysis—a marriage of convenience or a shotgun wedding. *Anal. Chim. Acta* 879, 10–23. <https://doi.org/10.1016/j.aca.2015.02.012>
- Guéraud, F., 2017. 4-Hydroxynonenal metabolites and adducts in pre-carcinogenic conditions and cancer. *Free Radic. Biol. Med.*, 4-Hydroxynonenal and Related Lipid Peroxidation Products 111, 196–208. <https://doi.org/10.1016/j.freeradbiomed.2016.12.025>
- Guéraud, F., Peiro, G., Bernard, H., Alary, J., Créminon, C., Debrauwer, L., Rathahao, E., Drumare, M.-F., Canlet, C., Wal, J.-M., Bories, G., 2006. Enzyme immunoassay for a urinary metabolite of 4-hydroxynonenal as a marker of lipid peroxidation. *Free Radic. Biol. Med.* 40, 54–62. <https://doi.org/10.1016/j.freeradbiomed.2005.08.011>

- Guéraud, F., Taché, S., Steghens, J.-P., Milkovic, L., Borovic-Sunjic, S., Zarkovic, N., Gaultier, E., Naud, N., Héliès-Toussaint, C., Pierre, F., Priymenko, N., 2015a. Dietary polyunsaturated fatty acids and heme iron induce oxidative stress biomarkers and a cancer promoting environment in the colon of rats. *Free Radic. Biol. Med.* 83, 192–200. <https://doi.org/10.1016/j.freeradbiomed.2015.02.023>
- Guéraud, F., Taché, S., Steghens, J.-P., Milkovic, L., Borovic-Sunjic, S., Zarkovic, N., Gaultier, E., Naud, N., Héliès-Toussaint, C., Pierre, F., Priymenko, N., 2015b. Dietary polyunsaturated fatty acids and heme iron induce oxidative stress biomarkers and a cancer promoting environment in the colon of rats. *Free Radic. Biol. Med.* 83, 192–200. <https://doi.org/10.1016/j.freeradbiomed.2015.02.023>
- Gunter, M.J., Probst-Hensch, N.M., Cortessis, V.K., Kulldorff, M., Haile, R.W., Sinha, R., 2005. Meat intake, cooking-related mutagens and risk of colorectal adenoma in a sigmoidoscopy-based case-control study. *Carcinogenesis* 26, 637–642. <https://doi.org/10.1093/carcin/bgh350>
- Gupta, A.K., Pretlow, T.P., Schoen, R.E., 2007. Aberrant crypt foci: what we know and what we need to know. *Clin. Gastroenterol. Hepatol. Off. Clin. Pract. J. Am. Gastroenterol. Assoc.* 5, 526–533. <https://doi.org/10.1016/j.cgh.2007.02.014>
- Hallberg, L., Brune, M., Erlandsson, M., Sandberg, A.S., Rossander-Hultén, L., 1991. Calcium: effect of different amounts on nonheme- and heme-iron absorption in humans. *Am. J. Clin. Nutr.* 53, 112–119. <https://doi.org/10.1093/ajcn/53.1.112>
- Hanna, P.E., Anders, M.W., 2019. The mercapturic acid pathway. *Crit. Rev. Toxicol.* 49, 819–929. <https://doi.org/10.1080/10408444.2019.1692191>
- Hao, L., Wang, J., Page, D., Asthana, S., Zetterberg, H., Carlsson, C., Okonkwo, O.C., Li, L., 2018. Comparative Evaluation of MS-based Metabolomics Software and Its Application to Preclinical Alzheimer’s Disease. *Sci. Rep.* 8, 9291. <https://doi.org/10.1038/s41598-018-27031-x>
- Harrill, J.A., Viant, M.R., Yauk, C.L., Sachana, M., Gant, T.W., Auerbach, S.S., Beger, R.D., Bouhifd, M., O’Brien, J., Burgoon, L., Caiment, F., Carpi, D., Chen, T., Chorley, B.N., Colbourne, J., Corvi, R., Debrauwer, L., O’Donovan, C., Ebbels, T.M.D., Ekman, D.R., Faulhammer, F., Gribaldo, L., Hilton, G.M., Jones, S.P., Kende, A., Lawson, T.N., Leite, S.B., Leonards, P.E.G., Luijten, M., Martin, A., Moussa, L., Rudaz, S., Schmitz, O., Sobanski, T., Strauss, V., Vaccari, M., Vijay, V., Weber, R.J.M., Williams, A.J., Williams, A., Thomas, R.S., Whelan, M., 2021. Progress towards an OECD reporting framework for transcriptomics and metabolomics in regulatory toxicology. *Regul. Toxicol. Pharmacol. RTP* 125, 105020. <https://doi.org/10.1016/j.yrtph.2021.105020>
- Heinen, C.D., 2010. Genotype to phenotype: analyzing the effects of inherited mutations in colorectal cancer families. *Mutat. Res.* 693, 32–45. <https://doi.org/10.1016/j.mrfmmm.2009.09.004>
- Helfer, A.G., Michely, J.A., Weber, A.A., Meyer, M.R., Maurer, H.H., 2017. LC-HR-MS/MS standard urine screening approach: Pros and cons of automated on-line extraction by turbulent flow chromatography versus dilute-and-shoot and comparison with established urine precipitation. *J. Chromatogr. B, Extraction and Sample Preparation Techniques in Bioanalysis* 1043, 138–149. <https://doi.org/10.1016/j.jchromb.2016.06.036>
- Helmus, D.S., Thompson, C.L., Zelenskiy, S., Tucker, T.C., Li, L., 2013. Red meat-derived heterocyclic amines increase risk of colon cancer: a population-based case-control study. *Nutr. Cancer* 65, 1141–1150. <https://doi.org/10.1080/01635581.2013.834945>
- Hemeryck, L.Y., Rombouts, C., De Paepe, E., Vanhaecke, L., 2018. DNA adduct profiling of in vitro colonic meat digests to map red vs. white meat genotoxicity. *Food Chem. Toxicol. Int. J. Publ. Br. Ind. Biol. Res. Assoc.* 115, 73–87. <https://doi.org/10.1016/j.fct.2018.02.032>
- Hemeryck, L.Y., Van Hecke, T., Vossen, E., De Smet, S., Vanhaecke, L., 2017. DNA adductomics to study the genotoxic effects of red meat consumption with and without added animal fat in rats. *Food Chem.* 230, 378–387. <https://doi.org/10.1016/j.foodchem.2017.02.129>
- Hemmer, S., Manier, S.K., Fischmann, S., Westphal, F., Wagmann, L., Meyer, M.R., 2020. Comparison of Three Untargeted Data Processing Workflows for Evaluating LC-HRMS Metabolomics Data. *Metabolites* 10, 378. <https://doi.org/10.3390/metabo10090378>

- Hercberg, S., Castetbon, K., Czernichow, S., Malon, A., Mejean, C., Kesse, E., Touvier, M., Galan, P., 2010. The Nutrinet-Santé Study: a web-based prospective study on the relationship between nutrition and health and determinants of dietary patterns and nutritional status. *BMC Public Health* 10, 242. <https://doi.org/10.1186/1471-2458-10-242>
- Hinchman, C.A., Ballatori, N., 1994. Glutathione conjugation and conversion to mercapturic acids can occur as an intrahepatic process. *J. Toxicol. Environ. Health* 41, 387–409. <https://doi.org/10.1080/15287399409531852>
- Honikel, K.-O., 2008. The use and control of nitrate and nitrite for the processing of meat products. *Meat Sci.* 78, 68–76. <https://doi.org/10.1016/j.meatsci.2007.05.030>
- Hooda, J., Shah, A., Zhang, L., 2014. Heme, an essential nutrient from dietary proteins, critically impacts diverse physiological and pathological processes. *Nutrients* 6, 1080–1102. <https://doi.org/10.3390/nu6031080>
- Huerta-Yépez, S., Tirado-Rodriguez, A.B., Hankinson, O., 2016. Role of diets rich in omega-3 and omega-6 in the development of cancer. *Bol. Med. Hosp. Infant. Mex.* 73, 446–456. <https://doi.org/10.1016/j.bmhmx.2016.11.001>
- Hughes, R., Cross, A.J., Pollock, J.R., Bingham, S., 2001. Dose-dependent effect of dietary meat on endogenous colonic N-nitrosation. *Carcinogenesis* 22, 199–202. <https://doi.org/10.1093/carcin/22.1.199>
- Ijssennagger, N., Belzer, C., Hooiveld, G.J., Dekker, J., van Mil, S.W.C., Müller, M., Kleerebezem, M., van der Meer, R., 2015. Gut microbiota facilitates dietary heme-induced epithelial hyperproliferation by opening the mucus barrier in colon. *Proc. Natl. Acad. Sci.* 112, 10038–10043. <https://doi.org/10.1073/pnas.1507645112>
- Ijssennagger, N., Derrien, M., Doorn, G.M. van, Rijnierse, A., Bogert, B. van den, Müller, M., Dekker, J., Kleerebezem, M., Meer, R. van der, 2012a. Dietary Heme Alters Microbiota and Mucosa of Mouse Colon without Functional Changes in Host-Microbe Cross-Talk. *PLOS ONE* 7, e49868. <https://doi.org/10.1371/journal.pone.0049868>
- Ijssennagger, N., Rijnierse, A., de Wit, N.J.W., Boeschoten, M.V., Dekker, J., Schonewille, A., Muller, M., van der Meer, R., 2013. Dietary heme induces acute oxidative stress, but delayed cytotoxicity and compensatory hyperproliferation in mouse colon. *Carcinogenesis* 34, 1628–1635. <https://doi.org/10.1093/carcin/bgt084>
- Ijssennagger, N., Rijnierse, A., Wit, N. de, Jonker-Termont, D., Dekker, J., Müller, M., Meer, R. van der, 2012b. Dietary haem stimulates epithelial cell turnover by downregulating feedback inhibitors of proliferation in murine colon. *Gut* 61, 1041–1049. <https://doi.org/10.1136/gutjnl-2011-300239>
- Itoh, K., Chiba, T., Takahashi, S., Ishii, T., Igarashi, K., Katoh, Y., Oyake, T., Hayashi, N., Satoh, K., Hatayama, I., Yamamoto, M., Nabeshima, Y., 1997. An Nrf2/Small Maf Heterodimer Mediates the Induction of Phase II Detoxifying Enzyme Genes through Antioxidant Response Elements. *Biochem. Biophys. Res. Commun.* 236, 313–322. <https://doi.org/10.1006/bbrc.1997.6943>
- Iyanagi, T., 2007. Molecular Mechanism of Phase I and Phase II Drug-Metabolizing Enzymes: Implications for Detoxification, in: *International Review of Cytology*. Academic Press, pp. 35–112. [https://doi.org/10.1016/S0074-7696\(06\)60002-8](https://doi.org/10.1016/S0074-7696(06)60002-8)
- Jacques, C., Jamin, E.L., Perdu, E., Duplan, H., Mavon, A., Zalko, D., Debrauwer, L., 2010. Characterisation of B(a)P metabolites formed in an ex vivo pig skin model using three complementary analytical methods. *Anal. Bioanal. Chem.* 396, 1691–1701. <https://doi.org/10.1007/s00216-009-3389-1>
- Jägerstad, M., Skog, K., Arvidsson, P., Solyakov, A., 1998. Chemistry, formation and occurrence of genotoxic heterocyclic amines identified in model systems and cooked foods. *Z. Für Leb. - Forsch. A* 207, 419–427. <https://doi.org/10.1007/s002170050355>
- Jamin, E.L., Bonvallot, N., Tremblay-Franco, M., Cravedi, J.-P., Chevrier, C., Cordier, S., Debrauwer, L., 2014. Untargeted profiling of pesticide metabolites by LC-HRMS: an exposomics tool for human exposure evaluation. *Anal. Bioanal. Chem.* 406, 1149–1161. <https://doi.org/10.1007/s00216-013-7136-2>

- Jamin, E.L., Costantino, R., Mervant, L., Martin, J.-F., Jouanin, I., Blas-Y-Estrada, F., Guéraud, F., Debrauwer, L., 2020. Global Profiling of Toxicologically Relevant Metabolites in Urine: Case Study of Reactive Aldehydes. *Anal. Chem.* 92, 1746–1754. <https://doi.org/10.1021/acs.analchem.9b03146>
- Jat, P.S., Noble, M.D., Ataliotis, P., Tanaka, Y., Yannoutsos, N., Larsen, L., Kioussis, D., 1991. Direct derivation of conditionally immortal cell lines from an H-2Kb-tsA58 transgenic mouse. *Proc. Natl. Acad. Sci. U. S. A.* 88, 5096–5100.
- Jeyakumar, A., Dissabandara, L., Gopalan, V., 2017. A critical overview on the biological and molecular features of red and processed meat in colorectal carcinogenesis. *J. Gastroenterol.* 52, 407–418. <https://doi.org/10.1007/s00535-016-1294-x>
- Ju, T., Kong, J.Y., Stothard, P., Willing, B.P., 2019. Defining the role of Parasutterella, a previously uncharacterized member of the core gut microbiota. *ISME J.* 13, 1520–1534. <https://doi.org/10.1038/s41396-019-0364-5>
- Kabat, G.C., Miller, A.B., Jain, M., Rohan, T.E., 2007. A cohort study of dietary iron and heme iron intake and risk of colorectal cancer in women. *Br. J. Cancer* 97, 118–122. <https://doi.org/10.1038/sj.bjc.6603837>
- Kapoore, R.V., Vaidyanathan, S., 2018. Quenching for Microalgal Metabolomics: A Case Study on the Unicellular Eukaryotic Green Alga *Chlamydomonas reinhardtii*. *Metabolites* 8, 72. <https://doi.org/10.3390/metabo8040072>
- Karlstetter, D., Mally, A., 2020. Biomonitoring of heat-induced food contaminants: Quantitative analysis of furan dependent glutathione- and lysine-adducts in rat urine as putative biomarkers of exposure. *Food Chem. Toxicol.* 143, 111562. <https://doi.org/10.1016/j.fct.2020.111562>
- Kato, T., Migita, H., Ohgaki, H., Sato, S., Takayama, S., Sugimura, T., 1989. Induction of tumors in the Zymbal gland, oral cavity, colon, skin and mammary gland of F344 rats by a mutagenic compound, 2-amino-3,4-dimethylimidazo[4,5-f]quinoline. *Carcinogenesis* 10, 601–603. <https://doi.org/10.1093/carcin/10.3.601>
- Keller, J., Baradat, M., Jouanin, I., Debrauwer, L., Guéraud, F., 2015. “Twin peaks”: Searching for 4-hydroxynonenal urinary metabolites after oral administration in rats. *Redox Biol.* 4, 136–148. <https://doi.org/10.1016/j.redox.2014.12.016>
- Keller, J., Chevolleau, S., Noguer-Meireles, M.-H., Pujos-Guillot, E., Delosière, M., Chantelauze, C., Joly, C., Blas-y-Estrada, F., Jouanin, I., Durand, D., Pierre, F., Debrauwer, L., Theodorou, V., Guéraud, F., 2020a. Heme-Iron-Induced Production of 4-Hydroxynonenal in Intestinal Lumen May Have Extra-Intestinal Consequences Through Protein-Adduct Formation. *Antioxidants* 9, 1293. <https://doi.org/10.3390/antiox9121293>
- Keller, J., Chevolleau, S., Noguer-Meireles, M.-H., Pujos-Guillot, E., Delosière, M., Chantelauze, C., Joly, C., Blas-Y-Estrada, F., Jouanin, I., Durand, D., Pierre, F., Debrauwer, L., Theodorou, V., Guéraud, F., 2020b. Heme-Iron-Induced Production of 4-Hydroxynonenal in Intestinal Lumen May Have Extra-Intestinal Consequences through Protein-Adduct Formation. *Antioxid. Basel Switz.* 9, E1293. <https://doi.org/10.3390/antiox9121293>
- Kellert, M., Scholz, K., Wagner, S., Dekant, W., Völkel, W., 2006. Quantitation of mercapturic acids from acrylamide and glycidamide in human urine using a column switching tool with two trap columns and electrospray tandem mass spectrometry. *J. Chromatogr. A* 1131, 58–66. <https://doi.org/10.1016/j.chroma.2006.07.026>
- Keum, N., Giovannucci, E.L., 2017. Epidemiology of Colorectal Cancer, in: Loda, M., Mucci, L.A., Mittelstadt, M.L., Van Hemelrijck, M., Cotter, M.B. (Eds.), *Pathology and Epidemiology of Cancer*. Springer International Publishing, Cham, pp. 391–407. [https://doi.org/10.1007/978-3-319-35153-7\\_21](https://doi.org/10.1007/978-3-319-35153-7_21)
- Khamis, M.M., Adamko, D.J., El-Aneed, A., 2017. Mass spectrometric based approaches in urine metabolomics and biomarker discovery. *Mass Spectrom. Rev.* 36, 115–134. <https://doi.org/10.1002/mas.21455>
- Khan, M.R., Busquets, R., Santos, F.J., Puignou, L., 2008. New method for the analysis of heterocyclic amines in meat extracts using pressurised liquid extraction and liquid chromatography–

- tandem mass spectrometry. *J. Chromatogr. A* 1194, 155–160. <https://doi.org/10.1016/j.chroma.2008.04.058>
- Khodorova, N.V., Rutledge, D.N., Oberli, M., Mathiron, D., Marcelo, P., Benamouzig, R., Tomé, D., Gaudichon, C., Pilard, S., 2019. Urinary Metabolomics Profiles Associated to Bovine Meat Ingestion in Humans. *Mol. Nutr. Food Res.* 63, 1700834. <https://doi.org/10.1002/mnfr.201700834>
- Kim, K., Mall, C., Taylor, S.L., Hitchcock, S., Zhang, C., Wettersten, H.I., Jones, A.D., Chapman, A., Weiss, R.H., 2014. Mealtime, Temporal, and Daily Variability of the Human Urinary and Plasma Metabolomes in a Tightly Controlled Environment. *PLOS ONE* 9, e86223. <https://doi.org/10.1371/journal.pone.0086223>
- Kirwan, J.A., Broadhurst, D.I., Davidson, R.L., Viant, M.R., 2013. Characterising and correcting batch variation in an automated direct infusion mass spectrometry (DIMS) metabolomics workflow. *Anal. Bioanal. Chem.* 405, 5147–5157. <https://doi.org/10.1007/s00216-013-6856-7>
- Kotronoulas, A., Pizarro, N., Serra, A., Robledo, P., Joglar, J., Rubió, L., Hernaéz, A., Tormos, C., Motilva, M.J., Fitó, M., Covas, M.-I., Solà, R., Farré, M., Saez, G., de la Torre, R., 2013. Dose-dependent metabolic disposition of hydroxytyrosol and formation of mercapturates in rats. *Pharmacol. Res.* 77, 47–56. <https://doi.org/10.1016/j.phrs.2013.09.001>
- Kowalczyk, M., Orłowski, M., Klepacki, Ł., Zinkiewicz, K., Kurpiewski, W., Kaczerska, D., Pesta, W., Zieliński, E., Siermontowski, P., 2020. Rectal aberrant crypt foci (ACF) as a predictor of benign and malignant neoplastic lesions in the large intestine. *BMC Cancer* 20, 133. <https://doi.org/10.1186/s12885-020-6590-4>
- Kuhnle, G.G.C., Story, G.W., Reda, T., Mani, A.R., Moore, K.P., Lunn, J.C., Bingham, S.A., 2007. Diet-induced endogenous formation of nitroso compounds in the GI tract. *Free Radic. Biol. Med.* 43, 1040–1047. <https://doi.org/10.1016/j.freeradbiomed.2007.03.011>
- Kuiper, H.C., Miranda, C.L., Sowell, J.D., Stevens, J.F., 2008. Mercapturic Acid Conjugates of 4-Hydroxy-2-nonenal and 4-Oxo-2-nonenal Metabolites Are in Vivo Markers of Oxidative Stress \*. *J. Biol. Chem.* 283, 17131–17138. <https://doi.org/10.1074/jbc.M802797200>
- Kuiper, H.C., Stevens, J.F., 2011. LC-MS/MS quantitation of mercapturic acid conjugates of lipid peroxidation products as markers of oxidative stress. *Curr. Protoc. Toxicol.* Chapter 17, Unit17.14.2. <https://doi.org/10.1002/0471140856.tx1714s45>
- Kumar, S., Bandyopadhyay, U., 2005. Free heme toxicity and its detoxification systems in human. *Toxicol. Lett.* 157, 175–188. <https://doi.org/10.1016/j.toxlet.2005.03.004>
- Lapins, M., Eklund, M., Spjuth, O., Prusis, P., Wikberg, J.E., 2008. Proteochemometric modeling of HIV protease susceptibility. *BMC Bioinformatics* 9, 181. <https://doi.org/10.1186/1471-2105-9-181>
- Larsson, S.C., Orsini, N., 2014. Red meat and processed meat consumption and all-cause mortality: a meta-analysis. *Am. J. Epidemiol.* 179, 282–289. <https://doi.org/10.1093/aje/kwt261>
- Lassale, C., Castetbon, K., Laporte, F., Camilleri, G.M., Deschamps, V., Vernay, M., Faure, P., Hercberg, S., Galan, P., Kesse-Guyot, E., 2015. Validation of a Web-based, self-administered, non-consecutive-day dietary record tool against urinary biomarkers. *Br. J. Nutr.* 113, 953–962. <https://doi.org/10.1017/S0007114515000057>
- Lassale, C., Castetbon, K., Laporte, F., Deschamps, V., Vernay, M., Camilleri, G.M., Faure, P., Hercberg, S., Galan, P., Kesse-Guyot, E., 2016. Correlations between Fruit, Vegetables, Fish, Vitamins, and Fatty Acids Estimated by Web-Based Nonconsecutive Dietary Records and Respective Biomarkers of Nutritional Status. *J. Acad. Nutr. Diet.* 116, 427-438.e5. <https://doi.org/10.1016/j.jand.2015.09.017>
- Lassale, C., Péneau, S., Touvier, M., Julia, C., Galan, P., Hercberg, S., Kesse-Guyot, E., 2013. Validity of web-based self-reported weight and height: results of the Nutrinet-Santé study. *J. Med. Internet Res.* 15, e152. <https://doi.org/10.2196/jmir.2575>
- Le Leu, R.K., Young, G.P., Hu, Y., Winter, J., Conlon, M.A., 2013. Dietary red meat aggravates dextran sulfate sodium-induced colitis in mice whereas resistant starch attenuates inflammation. *Dig. Dis. Sci.* 58, 3475–3482. <https://doi.org/10.1007/s10620-013-2844-1>

- Lee, T.-H., Park, D., Kim, Y.J., Lee, I., Kim, S., Oh, C.-T., Kim, J.-Y., Yang, J., Jo, S.-K., 2020. *Lactobacillus salivarius* BP121 prevents cisplatin-induced acute kidney injury by inhibition of uremic toxins such as indoxyl sulfate and p-cresol sulfate via alleviating dysbiosis. *Int. J. Mol. Med.* 45, 1130–1140. <https://doi.org/10.3892/ijmm.2020.4495>
- Leeming, E.R., Johnson, A.J., Spector, T.D., Le Roy, C.I., 2019. Effect of Diet on the Gut Microbiota: Rethinking Intervention Duration. *Nutrients* 11, E2862. <https://doi.org/10.3390/nu11122862>
- Leu, R.K.L., Winter, J.M., Christophersen, C.T., Young, G.P., Humphreys, K.J., Hu, Y., Gratz, S.W., Miller, R.B., Topping, D.L., Bird, A.R., Conlon, M.A., 2015. Butyrylated starch intake can prevent red meat-induced O6-methyl-2-deoxyguanosine adducts in human rectal tissue: a randomised clinical trial. *Br. J. Nutr.* 114, 220–230. <https://doi.org/10.1017/S0007114515001750>
- Lewin, M.H., Bailey, N., Bandaletova, T., Bowman, R., Cross, A.J., Pollock, J., Shuker, D.E.G., Bingham, S.A., 2006. Red Meat Enhances the Colonic Formation of the DNA Adduct O6-Carboxymethyl Guanine: Implications for Colorectal Cancer Risk. *Cancer Res.* 66, 1859–1865. <https://doi.org/10.1158/0008-5472.CAN-05-2237>
- Lewis, M.R., Pearce, J.T.M., Spagou, K., Green, M., Dona, A.C., Yuen, A.H.Y., David, M., Berry, D.J., Chappell, K., Horneffer-van der Sluis, V., Shaw, R., Lovestone, S., Elliott, P., Shockcor, J., Lindon, J.C., Cloarec, O., Takats, Z., Holmes, E., Nicholson, J.K., 2016. Development and Application of Ultra-Performance Liquid Chromatography-TOF MS for Precision Large Scale Urinary Metabolic Phenotyping. *Anal. Chem.* 88, 9004–9013. <https://doi.org/10.1021/acs.analchem.6b01481>
- Li, Z., Lu, Y., Guo, Y., Cao, H., Wang, Q., Shui, W., 2018. Comprehensive evaluation of untargeted metabolomics data processing software in feature detection, quantification and discriminating marker selection. *Anal. Chim. Acta* 1029, 50–57. <https://doi.org/10.1016/j.aca.2018.05.001>
- Lorrain, B., Dangles, O., Loonis, M., Armand, M., Dufour, C., 2012. Dietary Iron-Initiated Lipid Oxidation and Its Inhibition by Polyphenols in Gastric Conditions. *J. Agric. Food Chem.* 60, 9074–9081. <https://doi.org/10.1021/jf302348s>
- Lunn, J.C., Kuhnle, G., Mai, V., Frankenfeld, C., Shuker, D.E.G., Glen, R.C., Goodman, J.M., Pollock, J.R.A., Bingham, S.A., 2007. The effect of haem in red and processed meat on the endogenous formation of N-nitroso compounds in the upper gastrointestinal tract. *Carcinogenesis* 28, 685–690. <https://doi.org/10.1093/carcin/bgl192>
- Luo, H., Zhang, N.-Q., Huang, J., Zhang, X., Feng, X.-L., Pan, Z.-Z., Chen, Y.-M., Fang, Y.-J., Zhang, C.-X., 2019. Different forms and sources of iron in relation to colorectal cancer risk: a case-control study in China. *Br. J. Nutr.* 121, 735–747. <https://doi.org/10.1017/S0007114519000023>
- Lynch, H.T., de la Chapelle, A., 2003. Hereditary colorectal cancer. *N. Engl. J. Med.* 348, 919–932. <https://doi.org/10.1056/NEJMra012242>
- Lynch, H.T., Smyrk, T., 1996. Hereditary nonpolyposis colorectal cancer (Lynch syndrome). An updated review. *Cancer* 78, 1149–1167. [https://doi.org/10.1002/\(SICI\)1097-0142\(19960915\)78:6<1149::AID-CNCR1>3.0.CO;2-5](https://doi.org/10.1002/(SICI)1097-0142(19960915)78:6<1149::AID-CNCR1>3.0.CO;2-5)
- Lynch, S.R., Dassenko, S.A., Morck, T.A., Beard, J.L., Cook, J.D., 1985. Soy protein products and heme iron absorption in humans. *Am. J. Clin. Nutr.* 41, 13–20. <https://doi.org/10.1093/ajcn/41.1.13>
- Ma, Q., Kim, E.-Y., Lindsay, E.A., Han, O., 2011. Bioactive Dietary Polyphenols Inhibit Heme Iron Absorption in A Dose-Dependent Manner in Human Intestinal Caco-2 cells. *J. Food Sci.* 76, H143–H150. <https://doi.org/10.1111/j.1750-3841.2011.02184.x>
- Mafra, D., Borges, N.A., Cardozo, L.F.M. de F., Anjos, J.S., Black, A.P., Moraes, C., Bergman, P., Lindholm, B., Stenvinkel, P., 2018. Red meat intake in chronic kidney disease patients: Two sides of the coin. *Nutrition* 46, 26–32. <https://doi.org/10.1016/j.nut.2017.08.015>
- Mahé, F., Rognes, T., Quince, C., de Vargas, C., Dunthorn, M., 2014. Swarm: robust and fast clustering method for amplicon-based studies. *PeerJ* 2, e593. <https://doi.org/10.7717/peerj.593>
- Mahieu, N.G., Patti, G.J., 2017. Systems-Level Annotation of a Metabolomics Data Set Reduces 25 000 Features to Fewer than 1000 Unique Metabolites. *Anal. Chem.* 89, 10397–10406. <https://doi.org/10.1021/acs.analchem.7b02380>

- Majou, D., Christieans, S., 2018. Mechanisms of the bactericidal effects of nitrate and nitrite in cured meats. *Meat Sci.* 145, 273–284. <https://doi.org/10.1016/j.meatsci.2018.06.013>
- Markley, J.L., Brüschweiler, R., Edison, A.S., Eghbalnia, H.R., Powers, R., Raftery, D., Wishart, D.S., 2017. The future of NMR-based metabolomics. *Curr. Opin. Biotechnol., Analytical biotechnology* 43, 34–40. <https://doi.org/10.1016/j.copbio.2016.08.001>
- Martin, O.C.B., Naud, N., Taché, S., Debrauwer, L., Chevolleau, S., Dupuy, J., Chantelauze, C., Durand, D., Pujos-Guillot, E., Blas-Y-Estrada, F., Urbano, C., Kuhnle, G.G.C., Santé-Lhoutellier, V., Sayd, T., Viala, D., Blot, A., Meunier, N., Schlich, P., Attaix, D., Guéraud, F., Scislawski, V., Corpet, D.E., Pierre, F.H.F., 2018. Targeting Colon Luminal Lipid Peroxidation Limits Colon Carcinogenesis Associated with Red Meat Consumption. *Cancer Prev. Res. Phila. Pa* 11, 569–580. <https://doi.org/10.1158/1940-6207.CAPR-17-0361>
- Martin, O.C.B., Olier, M., Ellero-Simatos, S., Naud, N., Dupuy, J., Huc, L., Taché, S., Graillot, V., Levêque, M., Bézirard, V., Héliès-Toussaint, C., Estrada, F.B.Y., Tondereau, V., Lippi, Y., Naylies, C., Peyriga, L., Canlet, C., Davila, A.M., Blachier, F., Ferrier, L., Boutet-Robinet, E., Guéraud, F., Théodorou, V., Pierre, F.H.F., 2019a. Haem iron reshapes colonic luminal environment: impact on mucosal homeostasis and microbiome through aldehyde formation. *Microbiome* 7, 72. <https://doi.org/10.1186/s40168-019-0685-7>
- Martin, O.C.B., Olier, M., Ellero-Simatos, S., Naud, N., Dupuy, J., Huc, L., Taché, S., Graillot, V., Levêque, M., Bézirard, V., Héliès-Toussaint, C., Estrada, F.B.Y., Tondereau, V., Lippi, Y., Naylies, C., Peyriga, L., Canlet, C., Davila, A.M., Blachier, F., Ferrier, L., Boutet-Robinet, E., Guéraud, F., Théodorou, V., Pierre, F.H.F., 2019b. Haem iron reshapes colonic luminal environment: impact on mucosal homeostasis and microbiome through aldehyde formation. *Microbiome* 7, 72. <https://doi.org/10.1186/s40168-019-0685-7>
- Martins, M.C.M., Caldana, C., Wolf, L.D., de Abreu, L.G.F., 2018. The Importance of Experimental Design, Quality Assurance, and Control in Plant Metabolomics Experiments. *Methods Mol. Biol. Clifton NJ* 1778, 3–17. [https://doi.org/10.1007/978-1-4939-7819-9\\_1](https://doi.org/10.1007/978-1-4939-7819-9_1)
- Massey, R.C., Key, P.E., Mallett, A.K., Rowland, I.R., 1988. An investigation of the endogenous formation of apparent total N-nitroso compounds in conventional microflora and germ-free rats. *Food Chem. Toxicol. Int. J. Publ. Br. Ind. Biol. Res. Assoc.* 26, 595–600. [https://doi.org/10.1016/0278-6915\(88\)90230-x](https://doi.org/10.1016/0278-6915(88)90230-x)
- Mathias, P.I., B'Hymer, C., 2014. A survey of liquid chromatographic–mass spectrometric analysis of mercapturic acid biomarkers in occupational and environmental exposure monitoring. *J. Chromatogr. B, Analysis of Eicosanoids, Amino Acids, Organic Acids and microRNA* 964, 136–145. <https://doi.org/10.1016/j.jchromb.2014.02.057>
- Mervant, L., Tremblay-Franco, M., Jamin, E.L., Kesse-Guyot, E., Galan, P., Martin, J.-F., Guéraud, F., Debrauwer, L., 2021. Osmolality-based normalization enhances statistical discrimination of untargeted metabolomic urine analysis: results from a comparative study. *Metabolomics Off. J. Metabolomic Soc.* 17, 2. <https://doi.org/10.1007/s11306-020-01758-z>
- Michelmann, K., Silveira, J.A., Ridgeway, M.E., Park, M.A., 2015. Fundamentals of trapped ion mobility spectrometry. *J. Am. Soc. Mass Spectrom.* 26, 14–24. <https://doi.org/10.1007/s13361-014-0999-4>
- Milreu, P.V., Klein, C.C., Cottret, L., Acuña, V., Birmelé, E., Borassi, M., Junot, C., Marchetti-Spaccamela, A., Marino, A., Stougie, L., Jourdan, F., Crescenzi, P., Lacroix, V., Sagot, M.-F., 2014. Telling metabolic stories to explore metabolomics data: a case study on the yeast response to cadmium exposure. *Bioinforma. Oxf. Engl.* 30, 61–70. <https://doi.org/10.1093/bioinformatics/btt597>
- Mirvish, S.S., 1995. Role of N-nitroso compounds (NOC) and N-nitrosation in etiology of gastric, esophageal, nasopharyngeal and bladder cancer and contribution to cancer of known exposures to NOC. *Cancer Lett.* 93, 17–48. [https://doi.org/10.1016/0304-3835\(95\)03786-V](https://doi.org/10.1016/0304-3835(95)03786-V)
- Mistretta, V., Charlier, C., 2013. [Biomonitoring of exposure to chemical products in environmental and workplace toxicology]. *Ann. Biol. Clin. (Paris)* 71, 257–267. <https://doi.org/10.1684/abc.2013.0820>



- Morales, R., Guerrero, L., Aguiar, A.P.S., Guàrdia, M.D., Gou, P., 2013. Factors affecting dry-cured ham consumer acceptability. *Meat Sci.* 95, 652–657. <https://doi.org/10.1016/j.meatsci.2013.05.012>
- Muthusamy, S., Peng, C., Ng, J.C., 2018. Genotoxicity evaluation of multi-component mixtures of polyaromatic hydrocarbons (PAHs), arsenic, cadmium, and lead using flow cytometry based micronucleus test in HepG2 cells. *Mutat. Res. Genet. Toxicol. Environ. Mutagen.* 827, 9–18. <https://doi.org/10.1016/j.mrgentox.2018.01.002>
- Nascimento-Gonçalves, E., Mendes, B.A.L., Silva-Reis, R., Faustino-Rocha, A.I., Gama, A., Oliveira, P.A., 2021. Animal Models of Colorectal Cancer: From Spontaneous to Genetically Engineered Models and Their Applications. *Vet. Sci.* 8, 59. <https://doi.org/10.3390/vetsci8040059>
- Näthke, I., 2004. APC at a glance. *J. Cell Sci.* 117, 4873–4875. <https://doi.org/10.1242/jcs.01313>
- Nguyen, L.H., Goel, A., Chung, D.C., 2020. Pathways of Colorectal Carcinogenesis. *Gastroenterology* 158, 291–302. <https://doi.org/10.1053/j.gastro.2019.08.059>
- Nicholson, J.K., Lindon, J.C., Holmes, E., 1999. “Metabonomics”: understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data. *Xenobiotica Fate Foreign Compd. Biol. Syst.* 29, 1181–1189. <https://doi.org/10.1080/004982599238047>
- Obrenovich, M.E., Tima, M., Polinkovsky, A., Zhang, R., Emancipator, S.N., Donskey, C.J., 2017. Targeted Metabolomics Analysis Identifies Intestinal Microbiota-Derived Urinary Biomarkers of Colonization Resistance in Antibiotic-Treated Mice. *Antimicrob. Agents Chemother.* 61. <https://doi.org/10.1128/AAC.00477-17>
- Ohgaki, H., Hasegawa, H., Suenaga, M., Sato, S., Takayama, S., Sugimura, T., 1987. Carcinogenicity in mice of a mutagenic compound, 2-amino-3,8-dimethylimidazo[4,5-f]quinoxaline (MeIQx) from cooked foods. *Carcinogenesis* 8, 665–668. <https://doi.org/10.1093/carcin/8.5.665>
- Ohgaki, H., Matsukura, N., Morino, K., Kawachi, T., Sugimura, T., Takayama, S., 1984. Carcinogenicity in mice of mutagenic compounds from glutamic acid and soybean globulin pyrolysates. *Carcinogenesis* 5, 815–819. <https://doi.org/10.1093/carcin/5.6.815>
- Ohkawa, H., Ohishi, N., Yagi, K., 1979. Assay for lipid peroxides in animal tissues by thiobarbituric acid reaction. *Anal. Biochem.* 95, 351–358. [https://doi.org/10.1016/0003-2697\(79\)90738-3](https://doi.org/10.1016/0003-2697(79)90738-3)
- Oliveira, R.C., Abrantes, A.M., Tralhão, J.G., Botelho, M.F., 2020. The role of mouse models in colorectal cancer research—The need and the importance of the orthotopic models. *Anim. Models Exp. Med.* 3, 1–8. <https://doi.org/10.1002/ame2.12102>
- Ollberding, N.J., Wilkens, L.R., Henderson, B.E., Kolonel, L.N., Le Marchand, L., 2012. Meat consumption, heterocyclic amines and colorectal cancer risk: the Multiethnic Cohort Study. *Int. J. Cancer* 131, E1125–1133. <https://doi.org/10.1002/ijc.27546>
- Oz, F., Kotan, G., 2016. Effects of different cooking methods and fat levels on the formation of heterocyclic aromatic amines in various fishes. *Food Control* 67, 216–224. <https://doi.org/10.1016/j.foodcont.2016.03.013>
- Oz, F., Yuzer, M.O., 2016. The effects of cooking on wire and stone barbecue at different cooking levels on the formation of heterocyclic aromatic amines and polycyclic aromatic hydrocarbons in beef steak. *Food Chem.* 203, 59–66. <https://doi.org/10.1016/j.foodchem.2016.02.041>
- Paglia, G., Astarita, G., 2017. Metabolomics and lipidomics using traveling-wave ion mobility mass spectrometry. *Nat. Protoc.* 12, 797–813. <https://doi.org/10.1038/nprot.2017.013>
- Pan, A., Sun, Q., Bernstein, A.M., Schulze, M.B., Manson, J.E., Willett, W.C., Hu, F.B., 2011. Red meat consumption and risk of type 2 diabetes: 3 cohorts of US adults and an updated meta-analysis. *Am. J. Clin. Nutr.* 94, 1088–1096. <https://doi.org/10.3945/ajcn.111.018978>
- Panuwet, P., Nguyen, J.V., Kuklenyik, P., Udunka, S.O., Needham, L.L., Barr, D.B., 2008. Quantification of atrazine and its metabolites in urine by on-line solid-phase extraction–high-performance liquid chromatography–tandem mass spectrometry. *Anal. Bioanal. Chem.* 391, 1931–1939. <https://doi.org/10.1007/s00216-008-2102-0>
- Papadimitropoulos, M.-E.P., Vasilopoulou, C.G., Maga-Nteve, C., Klapa, M.I., 2018. Untargeted GC-MS Metabolomics, in: Theodoridis, G.A., Gika, H.G., Wilson, I.D. (Eds.), *Metabolic Profiling:*

- Methods and Protocols, *Methods in Molecular Biology*. Springer, New York, NY, pp. 133–147. [https://doi.org/10.1007/978-1-4939-7643-0\\_9](https://doi.org/10.1007/978-1-4939-7643-0_9)
- Patel, K.P., Luo, F.J.-G., Plummer, N.S., Hostetter, T.H., Meyer, T.W., 2012. The production of p-cresol sulfate and indoxyl sulfate in vegetarians versus omnivores. *Clin. J. Am. Soc. Nephrol. CJASN* 7, 982–988. <https://doi.org/10.2215/CJN.12491211>
- Patti, G.J., Yanes, O., Siuzdak, G., 2012. Metabolomics: the apogee of the omic trilogy. *Nat. Rev. Mol. Cell Biol.* 13, 263–269. <https://doi.org/10.1038/nrm3314>
- Pawlak, R., Berger, J., Hines, I., 2016. Iron Status of Vegetarian Adults: A Review of Literature. *Am. J. Lifestyle Med.* 12, 486–498. <https://doi.org/10.1177/1559827616682933>
- Perez De Souza, L., Alseekh, S., Brotman, Y., Fernie, A.R., 2020. Network-based strategies in metabolomics data analysis and interpretation: from molecular networking to biological interpretation. *Expert Rev. Proteomics* 17, 243–255. <https://doi.org/10.1080/14789450.2020.1766975>
- Perše, M., Cerar, A., 2011. Morphological and Molecular Alterations in 1,2 Dimethylhydrazine and Azoxymethane Induced Colon Carcinogenesis in Rats. *J. Biomed. Biotechnol.* 2011, 473964. <https://doi.org/10.1155/2011/473964>
- Philippe, C., Szabo de Edelenyi, F., Naudon, L., Druesne-Pecollo, N., Hercberg, S., Kesse-Guyot, E., Latino-Martel, P., Galan, P., Rabot, S., 2021. Relation between Mood and the Host-Microbiome Co-Metabolite 3-Indoxylsulfate: Results from the Observational Prospective NutriNet-Santé Study. *Microorganisms* 9, 716. <https://doi.org/10.3390/microorganisms9040716>
- Pierre, F., Freeman, A., Taché, S., Van der Meer, R., Corpet, D.E., 2004a. Beef meat and blood sausage promote the formation of azoxymethane-induced mucin-depleted foci and aberrant crypt foci in rat colons. *J. Nutr.* 134, 2711–2716. <https://doi.org/10.1093/jn/134.10.2711>
- Pierre, F., Freeman, A., Taché, S., Van der Meer, R., Corpet, D.E., 2004b. Beef meat and blood sausage promote the formation of azoxymethane-induced mucin-depleted foci and aberrant crypt foci in rat colons. *J. Nutr.* 134, 2711–2716. <https://doi.org/10.1093/jn/134.10.2711>
- Pierre, F., Peiro, G., Taché, S., Cross, A.J., Bingham, S.A., Gasc, N., Gottardi, G., Corpet, D.E., Guéraud, F., 2006. New Marker of Colon Cancer Risk Associated with Heme Intake: 1,4-Dihydroxynonane Mercapturic Acid. *Cancer Epidemiol. Prev. Biomark.* 15, 2274–2279. <https://doi.org/10.1158/1055-9965.EPI-06-0085>
- Pierre, F., Santarelli, R., Taché, S., Guéraud, F., Corpet, D.E., 2008a. Beef meat promotion of dimethylhydrazine-induced colorectal carcinogenesis biomarkers is suppressed by dietary calcium. *Br. J. Nutr.* 99, 1000–1006. <https://doi.org/10.1017/S0007114507843558>
- Pierre, F., Santarelli, R., Taché, S., Guéraud, F., Corpet, D.E., 2008b. Beef meat promotion of dimethylhydrazine-induced colorectal carcinogenesis biomarkers is suppressed by dietary calcium. *Br. J. Nutr.* 99, 1000–1006. <https://doi.org/10.1017/S0007114507843558>
- Pierre, F., Taché, S., Petit, C.R., Van der Meer, R., Corpet, D.E., 2003. Meat and cancer: haemoglobin and haemin in a low-calcium diet promote colorectal carcinogenesis at the aberrant crypt stage in rats. *Carcinogenesis* 24, 1683–1690. <https://doi.org/10.1093/carcin/bgg130>
- Pizarro, F., Olivares, M., Hertrampf, E., Mazariegos, D.I., Arredondo, M., 2003. Heme-iron absorption is saturable by heme-iron dose in women. *J. Nutr.* 133, 2214–2217. <https://doi.org/10.1093/jn/133.7.2214>
- Platz, E.A., Willett, W.C., Colditz, G.A., Rimm, E.B., Spiegelman, D., Giovannucci, E., 2000. Proportion of colon cancer risk that might be preventable in a cohort of middle-aged US men. *Cancer Causes Control CCC* 11, 579–588. <https://doi.org/10.1023/a:1008999232442>
- Pluskal, T., Castillo, S., Villar-Briones, A., Orešič, M., 2010. MZmine 2: Modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinformatics* 11, 395. <https://doi.org/10.1186/1471-2105-11-395>
- Poesen, R., Mutsaers, H.A.M., Windey, K., van den Broek, P.H., Verweij, V., Augustijns, P., Kuypers, D., Jansen, J., Evenepoel, P., Verbeke, K., Meijers, B., Masereeuw, R., 2015. The Influence of Dietary Protein Intake on Mammalian Tryptophan and Phenolic Metabolites. *PloS One* 10, e0140820. <https://doi.org/10.1371/journal.pone.0140820>

- Pöhö, P., Lipponen, K., Bespalov, M.M., Sikanen, T., Kotiaho, T., Kostianen, R., 2019. Comparison of liquid chromatography-mass spectrometry and direct infusion microchip electrospray ionization mass spectrometry in global metabolomics of cell samples. *Eur. J. Pharm. Sci. Off. J. Eur. Fed. Pharm. Sci.* 138, 104991. <https://doi.org/10.1016/j.ejps.2019.104991>
- Polycyclic Aromatic Hydrocarbons in Food - Scientific Opinion of the Panel on Contaminants in the Food Chain, 2008. *EFSA J.* 6, 724. <https://doi.org/10.2903/j.efsa.2008.724>
- Pourchet, M., Debrauwer, L., Klanova, J., Price, E.J., Covaci, A., Caballero-Casero, N., Oberacher, H., Lamoree, M., Damont, A., Fenaille, F., Vlaanderen, J., Meijer, J., Krauss, M., Sarigiannis, D., Barouki, R., Le Bizec, B., Antignac, J.-P., 2020. Suspect and non-targeted screening of chemicals of emerging concern for human biomonitoring, environmental health studies and support to risk assessment: From promises to challenges and harmonisation issues. *Environ. Int.* 139, 105545. <https://doi.org/10.1016/j.envint.2020.105545>
- Powell, S.M., Zilz, N., Beazer-Barclay, Y., Bryan, T.M., Hamilton, S.R., Thibodeau, S.N., Vogelstein, B., Kinzler, K.W., 1992. APC mutations occur early during colorectal tumorigenesis. *Nature* 359, 235–237. <https://doi.org/10.1038/359235a0>
- Psychogios, N., Hau, D.D., Peng, J., Guo, A.C., Mandal, R., Bouatra, S., Sinelnikov, I., Krishnamurthy, R., Eisner, R., Gautam, B., Young, N., Xia, J., Knox, C., Dong, E., Huang, P., Hollander, Z., Pedersen, T.L., Smith, S.R., Bamforth, F., Greiner, R., McManus, B., Newman, J.W., Goodfriend, T., Wishart, D.S., 2011. The Human Serum Metabolome. *PLOS ONE* 6, e16957. <https://doi.org/10.1371/journal.pone.0016957>
- Quintanilla, I., López-Cerón, M., Jimeno, M., Cuatrecasas, M., Zabalza, M., Moreira, L., Alonso, V., Rodríguez de Miguel, C., Muñoz, J., Castellvi-Bel, S., Llach, J., Castells, A., Balaguer, F., Camps, J., Pellisé, M., 2019. Rectal Aberrant Crypt Foci in Humans Are Not Surrogate Markers for Colorectal Cancer Risk. *Clin. Transl. Gastroenterol.* 10, e00047. <https://doi.org/10.14309/ctg.0000000000000047>
- Raffin, S.B., Woo, C.H., Roost, K.T., Price, D.C., Schmid, R., 1974. Intestinal absorption of hemoglobin iron-heme cleavage by mucosal heme oxygenase. *J. Clin. Invest.* 54, 1344–1352. <https://doi.org/10.1172/JCI107881>
- Rafiemanesh, H., Mohammadian-Hafshejani, A., Ghoncheh, M., Sepehri, Z., Shamlou, R., Salehiniya, H., Towhidi, F., Makhsosi, B.R., 2016. Incidence and Mortality of Colorectal Cancer and Relationships with the Human Development Index across the World. *Asian Pac. J. Cancer Prev. APJCP* 17, 2465–2473.
- Rajagopalan, H., Bardelli, A., Lengauer, C., Kinzler, K.W., Vogelstein, B., Velculescu, V.E., 2002. RAF / RAS oncogenes and mismatch-repair status. *Nature* 418, 934–934. <https://doi.org/10.1038/418934a>
- Ramirez, T., Daneshian, M., Kamp, H., Bois, F.Y., Clench, M.R., Coen, M., Donley, B., Fischer, S.M., Ekman, D.R., Fabian, E., Guillou, C., Heuer, J., Hogberg, H.T., Jungnickel, H., Keun, H.C., Krennrich, G., Krupp, E., Luch, A., Noor, F., Peter, E., Riefke, B., Seymour, M., Skinner, N., Smirnova, L., Verheij, E., Wagner, S., Hartung, T., van Ravenzwaay, B., Leist, M., 2013. Metabolomics in toxicology and preclinical research. *ALTEX* 30, 209–225. <https://doi.org/10.14573/altex.2013.2.209>
- Rathahao, E., Peiro, G., Martins, N., Alary, J., Guéraud, F., Debrauwer, L., 2005. Liquid chromatography-multistage tandem mass spectrometry for the quantification of dihydroxynonene mercapturic acid (DHN-MA), a urinary end-metabolite of 4-hydroxynonenal. *Anal. Bioanal. Chem.* 381, 1532–1539. <https://doi.org/10.1007/s00216-005-3095-6>
- Rawla, P., Sunkara, T., Barsouk, A., 2019. Epidemiology of colorectal cancer: incidence, mortality, survival, and risk factors. *Przegląd Gastroenterol.* 14, 89–103. <https://doi.org/10.5114/pg.2018.81072>
- Rico, E., González, O., Blanco, M.E., Alonso, R.M., 2014. Evaluation of human plasma sample preparation protocols for untargeted metabolic profiles analyzed by UHPLC-ESI-TOF-MS. *Anal. Bioanal. Chem.* 406, 7641–7652. <https://doi.org/10.1007/s00216-014-8212-y>

- Robertson, D.G., Watkins, P.B., Reily, M.D., 2011. Metabolomics in toxicology: preclinical and clinical applications. *Toxicol. Sci. Off. J. Soc. Toxicol.* 120 Suppl 1, S146-170. <https://doi.org/10.1093/toxsci/kfq358>
- Roca, M., Alcoriza, M.I., Garcia-Cañaveras, J.C., Lahoz, A., 2021. Reviewing the metabolome coverage provided by LC-MS: Focus on sample preparation and chromatography-A tutorial. *Anal. Chim. Acta* 1147, 38–55. <https://doi.org/10.1016/j.aca.2020.12.025>
- Rodriguez-Bigas, M.A., Boland, C.R., Hamilton, S.R., Henson, D.E., Jass, J.R., Khan, P.M., Lynch, H., Perucho, M., Smyrk, T., Sobin, L., Srivastava, S., 1997. A National Cancer Institute Workshop on Hereditary Nonpolyposis Colorectal Cancer Syndrome: meeting highlights and Bethesda guidelines. *J. Natl. Cancer Inst.* 89, 1758–1762. <https://doi.org/10.1093/jnci/89.23.1758>
- Rodríguez-Carrasco, Y., Narváez, A., Izzo, L., Gaspari, A., Graziani, G., Ritieni, A., 2020. Biomonitoring of Enniatin B1 and Its Phase I Metabolites in Human Urine: First Large-Scale Study. *Toxins* 12, E415. <https://doi.org/10.3390/toxins12060415>
- Rosen Vollmar, A.K., Rattray, N.J.W., Cai, Y., Santos-Neto, Á.J., Deziel, N.C., Jukic, A.M.Z., Johnson, C.H., 2019. Normalizing Untargeted Periconceptional Urinary Metabolomics Data: A Comparison of Approaches. *Metabolites* 9. <https://doi.org/10.3390/metabo9100198>
- Roux, A., Lison, D., Junot, C., Heilier, J.-F., 2011. Applications of liquid chromatography coupled to mass spectrometry-based metabolomics in clinical chemistry and toxicology: A review. *Clin. Biochem., Mass Spectrometry in Laboratory Medicine* 44, 119–135. <https://doi.org/10.1016/j.clinbiochem.2010.08.016>
- Russo, A., Bazan, V., Iacopetta, B., Kerr, D., Soussi, T., Gebbia, N., TP53-CRC Collaborative Study Group, 2005. The TP53 colorectal cancer international collaborative study on the prognostic and predictive significance of p53 mutation: influence of tumor site, type of mutation, and adjuvant treatment. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* 23, 7518–7528. <https://doi.org/10.1200/JCO.2005.00.471>
- Sabatini, L., Barbieri, A., Indiveri, P., Mattioli, S., Violante, F.S., 2008. Validation of an HPLC–MS/MS method for the simultaneous determination of phenylmercapturic acid, benzylmercapturic acid and o-methylbenzyl mercapturic acid in urine as biomarkers of exposure to benzene, toluene and xylenes. *J. Chromatogr. B* 863, 115–122. <https://doi.org/10.1016/j.jchromb.2008.01.022>
- Sakai, E., Takahashi, H., Kato, S., Uchiyama, T., Hosono, K., Endo, H., Maeda, S., Yoneda, M., Taguri, M., Nakajima, A., 2011. Investigation of the prevalence and number of aberrant crypt foci associated with human colorectal neoplasm. *Cancer Epidemiol. Biomark. Prev. Publ. Am. Assoc. Cancer Res. Cosponsored Am. Soc. Prev. Oncol.* 20, 1918–1924. <https://doi.org/10.1158/1055-9965.EPI-11-0104>
- Samadder, N.J., Vierkant, R.A., Tillmans, L.S., Wang, A.H., Weisenberger, D.J., Laird, P.W., Lynch, C.F., Anderson, K.E., French, A.J., Haile, R.W., Potter, J.D., Slager, S.L., Smyrk, T.C., Thibodeau, S.N., Cerhan, J.R., Limburg, P.J., 2013. Associations Between Colorectal Cancer Molecular Markers and Pathways With Clinicopathologic Features in Older Women. *Gastroenterology* 145, 348-356.e2. <https://doi.org/10.1053/j.gastro.2013.05.001>
- Samraj, A.N., Bertrand, K.A., Luben, R., Khedri, Z., Yu, H., Nguyen, D., Gregg, C.J., Diaz, S.L., Sawyer, S., Chen, X., Eliassen, H., Padler-Karavani, V., Wu, K., Khaw, K.-T., Willett, W., Varki, A., 2018. Polyclonal human antibodies against glycans bearing red meat-derived non-human sialic acid N-glycolylneuraminic acid are stable, reproducible, complex and vary between individuals: Total antibody levels are associated with colorectal cancer risk. *PLoS ONE* 13, e0197464. <https://doi.org/10.1371/journal.pone.0197464>
- Samraj, A.N., Pearce, O.M.T., Läubli, H., Crittenden, A.N., Bergfeld, A.K., Banda, K., Gregg, C.J., Bingman, A.E., Secrest, P., Diaz, S.L., Varki, N.M., Varki, A., 2015. A red meat-derived glycan promotes inflammation and cancer progression. *Proc. Natl. Acad. Sci.* 112, 542–547. <https://doi.org/10.1073/pnas.1417508112>
- Santarelli, R.L., Vendevre, J.-L., Naud, N., Taché, S., Guéraud, F., Viau, M., Genot, C., Corpet, D.E., Pierre, F.H.F., 2010. Meat processing and colon carcinogenesis: cooked, nitrite-treated, and

- oxidized high-heme cured meat promotes mucin-depleted foci in rats. *Cancer Prev. Res. Phila. Pa* 3, 852–864. <https://doi.org/10.1158/1940-6207.CAPR-09-0160>
- Saric, J., Want, E.J., Duthaler, U., Lewis, M., Keiser, J., Shockcor, J.P., Ross, G.A., Nicholson, J.K., Holmes, E., Tavares, M.F.M., 2012. Systematic Evaluation of Extraction Methods for Multiplatform-Based Metabotyping: Application to the *Fasciola hepatica* Metabolome. *Anal. Chem.* 84, 6963–6972. <https://doi.org/10.1021/ac300586m>
- Sasso, A., Latella, G., 2018. Dietary components that counteract the increased risk of colorectal cancer related to red meat consumption. *Int. J. Food Sci. Nutr.* 69, 536–548. <https://doi.org/10.1080/09637486.2017.1393503>
- Sawa, T., Akaike, T., Kida, K., Fukushima, Y., Takagi, K., Maeda, H., 1998. Lipid peroxyl radicals from oxidized oils and heme-iron: implication of a high-fat diet in colon carcinogenesis. *Cancer Epidemiol. Prev. Biomark.* 7, 1007–1012.
- Schepens, M.A.A., Vink, C., Schonewille, A.J., Dijkstra, G., van der Meer, R., Bovee-Oudenhoven, I.M.J., 2011. Dietary heme adversely affects experimental colitis in rats, despite heat-shock protein induction. *Nutrition* 27, 590–597. <https://doi.org/10.1016/j.nut.2010.05.002>
- Schiffman, C., Petrick, L., Perttula, K., Yano, Y., Carlsson, H., Whitehead, T., Metayer, C., Hayes, J., Rappaport, S., Dudoit, S., 2019. Filtering procedures for untargeted LC-MS metabolomics data. *BMC Bioinformatics* 20, 334. <https://doi.org/10.1186/s12859-019-2871-9>
- Schneider, C., Tallman, K.A., Porter, N.A., Brash, A.R., 2001. Two Distinct Pathways of Formation of 4-Hydroxynonenal: MECHANISMS OF NONENZYMATIC TRANSFORMATION OF THE 9- AND 13-HYDROPEROXIDES OF LINOLEIC ACID TO 4-HYDROXYALKENALS \*. *J. Biol. Chem.* 276, 20831–20838. <https://doi.org/10.1074/jbc.M101821200>
- Scholz, K., Dekant, W., Völkel, W., Pähler, A., 2005. Rapid detection and identification of N-acetyl-L-cysteine thioethers using constant neutral loss and theoretical multiple reaction monitoring combined with enhanced product-ion scans on a linear ion trap mass spectrometer. *J. Am. Soc. Mass Spectrom.* 16, 1976–1984. <https://doi.org/10.1016/j.jasms.2005.08.003>
- Seiwert, N., Heylmann, D., Hasselwander, S., Fahrner, J., 2020. Mechanism of colorectal carcinogenesis triggered by heme iron from red meat. *Biochim. Biophys. Acta BBA - Rev. Cancer* 1873, 188334. <https://doi.org/10.1016/j.bbcan.2019.188334>
- Serreli, G., Deiana, M., 2020. Extra Virgin Olive Oil Polyphenols: Modulation of Cellular Pathways Related to Oxidant Species and Inflammation in Aging. *Cells* 9, E478. <https://doi.org/10.3390/cells9020478>
- Sesink, A.L., Termont, D.S., Kleibeuker, J.H., Van der Meer, R., 2001a. Red meat and colon cancer: dietary haem-induced colonic cytotoxicity and epithelial hyperproliferation are inhibited by calcium. *Carcinogenesis* 22, 1653–1659. <https://doi.org/10.1093/carcin/22.10.1653>
- Sesink, A.L., Termont, D.S., Kleibeuker, J.H., Van der Meer, R., 2001b. Red meat and colon cancer: dietary haem-induced colonic cytotoxicity and epithelial hyperproliferation are inhibited by calcium. *Carcinogenesis* 22, 1653–1659. <https://doi.org/10.1093/carcin/22.10.1653>
- Sesink, A.L., Termont, D.S., Kleibeuker, J.H., Van der Meer, R., 2001c. Red meat and colon cancer: dietary haem-induced colonic cytotoxicity and epithelial hyperproliferation are inhibited by calcium. *Carcinogenesis* 22, 1653–1659. <https://doi.org/10.1093/carcin/22.10.1653>
- Sesink, A.L.A., Termont, D.S.M.L., Kleibeuker, J.H., Meer, R.V. der, 1999. Red Meat and Colon Cancer: The Cytotoxic and Hyperproliferative Effects of Dietary Heme. *Cancer Res.* 59, 5704–5709.
- Sharma, A., Singh, K., Almasan, A., 2012. Histone H2AX phosphorylation: a marker for DNA damage. *Methods Mol. Biol. Clifton NJ* 920, 613–626. [https://doi.org/10.1007/978-1-61779-998-3\\_40](https://doi.org/10.1007/978-1-61779-998-3_40)
- Shin, A., Shrubsole, M.J., Ness, R.M., Wu, H., Sinha, R., Smalley, W.E., Shyr, Y., Zheng, W., 2007. Meat and meat-mutagen intake, doneness preference and the risk of colorectal polyps: the Tennessee Colorectal Polyp Study. *Int. J. Cancer* 121, 136–142. <https://doi.org/10.1002/ijc.22664>
- Silva, J.C., Denny, R., Dorschel, C.A., Gorenstein, M., Kass, I.J., Li, G.-Z., McKenna, T., Nold, M.J., Richardson, K., Young, P., Geromanos, S., 2005. Quantitative proteomic analysis by accurate mass retention time pairs. *Anal. Chem.* 77, 2187–2200. <https://doi.org/10.1021/ac048455k>

- Silva, R.M., Campanholo, V.M. de L.P., Paiotti, A.P.R., Artigiani Neto, R., Oshima, C.T.F., Ribeiro, D.A., Forones, N.M., 2015. Chemopreventive activity of grape juice concentrate (G8000TM) on rat colon carcinogenesis induced by azoxymethane. *Environ. Toxicol. Pharmacol.* 40, 870–875. <https://doi.org/10.1016/j.etap.2015.09.012>
- Simopoulos, A.P., 2008. The importance of the omega-6/omega-3 fatty acid ratio in cardiovascular disease and other chronic diseases. *Exp. Biol. Med.* Maywood NJ 233, 674–688. <https://doi.org/10.3181/0711-MR-311>
- Sinha, R., Kulldorff, M., Gunter, M.J., Strickland, P., Rothman, N., 2005. Dietary benzo[a]pyrene intake and risk of colorectal adenoma. *Cancer Epidemiol. Biomark. Prev. Publ. Am. Assoc. Cancer Res. Cosponsored Am. Soc. Prev. Oncol.* 14, 2030–2034. <https://doi.org/10.1158/1055-9965.EPI-04-0854>
- Sitnikov, D.G., Monnin, C.S., Vuckovic, D., 2016. Systematic Assessment of Seven Solvent and Solid-Phase Extraction Methods for Metabolomics Analysis of Human Plasma by LC-MS. *Sci. Rep.* 6, 38885. <https://doi.org/10.1038/srep38885>
- Siu, I.M., Pretlow, T.G., Amini, S.B., Pretlow, T.P., 1997. Identification of dysplasia in human colonic aberrant crypt foci. *Am. J. Pathol.* 150, 1805–1813.
- Smith, C.A., Want, E.J., O'Maille, G., Abagyan, R., Siuzdak, G., 2006. XCMS: Processing Mass Spectrometry Data for Metabolite Profiling Using Nonlinear Peak Alignment, Matching, and Identification. *Anal. Chem.* 78, 779–787. <https://doi.org/10.1021/ac051437y>
- Smith, K.J., Johnson, K.A., Bryan, T.M., Hill, D.E., Markowitz, S., Willson, J.K., Paraskeva, C., Petersen, G.M., Hamilton, S.R., Vogelstein, B., 1993. The APC gene product in normal and tumor cells. *Proc. Natl. Acad. Sci. U. S. A.* 90, 2846–2850.
- Smith, L., Villaret-Cazadamont, J., Claus, S.P., Canlet, C., Guillou, H., Cabaton, N.J., Ellero-Simatos, S., 2020. Important Considerations for Sample Collection in Metabolomics Studies with a Special Focus on Applications to Liver Functions. *Metabolites* 10, 104. <https://doi.org/10.3390/metabo10030104>
- Sødring, M., Gunnes, G., Paulsen, J.E., 2016. Detection and Characterization of Flat Aberrant Crypt Foci (Flat ACF) in the Novel A/J Min/+ Mouse. *Anticancer Res.* 36, 2745–2750.
- Sofi, F., Dinu, M., Pagliai, G., Pierre, F., Gueraud, F., Bowman, J., Gerard, P., Longo, V., Giovannelli, L., Caderni, G., de Filippo, C., 2019. Fecal microbiome as determinant of the effect of diet on colorectal cancer risk: comparison of meat-based versus pesco-vegetarian diets (the MeaTIC study). *Trials* 20, 688. <https://doi.org/10.1186/s13063-019-3801-x>
- Song, M., Chan, A.T., Sun, J., 2020. Influence of the Gut Microbiome, Diet, and Environment on Risk of Colorectal Cancer. *Gastroenterology* 158, 322–340. <https://doi.org/10.1053/j.gastro.2019.06.048>
- Sonnenburg, E.D., Smits, S.A., Tikhonov, M., Higginbottom, S.K., Wingreen, N.S., Sonnenburg, J.L., 2016. Diet-induced extinctions in the gut microbiota compound over generations. *Nature* 529, 212–215. <https://doi.org/10.1038/nature16504>
- Spicer, R., Salek, R.M., Moreno, P., Cañueto, D., Steinbeck, C., 2017. Navigating freely-available software tools for metabolomics analysis. *Metabolomics* 13, 106. <https://doi.org/10.1007/s11306-017-1242-7>
- SPICER, S.S., 1965. DIAMINE METHODS FOR DIFFERENTIATING MUCOSUBSTANCES HISTOCHEMICALLY. *J. Histochem. Cytochem.* 13, 211–234. <https://doi.org/10.1177/13.3.211>
- Steinberg, P., 2019. Red Meat-Derived Nitroso Compounds, Lipid Peroxidation Products and Colorectal Cancer. *Foods* 8. <https://doi.org/10.3390/foods8070252>
- Steppeler, C., Sødring, M., Egelandsdal, B., Kirkhus, B., Oostindjer, M., Alvseike, O., Gangsei, L.E., Hovland, E.-M., Pierre, F., Paulsen, J.E., 2017. Effects of dietary beef, pork, chicken and salmon on intestinal carcinogenesis in A/J Min/+ mice. *PLoS One* 12, e0176001. <https://doi.org/10.1371/journal.pone.0176001>
- Stojiljkovic, N., Paris, A., Garcia, P., Popot, M.-A., Bonnaire, Y., Tabet, J.-C., Junot, C., 2014. Evaluation of horse urine sample preparation methods for metabolomics using LC coupled to HRMS. *Bioanalysis* 6, 785–803. <https://doi.org/10.4155/bio.13.324>

- Sugimura, T., Wakabayashi, K., Nakagama, H., Nagao, M., 2004. Heterocyclic amines: Mutagens/carcinogens produced during cooking of meat and fish. *Cancer Sci.* 95, 290–299. <https://doi.org/10.1111/j.1349-7006.2004.tb03205.x>
- Sumner, L.W., Amberg, A., Barrett, D., Beale, M.H., Beger, R., Daykin, C.A., Fan, T.W.-M., Fiehn, O., Goodacre, R., Griffin, J.L., Hankemeier, T., Hardy, N., Harnly, J., Higashi, R., Kopka, J., Lane, A.N., Lindon, J.C., Marriott, P., Nicholls, A.W., Reily, M.D., Thaden, J.J., Viant, M.R., 2007. Proposed minimum reporting standards for chemical analysis. *Metabolomics* 3, 211–221. <https://doi.org/10.1007/s11306-007-0082-2>
- Surya, R., Héliers-Toussaint, C., Martin, O.C., Gauthier, T., Guéraud, F., Taché, S., Naud, N., Jouanin, I., Chantelauze, C., Durand, D., Joly, C., Pujos-Guillot, E., Pierre, F.H., Huc, L., 2016. Red meat and colorectal cancer: Nrf2-dependent antioxidant response contributes to the resistance of preneoplastic colon cells to fecal water of hemoglobin- and beef-fed rats. *Carcinogenesis* 37, 635–645. <https://doi.org/10.1093/carcin/bgw035>
- Suzui, M., Inamine, M., Kaneshiro, T., Morioka, T., Yoshimi, N., Suzuki, R., Kohno, H., Tanaka, T., 2005. Indole-3-carbinol inhibits the growth of human colon carcinoma cells but enhances the tumor multiplicity and volume of azoxymethane-induced rat colon carcinogenesis. *Int. J. Oncol.* 27, 1391–1399.
- Takayama, S., Nakatsuru, Y., Masuda, M., Ohgaki, H., Sato, S., Sugimura, T., 1984. Demonstration of carcinogenicity in F344 rats of 2-amino-3-methyl-imidazo[4,5-f]quinoline from broiled sardine, fried beef and beef extract. *Gan* 75, 467–470.
- Tangvoranuntakul, P., Gagneux, P., Diaz, S., Bardor, M., Varki, N., Varki, A., Muchmore, E., 2003. Human uptake and incorporation of an immunogenic nonhuman dietary sialic acid. *Proc. Natl. Acad. Sci. U. S. A.* 100, 12045–12050. <https://doi.org/10.1073/pnas.2131556100>
- Theodoridis, G., Gika, H.G., Wilson, I.D., 2011. Mass spectrometry-based holistic analytical approaches for metabolite profiling in systems biology studies. *Mass Spectrom. Rev.* 30, 884–906. <https://doi.org/10.1002/mas.20306>
- Toden, S., Bird, A.R., Topping, D.L., Conlon, M.A., 2007. High red meat diets induce greater numbers of colonic DNA double-strand breaks than white meat in rats: attenuation by high-amylose maize starch. *Carcinogenesis* 28, 2355–2362. <https://doi.org/10.1093/carcin/bgm216>
- Touvier, M., Kesse-Guyot, E., Méjean, C., Pollet, C., Malon, A., Castetbon, K., Hercberg, S., 2011. Comparison between an interactive web-based self-administered 24 h dietary record and an interview by a dietitian for large-scale epidemiological studies. *Br. J. Nutr.* 105, 1055–1064. <https://doi.org/10.1017/S0007114510004617>
- Touvier, M., Méjean, C., Kesse-Guyot, E., Pollet, C., Malon, A., Castetbon, K., Hercberg, S., 2010. Comparison between web-based and paper versions of a self-administered anthropometric questionnaire. *Eur. J. Epidemiol.* 25, 287–296. <https://doi.org/10.1007/s10654-010-9433-9>
- Tremblay-Franco, M., Cabaton, N.J., Canlet, C., Gautier, R., Schaeberle, C.M., Jourdan, F., Sonnenschein, C., Vinson, F., Soto, A.M., Zalko, D., 2015. Dynamic Metabolic Disruption in Rats Perinatally Exposed to Low Doses of Bisphenol-A. *PLoS One* 10, e0141698. <https://doi.org/10.1371/journal.pone.0141698>
- Tricker, A.R., 1997. N-nitroso compounds and man: sources of exposure, endogenous formation and occurrence in body fluids. *Eur. J. Cancer Prev. Off. J. Eur. Cancer Prev. Organ. ECP* 6, 226–268.
- Trygg, J., Wold, S., 2002. Orthogonal projections to latent structures (O-PLS). *J. Chemom.* 16, 119–128. <https://doi.org/10.1002/cem.695>
- Tsugawa, H., Cajka, T., Kind, T., Ma, Y., Higgins, B., Ikeda, K., Kanazawa, M., VanderGheynst, J., Fiehn, O., Arita, M., 2015. MS-DIAL: data-independent MS/MS deconvolution for comprehensive metabolome analysis. *Nat. Methods* 12, 523–526. <https://doi.org/10.1038/nmeth.3393>
- Turesky, R.J., 2018. Mechanistic evidence for red meat and processed meat intake and cancer risk: A follow-up on the International Agency for Research on Cancer Evaluation of 2015. *Chimia* 72, 718–724. <https://doi.org/10.2533/chimia.2018.718>

- Turesky, R.J., Le Marchand, L., 2011. Metabolism and Biomarkers of Heterocyclic Aromatic Amines in Molecular Epidemiology Studies: Lessons Learned from Aromatic Amines. *Chem. Res. Toxicol.* 24, 1169–1214. <https://doi.org/10.1021/tx200135s>
- Umar, A., Boland, C.R., Terdiman, J.P., Syngal, S., de la Chapelle, A., Rüschoff, J., Fishel, R., Lindor, N.M., Burgart, L.J., Hamelin, R., Hamilton, S.R., Hiatt, R.A., Jass, J., Lindblom, A., Lynch, H.T., Peltomaki, P., Ramsey, S.D., Rodriguez-Bigas, M.A., Vasen, H.F.A., Hawk, E.T., Barrett, J.C., Freedman, A.N., Srivastava, S., 2004. Revised Bethesda Guidelines for hereditary nonpolyposis colorectal cancer (Lynch syndrome) and microsatellite instability. *J. Natl. Cancer Inst.* 96, 261–268. <https://doi.org/10.1093/jnci/djh034>
- Václavík, J., Coene, K.L.M., Vrobel, I., Najdekr, L., Friedecký, D., Karlíková, R., Mádrová, L., Petsalo, A., Engelke, U.F.H., van Wegberg, A., Kluijtmans, L.A.J., Adam, T., Wevers, R.A., 2018. Structural elucidation of novel biomarkers of known metabolic disorders based on multistage fragmentation mass spectra. *J. Inherit. Metab. Dis.* 41, 407–414. <https://doi.org/10.1007/s10545-017-0109-4>
- Valavanidis, A., Vlachogianni, T., Fiotakis, C., 2009. 8-hydroxy-2'-deoxyguanosine (8-OHdG): A critical biomarker of oxidative stress and carcinogenesis. *J. Environ. Sci. Health Part C Environ. Carcinog. Ecotoxicol. Rev.* 27, 120–139. <https://doi.org/10.1080/10590500902885684>
- Valle, L., de Voer, R.M., Goldberg, Y., Sjursen, W., Försti, A., Ruiz-Ponte, C., Caldés, T., Garré, P., Olsen, M.F., Nordling, M., Castellvi-Bel, S., Hemminki, K., 2019. Update on genetic predisposition to colorectal cancer and polyposis. *Mol. Aspects Med., New insights on the molecular aspects of colorectal cancer* 69, 10–26. <https://doi.org/10.1016/j.mam.2019.03.001>
- van de Merbel, N.C., 2008. Quantitative determination of endogenous compounds in biological samples using chromatographic techniques. *TrAC Trends Anal. Chem., Advanced MS Analysis of Metabolites and Degradation Products - I* 27, 924–933. <https://doi.org/10.1016/j.trac.2008.09.002>
- van den Berg, J.W., Koole-Lesuis, R., Edixhoven-Bosdijk, A., Brouwers, N., 1988. Automating the quantification of heme in feces. *Clin. Chem.* 34, 2125–2126.
- van der Hooft, J.J.J., Padmanabhan, S., Burgess, K.E.V., Barrett, M.P., 2016. Urinary antihypertensive drug metabolite screening using molecular networking coupled to high-resolution mass spectrometry fragmentation. *Metabolomics Off. J. Metabolomic Soc.* 12, 125. <https://doi.org/10.1007/s11306-016-1064-z>
- Van Hecke, T., Jakobsen, L.M.A., Vossen, E., Guéraud, F., De Vos, F., Pierre, F., Bertram, H.C.S., De Smet, S., 2016. Short-term beef consumption promotes systemic oxidative stress, TMAO formation and inflammation in rats, and dietary fat content modulates these effects. *Food Funct.* 7, 3760–3771. <https://doi.org/10.1039/c6fo00462h>
- van Ravenzwaay, B., Cunha, G.C.-P., Leibold, E., Looser, R., Mellert, W., Prokoudine, A., Walk, T., Wiemer, J., 2007. The use of metabolomics for the discovery of new biomarkers of effect. *Toxicol. Lett., Biomarkers in Children and Adults* 172, 21–28. <https://doi.org/10.1016/j.toxlet.2007.05.021>
- Vasilopoulou, C.G., Sulek, K., Brunner, A.-D., Meitei, N.S., Schweiger-Hufnagel, U., Meyer, S.W., Barsch, A., Mann, M., Meier, F., 2020. Trapped ion mobility spectrometry and PASEF enable in-depth lipidomics from minimal sample amounts. *Nat. Commun.* 11, 331. <https://doi.org/10.1038/s41467-019-14044-x>
- Veettil, S.K., Wong, T.Y., Loo, Y.S., Playdon, M.C., Lai, N.M., Giovannucci, E.L., Chaiyakunapruk, N., 2021. Role of Diet in Colorectal Cancer Incidence. *JAMA Netw. Open* 4, e2037341. <https://doi.org/10.1001/jamanetworkopen.2020.37341>
- Vogl, F.C., Mehrl, S., Heizinger, L., Schlecht, I., Zacharias, H.U., Ellmann, L., Nürnberger, N., Gronwald, W., Leitzmann, M.F., Rossert, J., Eckardt, K.-U., Dettmer, K., Oefner, P.J., GCKD Study Investigators, 2016. Evaluation of dilution and normalization strategies to correct for urinary output in HPLC-HRTOFMS metabolomics. *Anal. Bioanal. Chem.* 408, 8483–8493. <https://doi.org/10.1007/s00216-016-9974-1>



- Wagner, S., Scholz, K., Donegan, M., Burton, L., Wingate, J., Völkel, W., 2006. Metabonomics and biomarker discovery: LC-MS metabolic profiling and constant neutral loss scanning combined with multivariate data analysis for mercapturic acid analysis. *Anal. Chem.* 78, 1296–1305. <https://doi.org/10.1021/ac051705s>
- Walther, A., Johnstone, E., Swanton, C., Midgley, R., Tomlinson, I., Kerr, D., 2009. Genetic prognostic and predictive markers in colorectal cancer. *Nat. Rev. Cancer* 9, 489–499. <https://doi.org/10.1038/nrc2645>
- Wang, X., Lin, X., Ouyang, Y.Y., Liu, J., Zhao, G., Pan, A., Hu, F.B., 2016. Red and processed meat consumption and mortality: dose-response meta-analysis of prospective cohort studies. *Public Health Nutr.* 19, 893–905. <https://doi.org/10.1017/S1368980015002062>
- Wang, X., Zhang, J., Huang, Q., Alamdar, A., Tian, M., Liu, L., Shen, H., 2015. Serum metabolomics analysis reveals impaired lipid metabolism in rats after oral exposure to benzo(a)pyrene. *Mol. Biosyst.* 11, 753–759. <https://doi.org/10.1039/c4mb00565a>
- Want, E.J., 2018. LC-MS Untargeted Analysis. *Methods Mol. Biol. Clifton NJ* 1738, 99–116. [https://doi.org/10.1007/978-1-4939-7643-0\\_7](https://doi.org/10.1007/978-1-4939-7643-0_7)
- Warrack, B.M., Hnatyshyn, S., Ott, K.-H., Reily, M.D., Sanders, M., Zhang, H., Drexler, D.M., 2009. Normalization strategies for metabonomic analysis of urine samples. *J. Chromatogr. B* 877, 547–552. <https://doi.org/10.1016/j.jchromb.2009.01.007>
- WCRF, W., 2018. World Cancer Research Fund/American Institute for Cancer Research. Continuous Update Project Expert Report 2018. Meat, fish and dairy products and the risk of cancer.
- Wedekind, R., Kiss, A., Keski-Rahkonen, P., Viallon, V., Rothwell, J.A., Cross, A.J., Rostgaard-Hansen, A.L., Sandanger, T.M., Jakszyn, P., Schmidt, J.A., Pala, V., Vermeulen, R., Schulze, M.B., Kühn, T., Johnson, T., Trichopoulou, A., Peppas, E., La Vecchia, C., Masala, G., Tumino, R., Sacerdote, C., Wittenbecher, C., de Magistris, M.S., Dahm, C.C., Severi, G., Mancini, F.R., Weiderpass, E., Gunter, M.J., Huybrechts, I., Scalbert, A., 2020. A metabolomic study of red and processed meat intake and acylcarnitine concentrations in human urine and blood. *Am. J. Clin. Nutr.* 112, 381–388. <https://doi.org/10.1093/ajcn/nqaa140>
- West, A.R., Oates, P.S., 2008. Mechanisms of heme iron absorption: Current questions and controversies. *World J. Gastroenterol. WJG* 14, 4101–4110. <https://doi.org/10.3748/wjg.14.4101>
- White, A.J., Bradshaw, P.T., Herring, A.H., Teitelbaum, S.L., Beyea, J., Stellman, S.D., Steck, S.E., Mordukhovich, I., Eng, S.M., Engel, L.S., Conway, K., Hatch, M., Neugut, A.I., Santella, R.M., Gammon, M.D., 2016. Exposure to multiple sources of polycyclic aromatic hydrocarbons and breast cancer incidence. *Environ. Int.* 89–90, 185–192. <https://doi.org/10.1016/j.envint.2016.02.009>
- Wikoff, W.R., Anfora, A.T., Liu, J., Schultz, P.G., Lesley, S.A., Peters, E.C., Siuzdak, G., 2009. Metabolomics analysis reveals large effects of gut microflora on mammalian blood metabolites. *Proc. Natl. Acad. Sci. U. S. A.* 106, 3698–3703. <https://doi.org/10.1073/pnas.0812874106>
- Winter, C.K., Segall, H.J., Jones, A.D., 1987. Distribution of trans-4-hydroxy-2-hexenal and tandem mass spectrometric detection of its urinary mercapturic acid in the rat. *Drug Metab. Dispos. Biol. Fate Chem.* 15, 608–612.
- Winter, J., Nyskohus, L., Young, G.P., Hu, Y., Conlon, M.A., Bird, A.R., Topping, D.L., Leu, R.K.L., 2011. Inhibition by Resistant Starch of Red Meat-Induced Promutagenic Adducts in Mouse Colon. *Cancer Prev. Res. (Phila. Pa.)* 4, 1920–1928. <https://doi.org/10.1158/1940-6207.CAPR-11-0176>
- Winter, J., Young, G.P., Hu, Y., Gratz, S.W., Conlon, M.A., Leu, R.K., 2014. Accumulation of promutagenic DNA adducts in the mouse distal colon after consumption of heme does not induce colonic neoplasms in the western diet model of spontaneous colorectal cancer. *Mol. Nutr. Food Res.* 58, 550–558. <https://doi.org/10.1002/mnfr.201300430>
- Winter, J.M., Hu, Y., Young, G.P., Kohonen-Corish, M.R.J., Leu, R.K., 2014. Role of Red Meat and Resistant Starch in Promutagenic Adduct Formation, MGMT Repair, Thymic Lymphoma and

- Intestinal Tumourigenesis in Msh2 -Deficient Mice. *J. Nutr.* 7, 299–313. <https://doi.org/10.1159/000381675>
- Wiseman, M., 2008. The second World Cancer Research Fund/American Institute for Cancer Research expert report. Food, nutrition, physical activity, and the prevention of cancer: a global perspective. *Proc. Nutr. Soc.* 67, 253–256. <https://doi.org/10.1017/S002966510800712X>
- Wishart, D.S., Feunang, Y.D., Marcu, A., Guo, A.C., Liang, K., Vázquez-Fresno, R., Sajed, T., Johnson, D., Li, C., Karu, N., Sayeeda, Z., Lo, E., Assempour, N., Berjanskii, M., Singhal, S., Arndt, D., Liang, Y., Badran, H., Grant, J., Serra-Cayuela, A., Liu, Y., Mandal, R., Neveu, V., Pon, A., Knox, C., Wilson, M., Manach, C., Scalbert, A., 2018. HMDB 4.0: the human metabolome database for 2018. *Nucleic Acids Res.* 46, D608–D617. <https://doi.org/10.1093/nar/gkx1089>
- Wishart, D.S., Knox, C., Guo, A.C., Eisner, R., Young, N., Gautam, B., Hau, D.D., Psychogios, N., Dong, E., Bouatra, S., Mandal, R., Sinelnikov, I., Xia, J., Jia, L., Cruz, J.A., Lim, E., Sobsey, C.A., Shrivastava, S., Huang, P., Liu, P., Fang, L., Peng, J., Fradette, R., Cheng, D., Tzur, D., Clements, M., Lewis, A., De Souza, A., Zuniga, A., Dawe, M., Xiong, Y., Clive, D., Greiner, R., Nazyrova, A., Shaykhtudinov, R., Li, L., Vogel, H.J., Forsythe, I., 2009. HMDB: a knowledgebase for the human metabolome. *Nucleic Acids Res.* 37, D603-610. <https://doi.org/10.1093/nar/gkn810>
- Wolk, A., 2017. Potential health hazards of eating red meat. *J. Intern. Med.* 281, 106–122. <https://doi.org/10.1111/joim.12543>
- Wondrak, G.T., Villeneuve, N.F., Lamore, S.D., Bause, A.S., Jiang, T., Zhang, D.D., 2010. The Cinnamon-Derived Dietary Factor Cinnamic Aldehyde Activates the Nrf2-Dependent Antioxidant Response in Human Epithelial Colon Cells. *Molecules* 15, 3338–3355. <https://doi.org/10.3390/molecules15053338>
- World Cancer Research Fund International., 2018. Diet, Nutrition, Physical Activity and Cancer: a Global Perspective - The Third Expert Report. *World Cancer Res. Fund Int.* <https://www.wcrf.org/dietandcancer>.
- Wyness, L., 2016. The role of red meat in the diet: nutrition and health benefits. *Proc. Nutr. Soc.* 75, 227–232. <https://doi.org/10.1017/S0029665115004267>
- Xi, B., Gu, H., Baniyasadi, H., Raftery, D., 2014. Statistical Analysis and Modeling of Mass Spectrometry-Based Metabolomics Data, in: Raftery, D. (Ed.), *Mass Spectrometry in Metabolomics: Methods and Protocols*, Methods in Molecular Biology. Springer, New York, NY, pp. 333–353. [https://doi.org/10.1007/978-1-4939-1258-2\\_22](https://doi.org/10.1007/978-1-4939-1258-2_22)
- Xia, J., Wishart, D.S., 2016. Using MetaboAnalyst 3.0 for Comprehensive Metabolomics Data Analysis. *Curr. Protoc. Bioinforma.* 55, 14.10.1-14.10.91. <https://doi.org/10.1002/cpbi.11>
- Xia, W., Liu, G., Shao, Z., Xu, E., Yuan, H., Liu, J., Gao, L., 2020. Toxicology of tramadol following chronic exposure based on metabolomics of the cerebrum in mice. *Sci. Rep.* 10, 11130. <https://doi.org/10.1038/s41598-020-67974-8>
- Xiao, M., Xiao, S., Straaten, T. van der, Xue, P., Zhang, G., Zheng, X., Zhang, Q., Cai, Y., Jin, C., Yang, J., Wu, S., Zhu, G., Lu, X., 2016. Genetic polymorphisms in 19q13.3 genes associated with alteration of repair capacity to BPDE-DNA adducts in primary cultured lymphocytes. *Mutat. Res.* 812, 39–47. <https://doi.org/10.1016/j.mrgentox.2016.10.004>
- Yanes, O., Tautenhahn, R., Patti, G.J., Siuzdak, G., 2011. Expanding Coverage of the Metabolome for Global Metabolite Profiling. *Anal. Chem.* 83, 2152–2161. <https://doi.org/10.1021/ac102981k>
- Yang, Y., Cruickshank, C., Armstrong, M., Mahaffey, S., Reisdorph, R., Reisdorph, N., 2013. New sample preparation approach for mass spectrometry-based profiling of plasma results in improved coverage of metabolome. *J. Chromatogr. A* 1300, 217–226. <https://doi.org/10.1016/j.chroma.2013.04.030>
- Yiannakou, I., Barber, L.E., Li, S., Adams-Campbell, L.L., Palmer, J.R., Rosenberg, L., Petrick, J.L., 2021. A Prospective Analysis of Red and Processed Meat Intake in Relation to Colorectal Cancer in the Black Women’s Health Study. *J. Nutr.* nxab419. <https://doi.org/10.1093/jn/nxab419>
- Zaramela, L.S., Martino, C., Alisson-Silva, F., Rees, S.D., Diaz, S.L., Chuzel, L., Ganatra, M.B., Taron, C.H., Secret, P., Zuñiga, C., Huang, J., Siegel, D., Chang, G., Varki, A., Zengler, K., 2019. Gut bacteria responding to dietary change encode sialidases that exhibit preference for red meat-

- associated carbohydrates. *Nat. Microbiol.* 4, 2082–2089. <https://doi.org/10.1038/s41564-019-0564-9>
- Zhang, T., Watson, D.G., 2015. A short review of applications of liquid chromatography mass spectrometry based metabolomics techniques to the analysis of human urine. *Analyst* 140, 2907–2915. <https://doi.org/10.1039/C4AN02294G>
- Zhao, X., Liu, H., Zhou, X., Chen, X., Hu, N., Zhang, Y., Wang, S., 2021. 2-Amino-1-methyl-6-phenylimidazo[4,5-b]pyridine Induced Colon Injury by Disrupting the Intestinal Bacterial Composition and Lipid Metabolic Pathways in Rats. *J. Agric. Food Chem.* 69, 437–446. <https://doi.org/10.1021/acs.jafc.0c06588>
- Zheng, F., Zhao, X., Zeng, Z., Wang, L., Lv, W., Wang, Q., Xu, G., 2020. Development of a plasma pseudotargeted metabolomics method based on ultra-high-performance liquid chromatography-mass spectrometry. *Nat. Protoc.* 15, 2519–2537. <https://doi.org/10.1038/s41596-020-0341-5>
- Zheng, X., Xie, G., Zhao, A., Zhao, L., Yao, C., Chiu, N.H.L., Zhou, Z., Bao, Y., Jia, Weiping, Nicholson, J.K., Jia, Wei, 2011. The Footprints of Gut Microbial–Mammalian Co-Metabolism. *J. Proteome Res.* 10, 5512–5522. <https://doi.org/10.1021/pr2007945>
- Zhong, V.W., Van Horn, L., Greenland, P., Carnethon, M.R., Ning, H., Wilkins, J.T., Lloyd-Jones, D.M., Allen, N.B., 2020. Associations of Processed Meat, Unprocessed Red Meat, Poultry, or Fish Intake With Incident Cardiovascular Disease and All-Cause Mortality. *JAMA Intern. Med.* 180, 503–512. <https://doi.org/10.1001/jamainternmed.2019.6969>
- Zhou, L., Zahid, M., Anwar, M.M., Pennington, K.L., Cohen, S.M., Wisecarver, J.L., Shostrom, V., Mirvish, S.S., 2016. Suggestive evidence for the induction of colonic aberrant crypts in mice fed sodium nitrite. *Nutr. Cancer* 68, 105–112. <https://doi.org/10.1080/01635581.2016.1102298>
- Zhu, X., Chen, Y., Subramanian, R., 2014. Comparison of information-dependent acquisition, SWATH, and MS(All) techniques in metabolite identification study employing ultrahigh-performance liquid chromatography-quadrupole time-of-flight mass spectrometry. *Anal. Chem.* 86, 1202–1209. <https://doi.org/10.1021/ac403385y>
- Zhu, Y., Lin, X., Li, H., Li, Y., Shi, X., Zhao, F., Xu, X., Li, C., Zhou, G., 2016. Intake of Meat Proteins Substantially Increased the Relative Abundance of Genus *Lactobacillus* in Rat Feces. *PLOS ONE* 11, e0152678. <https://doi.org/10.1371/journal.pone.0152678>
- Zhu, Y., Lin, X., Zhao, F., Shi, X., Li, H., Li, Y., Zhu, W., Xu, X., Lu, C., Zhou, G., 2015. Meat, dairy and plant proteins alter bacterial composition of rat gut bacteria. *Sci. Rep.* 5, 15220. <https://doi.org/10.1038/srep15220>
- Zou, H., Hastie, T., 2005. Regularization and variable selection via the elastic net. *J. R. Stat. Soc. Ser. B Stat. Methodol.* 67, 301–320. <https://doi.org/10.1111/j.1467-9868.2005.00503.x>